Probabilidad

Un enfoque actuarial

Antonio V. González Fragoso

Probabilidad

Un enfoque actuarial

Antonio V. González Fragoso

Universidad de las Américas Puebla

Probabilidad. Un enfoque actuarial / Antonio Vicente González Fragoso. San Andrés Cholula: Fundación Universidad de las Américas Puebla, 2025.
p. 544
ISBN 978-607-26992-2-9
1. Probabilidades. 2. Probabilidad y estadística. 3. Probabilidades—Problemas, ejercicios, etc. I. González Fragoso, Antonio Vicente.
QA273 G66 2024

Ex hacienda Santa Catarina Mártir s/n
San Andrés Cholula, Puebla, México, C. P. 72810
Tel.: +52 222 229 20 00
www.udlap.mx · editorial.udlap@udlap.mx

ISBN: 978-607-26992-2-9
Primera edición: agosto de 2025

Ilustración de portada: Gerardo Arizmendi Echegaray

Edición a cargo del Departamento de Publicaciones de la UDLAP.
Diseño editorial y portada: Willy Daniel Sepúlveda Juárez
Corrección de estilo: Andrea Garza Carbajal
Ilustraciones de interiores: Angélica Cabañas Ramírez

Este libro fue sometido a un proceso de revisión por pares.

Agradecimientos:

A la Universidad de las Américas Puebla, por darme la oportunidad de pertenecer a su equipo académico y por su apoyo para escribir este libro.

A mis padres, Catalina y Rafael, por su gran cariño, por la educación que me brindaron y por sus grandes enseñanzas.

A mi esposa Angélica, por su comprensión, por su apoyo incondicional en todo momento y por estar siempre conmigo.

A mis hijos Angélica, Antonio y Pablo, por el apoyo y motivación que me brindaron.

A Andrea, Mariel y Manolo, por su apoyo.

A Isabel, por todas sus alegrías.

A mis profesores, que me motivaron, me enseñaron y contribuyeron a mi formación, en especial a Antonio Herrera Arévalo, Alejandro Mina Valdés, Víctor Pérez Abreu Carrión, Federico O 'Reilly Tongo y José Luis Fernández Muñiz.

A mis alumnos, por sus constantes comentarios, dudas y observaciones, su retroalimentación fue vital para mejorar este escrito, en especial a Alejandra Quintos Lima, por su constante participación desde el primer día.

A mis colegas de trabajo, por su orientación y sus observaciones, en especial a Francisco García Castillo y Gerardo Arizmendi Echegaray, que contribuyeron en forma importante a mejorar algunos temas.

A mis revisores, por sus observaciones, sugerencias y correcciones que me ayudaron a mejorar cada capítulo.

Índice general

Prefacio

El libro *Probabilidad. Un enfoque actuarial* cumple con varios propósitos, primero como apoyo para estudiantes de licenciatura de Actuaría, Ciencia de Datos y Matemáticas, que cursan por lo general dentro de sus programas, dos o tres materias de probabilidad. También, es de gran utilidad como libro de consulta, para estudiantes de otras licenciaturas o de algunos posgrados, donde tienen como fundamento algunos temas de probabilidad. Por otro lado, el material es de gran apoyo para aquellos estudiantes cuyo interés sea prepararse para los exámenes de The Society of Actuaries (SOA).

La razón de este proyecto surgió del autor después de haber impartido por varios años, en forma continua, los cursos de probabilidad correspondientes a la licenciatura en Actuaría en la Universidad de las Américas Puebla. El autor se dio cuenta de que los temas de probabilidad pueden ser explicados sin perder la formalidad de las matemáticas, con determinado detalle, orden y estructura, y en ocasiones ser apoyados con ilustraciones, gráficas y tablas, contribuyendo de esta manera a una mejor comprensión por parte del lector.

Uno de los principales objetivos es proporcionar al lector bases sólidas de probabilidad y que estas puedan ser un fundamento para otras áreas de estudio, que muy probablemente el estudiante enfrentará, como la inferencia estadística, las matemáticas actuariales y la simulación estocástica. Además, el material es la base para comprender varios temas de finanzas y de administración de riesgos, por mencionar algunas áreas.

En este libro, el lector notará que, en diversos resultados de probabilidad, se da una interpretación en lenguaje estadístico, con la finalidad de hacer una conexión entre estos dos campos de estudio. En este sentido, también se presenta una sección de aplicaciones de la probabilidad a diferentes problemas de inferencia estadística. De esta manera, el estudiante empezará a comprender la importancia que tiene la probabilidad en la teoría estadística.

Con el propósito de apoyar a diferentes áreas de la actuaría, se desarrollaron algunos temas que tradicionalmente en otros libros no se contemplan o no se explican con detalle, como la mezcla de dos o más distribuciones, las distribuciones truncadas y las distribuciones de cola pesada, considerando también algunos modelos continuos que pueden ser ejemplos de estas últimas, como las distribuciones de Weibull, lognormal y de Pareto. Además, el lector podrá observar que algunos ejemplos y ejercicios para el estudiante, están desarrollados con un enfoque actuarial, principalmente, en el área de seguros y finanzas.

Como ya se señaló, el material de este libro cumple con otro propósito, la teoría, los ejemplos resueltos y el nivel de los ejercicios propuestos son un gran apoyo para aquellos estudiantes que tengan interés en presentar los exámenes de The Society of Actuaries (SOA), principalmente el examen de Probabilidad (Exam P).

Por todo lo anterior, el autor creyó conveniente que el nombre de este libro fuera: *Probabilidad. Un enfoque actuarial*.

El lector podrá observar que, en algunos temas, después de haberse realizado algunos cálculos de probabilidad, de cuantiles o de algunas simulaciones, se dan a conocer las indicaciones para realizar los cálculos por medio del *software* R. De esta manera, se motiva al lector a utilizar R para sus futuros proyectos de estudio o trabajo. En este sentido, es importante también destacar que todas las gráficas, diagramas y simulaciones que se exponen en los capítulos fueron realizadas por medio de R.

Las bases matemáticas que el estudiante debe tener, para leer y estudiar con cierta fluidez los diferentes temas de este libro, son, principalmente, teoría de conjuntos; solución de ecuaciones, lineales y no lineales. En general: álgebra elemental y solución de algunas sumas y series. En particular: series de Taylor; límites de sucesiones y de funciones; derivadas, tanto de una variable como derivadas parciales, e integrales, tanto simples como múltiples.

Se sugiere al lector que les dé seguimiento a las demostraciones de los teoremas, corolarios y algunos ejemplos que se presentan. La finalidad es que logre un mejor dominio sobre los diferentes temas y que pueda resolver diversos ejercicios teóricos, que siempre se presentan al final de cada capítulo. En general, el estudiante va a lograr una mejor madurez y dominio en matemáticas si le da importancia a este tipo de ejercicios.

El libro está conformado por 7 capítulos, y en cada uno se presenta el material correspondiente, apoyado con diversos ejemplos, tablas, gráficas o simulaciones, con la idea de lograr una lectura fluida y comprender mejor cada tema. Al final de cada capítulo, hay una lista de ejercicios propuestos para el estudiante con la idea de fortalecer sus conocimientos. Las soluciones se presentan al final del libro en un apéndice.

Al principio del capítulo 1 se abordan algunos temas preliminares como la teoría de conjuntos y la definición de función indicadora. Después, se dan los primeros conceptos de probabilidad, presentándose la definición de probabilidad (axiomas de Kolmogorov) y sus principales propiedades básicas. Se muestran los primeros cálculos de probabilidad, con apoyo de las técnicas de conteo. Se expone la definición de probabilidad condicional, y a partir de esta, se presenta la regla de multiplicación de probabilidades. Posteriormente, se exponen los teoremas de probabilidad total y teorema de Bayes. Finalmente, se muestra la definición de eventos independientes, explicando su importancia.

Posteriormente, en el capítulo 2 se da primero la definición de variable aleatoria, y a partir de esta, se presenta la definición de función de distribución. En seguida, para cada tipo de variables aleatorias, se presentan los conceptos: función de densidad, valor esperado, momentos y función generadora momentos. También, se da a conocer el significado de un cuantil y cómo se calcula, además, se presentan los conceptos de mediana y de moda. Después, se exponen las desigualdades: de Chebyshev y de Jensen. Finalmente, se da a conocer lo que es una mezcla de distribuciones, una distribución de cola pesada y cómo se calcula una distribución truncada, las cuales son de gran importancia en temas de actuaría.

A continuación, en los capítulos 3 y 4 se dan a conocer, respectivamente, las distribuciones discretas y continuas más comunes que existen en la literatura. Cabe mencionar, que la importancia de estos modelos se debe a la aplicación que tienen en la vida real. Para cada distribución, se presentan las definiciones y propiedades de más relevancia y en cada caso, se muestran diversos ejemplos, dejando en claro la importancia que tienen para ser aplicadas en diversas áreas.

En el capítulo 5 se presentan las definiciones y propiedades correspondientes a los vectores aleatorios. Se dan a conocer las definiciones de funciones de distribución y de densidad conjuntas. A partir de la función de densidad, se explica la forma de obtener las funciones de densidad marginales y condicionales. Posteriormente, se tratan los valores esperados, destacando las medias y las varianzas individuales, así como la función generadora de momentos conjunta. Se exponen los conceptos de covarianza y coeficiente de correlación, explicando la utilidad que tienen en varias aplicaciones. Después, se da la definición de variables aleatorias independientes y se explica la importancia que tiene en diferentes áreas. Finalmente, se presentan las distribuciones trinomial y normal bivariada con sus principales propiedades.

Después, en el capítulo 6, se exponen los diferentes métodos para encontrar la distribución de una función de una variable o vector aleatorio. Se explica en qué consiste cada método, presentándose ejemplos con diversos grados de dificultad. Se explica la importancia de la solución de este tipo de problemas en diversas áreas de aplicación, como la estadística, simulación y la matemática actuarial. Posteriormente, se presentan las distribuciones de las estadísticas de orden, explicando la importancia que tienen en diversas aplicaciones. Con la finalidad de conectar la probabilidad con la estadística, se considera una sección de aplicaciones de la probabilidad a la estadística.

En el último capítulo, se da una introducción a la convergencia de sucesiones de variables aleatorias. Se presentan los diferentes modos de convergencia y las relaciones que existen entre ellos. Se da a conocer las dos versiones de la ley de los grandes números: la débil y la fuerte. Se muestra la importancia que tiene la función generadora de momentos en varios resultados límites. Se presenta y se demuestra el teorema de límite central, explicando la importancia que tiene este resultado en diferentes temas, tanto de probabilidad como de estadística. Finalmente, se dan a conocer diferentes teoremas de convergencia, los cuales son de gran apoyo para la solución de varios problemas.

Como apoyo al lector, se consideraron varios apéndices. En el apéndice A se presentan las tablas de las distribuciones: normal estándar, ji cuadrada, t de student y distribución F, donde se muestran diferentes probabilidades y cuantiles. En el apéndice B se muestran las tablas de las diferentes distribuciones discretas y continuas, donde para cada modelo se exponen la función de densidad, el valor esperado, la varianza y la función generadora de momentos. Después, en el apéndice C se exponen, algunos resultados de matemáticas que son importantes para determinados temas de probabilidad. Luego, en el apéndice D, se presenta una pequeña introducción a la teoría de la medida, para aquellos lectores que les interese el tema y lo puedan relacionar con diferentes temas de probabilidad. En el apéndice E, se presentan algunos conceptos básicos de seguros, finanzas y administración de riesgos, con la idea de familiarizar al lector con estos temas. Finalmente, en el apéndice F, se muestran las soluciones de cada uno de los ejercicios propuestos en los diferentes capítulos.

Antonio V. González Fragoso
San Andrés Cholula, Puebla

Introducción
a la probabilidad

Capítulo

Introducción

En este capítulo, se presentarán los primeros conceptos y las propiedades más elementales de la probabilidad, siendo, en conjunto, los fundamentos para los temas que más adelante se tratarán.

Se empezará con un repaso a la teoría de conjuntos, dando las definiciones y propiedades más importantes. Cabe mencionar que los conjuntos son parte fundamental para el desarrollo de la probabilidad, el cálculo de una probabilidad se realiza sobre un evento, y el evento es un conjunto.

Posteriormente, se introducirá la definición de función indicadora y sus principales propiedades, también se comentará y se ejemplificará la importancia que tiene esta función para describir otras funciones de una manera práctica, en particular, será de gran utilidad para describir varias de las funciones que se utilizan en probabilidad, principalmente, las de densidad y de distribución.

En materia de probabilidad, se empezará explicando las diferentes formas en que algunos autores describen lo que es la aleatoriedad. Luego se expondrán algunas «definiciones» o interpretaciones tradicionales de la probabilidad, explicando sus ventajas y desventajas. Posteriormente, se expondrán los primeros conceptos de la probabilidad, presentando sus axiomas (la definición de probabilidad). Después, se presentarán y se demostrarán sus propiedades más importantes.

Una vez aclarados los fundamentos de la probabilidad, se realizarán los primeros cálculos de probabilidad, primero para espacios muestrales finitos y equiprobables, apoyándose por lo general del cálculo combinatorio.

Posteriormente, se introducirá la definición de probabilidad condicional, aclarando la importancia que tiene este concepto en diversas áreas de aplicación, entre las más importantes, en procesos estocásticos. También se explicará que es la base para otras propiedades de probabilidad que se tratarán posteriormente, como la regla de multiplicación de probabilidades, el teorema de la probabilidad total y el teorema de Bayes. Este último teorema cobra gran importancia en varias aplicaciones y en particular es el fundamento para la teoría de la estadística bayesiana. Cada uno de estos conceptos será ilustrado con diversos ejemplos y ejercicios.

Finalmente, se presentará la definición de independencia de dos o más eventos y se comentará la importancia que tiene este concepto en diversos temas de probabilidad, además, en diversas áreas de aplicación, en particular, en algunos temas de inferencia estadística, y en varios resultados de la matemática actuarial.

1.1. Teoría de conjuntos

En esta primera sección serán presentados los conceptos y las propiedades más importantes en lo que se refiere al tema de conjuntos. Esta teoría es de suma importancia para el desarrollo de la teoría de probabilidad, ya que los eventos o sucesos usados en probabilidad son conjuntos. Se sugiere que el lector domine, tanto en forma práctica como en forma teórica, la teoría de conjuntos, antes de comenzar con los temas de probabilidad.

Definición 1.1. Un *conjunto* es una colección de objetos, donde cada uno es llamado *elemento*.

Las letras mayúsculas $A, B, C, ..., Z$ denotarán conjuntos, las letras minúsculas $a, b, c, ..., z$ elementos. Por otro lado, $a \in A$ significa, a está en A o a pertenece a A y $a \notin A$, significa a no está en A o a no pertenece a A.

El conjunto *universal o universo* es el conjunto que contiene a todos los elementos y es denotado por Ω. El conjunto *vacío* es denotado por $\emptyset$.

Definición 1.2. Si cada elemento de A también está en B, se dice que A está *contenido* en B o que A es *subconjunto* de B, y se denota como $A \subset B$.

En diagrama de Venn, A subconjunto de B queda representado de la siguiente manera:

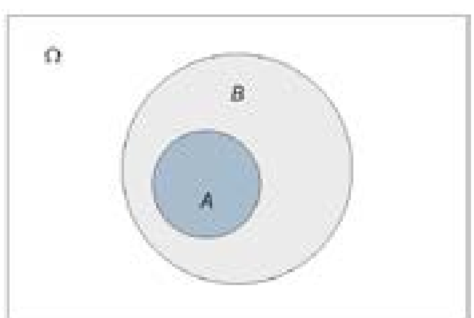


Figura 1.1. *A subconjunto de B*

Definición 1.3. Dos conjuntos A y B son *iguales*, si $A \subset B$ y $B \subset A$, y se denota como $A = B$.

A continuación, se presentarán las operaciones básicas de dos o más conjuntos.

Definición 1.4. La *unión* de dos conjuntos A y B es el conjunto formado por todos los elementos que están en A o que están en B. Esta operación es denotada como $A \cup B$.

Dicho de otra manera,

$$A \cup B = \{x \in \Omega | x \in A \text{ } o \text{ } x \in B\}.$$

Observación. El *o*, en este caso, es en un sentido inclusivo. Esto es, el *o* considera también la ocurrencia de ambas situaciones.

Es posible generalizar esta definición para un número finito de conjuntos,

$$\bigcup_{i=1}^{n} A_i = \{x \in \Omega | x \in A_i \text{ para alguna } i \in \{1, 2, ..., n\}\},$$

y también se puede generalizar para un número infinito de conjuntos,

$$\bigcup_{i=1}^{\infty} A_i = \{x \in \Omega | x \in A_i \text{ para alguna } i \in \{1, 2, ...\}\}.$$

La unión de los conjuntos A y B es la parte rayada en el siguiente diagrama

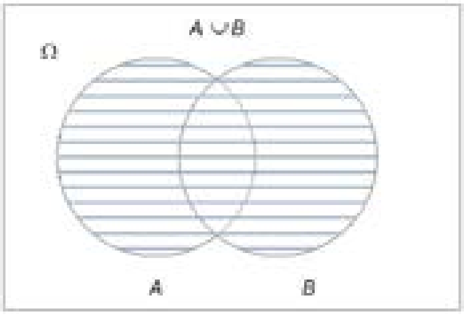


Figura 1.2. *Unión de A y B*

Definición 1.5. La *intersección* de dos conjuntos A y B es el conjunto que consta de todos los elementos que están en A y que están en B. Este conjunto es denotado como $A \cap B$.

Esto es,

$$A \cap B = \{x \in \Omega | x \in A \text{ } y \text{ } x \in B\}.$$

Esta operación se puede generalizar para un número finito de conjuntos,

$$\bigcap_{i=1}^{n} A_i = \{x \in \Omega | x \in A_i \text{ para toda } i \in \{1, 2, ..., n\}\},$$

y para un número infinito de conjuntos,

$$\bigcap_{i=1}^{\infty} A_i = \{x \in \Omega | x \in A_i \text{ para toda } i \in \{1, 2, ...\}\}.$$

La intersección de los conjuntos A y B es la zona rayada en el siguiente diagrama

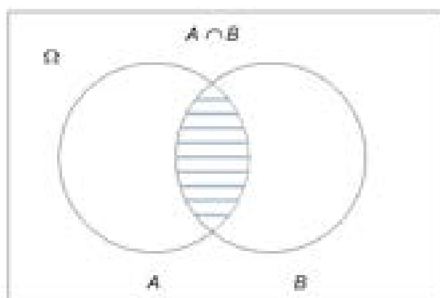


Figura 1.3. *Intersección de A y B*

Definición 1.6. El *complemento* del conjunto A se define como el conjunto que tiene como elementos aquellos que están en el conjunto universal Ω y que no están en A. Este conjunto se denota como A^c.

Así que

$$A^c = \{x \in \Omega | x \notin A\}.$$

En diagrama de Venn, el complemento de A es la parte sombreada

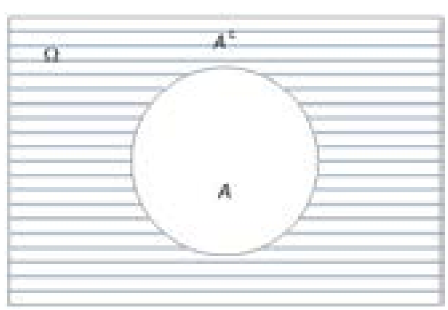


Figura 1.4. *Complemento de A*

Definición 1.7. La *diferencia* del conjunto A menos el conjunto B es el conjunto formado por todos los elementos que pertenecen a A pero que no pertenecen a B. Este conjunto se denota como $A - B$.

Esto es,

$$A - B = \{x \in \Omega | x \in A \text{ y } x \notin B\}.$$

La diferencia del conjunto A menos el conjunto B es la parte rayada en el siguiente diagrama

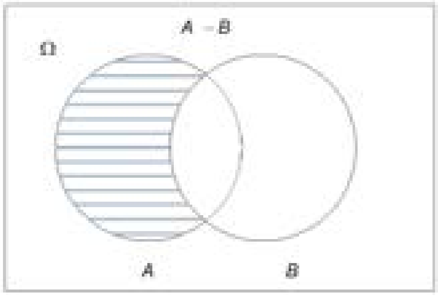


Figura 1.5. *Diferencia de A menos B*

Obsérvese que

$$A - B = A \cap B^c.$$

Definición 1.8. La *diferencia simétrica* del conjunto A y del conjunto B es el conjunto formado por todos los elementos que están en A pero no están en B, o todos los elementos que están en B pero no están en A. Este conjunto se denota como $A \triangle B$.

Dicho de otra manera,

$$A \triangle B = (A - B) \cup (B - A) = (A \cap B^c) \cup (B \cap A^c).$$

Obsérvese que la diferencia simétrica de dos conjuntos es una operación conmutativa.

En el siguiente diagrama, la parte rayada representa la diferencia simétrica de los conjuntos A y B

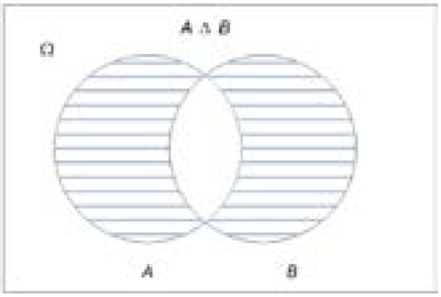


Figura 1.6. *Diferencia simétrica de A y B*

Definición 1.9. Se dice que dos conjuntos A y B son *mutuamente excluyentes (ajenos o disjuntos)* si $A \cap B = \emptyset$. Por otro lado, se dice que $A_1, A_2, \ldots, A_n, \ldots$ son eventos mutuamente excluyentes (ajenos o disjuntos), si para cada pareja de conjuntos A_i y A_j, $A_i \cap A_j = \emptyset$, donde $i \neq j$.

En otras palabras, dos o más conjuntos son mutuamente excluyentes si no existen elementos en común.

En diagrama de Venn, dos conjuntos A y B son ajenos si

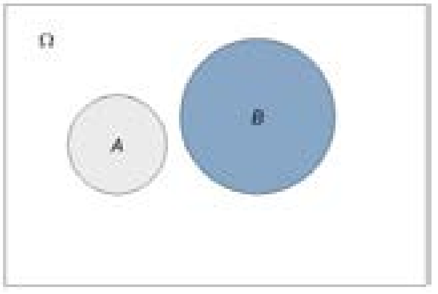


Figura 1.7. *Conjuntos disjuntos*

Obsérvese que si los conjuntos A y B son ajenos, entonces,

$$A \triangle B = A \cup B.$$

Ejemplo 1.1. Sea $\Omega = \{(x, y) \in \mathbb{R}^2 | x \geq 0, y \geq 0\}$ el conjunto universal, y considérense los siguientes conjuntos: $A = \{(x, y) \in \Omega | x^2 + y^2 \leq 1\}$; $B = \{(x, y) \in \Omega | x \leq 1, y \leq 1\}$; $C = \{(x, y) \in \Omega | x + y \leq 1\}$ y $D = \{(x, y) \in \Omega | x \leq y\}$.

a) Encontrar A^c ; B^c ; C^c; D^c.
b) Encontrar $A^c \cap C$; $B - C$; $B - D$; $A \cap D$; $C \cap D$; $B \cup D$ y $B \triangle C$.
c) Mostrar que $C \subset A \subset B$.

a) En la siguiente gráfica 1.8 se puede apreciar el conjunto D y en la gráfica 1.9 se presentan simultáneamente los conjuntos A, B y C.

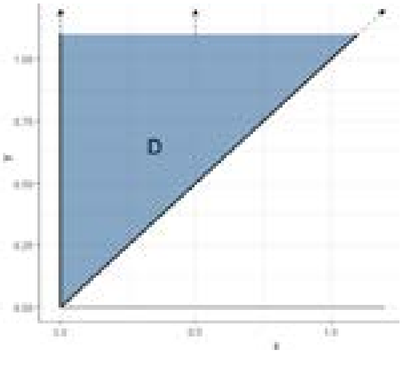


Figura 1.8. *Subconjunto D*

Las respuestas de los conjuntos complementos son, respectivamente,
$A^c = \{(x, y) \in \Omega | x^2 + y^2 > 1\}$,
$B^c = \{(x, y) \in \Omega | x > 1 \;\; o \; y > 1\}$,
$C^c = \{(x, y) \in \Omega | x + y > 1\}$,
$D^c = \{(x, y) \in \Omega | x > y\}$.

b) Los conjuntos solicitados se presentan a continuación:
$A^c \cap C = \emptyset$,
$B - C = \{(x, y) \in \Omega | x + y > 1, x \leq 1 \text{ y } y \leq 1\}$,
$B - D = \{(x, y) \in \Omega | x > y \text{ y } x \leq 1\}$,
$A \cap D = \{(x, y) \in \Omega | x^2 + y^2 \leq 1 \text{ y } x \leq y\}$,
$C \cap D = \{(x, y) \in \Omega | x \leq y \leq 1 - x \text{ y } 0 \leq x \leq 0.5\}$,
$B \cup D = \{(x, y) \in \Omega | (x \leq 1) \text{ o } (x \leq y \text{ para } x > 1)\}$,
$B \triangle C = (B - C) \cup (C - B) = B - C$.

c) En la siguiente gráfica se observa cómo $C \subset A \subset B$.

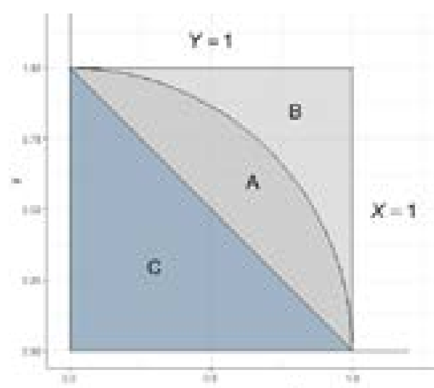


Figura 1.9. *Subconjuntos A, B y C*

A continuación, se presentarán algunas propiedades básicas de conjuntos.

Propiedades conmutativas

1) $A \cup B = B \cup A$,
2) $A \cap B = B \cap A$.

Propiedades asociativas

1) $A \cup (B \cup C) = (A \cup B) \cup C$,
2) $A \cap (B \cap C) = (A \cap B) \cap C$.

Propiedades distributivas

1) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$,
2) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

Las propiedades distributivas se pueden generalizar

1) $A \cap (\bigcup_{i=1}^{n} A_i) = \bigcup_{i=1}^{n} (A \cap A_i)$,
2) $A \cup (\bigcap_{i=1}^{n} A_i) = \bigcap_{i=1}^{n} (A \cup A_i)$.

Leyes de De Morgan

1) $(A \cup B)^c = A^c \cap B^c$,
2) $(A \cap B)^c = A^c \cup B^c$.

Las leyes de De Morgan se pueden generalizar

1) $(\bigcup_{i=1}^{n} A_i)^c = \bigcap_{i=1}^{n} A_i^c$,
2) $(\bigcap_{i=1}^{n} A_i)^c = \bigcup_{i=1}^{n} A_i^c$.

Otras propiedades

$(A^c)^c = A$,
$A \cap A = A$,
$A \cup A = A$,
$A \cup \emptyset = A$,
$A \cup \Omega = \Omega$,
$A \cap \emptyset = \emptyset$,
$A \cup A^c = \Omega$,
$A \cap A^c = \emptyset$,
$A - B = A \cap B^c$,
$A \cup (B \cap A^c) = A \cup B$,
$(A \cap B) \cup (B \cap A^c) = B$,
$(A \cap B) \cap (B^c \cap C) = \emptyset$,
$A \subset B \Rightarrow B^c \subset A^c$,
$A \subset B$ y $B \subset C \Rightarrow A \subset C$,
$A \subset B \Rightarrow A \cup B = B$ y $A \cap B = A$,
$A \subset B \Rightarrow A \cup C \subset B \cup C$,
$A \cap B = \emptyset$ y $C \subset A \Rightarrow B \cap C = \emptyset$.

El estudiante deberá de ser capaz de demostrar cada una de las propiedades previamente mencionadas. Se recomienda que lo haga con la finalidad de fortalecer sus conocimientos en teoría de conjuntos. Es importante aclarar que una demostración formal no se realiza por medio de diagramas de Venn, estos son de gran utilidad para verificar rápidamente si alguna propiedad es verdadera o no, mas estas verificaciones no son formales. A continuación, se realizará la demostración de una de las igualdades de De Morgan.

Ejemplo 1.2. Demostrar $(A \cap B)^c = A^c \cup B^c$.

Sea $x \in (A \cap B)^c$, entonces, $x \notin (A \cap B)$, esto es, $x \notin A$ o $x \notin B$, de esta manera, $x \in A^c$ o $x \in B^c$, entonces, $x \in A^c \cup B^c$. Por lo tanto, $(A \cap B)^c \subset A^c \cup B^c$.

Sea $x \in A^c \cup B^c$, entonces, $x \in A^c$ o $x \in B^c$, esto es, $x \notin A$ o $x \notin B$, de aquí se deduce que $x \notin A \cap B$, esto es, $x \in (A \cap B)^c$. De este modo, $A^c \cup B^c \subset (A \cap B)^c$.

Se concluye que

$$(A \cap B)^c = A^c \cup B^c.$$

Las siguiente definiciones y propiedades tienen que ver con colecciones de conjuntos. Entenderemos por una *colección, clase o familia* de conjuntos, aquellos conjuntos donde sus elementos a su vez son conjuntos. Las familias de conjuntos las denotaremos con las letras: $\mathcal{A}, \mathcal{B}, ..., \mathcal{Z}$.

A continuación, se presentarán algunas familias de conjuntos que son importantes en temas de probabilidad.

Definición 1.10. El *conjunto potencia* de un conjunto Ω es denotado como $\mathbb{P}(\Omega)$ y se define como la familia de conjuntos que tiene como elementos a todos los subconjuntos que se pueden formar de Ω.

Definición 1.11. Un *álgebra* sobre un conjunto Ω es una familia de conjuntos no vacía de Ω, denotada como $\mathcal{A}$, la cual cumple con las siguientes propiedades:

(I) $\emptyset \in \mathcal{A}$.
(II) Si $A \in \mathcal{A}$, entonces, $A^c \in \mathcal{A}$.
(III) Si A_1 y $A_2 \in \mathcal{A}$, entonces, $A_1 \cup A_2 \in \mathcal{A}$.

Definición 1.12. Un *σ-álgebra* sobre un conjunto Ω es una familia de conjuntos no vacía de Ω, denotada como $\mathcal{A}$, la cual cumple con las siguientes propiedades:

(I) $\emptyset \in \mathcal{A}$.
(II) Si $A \in \mathcal{A}$, entonces, $A^c \in \mathcal{A}$.
(III) Si $A_1, A_2, A_3, ... \in \mathcal{A}$, entonces, $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$.

La diferencia entre álgebra y σ-álgebra es la tercera propiedad de ambas definiciones. Es posible demostrar que un σ-álgebra es un álgebra, pero inversamente no necesariamente es cierto.

Los conceptos de álgebra y σ-álgebra sirven para definir medidas sobre familias de un conjunto Ω, véase apéndice sección D, donde se proporciona una breve introducción a la teoría de la medida. Como se verá más adelante, la probabilidad es un caso particular de medida.

Los siguientes ejemplos ayudarán a entender los anteriores conceptos.

Ejemplo 1.3. Supongamos el siguiente conjunto $\Omega = \{1, 2, 3\}$. Para las siguientes colecciones de conjuntos, indicar qué tipo de familias son
$\mathcal{A}_1 = \{\emptyset, \{1\}, \{2\}, \Omega\}$,
$\mathcal{A}_2 = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \Omega\}$,
$\mathcal{A}_3 = \{\emptyset, \Omega\}$,
$\mathcal{A}_4 = \{\{1\}, \{2\}, \Omega\}$,
$\mathcal{A}_5 = \{\emptyset, \{1\}, \{2,3\}, \Omega\}$.

La familia $\mathcal{A}_2$ es el conjunto potencia de Ω, además es un álgebra, y en este caso, por ser Ω un conjunto finito, $\mathcal{A}_2$ también es un σ-álgebra.

Las familias $\mathcal{A}_1$ y $\mathcal{A}_4$ no son álgebras (tampoco σ-álgebras).

Las familias $\mathcal{A}_3$ y $\mathcal{A}_5$ son álgebras y también σ-álgebras.

Ejemplo 1.4. Supongamos que Ω es un conjunto infinito numerable. Demostrar que las siguientes colecciones de conjuntos son σ-álgebras
a) $\mathcal{A}_1 = \mathbb{P}(\Omega)$.
b) $\mathcal{A}_2 = \{\emptyset, \Omega\}$.

a) Se sabe que $\emptyset \in \mathcal{A}_1$.

Además, si $A \in \mathcal{A}_1$, entonces, $A^c \in \mathcal{A}_1$, ya que $\mathcal{A}_1$ es el conjunto potencia.

Finalmente, para cualquier colección de conjuntos $A_1, A_2, \ldots$ que pertenece a $\mathcal{A}_1$, la unión $\bigcup_{i=1}^{\infty} A_i$, también pertenece a $\mathcal{A}_1$, ya que este último es parte del conjunto potencia.

Por lo tanto, $\mathcal{A}_1$ es un σ-álgebra.

b) Primero, $\emptyset \in \mathcal{A}_2$.

Por otro lado, se sabe que $\Omega^c = \emptyset$ y $\emptyset^c = \Omega$ están en $\mathcal{A}_2$.

Cualquier unión de conjuntos que se considere en esta familia, sea una unión finita o infinita, siempre se tendrá como resultado el conjunto $\emptyset$ o el conjunto Ω. En cualquier caso, la unión pertenece a $\mathcal{A}_2$.

En conclusión, $\mathcal{A}_2$ es σ-álgebra.

$\mathcal{A}_1$ y $\mathcal{A}_2$ son conocidas como la *máxima y mínima σ-álgebra* de Ω respectivamente.

Como se había comentado, las definiciones de álgebra y σ-álgebra son de mucha importancia en la teoría de la medida (véase apéndice sección D), el objetivo principal de estos conceptos es lograr que las operaciones de unión, intersección y complemento de conjuntos sean cerradas. En particular, la probabilidad es un caso particular de medida, de hecho, el concepto de álgebra (σ-álgebra) aparece, en primera instancia, en la definición de probabilidad (definición 1.17).

1.2. Función indicadora

El concepto de función indicadora es de mucha utilidad en probabilidad y estadística, esta se utiliza para describir en una forma práctica funciones, en particular ayuda a expresar funciones de densidad y distribución.

Definición 1.13. Sea Ω el conjunto universal y A un subconjunto de Ω. La *función indicadora* de A denotada por $I_A(x)$ es una función que tiene como dominio Ω y contradominio el conjunto $\{0, 1\}$, con regla de correspondencia dada por

$$I_A(x) = \begin{cases} 1 & si \quad x \in A, \\ 0 & si \quad x \notin A. \end{cases}$$

Ejemplo 1.5. (Aplicación en la expresión de una función). Sea $f(x)$ una función definida de la siguiente manera:

$$f(x) = \begin{cases} 0 & si \quad x < 0, \\ 2 & si \quad 0 \leq x < 1, \\ 3 - x & si \quad x \geq 1. \end{cases}$$

Describir esta función con funciones indicadoras.

Esta función descrita por funciones indicadoras se da a continuación:

$$f(x) = 2I_{[0,1)}(x) + (3 - x)I_{[1,\infty)}(x).$$

A continuación, se presentarán las propiedades más importantes de la función indicadora.

Propiedades de las funciones indicadoras

(I) $I_{A^c}(x) = 1 - I_A(x)$.
(II) $I_{A_1 \cap A_2 \cap ... \cap A_n}(x) = I_{A_1}(x)I_{A_2}(x)...I_{A_n}(x)$.
(III) $I_{A_1 \cup A_2 \cup ... \cup A_n}(x) = Max\{I_{A_1}(x), I_{A_2}(x), ..., I_{A_n}(x)\}$.

Ejemplo 1.6. Realizar la demostración de la tercera propiedad.

Demostración. La función indicadora $I_{A_1 \cup A_2 \cup ... \cup A_n}(x)$, es 1 o es 0.

Obsérvese que si

$$I_{A_1 \cup A_2 \cup ... \cup A_n}(x) = 1,$$

si y solo si

$$x \in A_1 \cup A_2 \cup ... \cup A_n,$$

lo anterior equivale a

$$x \in A_i, \text{ para alguna } i \in \mathbb{N},$$

esto es, si y solo si

$$I_{A_i} = 1, \text{ para alguna } i \in \mathbb{N},$$

esto último sucede, si y solo si

$$Max\{I_{A_1}(x), I_{A_2}(x), ..., I_{A_n}(x)\} = 1.$$

En el caso

$$I_{A_1 \cup A_2 \cup ... \cup A_n}(x) = 0,$$

se cumple, si y solo si

$$x \notin A_1 \cup A_2 \cup ... \cup A_n,$$

esto equivale a

$$x \notin A_i \text{ para toda } i \in \mathbb{N},$$

lo anterior se cumple, si y solo si

$$I_{A_i} = 0 \text{ para toda } i \in \mathbb{N},$$

esto último sucede, si y solo si

$$Max\{I_{A_1}(x), I_{A_2}(x), ..., I_{A_n}(x)\} = 0. \qquad \square$$

Las demostraciones de las otras dos propiedades se dejan como ejercicio para el estudiante.

1.3. Aleatoriedad

En esta sección se presentarán algunas opiniones de diversos autores y en diversas épocas sobre qué es la aleatoriedad. No existe una definición única y consistente sobre este concepto. A pesar de esta situación, la aleatoriedad es un elemento de mucha importancia en el cálculo de probabilidades.

Según el diccionario de la Real Academia Española (RAE), *aleatoriedad* es cualidad de *aleatorio*, y aleatorio significa que depende del *azar*. Del mismo modo, el diccionario le da varios significados a la palabra azar, entre ellos: casualidad; caso fortuito; sin rumbo ni orden.

De acuerdo con el significado del diccionario de la RAE, aleatorio depende del azar, de la casualidad; dicho de otra manera, depende de la suerte. Lo aleatorio sería lo contrario de aquello que puede ser explicado por sus causas.

A continuación, se presenta diferentes significados de aleatoriedad, dados por algunos autores. Estos son encontrados en el trabajo realizado por Batanero y Serrano [2]. Lo que se hace notar en este escrito, que en parte es una recopilación de significados de aleatoriedad y algunos derivados, es que no se ha logrado, a través del tiempo, una definición concreta del término aleatorio o aleatoriedad.

Poincaré [23] indicó que, en la antigüedad, se diferenciaban los fenómenos entre aquellos que obedecían a leyes armónicas y aquellos fenómenos que dependían del azar y que no podían preverse, ya que no había ley que los determinara. Lo que algunos autores llamaron fenómenos deterministas y fenómenos aleatorios respectivamente.

Por otro lado, Bennett [3] menciona que, desde tiempo atrás, se decía que todo fenómeno tiene una causa, nada ocurre por el azar. Que el azar se debe a nuestra ignorancia. Dicho de otra manera, el azar no es más que una medida de nuestra ignorancia. Los fenómenos que se consideran aleatorios son aquellos cuyas leyes (que determinan sus resultados) no conocemos.

Poincaré no está totalmente de acuerdo con la definición anterior, pues con ejemplos, trata de explicar que de todas maneras existen los fenómenos aleatorios.

La teoría de la probabilidad busca modelos matemáticos, no para determinar resultados de las variables de algún fenómeno, sino para determinar las probabilidades de algunos sucesos de interés.

En la explicación anterior, de acuerdo con Poincaré, los fenómenos serán los aleatorios, y según a lo expuesto por Bennett, los fenómenos serán aquellos cuyas causas ignoramos.

Cabe mencionar que la teoría de probabilidad no se ve afectada en cualquiera de las dos formas de exponer el concepto de aleatoriedad. Un modelo (una distribución) de probabilidad no cambiará, en cualquiera de las dos formas de interpretar la aleatoriedad.

Otra forma de explicar la aleatoriedad es usando el concepto de probabilidad. Se dice que un elemento de un conjunto es aleatorio, si la probabilidad de este elemento es igual a la probabilidad de cualquier otro elemento. En principio varios ejemplos cumplirían con esta condición de aleatoriedad, como el lanzamiento de una moneda o de un dado, como sacar una carta de una baraja, etc.

Kyburg [19] menciona que esta última forma de definir la aleatoriedad impone restricciones muy fuertes y no naturales a la hora de aplicarse, no obstante, la idea de considerar la equiprobabilidad para explicar la aleatoriedad es interesante para varios problemas prácticos, entre ellos, para definir una muestra aleatoria o la asignación aleatoria de objetos a experimentos, en problemas de estadística aplicada.

Una forma práctica de explicar lo que es una muestra aleatoria, es la selección de elementos de una población, donde cada elemento tiene la misma posibilidad de ser seleccionado, no obstante, se da una definición formal en el capítulo 6, véase definición 6.1.

Cabe mencionar que cada una de las citas realizadas previamente fueron hechas en el trabajo de Batanero y Serrano. Se le recomienda al estudiante, leer el escrito realizado por estos autores.

A continuación, se presentarán algunos ejemplos, donde está presente la aleatoriedad.

Ejemplo 1.7. Sin haberse realizado u observado el fenómeno no es posible determinar:

- El resultado de una moneda al ser lanzada.
- El resultado de una carta al ser sacada aleatoriamente de la baraja.
- El número ganador en una rifa.
- La vida (duración) de un componente electrónico.
- El número o proporción de artículos defectuosos obtenidos de un lote específico de artículos seleccionado al azar.
- El diámetro (peso o longitud) de una pieza fabricada.
- El tiempo necesario para que una máquina requiera mantenimiento.
- La satisfacción de una persona por un servicio recibido.
- El tiempo de espera de un cliente para ser atendido.

- La cantidad de lluvia que cae en un día determinado en el que llueve.
- La tasa de inflación para el próximo mes.

En el área de la actuaría se puede encontrar varios fenómenos donde está presente la aleatoriedad, véase el siguiente ejemplo.

Ejemplo 1.8. De igual manera, sin haberse realizado u observado el fenómeno, no se puede determinar:

- El número de defunciones de un grupo de personas aseguradas en un periodo dado.
- El número de personas que llegan a edad de 65, de un grupo de trabajadores candidatos a jubilarse.
- El número de accidentes automovilísticos ocurridos en una ciudad y en un periodo dado.
- El tiempo entre siniestro y siniestro.
- El número de personas aseguradas con seguro médico que requerirán hospitalización, en un periodo dado.
- Los gastos de honorarios médicos por cada paciente asegurado, en un periodo determinado.
- El número de personas que contraen cáncer por fumar.
- El número de personas que sufren un accidente en un mes determinado.
- El número de reclamaciones que le hacen a una compañía de seguros en un periodo dado.
- El total de dinero reclamado en una compañía de seguros en determinado periodo de tiempo.
- Las pérdidas por invertir en la bolsa de valores, en un conjunto de acciones de varias empresas (conocido como riesgo de mercado).
- Las pérdidas de una institución financiera por el incumplimiento en la cartera de crédito (conocido como riesgo de crédito).
- Las pérdidas en los activos de una compañía aseguradora por el incremento en la tasa de interés soberana (conocido como riesgo de mercado).

1.4. Clases de probabilidad

La palabra *probabilidad* o derivados de esta son usados comúnmente por la gente, por ejemplo, « es muy probable que llueva hoy por la tarde»; «con alta probabilidad aprobaré la materia», etc. A pesar de esta familiaridad, no existe una interpretación única del término que se haya aceptado por científicos, filósofos, actuarios, etc. A continuación, se presentarán tres interpretaciones («definiciones») comunes de la probabilidad:

(i) Clásica o *a priori*.
(ii) Frecuentista o *a posteriori*.
(iii) Subjetiva o personal.

La interpretación *clásica*, se basa en que todos los resultados posibles de un experimento son igualmente verosímiles.

Ejemplo 1.9. En el lanzamiento de una moneda «honesta» (bien balanceada), los dos resultados posibles, águila y sol, son igualmente verosímiles. Como la suma de las dos probabilidades tiene que dar el 100 %, la probabilidad de cada resultado es del 50 %. En este caso, los dos resultados posibles son igualmente probables.

Ejemplo 1.10. En el nacimiento de un bebé, existen también dos resultados posibles: niña y niño. Pero en este caso, a diferencia del anterior, puede ser que dichos resultados no sean igualmente verosímiles. En algunas poblaciones se parecen mucho las dos probabilidades, pero no son iguales. Estos valores también dependen de la población en estudio.

En términos generales, en la definición clásica de probabilidad, si el resultado de un suceso es alguno de los n_a posibles de un total de n resultados, la probabilidad (clásica) del suceso es igual a n_a/n.

En el ejemplo 1.9, la probabilidad de águila (y la de sol) es $1/2$.

En la interpretación *frecuentista*, suponiendo que el experimento se pueda repetir de igual manera «un gran número de veces», la probabilidad de un suceso es considerada como la frecuencia relativa «aproximada» del suceso (fracción de veces que ocurre el suceso).

Ejemplo 1.11. Aplicando la interpretación frecuentista, la probabilidad de que resulte águila al lanzar una moneda bien balanceada es aproximadamente igual a 0.5. Esto es, si se lanza un gran número de veces la misma moneda de manera idéntica, seguro que la fracción de veces que resulta águila será aproximadamente al 50 %.

En el ejemplo 1.10, se podrán encontrar aproximaciones de las probabilidades de nacimiento de niña y de niño, siempre y cuando se cuente con información suficiente. Es importante señalar que esta información tiene que ser veraz, sin errores, que no cambien las condiciones del experimento, y entre más información, mejor será la aproximación.

En la interpretación *subjetiva o personal*, una persona dará su opinión acerca de la verosimilitud del suceso, véase el siguiente ejemplo.

Ejemplo 1.12. Una persona, «de buenas a primeras», sin información de una moneda en particular, podría considerar que la probabilidad de águila (y la de sol) es igual a 0.5. Otra persona, teniendo más información de los resultados de esta moneda, opina que águila es más probable que sol, y que a su juicio, la probabilidad de águila es $\frac{3}{5}$ y la de sol $\frac{2}{5}$.

No es posible generalizar alguna de estas interpretaciones como una definición general y formal de probabilidad. Dependiendo de la situación, existirán ventajas y desventajas de usar cada una de estas definiciones.

La definición clásica no puede aplicar en situaciones donde los resultados no son igualmente verosímiles y/o donde el conjunto de resultados posibles es infinito. Por ejemplo, si se quiere observar el número de accidentes automovilísticos en cierta ciudad, en un lapso de tiempo determinado, puede notarse que el número de accidentes posibles puede ser cualquier número entero no negativo (número de posibilidades infinitas), es lógico intuir que 0 accidentes no es igualmente verosímil que 500 accidentes y, en general, los resultados posibles no son igualmente probables.

La definición frecuentista carece de formalidad, ¿qué significa número grande de repeticiones, o que la moneda sea lanzada siempre de la misma manera?, también, no queda claro cuándo se tendrá una buena aproximación, esto es, nunca nos indican el número de repeticiones para asegurar una buena aproximación. Además, en ciertos fenómenos, la repetición será muy compleja, y en algunos casos será hasta imposible, esto es, la probabilidad frecuentista no siempre es viable.

En la definición subjetiva, el que los juicios de una persona sobre las verosimilitudes de los sucesos de interés sean completamente consistentes y libres de contradicciones es humanamente imposible. No existirán bases objetivas para que dos o más científicos o personas que trabajan sobre el mismo problema lleguen a las mismas evaluaciones o criterios sobre la probabilidad de interés.

No discutiremos qué interpretación (definición) es la más adecuada, como se había ya comentado, esto depende de la situación (del fenómeno) que se esté tratando. Las interpretaciones *a priori* y *a posteriori* son las más usadas en la práctica.

La primera nos servirá para calcular probabilidades, en situaciones donde el número de posibilidades del experimento es finito y además equiprobable. De hecho, las primeras probabilidades se calcularán usando esta definición.

La segunda nos ayudará en situaciones donde se cuenta con información veraz y suficiente del evento, es usada comúnmente para aproximar probabilidades, cuando la definición clásica no puede utilizarse. La probabilidad frecuentista también sirve para interpretar probabilidades en la vida real.

Parecería que la probabilidad subjetiva es la más informal de las tres definiciones, más cabe señalar, con base en cierto conocimiento bien fundamentado del evento en estudio, que estas probabilidades son usadas en la estadística bayesiana, y en los últimos años, esta área de la estadística se ha aplicado de manera muy importante.

Finalmente, reconsiderando el ejemplo de lanzar una moneda «honesta», se sabe que si se lanza un gran número de veces en forma idéntica, la frecuencia relativa de águila tiende a ser igual a la probabilidad de este evento (definición frecuentista). La aproximación entre la frecuencia relativa y la probabilidad tiende a mejorar cuando se aumenta el número de repeticiones del experimento. Dicho fenómeno de estabilización de la frecuencia relativa a la probabilidad, es conocido como *regularidad estadística*.

1.5. Espacio muestral y eventos

Es importante aclarar que en los experimentos considerados de aquí en adelante está presente la aleatoriedad.

Las siguientes definiciones son propias de la teoría de probabilidad y son necesarias antes de presentar los axiomas y cada una de las propiedades de probabilidad. A partir de estos conceptos, será notorio la importancia que tiene los conceptos y propiedades de conjuntos.

Definición 1.14. El conjunto de todos los resultados posibles de un experimento es llamado *espacio muestral* y es denotado como Ω. Los elementos del espacio muestral Ω son llamados *puntos muestrales.*

Definición 1.15. Un subconjunto del espacio muestral asociado a un experimento es llamado *evento o suceso*. Un evento que consiste de un solo punto muestral es llamado *evento simple*. Los eventos son denotados por las letras mayúsculas $A, B, \ldots, X, Y, Z$.

Definición 1.16. La clase de los eventos asociados a un experimento es definido como *espacio de eventos*, y es denotado por $\mathcal{A}$.

A continuación, se darán algunos ejemplos de los anteriores conceptos.

Ejemplo 1.13. Se lanza un tetraedro numerado bien balanceado para observar el número resultante. Determinar el espacio muestral, los eventos simples y el espacio de eventos.

Obsérvese que el espacio muestral es

$$\Omega = \{1, 2, 3, 4\}.$$

Algunos eventos son

$$\emptyset, \{1\}, \{2\}, \{2, 3, 4\}, \{2, 4\}, \{1, 3\},$$

donde los eventos simples son $\{1\}, \{2\}, \{3\}, \{4\}$.

El espacio de eventos (conjunto potencia de Ω) es

$$\mathcal{A} = \{\emptyset, \{1\}, \ldots, \{4\}, \{1, 2\}, \ldots, \{3, 4\}, \{1, 2, 3\}, \ldots, \{2, 3, 4\}, \Omega\}.$$

Obsérvese que, estrictamente hablando, no es lo mismo punto muestral que evento simple.

En muchos casos, cuando el espacio muestral es finito o infinito numerable, una propuesta común del espacio de eventos es el conjunto potencia del espacio muestral, pero esto depende también de los resultados que nos interesa observar.

Lo que es importante, es pedir que el espacio de eventos sea un álgebra o σ-álgebra, véase definición de probabilidad (definición 1.17).

Ejemplo 1.14. Se arrojan tres monedas al aire. Determinar el espacio muestral e indicar la cardinalidad del espacio de eventos.

Al lanzar una moneda, consideremos la siguiente notación, A resulte águila y S resulte sol.

Entonces, el espacio muestral es

$$\Omega = \{SSS, SSA, ASS, SAS, AAS, SAA, ASA, AAA\}.$$

Observemos algunos eventos,

$$B = \{\text{un águila}\} = \{SSA, ASS, SAS\},$$

$$C = \{\text{por lo menos un águila}\} = \Omega - \{SSS\}.$$

En este caso, la cardinalidad del espacio de eventos $\mathcal{A}$ (conjunto potencia), es de 256 eventos. El cálculo de esta cardinalidad se comprenderá más adelante, en la sección de análisis combinatorio de este capítulo.

Ejemplo 1.15. Dentro de una caja hay 6 fichas numeradas del 1 al 6. Se escogen dos fichas aleatoriamente, una tras otra sin sustitución. Determinar el espacio muestral e indicar la cardinalidad del espacio de eventos.

El espacio muestral en este caso es

$$\begin{aligned}\Omega = \{&(1,2),(1,3),(1,4),(1,5),(1,6),(2,1),(2,3),(2,4),(2,5),(2,6),\ldots,\\ &(6,1),(6,2),(6,3),(6,4),(6,5)\}.\end{aligned}$$

Algunos eventos se presentan a continuación:

$$\begin{aligned}A &= \{\text{Resulte la ficha 1}\}\\ &= \{(1,2),(1,3),(1,4),(1,5),(1,6),(2,1),(3,1),(4,1),(5,1),(6,1)\}.\end{aligned}$$

$$\begin{aligned}B &= \{\text{La suma de ambas fichas es un número par}\}\\ &= \{(1,3),(3,1),(1,5),(5,1),(2,4),(4,2),(2,6),(6,2),(3,5),(5,3),\\ &\quad (4,6),(6,4)\}.\end{aligned}$$

Considerando que el espacio de eventos es el conjunto potencia, existen $2^{30} = 1{,}073{,}741{,}824$ eventos posibles.

El cálculo del número de eventos posibles se comprenderá en el tema de análisis combinatorio.

Como se había comentado, si el espacio muestral es finito (también infinito numerable) es muy común que el espacio de eventos se considere como el conjunto potencia de Ω. Para espacio muestrales infinitos no numerables, ya no será posible considerar el conjunto potencia como el espacio de eventos, véase el siguiente ejemplo.

Ejemplo 1.16. Se observa el tiempo de duración de cierto componente electrónico. Determinar el espacio muestral y proponer un espacio de eventos adecuado.

En este caso, el espacio muestral es $\Omega = \{x \in \mathbb{R} | x \geq 0\} = [0, \infty)$.

Cualquier intervalo, subconjunto de $[0, \infty)$ es un suceso, por ejemplo, suponiendo que $0 < a < b$, los intervalos (a, b), $(0, a)$, (a, ∞), entre otros conjuntos son sucesos.

Como espacio de eventos, se puede sugerir a la *clase de conjuntos de Borel* sobre el conjunto $[0, \infty)$. Cabe señalar, que la clase de conjuntos de Borel, es el σ-álgebra generada por todos los intervalos abiertos en $[0, \infty)$. Véase apéndice D, para tener una mejor idea de qué subconjuntos son parte de la familia de conjuntos de Borel.

Se puede construir un subconjunto de $[0, \infty)$, que no es medible (tampoco es conjunto Borel), para el cual no es posible asignarle una probabilidad. No corresponde a este curso la construcción de este subconjunto, pero se puede consultar el apéndice D, también Capinski y Kopp [5] o Galaz F. [9]. De esta manera, se aclara que el conjunto potencia no es adecuado para ser el espacio de eventos.

1.6. Axiomas de probabilidad

A continuación, se presentarán las definiciones de función de probabilidad y probabilidad de un evento.

Definición 1.17. Sea Ω el espacio muestral asociado a un experimento, la función denotada como $P[A]$, que es definida sobre un σ-álgebra de eventos $\mathcal{A}$, con valores reales, es conocida como *función de probabilidad (medida de probabilidad)* si cumple los siguientes tres axiomas:

(I) $P[A] \geq 0$, para todo $A \in \mathcal{A}$.

(II) $P[\Omega] = 1$.

(III) Si $A_1, A_2, ...$ son eventos mutuamente excluyentes, entonces,

$$P\left[\bigcup_{i=1}^{\infty} A_i\right] = \sum_{i=1}^{\infty} P[A_i].$$

La definición anterior, también es conocida como los *axiomas de probabilidad o axiomas de Kolmogorov.*

El valor asignado a un evento específico A, bajo la función de probabilidad $P[A]$, es llamado *probabilidad* de A .

Es importante aclarar que en la definición de probabilidad no se indica en ningún momento cómo se debe de calcular la probabilidad de un evento. No existe una fórmula general para calcular probabilidades de eventos, y dependiendo de la circunstancia, el cálculo de probabilidad se realizará por el procedimiento adecuado. Cualquier procedimiento de cálculo de probabilidad que se aplique deberá de cumplir los tres axiomas de probabilidad.

Observación. Algunos autores, en lugar de definir la función $P[A]$ en un σ-álgebra de eventos, lo hacen en un álgebra de eventos $\mathcal{A}$, en este caso, el tercer axioma cambia de la siguiente manera:

Si $A_1, A_2, ...$ son eventos mutuamente excluyentes y además, $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$, entonces,

$$P\left[\bigcup_{i=1}^{\infty} A_i\right] = \sum_{i=1}^{\infty} P[A_i].$$

Por otro lado, de aquí en adelante, las palabras porcentaje y proporción de un evento A serán usados como sinónimos de probabilidad de A.

1.7. Espacio de probabilidad

En esta sección se tratará con la definición de espacio de probabilidad.

Definición 1.18. Un *espacio de probabilidad* es una terna $(\Omega, \mathcal{A}, P[A])$, donde Ω es el espacio muestral, $\mathcal{A}$ es el espacio de eventos y $P[A]$ es una función de probabilidad.

Observación. A todo experimento le corresponde un espacio de probabilidad. Dicho de otra manera, en todo resultado o problema de probabilidad, debe quedar bien definido cada elemento del espacio de probabilidad.

A partir de este momento, el espacio de probabilidad queda bien determinado para toda proposición, ejercicio o ejemplo de probabilidad.

Ejemplo 1.17. Se lanza un tetraedro numerado y bien balanceado. Determinar el espacio de probabilidad.

El espacio muestral es $\Omega = \{1, 2, 3, 4\}$.

El espacio de eventos, en este caso, es el conjunto potencia de Ω, esto es,

$$\mathcal{A} = \{\emptyset, \{1\}, \{2\}, \{3\}, \{4\}, \{1, 2\}, \ldots, \{3, 4\}, \{1, 2, 3\}, \ldots, \{2, 3, 4\}, \Omega\}.$$

Al ser un tetraedro bien balanceado, es razonable definir la función de probabilidad de la siguiente forma:

$$P[\{i\}] = \frac{1}{4} = 0.25, \quad \text{para } i = 1, 2, 3, 4.$$

De este modo, queda bien definido el espacio de probabilidad por la terna $(\Omega, \mathcal{A}, P[A])$, donde cada uno de ellos previamente ya fue definido.

1.8. Propiedades de probabilidad

A partir de los axiomas de probabilidad es posible demostrar las propiedades más importantes. Para cada una de estas propiedades se supone que queda bien definido el espacio de probabilidad $(\Omega, \mathcal{A}, P[A])$, el cual está asociado a un experimento.

Teorema 1.1. $P[\emptyset] = 0$.

Demostración. Definimos $A_1, A_2, A_3, \ldots$ eventos tal que $A_i = \emptyset$, para toda $i \in \{1, 2, \ldots\}$.

Se sabe $\bigcup_{i=1}^{\infty} A_i = \emptyset$, y además $A_i \cap A_j = \emptyset$ para toda $i \neq j$. Por lo tanto,

$$\begin{aligned}
P[\emptyset] &= P\left[\bigcup_{i=1}^{\infty} A_i\right] \\
&= \sum_{i=1}^{\infty} P[A_i] \quad \text{(por el axioma 3)} \\
&= \sum_{i=1}^{\infty} P[\emptyset] = \lim_{n\to\infty} \sum_{i=1}^{n} P[\emptyset] \\
&= P[\emptyset] \lim_{n\to\infty} \sum_{i=1}^{n} 1 \\
&= P[\emptyset] \lim_{n\to\infty} n.
\end{aligned}$$

Como $P[\emptyset] \geq 0$ (por el primer axioma), entonces,

$$P[\emptyset] > 0 \quad \text{o} \quad P[\emptyset] = 0.$$

Si $P[\emptyset] > 0$, entonces,

$$P[\emptyset] = P[\emptyset] \lim_{n\to\infty} n \ \text{ diverge},$$

lo cual es una contradicción, por definición, $P[\emptyset]$ tiene que ser un número real.

Entonces, no existe más opción que $P[\emptyset] = 0$, observando que, en este caso, no existe contradicción alguna. $\square$

Teorema 1.2. *Sean $A_1, A_2, ..., A_n$ eventos mutuamente excluyentes, entonces,*

$$P\left[\bigcup_{i=1}^{n} A_i\right] = \sum_{i=1}^{n} P[A_i].$$

Demostración. Definimos $A_j = \emptyset$, para $j = n+1, n+2, \ldots$

Entonces, podemos afirmar, $\bigcup_{i=1}^{n} A_i = \bigcup_{i=1}^{\infty} A_i$, además $A_i \cap A_j = \emptyset$ para toda $i \neq j$.

Por lo tanto,

$$\begin{aligned}
P\left[\bigcup_{i=1}^{n} A_i\right] &= P\left[\bigcup_{i=1}^{\infty} A_i\right] \\
&= \sum_{i=1}^{\infty} P[A_i] \quad \text{(por el axioma 3)} \\
&= \sum_{i=1}^{n} P[A_i] + \sum_{j=n+1}^{\infty} P[A_j] \\
&= \sum_{i=1}^{n} P[A_i] + \sum_{j=n+1}^{\infty} P[\emptyset] \\
&= \sum_{i=1}^{n} P[A_i] + \sum_{j=n+1}^{\infty} 0 \quad \text{(por el teorema 1.1)} \\
&= \sum_{i=1}^{n} P[A_i].
\end{aligned}$$

□

Teorema 1.3. *Sea A un evento, entonces,*

$$P[A] = 1 - P[A^c].$$

Demostración. Obsérvese

$$\begin{aligned}
1 &= P[\Omega] \quad \text{(por el axioma 2)} \\
&= P[A \cup A^c] \\
&= P[A] + P[A^c]. \quad \text{(por el teorema 1.2)}
\end{aligned}$$

Por lo tanto,

$$P[A] = 1 - P[A^c].$$

□

Teorema 1.4. *Sean A y B dos eventos, entonces,*

$$P[A \cap B^c] = P[A] - P[A \cap B].$$

Demostración. Obsérvese

$$A = (A \cap B) \cup (A \cap B^c), \text{ donde } (A \cap B) \cap (A \cap B^c) = \emptyset.$$

De esta manera,

$$P[A] = P[A \cap B] + P[A \cap B^c].$$

Por lo tanto,

$$P[A \cap B^c] = P[A] - P[A \cap B].$$

□

Obsérvese, $P[A \cap B^c] = P[A - B]$.

Teorema 1.5. *Sean A y B dos eventos, entonces,*

$$P[A \cup B] = P[A] + P[B] - P[A \cap B].$$

Demostración. Dado que

$$A \cup B = B \cup (A \cap B^c), \text{ donde } B \cap (A \cap B^c) = \emptyset,$$

se cumple

$$\begin{aligned} P[A \cup B] &= P[B] + P[A \cap B^c] \\ &= P[B] + P[A] - P[A \cap B], \quad \text{(por el teorema 1.4)} \end{aligned}$$ □

Corolario 1.1. *Sean A, B y C eventos cualesquiera, entonces,*

$$\begin{aligned} P[A \cup B \cup C] = P[A] + P[B] + P[C] - P[A \cap B] - P[A \cap C] - P[B \cap C] \\ + P[A \cap B \cap C]. \end{aligned}$$

La demostración de este resultado se deja al estudiante.

Corolario 1.2. *Sean $A_1, A_2, ..., A_n$ eventos cualesquiera, entonces,*

$$\begin{aligned} P\left[\bigcup_{i=1}^{n} A_i\right] &= \sum_{i=1}^{n} P[A_i] - \sum_{i,j;i<j} P[A_i \cap A_j] \\ &+ \sum_{i,j,k;i<j<k} P[A_i \cap A_j \cap A_k] - ... + (-1)^{n+1} P[A_1 \cap A_2 \cap ... \cap A_n]. \end{aligned}$$

La demostración de este corolario se deja al estudiante, y se recomienda utilizar inducción matemática.

Teorema 1.6. *Sean A y B dos eventos tales que $A \subset B$, entonces,*

$$P[A] \leq P[B].$$

Demostración. Como $A \subset B$, entonces, B se puede expresar de la siguiente manera:

$$B = A \cup (B \cap A^c), \text{ donde } A \cap (B \cap A^c) = \emptyset.$$

Así que

$$P[B] = P[A] + P[B \cap A^c].$$

Se deduce que

$$P[B] - P[A] = P[B \cap A^c] \geq 0, \quad \text{(por el axioma 1).}$$

Por consiguiente,

$$P[A] \leq P[B].$$ □

Corolario 1.3. *Sea A un evento, entonces,*

$$P[A] \leq 1.$$

Demostración. Se sabe $A \subset \Omega$.

Entonces,

$$P[A] \leq P[\Omega], \quad \text{(por el teorema 1.6)},$$

Se concluye que

$$P[A] \leq 1. \qquad \square$$

Teorema 1.7 (Desigualdad de Boole)**.** *Sean $A_1, A_2, ..., A_n$ eventos cualesquiera, entonces,*

$$P\left[\bigcup_{i=1}^{n} A_i\right] \leq \sum_{i=1}^{n} P[A_i].$$

Demostración. Se realizará por inducción matemática.

Primero para $n = 2$, se sabe

$$P[A_1 \cup A_2] = P[A_1] + P[A_2] - P[A_1 \cap A_2].$$

De esta manera,

$$P[A_1] + P[A_2] - P[A_1 \cup A_2] = P[A_1 \cap A_2] \geq 0, \quad \text{(por el axioma 1)}.$$

Por lo tanto,

$$P[A_1] + P[A_2] \geq P[A_1 \cup A_2].$$

Suponemos válido para $n = k$, esto es,

$$P\left[\bigcup_{i=1}^{k} A_i\right] \leq \sum_{i=1}^{k} P\left[A_i\right].$$

Se demostrará para $n = k + 1$,

$$\begin{aligned}
P\left[\bigcup_{i=1}^{k+1} A_i\right] &= P\left[\bigcup_{i=1}^{k} A_i \cup A_{k+1}\right] \\
&\leq P\left[\bigcup_{i=1}^{k} A_i\right] + P\left[A_{k+1}\right] \quad \text{(válido para } n = 2\text{)} \\
&\leq \sum_{i=1}^{k} P\left[A_i\right] + P\left[A_{k+1}\right] \quad \text{(válido para } n = k\text{)} \\
&= \sum_{i=1}^{k+1} P\left[A_i\right].
\end{aligned} \qquad \square$$

Observación. Las propiedades anteriores tampoco indican la forma de calcular la probabilidad de un evento. El cálculo de probabilidades, como ya se había mencionado, es un tema que se empezará a tratar más adelante, y más bien, cualquier forma de calcular una probabilidad deberá de cumplir tanto los axiomas de probabilidad como las propiedades.

1.9. Espacios muestrales finitos y equiprobables

Hasta este momento no se ha tratado el cálculo de probabilidad de un evento. Es importante aclarar que este no es posible hacerlo de una forma general. Como se verá a través de los capítulos, hay diferentes formas de hacerlo, depende mucho del experimento.

A continuación, se analizará el cálculo de probabilidades en forma general para experimentos donde el espacio muestral es finito y equiprobable, entendiendo como un espacio muestral equiprobable, aquel espacio donde cada uno de sus eventos simples tiene la misma probabilidad de ocurrir.

Consideremos que el espacio muestral Ω asociado a un experimento es finito, digamos que está conformado por n puntos muestrales, supongamos que $\Omega = \{a_1, a_2, \ldots, a_n\}$.

Por otro lado, sea A un evento con k puntos muestrales, supongamos que $A = \{a\acute{}_1, a\acute{}_2, ..., a\acute{}_k\}$.

Ahora, si el espacio muestral Ω es equiprobable, se cumplirá

$$P[\{a_i\}] = \frac{1}{n}.$$

Entonces,

$$\begin{aligned} P[A] &= P[\{a\acute{}_1\} \cup \{a\acute{}_2\} \cup ... \cup \{a\acute{}_k\}] \\ &= P[\{a\acute{}_1\}] + P[\{a\acute{}_2\}] + ... + P[\{a\acute{}_k\}] \\ &= \underbrace{\frac{1}{n} + \cdots + \frac{1}{n}}_{k \quad \text{veces}} = \frac{k}{n}. \end{aligned}$$

Observación. La anterior fórmula corresponde a la *probabilidad clásica* de un evento, y se podrá aplicar para aquellos fenómenos donde el espacio muestral Ω sea finito y equiprobable. Si alguno de los dos supuestos no se cumple, esta fórmula no se puede usar o se estaría usando en forma incorrecta.

A continuación, se presentarán los primeros cálculos de probabilidad.

Ejemplo 1.18. Para los siguientes casos, calcular la probabilidad de que resulte

a) Un número par, en el lanzamiento de un dado.
b) Un total de diez, en el lanzamiento de dos dados.
c) Uno de los dados sea el doble del otro, en el lanzamiento de dos dados.

a) El espacio muestral es $\Omega = \{1, 2, 3, 4, 5, 6\}$. Sea $A = \{\text{par}\}$, entonces,

$$P[A] = \frac{3}{6} = \frac{1}{2}.$$

b) En este caso $\Omega = \{(1,1), (1,2), \ldots, (1,6), \ldots, (6,1), (6,2), \ldots, (6,6)\}$. Sea $B = \{\text{un total igual a } 10\}$, entonces,

$$P[B] = \frac{3}{36} = \frac{1}{12}.$$

c) El espacio muestral coincide con el espacio del inciso anterior. Sea $C = \{\text{uno dado sea el doble del otro}\}$, por lo tanto,

$$P[C] = \frac{6}{36} = \frac{1}{6}.$$

Ejemplo 1.19. En una urna hay n bolas numeradas, del 1 al n. Se seleccionan dos bolas aleatoriamente (al mismo tiempo). Calcular la probabilidad de que resulte el número más pequeño (el 1) y el más grande (el n), en los siguientes casos:
a) $n = 4$.
b) $n = 5$.
c) $n = 10$.

Sea $A = \{\text{Resulte el número más pequeño y el más grande}\}$.
a) El espacio muestral es

$$\Omega = \{\{1,2\}, \{1,3\}, \{1,4\}, \{2,3\}, \{2,4\}, \{3,4\}\}.$$

De este modo,

$$P[A] = \frac{1}{6}.$$

b) En este caso

$$\Omega = \{\{1,2\}, \{1,3\}, \{1,4\}, \{2,3\}, \{2,4\}, \{3,4\}, \{1,5\}, \{2,5\}, \{3,5\}, \{4,5\}\}.$$

Por lo tanto,

$$P[A] = \frac{1}{10}.$$

c) Se puede observar que el número de puntos muestrales en este caso es 45, de los cuales uno de ellos favorece al evento, por lo tanto,

$$P[A] = \frac{1}{45}.$$

Se puede notar que, conforme el número de bolas en la urna es más grande, el cálculo del denominador se va complicando, aunque este cálculo todavía se pudo realizar, enlistando los puntos muestrales.

Considerando el ejemplo anterior, si las bolas son seleccionadas al azar, una bola tras otra sin reemplazo, las probabilidades anteriores son respectivamente: $\frac{2}{12} = \frac{1}{6}$; $\frac{2}{20} = \frac{1}{10}$ y $\frac{2}{90} = \frac{1}{45}$.

Las probabilidades no cambiaron, con respecto a la forma de seleccionar las bolas. Se puede observar que el número de casos, tanto en el numerador, como en el denominador, sí resultaron diferentes, pero los cocientes, respectivamente, son iguales.

Por otro lado, si las bolas del ejemplo anterior se sacan una tras otra con reemplazo, las probabilidades correspondientes se calculan de la siguiente manera: $\frac{2}{16} = \frac{1}{8}$; $\frac{2}{25}$ y $\frac{2}{100} = \frac{1}{50}$.

En este caso, las probabilidades respectivamente sí son diferentes.

Es importante estar atentos en la forma como se seleccionan los elementos en un experimento, pues, como se puede observar, la forma de seleccionar puede influir en el cálculo de probabilidades.

Finalmente, supongamos que en la urna del ejemplo anterior hay 20 bolas numeradas del 1 al 20, y de estas se sacan 5 bolas al azar. El número de puntos muestrales difícilmente se podrá calcular enlistando los puntos muestrales, por lo que será necesario usar técnicas de conteo. Como dato, el número de puntos muestrales en este caso es 15,504.

En general, contar el número de resultados posibles en un experimento y el número de resultados a favor de un evento no siempre es un problema fácil. En ocasiones, es necesario apoyarse del análisis combinatorio o técnicas de conteo. A continuación, se presentarán estas técnicas de apoyo para el conteo, y por lo tanto, también serán de gran apoyo para el cálculo de probabilidades.

1.10. Análisis combinatorio

El análisis combinatorio o técnicas de conteo se apoya básicamente en el principio fundamental de conteo. El principio fundamental de conteo consiste en dos reglas básicas: la de multiplicación y la de adición. A partir de estas reglas, se presentan las técnicas de conteo más importantes, las diferentes maneras de contar permutaciones y la forma de contar combinaciones.

Regla básica de multiplicación

Supongamos que existen k procedimientos, donde el número de maneras en que se puede realizar el procedimiento i es n_i, $i = 1, 2, \ldots, k$. Además, asumimos que el procedimiento i, para $i = 2, \ldots, k$, se puede realizar en cualquiera de sus maneras, después de haberse realizado el procedimiento $i-1$, no importando en que manera se llevó a cabo este último. Entonces, el proceso que consiste en realizar el procedimiento 1 y después el procedimiento 2 y después el procedimiento 3,..., y después el procedimiento k, se puede efectuar de $n_1 n_2 \cdots n_k$ formas diferentes.

Ejemplo 1.20. Un menú de comida consta de 5 tiempos. El primer tiempo se puede seleccionar de 3 formas diferentes, el segundo de 2, el tercero de 1, el cuarto de 2 y el quinto de 6 maneras diferentes. ¿De cuántas formas diferentes se puede seleccionar el menú?

El número de maneras diferentes es $(3)(2)(1)(2)(6) = 72$.

Regla básica de adición

Supongamos que existen k procedimientos, donde el número de maneras diferentes en que se puede realizar el procedimiento i es $n_i, i = 1, 2, \ldots, k$. Además, suponemos que los procedimientos no se pueden realizar simultáneamente. Entonces, el proceso que consiste en realizar el procedimiento 1, o el procedimiento 2, o el procedimiento 3, ... , o el procedimiento k, se puede efectuar de $n_1 + n_2 + \cdots + n_k$ maneras diferentes.

Ejemplo 1.21. Supongamos que existe solo 1 manera de comer en casa, 5 formas diferentes en la universidad, 3 en cierto restaurante y 2 maneras diferentes de comer en el mercado. ¿De cuántas maneras diferentes se puede comer en un día dado?

El número de maneras diferentes es $1 + 5 + 3 + 2 = 11$.

En la regla de multiplicación está involucrada la letra «y», y en la de adición la letra «o» en un sentido exclusivo.

A continuación, se verán las técnicas de conteo más importantes, las cuales se dividen en la forma de contar las permutaciones y la manera de contar las combinaciones. Estas técnicas están sustentadas en los dos principios básicos de conteo, pero principalmente en el de multiplicación.

Permutaciones

Una *permutación u ordenación* de un conjunto es un arreglo hecho por los elementos o por algunos de los elementos del conjunto.

Ejemplo 1.22. Sea $A = \{a, b, c, d\}$ un conjunto. Indicar algunas permutaciones de diferente tamaño.

Algunas permutaciones sin permitir la repetición de sus elementos son: ab, ac, bc, da (tamaño 2); abc, acd, bcd, cda (tamaño 3); $abcd, abdc, cbad$ (tamaño 4).

Algunas permutaciones de diferente tamaño, permitiendo la repetición de algunos de sus elementos son: $abbdc, aaabb, abcc, add$.

Las permutaciones u ordenaciones de un conjunto se clasifican en:

(I) Permutaciones sin repetición.
(II) Permutaciones con repetición.
(III) Permutaciones con elementos iguales.

El objetivo de esta sección es calcular el número de permutaciones de cierto tipo y de determinado tamaño que se pueden formar con los elementos de un conjunto. A continuación, se mostrarán estos cálculos.

(I) *Permutaciones sin repetición*
El número de maneras de ordenar a r elementos diferentes de un conjunto de n elementos, donde $r \leq n$, se denota como P_r^n y es igual a

$$P_r^n = n(n-1)\cdots(n-(r-1)).$$

Ejemplo 1.23. ¿De cuántas maneras diferentes se pueden ordenar los elementos del conjunto $\{a, b, c\}$?

El número de maneras es $P_3^3 = (3)(2)(1) = 6$.
Las ordenaciones son $abc, acb, bac, bca, cba, cab$.

Ejemplo 1.24. Calcular el número de arreglos de tamaño dos del conjunto $\{a, b, c, d\}$.

La respuesta es $P_2^4 = (4)(3) = 12$.
Los arreglos son $ab, ac, ad, ba, ca, da, bc, bd, cb, db, cd, dc$.
Ahora obsérvese

$$P_n^n = n(n-1)\cdots(2)(1) = n!,$$

donde $n!$, es conocido como *factorial de* n, véase la sección C.1 del apéndice.
Por otro lado, también es cierto

$$P_r^n = \frac{n!}{(n-r)!}.$$

Considerando el último ejemplo, el número de permutaciones se puede calcular usando la anterior fórmula, esto es,

$$P_2^4 = \frac{4!}{(4-2)!} = 12.$$

Ejemplo 1.25. ¿Cuántas placas diferentes se pueden formar de 3 letras diferentes y 2 dígitos diferentes, si las letras van primero y pueden ser de la A a la C y los dígitos pueden ser el 1 o 2?

La respuesta es $3(2)(1) \times (2)(1) = 12$.
Las placas son $ABC12, ACB12, BAC12, BCA12, CAB12, CBA12,$
$ABC21, ACB21, BAC21, BCA21, CAB21, CBA21$.

Ejemplo 1.26. ¿Cuántas placas diferentes se pueden formar de 3 letras diferentes y 3 dígitos diferentes, si las letras van primero y pueden ser escogidas de la A a la E?

La respuesta es $5(4)(3) \times (10)(9)(8)$ =43,200.

(II) *Permutaciones con repetición*
El número de permutaciones de r elementos cada una, donde los elementos se toman de un conjunto de n elementos y donde se permite la repetición de los elementos es igual a

$$\underbrace{n \times n \times \cdots \times n}_{r \quad \text{veces}} = n^r.$$

Ejemplo 1.27. Calcular el número de permutaciones de tamaño 2 que se pueden formar con los elementos del conjunto $A = \{a, b, c\}$, permitiendo la repetición de elementos.

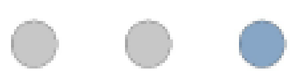
El resultado es $(3)(3) = 3^2 = 9$.
Las permutaciones son $aa, ab, ac, ba, ca, bc, cb, bb, cc$.

Ejemplo 1.28. ¿Cuántas placas diferentes se pueden formar de 3 letras y 3 cifras, si las letras van primero y solamente pueden ser de la A a la E? ¿Cuántas placas, si no se permite que la placa termine en 000?

El número de placas diferentes es $(5)(5)(5) \times (10)(10)(10)$ =125,000.
El número de placas que terminan en 000 es $(5)(5)(5) \times (1)(1)(1) = 125$.
Por lo tanto, el número de placas diferentes que no terminan en 000 es igual a 125,000-125 =124,875.

(III) *Permutaciones con elementos iguales*
Introduciremos la idea de permutaciones con elementos iguales por medio de un ejemplo.

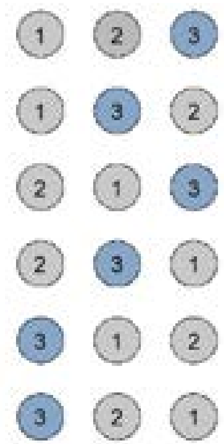
Ejemplo 1.29. Supongamos que tenemos los siguientes 3 elementos (2 de ellos son iguales).



¿Cuántas permutaciones diferentes existen?

Para realizar este cálculo, supondremos por el momento, que todos los elementos son diferentes, para esto les asociamos un número diferente a cada elemento.

Bajo este supuesto, de que los tres elementos son diferentes, sabemos que el número de permutaciones se calcula como: $3! = 6$. Estas se presentan a continuación:



Si se considera que son iguales aquellos elementos del mismo color, el número de permutaciones son únicamente 3. A continuación, se presentarán estas permutaciones

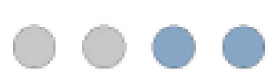
Las cuales pueden ser calculadas de la siguiente forma

$$\frac{3!}{2!1!} = 3.$$

En este caso, hay dos subgrupos de colores diferentes, uno de tamaño 2 y el otro de tamaño 1, los factoriales de estos números son los que aparecen en el denominador, en el anterior cálculo.

Para acabar de ilustrar las permutaciones con elementos iguales, consideremos otro ejemplo.

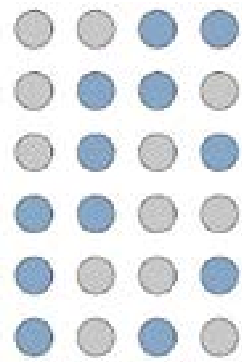
Ejemplo 1.30. Calcular el número de permutaciones que se pueden realizar con los siguientes elementos



Si los 4 elementos fueran diferentes, el número de permutaciones es

$$4! = 24.$$

Pero realmente, las permutaciones diferentes se dan a continuación:



Considerando que existen dos subgrupos, cada uno con 2 elementos iguales, se puede observar que el número de permutaciones diferentes pueden ser calculadas como

$$\frac{4!}{2!2!} = 6.$$

En general, el número de permutaciones diferentes de un conjunto que tiene n elementos donde este conjunto se divide en k subgrupos, el primer subgrupo tiene r_1 elementos iguales, el segundo subgrupo tiene r_2 elementos iguales,..., y donde el k-ésimo subgrupo tiene r_k elementos iguales, y además cualesquiera dos elementos de subgrupos diferentes son diferentes, se calcula de la siguiente manera:

$$\frac{n!}{r_1!r_2!\cdots r_k!}.$$

Obsérvese que $r_1 + r_2 + \ldots + r_k = n$.

Ejemplo 1.31. ¿Cuántas palabras diferentes pueden formarse con las letras de la palabra ESTADÍSTICA?

Suponiendo que las letras fueran diferentes, el número de permutaciones sería igual a 11! Pero las letras S, T, A e I, aparecen dos veces, por lo tanto, el número de palabras diferentes es igual a

$$\frac{11!}{2!2!2!2!} = 2{,}494{,}800.$$

Combinaciones

Consideremos un conjunto con un número de elementos finito, una *combinación* es un subconjunto formado con elementos de este conjunto.

El objetivo es contar el número de combinaciones (subconjuntos) diferentes de cierto tamaño que pueden formarse con los elementos de un determinado conjunto. A continuación, se ilustrará este cálculo con un ejemplo.

Ejemplo 1.32. Sea $A = \{a, b, c, d\}$, indicar las permutaciones y combinaciones diferentes de tamaño dos. Para este ejemplo, dar una relación entre el número de permutaciones y el número de combinaciones.

En la parte izquierda se mostrarán las permutaciones, y en la parte derecha las combinaciones.

$$ab, ba \to \{a, b\}$$

$$ac, ca \to \{a, c\}$$

$$ad, da \to \{a, d\}$$

$$bc, cb \to \{b, c\}$$

$$bd, db \to \{b, d\}$$

$$cd, dc \to \{c, d\}$$

Observemos que el número de permutaciones es

$$P_2^4 = 4(3) = 12.$$

Y el número de combinaciones es

$$\frac{P_2^4}{2!} = 6.$$

Esto es, por cada combinación pueden formarse 2! permutaciones, dicho de otra manera, por cada 2! permutaciones hay una combinación.

Para acabar de ilustrar el cálculo del número de combinaciones, se presentará otro ejemplo.

Ejemplo 1.33. Considerando de nuevo el conjunto $A = \{a, b, c, d\}$, considerar las permutaciones y las combinaciones de tamaño tres. Para este caso, dar una relación entre el número de permutaciones y el número de combinaciones.

En la parte izquierda se mostrarán las permutaciones, y en la parte derecha las combinaciones.

$$abc, acb, bac, bca, cab, cba \to \{a, b, c\}$$

$$abd, adb, bad, bda, dab, dba \to \{a, b, d\}$$

$$acd, adc, cad, cda, dac, dca \to \{a, c, d\}$$

$$bcd, bdc, cbd, cdb, dbc, dcb \to \{b, c, d\}$$

El número de permutaciones, en este caso es

$$P_3^4 = 4(3)(2) = 24.$$

El número de combinaciones es

$$\frac{P_3^4}{3!} = 4.$$

En este caso, por cada 3! permutaciones existe una combinación.

En términos generales, el número de combinaciones de tamaño r de un conjunto con n elementos se denota como $C_r^n = \binom{n}{r}$ y se calcula como

$$C_r^n = \binom{n}{r} = \frac{n!}{r!(n-r)!}.$$

La interpretación general es, por cada $r!$ permutaciones diferentes y con los mismos elementos, existe una combinación o por cada combinación, se pueden formar $r!$ permutaciones diferentes.

Los números $\binom{n}{r}$ son conocidos como *números combinatorios o coeficientes binomiales*. Véase la sección C.1 del apéndice para revisar algunas propiedades de estos números.

Es importante mencionar que los números combinatorios aparecen en el *teorema del binomio*, véase ecuación C.7, del apéndice.

También, los números combinatorios conforman el *triángulo de Pascal*, véase una parte de este triángulo en la sección C.1 del apéndice.

Regresando al ejemplo 1.19, las probabilidades se pueden calcular respectivamente de la siguiente manera:

$$\frac{C_2^2 C_0^2}{C_2^4} = \frac{1}{6},$$

$$\frac{C_2^2 C_0^3}{C_2^5} = \frac{1}{10},$$

$$\frac{C_2^2 C_0^8}{C_2^{10}} = \frac{1}{45}.$$

Después, del ejemplo 1.19 se consideró que la urna de este ejemplo tenía 20 bolas numeradas del 1 al 20 y se seleccionaron al azar 5 bolas, el número de puntos muestrales se calcula como C_5^{20} =15,504.

A continuación, se presentarán algunos ejemplos de cálculo de probabilidades, donde se aplicarán las diferentes técnicas de conteo.

Ejemplo 1.34. En una urna hay 15 bolas numeradas del 1 al 15. Se sacan 5 bolas en forma aleatoria, calcular la probabilidad de que las bolas obtenidas sean entre los primeros 10 números del 1 al 10.

La solución es

$$\frac{\binom{10}{5}\binom{5}{0}}{\binom{15}{5}} = 0.08392.$$

Ejemplo 1.35. ¿Cuántos subconjuntos de diferente tamaño se pueden formar de un conjunto con n elementos?

El número de subconjuntos de todos los tamaños es

$$\begin{aligned}\binom{n}{0}+\binom{n}{1}+\cdots+\binom{n}{n-1}+\binom{n}{n}&=\sum_{i=0}^{n}\binom{n}{i}\\&=\sum_{i=0}^{n}\binom{n}{i}1^i1^{n-i}\\&=2^n,\end{aligned}$$

aplicando el teorema del binomio en la última igualdad (véase ecuación C.7 del apéndice).

Este último resultado se puede usar para calcular la cardinalidad de un espacio de eventos. Considerando los ejemplos 1.14 y 1.15, la cardinalidad de cada espacio de eventos, se calcula como $2^8 = 256$ y 2^{30} =1,073,741,824, respectivamente.

Ejemplo 1.36. Se forman números de tres cifras utilizando los dígitos $0, 1, 2, 3, 4, 5, 6$, sin permitir la repetición de dígitos.
a) ¿Cuántos números se pueden formar?
b) ¿Cuántos de estos números son pares?
c) ¿Cuántos de estos números son mayores que 440?

a) Considerando que el primer dígito no puede ser 0, la cantidad de números de tres cifras que se pueden formar son

$$6 \times 6 \times 5 = 180.$$

b) Para contar el número de elementos pares, existen dos casos:
Caso 1. Cuando el número termina en cero, el número de maneras es

$$6 \times 5 \times 1 = 30.$$

Caso 2. Cuando el número termina en $2, 4, 6$, el número de casos es

$$5 \times 5 \times 3 = 75.$$

Por lo tanto, la respuesta es

$$30 + 75 = 105.$$

c) En este conteo existen dos casos:
Caso 1. Cuando el número comienza en 4, el número de maneras es

$$1 \times 2 \times 5 = 10.$$

Caso 2. Cuando el número comienza en 5 o 6, el número de maneras es

$$2 \times 6 \times 5 = 60.$$

Entonces, la respuesta es

$$10 + 60 = 70.$$

Se deja al estudiante contestar las mismas preguntas, pero considerando que sí se permite la repetición de los dígitos.

Ejemplo 1.37. Un estudiante tiene que contestar 12 preguntas de 15 en un examen. Si el estudiante selecciona las 12 preguntas a contestar en forma aleatoria, calcular la probabilidad de que el estudiante conteste
a) Las 5 primeras preguntas.
b) Por lo menos 4 preguntas de las últimas 5 preguntas.
c) A lo más 4 de las 5 preguntas intermedias (de la 6 a la 10).

a) La probabilidad pedida se calcula de la siguiente manera:

$$\frac{\binom{5}{5}\binom{10}{7}}{\binom{15}{12}} = 0.26374.$$

b) La respuesta es

$$\frac{\binom{5}{4}\binom{10}{8}}{\binom{15}{12}} + \frac{\binom{5}{5}\binom{10}{7}}{\binom{15}{12}} = 0.75824.$$

c) La probabilidad es

$$1 - \frac{\binom{5}{5}\binom{10}{7}}{\binom{15}{12}} = 0.73626.$$

Ejemplo 1.38. Se quieren formar aleatoriamente a 3 niñas y 2 niños. Calcular la probabilidad de que
a) Resulten intercalados los niños con las niñas.
b) Resulten los niños juntos.
c) Resulten los niños juntos y las niñas juntas.
d) De que no resulten Rosa y Pepe juntos.

El número de maneras de formar a los 5 niños, sin ninguna condición es

$$5 \times 4 \times 3 \times 2 \times 1 = 120.$$

a) El número de maneras de formar a los niños de tal manera que las niñas y niños queden intercalados es

$$3 \times 2 \times 2 \times 1 \times 1 = 12.$$

Por lo tanto,

$$P[\text{Intercalados niños y niñas}] = \frac{12}{120} = 0.1.$$

b) Se usará la siguiente notación: A que resulte niña y O que resulte niño. Existen 4 casos:
Caso 1. Cuando se forman O, O, A, A, A.
El número de maneras es

$$2 \times 1 \times 3 \times 2 \times 1 = 12.$$

Caso 2. Cuando se acomodan A, O, O, A, A.
El número de formas es

$$3 \times 2 \times 1 \times 2 \times 1 = 12.$$

Caso 3. Si se forman de la siguiente manera: A, A, O, O, A.
El número de maneras diferentes es

$$3 \times 2 \times 2 \times 1 \times 1 = 12.$$

Caso 4. Si se acomodan de la siguiente forma A, A, A, O, O.
El número de casos es

$$3 \times 2 \times 1 \times 2 \times 1 = 12.$$

El número total de maneras es

$$12 + 12 + 12 + 12 = 48.$$

Por lo tanto,

$$P[\text{Niños juntos }] = \frac{48}{120} = 0.4.$$

c) El número de maneras diferentes de sentar a las niñas juntas y a los niños juntos es 12+12=24 (casos 1 y 4 del inciso anterior), por lo tanto,

$$P[\text{Niñas juntas y Niños juntos}] = \frac{24}{120} = 0.2.$$

d) El número de maneras en que Rosa y Pepe queden juntos es igual a 48 (misma respuesta que el inciso b de este ejercicio), entonces, el número de maneras en que Rosa y Pepe no queden juntos es $120 - 48 = 72$, por lo tanto,

$$P[\text{Rosa y Pepe no queden juntos}] = \frac{72}{120} = 0.6.$$

Ejemplo 1.39. Se escoge aleatoriamente una palabra de todas las que se pueden formar con las letras de la palabra MATEMÁTICA (con todas las letras). ¿Cuál es la probabilidad de que la palabra escogida empiece con la letra I?

El número de palabras diferentes es

$$\frac{10!}{2!3!2!} = 151{,}200.$$

El número de palabras diferentes que empiezan con la letra I, suponiendo que las todas las letras son diferentes, es

$$1 \times 9 \times 8 \times \cdots \times 1 = 9!$$

Así, el número de palabras diferentes que inician con la letra I es

$$\frac{9!}{2!3!2!} = 15{,}120.$$

Por lo tanto,

$$P[\text{Empiece con I}] = \frac{9!}{10!} = 0.1.$$

Ejemplo 1.40. Una delegación de 4 estudiantes de un posgrado es seleccionada todos los años. Si esta se escoge en forma aleatoria de un grupo de 12 estudiantes, donde se sabe que dos de ellos están casados. Calcular las siguientes probabilidades:
a) Que no sean escogidos al mismo tiempo los estudiantes casados.
b) Que sean escogidos al mismo tiempo los estudiantes casados.

a) La respuesta es

$$\frac{\binom{10}{4}\binom{2}{0} + \binom{10}{3}\binom{2}{1}}{\binom{12}{4}} = 0.909.$$

b) La probabilidad es

$$\frac{\binom{10}{2}\binom{2}{2}}{\binom{12}{4}} = 0.090909.$$

Ejemplo 1.41. Seis parejas de casados, formadas cada una por una mujer y un hombre, se encuentran en un salón.

a) Si se escogen dos personas al azar del salón, calcular la probabilidad de que resulten
   I) Esposos.
   II) Un hombre y una mujer.

b) Si se escogen cuatro personas al azar del salón, calcular la probabilidad de que resulten
   I) Dos parejas de casados.
   II) Ninguna pareja de casados se encuentre entre los cuatro.
   III) Exactamente una pareja de casados entre los cuatro.

a) I) El resultado es

$$P[\text{Esposos}] = \frac{\binom{2}{2}\binom{10}{0}\binom{6}{1}}{\binom{12}{2}} = 0.09091.$$

a) II) La probabilidad pedida es

$$P[\text{Un hombre y una mujer}] = \frac{\binom{6}{1}\binom{6}{1}}{\binom{12}{2}} = 0.54545.$$

b) I) La probabilidad se calcula de la siguiente manera:

$$P[\text{Dos parejas de casados}] = \frac{\binom{2}{2}\binom{2}{2}\binom{8}{0}\binom{6}{2}}{\binom{12}{4}} = 0.0303.$$

b) II) El resultado es

$$P[\text{Ninguna pareja de casados}] = \frac{\binom{2}{1}\binom{2}{1}\binom{2}{1}\binom{2}{1}\binom{4}{0}\binom{6}{4}}{\binom{12}{4}} = 0.48485.$$

b) III) El cálculo de la probabilidad es

$$P[\text{Una pareja de casados}] = \frac{\binom{2}{2}\binom{2}{1}\binom{2}{1}\binom{4}{0}\binom{6}{3} \times 3}{\binom{12}{4}} = 0.48485.$$

Ejemplo 1.42. De una baraja ordinaria de 52 cartas se sacan en forma aleatoria 5 cartas. Calcular la probabilidad de obtener:
a) A lo más dos ases.
b) Exactamente dos reyes y un as.
c) Cuatro cartas del mismo valor.
d) *Full house* (un par y una tercia).
e) Exactamente dos pares.

a) La probabilidad pedida se calcula

$$\begin{aligned} P[\text{A lo más dos ases}] &= 1 - P[3 \text{ o } 4 \text{ ases}] \\ &= 1 - \left\{ \frac{\binom{4}{3}\binom{48}{2} + \binom{4}{4}\binom{48}{1}}{\binom{52}{5}} \right\} = 0.99825. \end{aligned}$$

b) El cálculo de la probabilidad pedida se da continuación:

$$P[\text{Dos reyes y un as}] = \frac{\binom{4}{2}\binom{4}{1}\binom{44}{2}}{\binom{52}{5}} = 0.00041.$$

c) El resultado es

$$P[\text{Cuatro cartas del mismo valor}] = \frac{\binom{4}{4}\binom{48}{1}\binom{13}{1}}{\binom{52}{5}} = 0.00024.$$

d) La probabilidad se calcula de la siguiente manera:

$$P[\text{Un par y una tercia}] = \frac{\binom{4}{2}\binom{4}{3} \times 2 \times \binom{13}{2}}{\binom{52}{5}} = 0.00144.$$

e) La respuesta es

$$P[\text{Exactamente dos pares}] = \frac{\binom{4}{2}\binom{4}{2}\binom{44}{1}\binom{13}{2}}{\binom{52}{5}} = 0.04754.$$

Ejemplo 1.43. Una persona distribuye n bolas aleatoriamente en n cajas. Cada bola puede caer en cualquiera de la n cajas. Calcular la probabilidad de que la caja 1 y solamente la caja 1 quede vacía.

El número de resultados posibles es $\underbrace{n \times n \times \cdots \times n}_{n \quad \text{veces}} = n^n$.

Fijamos momentáneamente la caja 2 para que resulten dos bolas y en el resto de las cajas una bola, excepto en la caja 1 que quedará vacía. Por lo tanto, el número de maneras diferentes de que resulten dos bolas en la caja 2 es $\binom{n}{2}$.

De esta forma, el número de maneras de que resulte una bola en la tercera caja es $n-2$, y que resulte una bola en la cuarta caja es $n-3, \ldots$, y resulte una bola en la última caja es 1. Por lo tanto, el número de resultados diferentes con 2 bolas en la segunda caja y una bola en cada una de las cajas restantes, excepto en la primera caja es

$$\binom{n}{2} \times (n-2) \times (n-3) \times \cdots \times (1) = \binom{n}{2} \times (n-2)!$$

Pero como son $n-1$ cajas (no solamente la caja 2) para que resulten 2 bolas, por lo tanto, la probabilidad pedida es igual a

$$\frac{(n-1)\binom{n}{2}(n-2)!}{n^n} = \frac{\binom{n}{2}(n-1)!}{n^n}.$$

1.11. Probabilidad condicional

El concepto de probabilidad condicional es sumamente importante en temas de procesos estocásticos, de estadística y de la misma probabilidad. Se motivará la definición de probabilidad condicional por medio de un ejemplo.

Ejemplo 1.44. Se lanzan dos dados, uno rojo y otro azul. Se definen los siguientes eventos $A = \{\text{Ambos números son impares}\}; B = \{\text{El número del dado rojo es menor que el número del dado azul}\}$. Encontrar
a) $P[A]$, $P[B]$ y $P[A \cap B]$.
b) $P[B$ dado que se sabe que A ocurrió $] = P[B|A]$.
c) $P[A$ dado que se sabe que B ocurrió $] = P[A|B]$.

a) Obsérvese que
$A = \{(1,1),(1,3),(1,5),(3,1),(3,1),(3,3),(3,5),(5,1),(5,3),(5,5)\}$;
$B = \{(1,2),(1,3),(1,4),(1,5),(1,6),(2,3),(2,4),(2,5),(2,6),(3,4),$
$(3,5),(3,6),(4,5),(4,6),(5,6)\}$ y $A \cap B = \{(1,3),(1,5),(3,5)\}$.
Por lo tanto,

$$P[A] = \frac{9}{36} = \frac{1}{4}; P[B] = \frac{15}{36} = \frac{5}{12}.$$

Además,

$$P[A \cap B] = \frac{3}{36} = \frac{1}{12}.$$

b) Ahora, si se supone que el evento A ocurrió, entonces, los resultados posibles se reducen únicamente a los nueve elementos del evento A, y tres de estos nueve favorecen al evento B, entonces,

$$P[B|A] = \frac{3}{9} = \frac{1}{3}.$$

c) Por otro lado,

$$P[A|B] = \frac{3}{15} = \frac{1}{5}.$$

Observación. Las probabilidades $P[A|B]$ y $P[B|A]$ (probabilidades condicionales) se pueden recalcular de la siguiente manera:

$$P[B|A] = \frac{\frac{3}{36}}{\frac{9}{36}} = \frac{P[A \cap B]}{P[A]},$$

$$P[A|B] = \frac{\frac{3}{36}}{\frac{15}{36}} = \frac{P[A \cap B]}{P[B]}.$$

A continuación, se presentará la definición de probabilidad condicional.

Definición 1.19. Sean A y B dos eventos, la *probabilidad condicional* del evento A dado el evento B es denotada por $P[A|B]$, y definida

$$P[A|B] = \frac{P[A \cap B]}{P[B]}, \quad \text{si} \quad P[B] > 0.$$

Si $P[B] = 0$, la probabilidad condicional queda indefinida.

La *probabilidad condicional* del evento B dado el evento A es denotada por $P[B|A]$, y es definida

$$P[B|A] = \frac{P[A \cap B]}{P[A]}, \quad \text{si} \quad P[A] > 0.$$

Si $P[A] = 0$, la probabilidad condicional queda indefinida.

Observación. En una probabilidad condicional, el espacio muestral Ω se reduce a los puntos muestrales del evento dado. Lo que significa que una probabilidad condicional sí es una probabilidad, pero con espacio muestral diferente, y por lo tanto, los axiomas de probabilidad se cumplen, esto es,

(i) $P[A|B] \geq 0$.
(ii) $P[\Omega|B] = 1$.
(iii) $P[\bigcup_{i=1}^{\infty} A_i|B] = \sum_{i=1}^{\infty} P[A_i|B]$, si $A_i \cap A_j = \emptyset$, para $i \neq j$.

Así mismo, las probabilidades condicionales, cumplen también cada una de las propiedades de probabilidad, esto es,

a) $P[\emptyset|B] = 0$.
b) $P[\bigcup_{i=1}^{n} A_i|B] = \sum_{i=1}^{n} P[A_i|B]$, si $A_i \cap A_j = \emptyset$, para $i \neq j$.
c) $P[A^c|B] = 1 - P[A|B]$.
d) $P[A_1 \cup A_2|B] = P[A_1|B] + P[A_2|B] - P[A_1 \cap A_2|B]$.
e) Si $A_1 \subset A_2$, entonces, $P[A_1|B] \leq P[A_2|B]$.
f) $P[\bigcup_{i=1}^{n} A_i|B] \leq \sum_{i=1}^{n} P[A_i|B]$.

Se deja al estudiante la comprobación de cada una de las propiedades anteriores.

Ejemplo 1.45. En una urna hay 3 bolas blancas y 4 rojas. Se sacan en forma aleatoria dos de ellas. Calcular la probabilidad de que resulten ambas bolas rojas dado que se sabe que por lo menos una bola roja resultó en la muestra.

Definamos los siguientes eventos: $A = \{\text{ambas bolas son rojas}\}$ y $B = \{\text{por lo menos una bola roja}\}$.

Por lo tanto,

$$P[A] = \frac{\binom{4}{2}\binom{3}{0}}{\binom{7}{2}} = 0.28571,$$

$$P[B] = 1 - \frac{\binom{4}{0}\binom{3}{2}}{\binom{7}{2}} = 0.85714,$$

$$P[A|B] = \frac{P[A \cap B]}{P[B]} = \frac{P[A]}{P[B]} = 0.33333.$$

La solución de esta probabilidad, también se puede realizar por espacio reducido, esto es, el espacio muestral se reduce al evento dado B.

Para describir el espacio muestral y los eventos de este problema, usaremos la siguiente notación: r_i = bola i de color rojo, $i = 1, 2, 3, 4$ y b_i=bola i de color blanco, $i = 1, 2, 3$.

De esta manera, se puede describir el espacio muestral

$$\Omega = \{\{r_1, r_2\}, \{r_1, r_3\}, \{r_1, r_4\}, \{r_1, b_1\}, \{r_1, b_2\}, \{r_1, b_3\}, \{r_2, r_3\}, \{r_2, r_4\},$$

$$\{r_2, b_1\}, \{r_2, b_2\}, \{r_2, b_3\}, \{r_3, r_4\}, \{r_3, b_1\}, \{r_3, b_2\}, \{r_3, b_3\}, \{r_4, b_1\}, \{r_4, b_2\},$$

$$\{r_4, b_3\}, \{b_1, b_2\}, \{b_1, b_3\}, \{b_2, b_3\}\}.$$

Los eventos A y B se describen a continuación:

$$A = \{\{r_1, r_2\}, \{r_1, r_3\}, \{r_1, r_4\}, \{r_2, r_3\}, \{r_2, r_4\}, \{r_3, r_4\}\},$$

$$B = \{\{r_1, r_2\}, \{r_1, r_3\}, \{r_1, r_4\}, \{r_1, b_1\}, \{r_1, b_2\}, \{r_1, b_3\}, \{r_2, r_3\}, \{r_2, r_4\},$$
$$\{r_2, b_1\}, \{r_2, b_2\}, \{r_2, b_3\}, \{r_3, r_4\}, \{r_3, b_1\}, \{r_3, b_2\}, \{r_3, b_3\}, \{r_4, b_1\}, \{r_4, b_2\},$$
$$\{r_4, b_3\}\}.$$

Por lo tanto,

$$P[A|B] = \frac{6}{18} = 0.33333.$$

1.12. Regla de la multiplicación

A continuación, se presentará la regla de multiplicación de probabilidades, esta se obtiene a partir de la definición de probabilidad condicional y es de utilidad para calcular la probabilidad de la intersección de dos eventos.

$$P[A \cap B] = P[A]P[B|A], \quad \text{si } P[A] > 0,$$

$$P[A \cap B] = P[B]P[A|B], \quad \text{si } P[B] > 0.$$

Las igualdades anteriores se les conoce como la *regla de la multiplicación de probabilidades.* A continuación, se presenta la generalización para n eventos.

Teorema 1.8. *Sean $A_1, \ldots, A_n$ eventos tal que $P[A_1 \cap A_2 \cap \cdots \cap A_{n-1}] > 0$, entonces,*

$$P[A_1 \cap \cdots \cap A_n] = P[A_1]P[A_2|A_1]P[A_3|A_1 \cap A_2] \cdots P[A_n|A_1 \cap \cdots \cap A_{n-1}].$$

La demostración se realiza por inducción matemática y se queda como ejercicio para el estudiante.

Ejemplo 1.46. En una urna se tienen tres bolas blancas y dos negras. Se sacan dos bolas en forma aleatoria una tras otra, sin reemplazo. ¿Cuál es la probabilidad de que las dos bolas sean blancas?

Sean $A = \{\text{Bola 1 sea blanca}\}$ y $B = \{\text{Bola 2 sea blanca}\}$.

De esta manera,

$$P[A \cap B] = P[A]P[B|A] = \frac{3}{5} \times \frac{2}{4} = 0.3.$$

Ejemplo 1.47. Considerando la misma urna del ejemplo 1.46, ahora se sacan cuatro bolas en forma aleatoria, una tras otra, sin reemplazo. ¿Cuál es la probabilidad de que las tres primeras sean blancas y la última sea negra?

Sean $A_1 = \{\text{Bola 1 blanca}\}$; $A_2 = \{\text{Bola 2 blanca}\}$; $A_3 = \{\text{Bola 3 blanca}\}$ y $A_4 = \{\text{Bola 4 negra}\}$.

De este modo,

$$\begin{aligned} P[A_1 \cap A_2 \cap A_3 \cap A_4] &= P[A_1]P[A_2|A_1]P[A_3|A_1 \cap A_2]P[A_4|A_1 \cap A_2 \cap A_3] \\ &= \frac{3}{5} \times \frac{2}{4} \times \frac{1}{3} \times \frac{2}{2} = 0.1. \end{aligned}$$

Ejemplo 1.48. De una urna que contiene 5 bolas negras y 3 verdes se sacan sucesivamente 3 bolas, sin reemplazo. Calcular la probabilidad de que
a) Las bolas sean del mismo color.
b) Cada color quede representado.

Sean $C = \{\text{3 del mismo color}\}$; $D = \{\text{Cada color quede representado}\}$.

a) De esta manera,

$$\begin{aligned} P[C] &= P[\text{3 bolas negras o 3 bolas verdes}] \\ &= P[\text{3 bolas negras}] + P[\text{3 bolas verdes}] \\ &= \frac{5}{8} \times \frac{4}{7} \times \frac{3}{6} + \frac{3}{8} \times \frac{2}{7} \times \frac{1}{6} = 0.19643. \end{aligned}$$

b) El resultado es

$$\begin{aligned} P[D] &= P[(\text{1 negra y 2 verdes}) \text{ o } (\text{2 negras y 1 verde})] \\ &= P[\text{1 negra y 2 verdes }] + P[\text{ 2 negras y 1 verde}] \\ &= \frac{5}{8} \times \frac{3}{7} \times \frac{2}{6} + \frac{3}{8} \times \frac{5}{7} \times \frac{2}{6} + \frac{3}{8} \times \frac{2}{7} \times \frac{5}{6} \\ &+ \frac{3}{8} \times \frac{5}{7} \times \frac{4}{6} + \frac{5}{8} \times \frac{3}{7} \times \frac{4}{6} + \frac{5}{8} \times \frac{4}{7} \times \frac{3}{6} = 0.80357. \end{aligned}$$

Observación. La segunda respuesta también se puede calcular por medio del uso de combinaciones, obsérvese en la última respuesta, los tres primeros sumandos se refieren a la probabilidad de obtener una bola negra y dos bolas verdes, y los tres últimos sumandos se refieren a la probabilidad de obtener dos negras y una verde; entonces, la probabilidad del evento D es posible calcularla también de la siguiente manera:

$$P[D] = \frac{\binom{5}{1}\binom{3}{2}}{\binom{8}{3}} + \frac{\binom{5}{2}\binom{3}{1}}{\binom{8}{3}} = 0.80357.$$

1.13. Teorema de la probabilidad total

En esta sección se tratará el teorema de la probabilidad total, el cual será de gran utilidad para el cálculo de algunas probabilidades. Previamente se va a presentar el concepto de partición de un espacio muestral, el cual es importante para este teorema y también para el teorema de Bayes.

Definición 1.20. Los eventos $B_1, B_2, \ldots, B_n$ forman una *partición* del espacio muestral Ω, si cumplen las siguientes 3 condiciones:

(I) $P[B_i] > 0$, para toda $i = 1, 2, \ldots, n$,

(II) $B_i \cap B_j = \emptyset$, para toda $i \neq j$,

(III) $\displaystyle\bigcup_{i=1}^{n} B_i = \Omega$.

Observación. Si los eventos $B_1, B_2, \ldots, B_n$ cumplen (III), entonces, se dice que los eventos son *colectivamente exhaustivos.*

A continuación, se presentará el teorema de la probabilidad total, el cual se puede aplicar para situaciones donde existe una partición del espacio muestral.

Teorema 1.9. *Sean $B_1, B_2, \ldots, B_n$ eventos que forman una partición del espacio muestral Ω y sea A un evento cualesquiera, entonces,*

$$P[A] = \sum_{i=1}^{n} P[A|B_i]P[B_i].$$

Demostración. Al ser $B_1, B_2, \ldots, B_n$ una partición de Ω, el evento A puede describirse de la siguiente manera:

$$A = \bigcup_{i=1}^{n} (A \cap B_i), \quad \text{donde} \quad (A \cap B_i) \cap (A \cap B_j) = \emptyset, \text{para toda} i \neq j.$$

Por lo tanto,

$$\begin{aligned} P[A] &= \sum_{i=1}^{n} P[A \cap B_i] \\ &= \sum_{i=1}^{n} P[A|B_i]P[B_i] \quad \text{(por el teorema 1.8).} \end{aligned}$$ □

Ejemplo 1.49. De una urna con 3 bolas blancas y 2 negras, se seleccionan en forma aleatoria

a) Dos bolas una tras otra sin sustitución. ¿Cuál es la probabilidad de que la segunda bola sea negra?

b) Tres bolas sin reemplazo. ¿Cuál es probabilidad de que la tercera bola sea blanca?

a) Consideremos $A = \{\text{Bola 2 sea negra}\}$; $B_1 = \{\text{Bola 1 sea blanca}\}$ y $B_2 = \{\text{Bola 1 sea negra}\}$.

De esta manera,

$$\begin{aligned} P[A] &= P[A|B_1]P[B_1] + P[A|B_2]P[B_2] \\ &= \frac{1}{2} \times \frac{3}{5} + \frac{1}{4} \times \frac{2}{5} = 0.4. \end{aligned}$$

b) Consideremos la siguiente notación: $A = \{\text{Bola 3 sea blanca}\}$; $B_1 = \{\text{Bola 1 y bola 2 sean blancas}\}$; $B_2 = \{\text{Bola 1 y bola 2 sean negras}\}$; $B_3 = \{\text{Bola 1 sea blanca y bola 2 sea negra}\}$ y $B_4 = \{\text{Bola 1 sea negra y bola 2 sea blanca}\}$.

Entonces,

$$\begin{aligned} P[A] &= P[A|B_1]P[B_1] + P[A|B_2]P[B_2] + P[A|B_3]P[B_3] + P[A|B_4]P[B_4] \\ &= \frac{1}{3} \times \frac{3}{5} \times \frac{2}{4} + \frac{3}{3} \times \frac{2}{5} \times \frac{1}{4} + \frac{2}{3} \times \frac{3}{5} \times \frac{2}{4} + \frac{2}{3} \times \frac{2}{5} \times \frac{3}{4} = \frac{3}{5} = 0.6. \end{aligned}$$

Observación. En la probabilidad anterior, se puede observar que cada evento B_i es la intersección de dos eventos, y para el cálculo de su probabilidad, se aplicó la regla de multiplicación de probabilidades.

Ejemplo 1.50. Para cierta acción financiera, se sabe que, inicialmente, la acción sube con una probabilidad de 0.6 y baja con una probabilidad de 0.4. Para cualquier otro periodo se sabe que la acción subirá con una probabilidad de 0.7 si en el anterior periodo la acción subió, y la acción subirá con una probabilidad de 0.4, si en el anterior periodo bajó. Calcular la probabilidad de que la acción suba, en el tercer periodo.

Consideremos la notación: $A_i = \{\text{La acción sube en el periodo } i\}$, donde $i = 1, 2, 3$.

La probabilidad se calcula de la siguiente manera:

$$\begin{aligned} P[A_3] &= P[A_3|A_1 \cap A_2]P[A_1 \cap A_2] + P[A_3|A_1 \cap A_2^c]P[A_1 \cap A_2^c] \\ &\quad + P[A_3|A_1^c \cap A_2]P[A_1^c \cap A_2] + P[A_3|A_1^c \cap A_2^c]P[A_1^c \cap A_2^c] \\ &= P[A_3|A_2]P[A_1]P[A_2|A_1] + P[A_3|A_2^c]P[A_1]P[A_2^c|A_1] \\ &\quad + P[A_3|A_2]P[A_1^c]P[A_2|A_1^c] + P[A_3|A_2^c]P[A_1^c]P[A_2^c|A_1^c] \\ &= (0.7)(0.6)(0.7) + (0.4)(0.6)(0.3) + (0.7)(0.4)(0.4) \\ &\quad + (0.4)(0.4)(0.6) = 0.574. \end{aligned}$$

1.14. Teorema de Bayes

En esta sección se presentará y se demostrará el teorema de Bayes, el cual es de mucha utilidad para calcular probabilidades condicionales, además es el fundamento de la teoría de la estadística bayesiana.

Teorema 1.10. *Sean $B_1, B_2, \ldots, B_n$ eventos que forman una partición del espacio muestral Ω y sea A un evento cualquiera, entonces,*

$$P[B_i|A] = \frac{P[A|B_i]P[B_i]}{\sum_{i=1}^{n} P[A|B_i]P[B_i]}, \quad \textit{para toda } i = 1, \ldots, n.$$

Demostración.

$$
\begin{aligned}
P[B_i|A] &= \frac{P[A \cap B_i]}{P[A]} \\
&= \frac{P[A|B_i]P[B_i]}{\sum_{i=1}^{n} P[A|B_i]P[B_i]}. \qquad \square
\end{aligned}
$$

En la demostración anterior, en el numerador se aplicó la regla de multiplicación y en el denominador el teorema de la probabilidad total.

Ejemplo 1.51. Considerando el ejemplo 1.49. Se sabe que la segunda bola fue negra, calcular la probabilidad de que la primera bola sea negra.

Consideremos la notación: $A = \{\text{Bola 2 negra}\}$; $B_1 = \{\text{Bola 1 blanca}\}$ y $B_2 = \{\text{Bola 1 negra}\}$.

Entonces,

$$
\begin{aligned}
P[B_2|A] &= \frac{P[A|B_2]P[B_2]}{P[A|B_1]P[B_1] + P[A|B_2]P[B_2]} \\
&= \frac{\frac{1}{4} \times \frac{2}{5}}{\frac{2}{4} \times \frac{3}{5} + \frac{1}{4} \times \frac{2}{5}} = 0.25.
\end{aligned}
$$

Ejemplo 1.52. El departamento de créditos en un banco realizó un análisis de sus clientes que tienen asignada una tarjeta de crédito. De este análisis, se llegaron a las siguientes conclusiones: el 7.5 % de sus clientes se han demorado con sus pagos mensuales en dos meses seguidos o más en alguna ocasión. De sus clientes que no han incumplido o se han demorado solo por una mensualidad, el 95 % realizan su siguiente mensualidad sin demora alguna, y de sus clientes que sean demorado 2 o más veces seguidas, el 55 % de las veces se demoran en su siguiente mensualidad. Se selecciona al azar un cliente de tarjeta de crédito. Calcular la probabilidad

a) De que el cliente no cumpla con su siguiente pago.

b) De que sea un cliente que se ha demorado dos o más veces seguidas en alguna ocasión, dado que se observó que se demoró en su último pago.

Consideremos la siguiente notación: $C = \{\text{Cliente cumplido o a lo más una mensualidad con demora}\}$; $I = \{\text{Cliente incumplido con dos o más mensualidades vencidas}\}$; $MC = \{\text{Cumple con la siguiente mensualidad}\}$ y $MI = \{\text{No cumple con la siguiente mensualidad}\}$.

a) La probabilidad de que el cliente no cumpla su siguiente mensualidad

$$
\begin{aligned}
P[MI] &= P[MI|C]P[C] + P[MI|I]P[I] \\
&= (0.05)(0.925) + (0.55)(0.075) = 0.0875.
\end{aligned}
$$

b) La probabilidad solicitada se calcula de la siguiente manera:

$$
\begin{aligned}
P[I|MI] &= \frac{P[MI|I]P[I]}{P[MI|I]P[I] + P[MI|C]P[C]} \\
&= \frac{0.55 \times 0.075}{0.0875} = 0.47143.
\end{aligned}
$$

Ejemplo 1.53. En una población se sabe que el 0.01 de la gente tiene una enfermedad. Una prueba para detectar dicha enfermedad tiene las siguientes características: si una persona tiene la enfermedad, la prueba lo indicará con una probabilidad de 0.99, y si no la tiene, la prueba indicará lo contrario con una probabilidad del 0.02. Se selecciona una persona en forma aleatoria de dicha población y se le realiza la prueba. La prueba indica que tiene la enfermedad. Calcular la probabilidad de que realmente la persona esté enferma.

Consideremos la siguiente notación: $E = \{\text{La persona está enferma}\}$; $S = \{\text{La persona está sana}\}$ y $P = \{\text{La prueba indica positivo}\}$.

$$
\begin{aligned}
P[E|P] &= \frac{P[P|E]P[E]}{P[P|E]P[E] + P[P|S]P[S]} \\
&= \frac{0.99 \times 0.01}{0.99 \times 0.01 + 0.02 \times 0.99} = 0.33333.
\end{aligned}
$$

1.15. Independencia

En esta sección se tratará el concepto de independencia de dos o más eventos. Cabe señalar que la independencia de eventos es de suma importancia para el cálculo de probabilidades y, en temas de inferencia estadística, es un concepto fundamental.

A continuación, se dará la definición de independencia de dos eventos.

Definición 1.21. Dos eventos A y B son *independientes* si cumplen alguna de las siguientes igualdades:

(i) $P[A \cap B] = P[A]P[B]$.

(ii) $P[A|B] = P[A]$, si $P[B] > 0$.

(iii) $P[B|A] = P[B]$, si $P[A] > 0$.

Observación. Si los eventos A y B cumplen alguna de las tres condiciones anteriores, es más que suficiente para poder afirmar que los eventos son independientes. En otras palabras, las tres igualdades anteriores son definiciones equivalentes de independencia de dos eventos. La demostración queda como ejercicio para el estudiante.

Ejemplo 1.54. Se lanzan dos dados y se definen los siguientes eventos: $A = \{\text{El primer dado es par}\}$ y $B = \{\text{La suma de los dados es par}\}$. ¿Los eventos A y B son independientes?

Obsérvese

$$P[A] = \frac{18}{36} = \frac{1}{2}; P[B] = \frac{18}{36} = \frac{1}{2} \text{ y } P[A \cap B] = \frac{9}{36} = \frac{1}{4}.$$

De este modo,

$$P[A \cap B] = P[A]P[B].$$

Por lo tanto, A y B son eventos independientes.

La definición de independencia de eventos se puede generalizar para más de dos eventos, véase la siguiente definición.

Definición 1.22. Los eventos $A_1, A_2, \ldots A_n$ son *independientes*, si cumplen las siguientes igualdades:

$$\begin{aligned} P[A_{i_1} \cap A_{i_2}] &= P[A_{i_1}]P[A_{i_2}], \text{ para } i_1 \neq i_2, \\ P[A_{i_1} \cap A_{i_2} \cap A_{i_3}] &= P[A_{i_1}]P[A_{i_2}]P[A_{i_3}], \text{ para } i_1 \neq i_2 \neq i_3, \\ &\vdots \\ P[A_1 \cap A_2 \cap \cdots \cap A_n] &= P[A_1]P[A_2] \cdots P[A_n]. \end{aligned}$$

Observación. Para verificar que los eventos A, B y C sean independientes, se tienen que cumplir las siguientes 4 igualdades:

$$P[A \cap B] = P[A]P[B], \tag{1.1}$$

$$P[A \cap C] = P[A]P[C], \tag{1.2}$$

$$P[B \cap C] = P[B]P[C], \tag{1.3}$$

$$P[A \cap B \cap C] = P[A]P[B]P[C]. \tag{1.4}$$

Obsérvese que para verificar que los eventos $A_1, A_2, \ldots A_n$ sean independientes, se tienen que verificar $2^n - n - 1$ igualdades.

Ejemplo 1.55. Se arrojan dos dados al aire y se definen los siguientes eventos $A = \{\text{El primer dado es par}\}$, $B = \{\text{El segundo dado es par}\}$ y $C = \{\text{El segundo dado es mayor que } 4\}$. ¿Son los eventos A, B y C independientes?

Se calcularán las probabilidades correspondientes

$$P[A] = \frac{1}{2}; P[B] = \frac{1}{2}; P[C] = \frac{1}{3}; P[A \cap B] = \frac{9}{36} = \frac{1}{4};$$

$$P[A \cap C] = \frac{6}{36} = \frac{1}{6}; P[B \cap C] = \frac{6}{36} = \frac{1}{6} \text{ y } P[A \cap B \cap C] = \frac{3}{36} = \frac{1}{12}.$$

Así que,

$$P[A \cap B] = P[A]P[B] = \frac{1}{4},$$
$$P[A \cap C] = P[A]P[C] = \frac{1}{6},$$
$$P[B \cap C] = P[B]P[C] = \frac{1}{6},$$
$$P[A \cap B \cap C] = P[A]P[B]P[C] = \frac{1}{12}.$$

Por lo tanto, los eventos A, B y C son independientes.

Observación. El hecho de que se cumpla la igualdad 1.4, no garantiza de que los eventos A, B y C sean independientes, véase el siguiente ejemplo.

Ejemplo 1.56. Se lanzan dos dados al aire, y se definen los siguientes eventos: $A = \{$El primer dado es mayor que el segundo dado$\}$, $B = \{$La suma de los dos dados es igual a 10$\}$, $C = \emptyset$. Verificar si los eventos A, B y C son independientes.

Obsérvese

$$P[A \cap B \cap C] = P[\emptyset] = 0 = P[A]P[B]P[C].$$

Esto es, se cumple la igualdad 1.4.
Ahora, obsérvese

$$P[A] = \frac{15}{36} = \frac{5}{12}; P[B] = \frac{3}{36} = \frac{1}{12} \text{ y } P[A \cap B] = \frac{1}{36}.$$

De este modo,

$$P[A \cap B] \neq P[A]P[B].$$

Se concluye que los eventos A, B y C no son independientes.

Observación. La independencia por parejas de los eventos A, B y C no implica la independencia de los tres eventos, véase el siguiente ejemplo.

Ejemplo 1.57. Se lanzan dos dados al aire y se definen los siguientes eventos $A = \{$El primer dado es par$\}$; $B = \{$El segundo dado es par$\}$ y $C = \{$El total de ambos resultados es par$\}$.

Obsérvese

$$P[A] = \frac{1}{2}; P[B] = \frac{1}{2}; P[C] = \frac{1}{2};$$
$$P[A \cap B] = \frac{1}{4}; P[A \cap C] = \frac{1}{4} \text{ y } P[B \cap C] = \frac{1}{4}.$$

Se puede notar que los eventos A, B y C son independientes por parejas, pero

$$P[A \cap B \cap C] = \frac{1}{4} \neq \frac{1}{8} = P[A]P[B]P[C].$$

En general, los tres eventos no son mutuamente independientes.

La interpretación de la anterior situación es que la ocurrencia de algún evento no afecta a la probabilidad de algún otro evento. Pero la ocurrencia de dos eventos posiblemente afecte la ocurrencia del tercer evento, o también, la ocurrencia de un evento posiblemente afecte la ocurrencia simultánea de los otros dos eventos. Para aclarar este comentario, del anterior ejemplo, observemos las siguientes probabilidades

$$P[A] = \frac{1}{2};\ P[A|B] = \frac{1}{2} \text{ y } P[A|B \cap C] = 1.$$

Esto es, la ocurrencia de B no afectó la probabilidad de A, pero la ocurrencia de B y C sí modificó la probabilidad de A.

Teorema 1.11. *Si A y B son eventos independientes, entonces, los eventos A y B^c también son independientes.*

Demostración.

$$\begin{aligned} P[A \cap B^c] &= P[A] - P[A \cap B] \\ &= P[A] - P[A]P[B] \\ &= P[A](1 - P[B]) \\ &= P[A]P[B^c]. \end{aligned}$$

Por lo tanto, los eventos A y B^c son independientes. □

El teorema anterior indica que si los eventos A y B son independientes, entonces, los eventos A y B^c también lo son, y si estos dos últimos son independientes, aplicando de nuevo el teorema, entonces, los eventos A^c y B^c también son independientes, y aplicando otra vez el teorema, podemos afirmar que los eventos A^c y B también lo son.

Ejemplo 1.58. Si la probabilidad de que Pepe todavía esté vivo al cabo de 10 años es de 0.8 y la probabilidad de que María sobreviva el mismo lapso es de 0.95. ¿Cuál es la probabilidad de que ninguno sobreviva 10 años?

Es importante aclarar que las vidas de Pepe y de María son eventos independientes, supuesto que en la matemática actuarial también se hace, las vidas de dos personas son independientes. Este supuesto se hace para calcular la probabilidad conjunta de muerte (o de vida) de dos o más personas.

Consideremos la siguiente notación: $P = \{\text{Pepe sobreviva 10 años}\}$ y $M = \{\text{María sobreviva 10 años}\}$.

Por lo tanto, la probabilidad solicitada se calcula de la siguiente manera:

$$\begin{aligned} P[P^c \cap M^c] &= P[P^c]P[M^c] \\ &= 0.2 \times 0.05 = 0.01. \end{aligned}$$

Ejemplo 1.59. Un jugador de baloncesto es exitoso en el 50 % de sus tiros, además, se supone independencia entre los diferentes tiros que él realiza. De sus cuatro tiros siguientes, calcular la probabilidad de que enceste
a) Tres tiros.
b) Por lo menos un tiro.
c) Dos tiros.

a) Encestar tres tiros de cuatro, significa no encestar uno, este tiro puede ser el primero, el segundo, el tercero o el último (4 casos). La probabilidad de que falle el primer tiro y los otros tres sean exitosos es

$$0.5 \times 0.5 \times 0.5 \times 0.5 = 0.5^4.$$

La multiplicación de estas probabilidades es posible ya que los tiros que realiza el jugador son independientes entre sí.

En general, la probabilidad de que falle un tiro específico y los demás los enceste es 0.5^4. Entonces,

$$P[\text{Enceste tres tiros }] = 4 \times 0.5^4 = 0.25.$$

b) La siguiente probabilidad se calculará por complemento.

$$\begin{aligned} P[\text{Encestar por lo menos un tiro}] &= 1 - P[\text{Encestar cero tiros }] \\ &= 1 - 0.5^4 = 0.9375. \end{aligned}$$

c) Dos tiros encestados de los 4, significa alguno de los siguientes casos: $EEFF; EFEF; EFFE; FFEE; FEFE$ y $FEEF$, donde E significa encestar y F fallar un tiro. La probabilidad de que ocurra cualquiera de estos 6 casos, es igual a 0.5^4, por lo tanto,

$$P[\text{Encestar dos tiros }] = 6 \times 0.5^4 = 0.375.$$

Otra manera de calcular esta probabilidad es

$$\binom{4}{2}(0.5)^2(0.5)^2 = 0.375.$$

Ejemplo 1.60. Los empleados de cierta universidad se encuentran separados en tres divisiones: Administración (**AD**), Académicos (**AC**) y Sindicalizados (**S**). La siguiente tabla indica el número de empleados en cada división, clasificados por mujeres (**M**) y hombres (**H**).

	M	**H**
AD	20	30
AC	60	140
S	100	50

Se selecciona en forma aleatoria un empleado de esta universidad,

a) ¿Cuál es la probabilidad de que el empleado sea académico si se sabe que es mujer?

b) ¿Los eventos hombre y sindicalizado son independientes?

c) ¿Cuál es la probabilidad de que el empleado no sea mujer y no sea académico?

a) La solución es

$$P[AC|M] = \frac{60}{180} = 0.33333.$$

b) Obsérvese

$P[H \cap S] = \dfrac{50}{400} = 0.125$; $P[H] = \dfrac{220}{400} = 0.55$ y $P[S] = \dfrac{150}{400} = 0.375.$

De este modo,

$$0.55 \times 0.375 = 0.20625 \neq 0.125.$$

Por lo tanto, los eventos **H** y **S** no son independientes (son dependientes).

c) La probabilidad se calcula como

$$\begin{aligned} P[M^c \cap AC^c] &= P[(M \cup AC)^c] \\ &= 1 - P[M \cup AC] \\ &= 1 - P[M] - P[AC] + P[M \cap AC] \\ &= 1 - \frac{180}{400} - \frac{200}{400} + \frac{60}{400} = 0.2. \end{aligned}$$

Ejemplo 1.61. Una urna contiene tres monedas, una de las cuales está acuñada con dos águilas, mientras que las otras dos son monedas normales y no son sesgadas. Se escoge una moneda al azar de la urna y se lanza cuatro veces en forma sucesiva, resultando todas las veces águila. Calcular la probabilidad de que la moneda seleccionada se trate de la moneda anormal.

Usaremos la siguiente notación: $A = \{\text{Resulten cuatro águilas sucesivas}\}$; $B_1 = \{\text{Resulte una moneda normal}\}$; $B_2 = \{\text{Resulte la moneda anormal}\}$. De este modo,

$$\begin{aligned} P[B_2|A] &= \frac{P[A|B_2]P[B_2]}{P[A|B_1]P[B_1] + P[A|B_2]P[B_2]} \\ &= \frac{1 \times \frac{1}{3}}{1 \times \frac{1}{3} + \left(\frac{1}{2}\right)^4 \times \frac{2}{3}} = 0.89. \end{aligned}$$

Obsérvese que se aplicó el teorema de Bayes e independencia de eventos.

Ejemplo 1.62. De una urna que contiene 5 bolas negras y 3 verdes se sacan sucesivamente 3 bolas con reemplazo. Calcular la probabilidad de que
a) Las tres bolas sean del mismo color.
b) Cada color quede representado.

Obsérvese que este ejercicio es como el ejemplo 1.48, lo que cambia es la forma de seleccionar las bolas, en lugar de obtener las bolas sin reemplazo, ahora son seleccionadas con reemplazo. Obsérvese que este cambio influye en los resultados.

a) El hecho de que la selección de las bolas se realice con reemplazo, significa que existe independencia entre la selección de una bola y otra.

De esta manera,

$$\begin{aligned}
P[\text{Resulten 3 bolas del mismo color}] &= P[\text{Resulten 3 bolas negras}] \\
&\quad + P[\text{Resulten 3 bolas verdes}] \\
&= \left(\frac{5}{8}\right)^3 + \left(\frac{3}{8}\right)^3 = 0.29688.
\end{aligned}$$

b) La probabilidad pedida se calcula como

$$\begin{aligned}
P[\text{Cada color quede representado}] &= P[\text{1 negra y 2 verdes}] \\
&\quad + P[\text{2 negras y 1 verde}] \\
&= \frac{5}{8} \times \frac{3}{8} \times \frac{3}{8} \times 3 + \frac{3}{8} \times \frac{5}{8} \times \frac{5}{8} \times 3 \\
&= 0.70313.
\end{aligned}$$

Ejemplo 1.63. Dos personas lanzan 3 monedas regulares cada una. Calcular la probabilidad de que ambas personas obtengan el mismo número de águilas.

Sea $A = \{\text{Obtengan el mismo número de águilas}\}$, entonces,

$$\begin{aligned}
P[A] &= P[\text{Ambos obtengas 0 águilas}] + P[\text{Ambos obtengan 1 águila}] \\
&\quad + P[\text{Ambos obtengas 2 águilas }] + P[\text{Ambos obtengan 3 águilas }] \\
&= P[\text{Persona 1 obtenga 0 águilas}]P[\text{Persona 2 obtenga 0 águilas}] \\
&\quad + P[\text{ Persona 1 obtenga 1 águila }]P[\text{Persona 2 obtenga 1 águila}] \\
&\quad + P[\text{ Persona 1 obtenga 2 águilas }]P[\text{Persona 2 obtenga 2 águilas}] \\
&\quad + P[\text{ Persona 1 obtenga 3 águilas }]P[\text{Persona 2 obtenga 3 águilas}] \\
&= \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^3 + 3\left(\frac{1}{2}\right)^3 3\left(\frac{1}{2}\right)^3 + 3\left(\frac{1}{2}\right)^3 3\left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^3 \\
&= 0.3125.
\end{aligned}$$

Ejercicios del capítulo 1

1.- Sean A, B y C tres eventos asociados a un experimento. Expresar en notación de conjuntos las siguientes expresiones verbales y, a la vez, realizar el correspondiente diagrama de Venn.

a) No ocurre el evento A.
b) No ocurre ninguno de los tres eventos.
c) Solo ocurre el evento A.
d) Al menos uno de los eventos ocurre.
e) Exactamente uno de los eventos ocurre.
f) Exactamente dos de los eventos ocurren.
g) No ocurren más de dos sucesos simultáneamente.

2.- Demostrar que $I_{A^c}(x) = 1 - I_A(x)$.

3.- De un conjunto de 7,000 profesionistas. Cada profesionista se clasifica en

I) Administrativo o académico.
II) Hombre o mujer.
III) Casado o soltero.

De estos profesionistas, 4,000 son académicos, 2,300 son mujeres y la mitad del total son profesionistas que están casados. Por otro lado, se sabe que 660 son mujeres y administrativas, 1,506 son mujeres solteras y 700 son administrativos y solteros. Además, se sabe que 300 profesionistas son mujeres solteras y que trabajan en el área administrativa. Calcular el número de profesionistas de la empresa que son académicos, hombres y casados.

4.- Sean $A_1, A_2, \ldots, A_n$ eventos, demostrar que $I_{\{A_1 \cap A_2 \cap \cdots \cap A_n\}}(x) = I_{A_1}(x) I_{A_2}(x) \cdots I_{A_n}(x)$.

5.- Demostrar el corolario 1.1.

6.- Supóngase que A, B y C son eventos tales que $P[A] = P[B] = \frac{1}{3}$; $P[C] = \frac{2}{3}$; $P[A \cap B] = \frac{1}{6}$; $P[B \cap C] = \frac{1}{4}$ y $P[A \cap C] = 0$. Justificando bien la respuesta, calcular la probabilidad de que

a) Al menos uno de los 3 eventos A, B o C ocurra.
b) A lo más dos eventos de A, B y C ocurran.

7.- Sean A y B eventos cualesquiera, demostrar $P[A \triangle B] = P[A] + P[B] - 2P[A \cap B]$.

8.- Demostrar el corolario 1.2.

9.- Supóngase que A y B son dos eventos independientes asociados a un experimento. Se sabe que la probabilidad de que ocurra A o B es igual a

0.7 y de que ocurra A es igual a 0.4, determinar la probabilidad de que B ocurra.

10.- Entre los números $1, 2, ..., 90$ se escoge un número aleatoriamente. Calcular la probabilidad de que el número escogido sea divisible por 4, 5 o por 6.

11.- Sean A y B dos eventos asociados con un experimento. Supóngase que $P[A \cup B] = 0.7$ y $P[B] = 0.4$.

a) ¿Para qué elección de $P[A]$, los eventos A y B pueden ser mutuamente excluyentes?

b) ¿Para qué elección de $P[A]$, los eventos A y B son independientes?

12.- De 8 números positivos y 9 negativos, se eligen 10 números al azar (sin sustitución) y se multiplican. Calcular la probabilidad de que el producto sea un número positivo.

13.- En una caja hay 12 números del 1 al 12. Se eligen tres números al azar. Calcular la probabilidad de que

a) El número menor de los seleccionados sea 6.

b) El número mayor de los seleccionados sea 6.

14.- Se lanza un dado 10 veces. Calcular la probabilidad de que

a) El número 4 salga por lo menos una vez.

b) El número 4 resulte exactamente una vez.

15.- Se lanzan dos dados. Puesto que las caras muestran números diferentes, calcular la probabilidad de que una de las caras sea el número 6.

16.- Supóngase que tenemos 2 pantalones: pantalón A y pantalón B, cada uno con dos bolsas. El pantalón A tiene una moneda de \$10 en una bolsa y una de \$5 en la otra, mientras que el pantalón B tiene una moneda de \$10 en cada una de las bolsas. Se escoge un pantalón al azar, y de ésta se escoge una bolsa al azar. La moneda encontrada en esta bolsa seleccionada resultó ser de \$10. Calcular la probabilidad de que la moneda provenga del pantalón B.

17.- Una urna contiene tres monedas de \$10, dos monedas de \$5 y una moneda de \$1. Otra urna contiene una moneda de \$10, dos monedas de \$5 y tres monedas de \$1. Una moneda es seleccionada al azar de cada urna.

a) Describir el espacio muestral.

b) Encontrar la probabilidad de que ambas monedas sean del mismo valor.

c) ¿Es la probabilidad de que ambas monedas sean de \$5, más grande que la probabilidad de que ambas monedas sean de \$1?

18.- Las monedas de las dos urnas del ejercicio anterior, se juntan en una sola urna, entonces, una muestra aleatoria de 3 monedas es obtenida. Calcular la probabilidad de que los tres valores estén representados, si la muestra es realizada

a) Con reemplazo.
b) Sin reemplazo.

19.- Suponiendo que $P[B] > 0$, ¿es cierto que $2P[A|B] - \dfrac{P[A]}{P[B]} \leq 1$? En caso afirmativo, probarlo, en caso contrario, dar un contraejemplo.

20.- ¿Es cierto que si $P[A] = 1 - P[B]$, entonces, $A = B^c$? En caso afirmativo, probarlo, en caso contrario, dar un contraejemplo.

21.- Si $P[A \cup B] = 0$, ¿es posible que $P[A \cap B] > 0$? Justificar bien la respuesta.

22.- ¿Es cierto, que si se cumple $P[A] = 1 - x$ y $P[B] = 1 - y$, entonces, $P[A \cap B] \geq 1 - x - y$ y $P[A \cup B] \geq 1 - x - y$? En caso afirmativo, probarlo, en caso negativo, dar un contraejemplo.

23.- De un seguro de auto se sabe que el 12 % de los asegurados, tuvieron al menos un siniestro el año pasado. Además, se sabe que la probabilidad de que una persona tenga un accidente en el presente año, dado que el año pasado tuvo al menos un siniestro de automóvil es del 38 %, por otro lado, se sabe que la probabilidad de que una persona no tenga un accidente en el presente año, dado que no tuvo algún accidente el año pasado, es del 93 %. Se selecciona un asegurado al azar y se observa que tuvo accidente en el presente año, ¿cuál es la probabilidad de que el asegurado no tuvo algún accidente el año pasado?

24.- Sean A y B dos eventos, tal que $P[A] > 0$ y $P[B] > 0$, ¿es cierto que si $P[A|B] > P[A]$, entonces, $P[B|A] > P[B]$? En caso afirmativo, probarlo y en caso negativo, dar un contraejemplo.

25.- Suponiendo $P[C] > 0$, ¿es cierto que si $A \subset B$, entonces, se cumple $P[B|C] \geq P[A|C]$? En caso afirmativo, probarlo y en caso negativo, dar un contraejemplo.

26.- Suponiendo $P[C] > 0$, ¿es cierto que si $P[A] \leq P[B]$, entonces, se cumple $P[A|C] \leq P[B|C]$? En caso afirmativo, probarlo, en caso negativo, dar un ejemplo.

27.- Demostrar que si $P[B|A^c] = P[B|A]$, entonces, A y B son eventos independientes.

28.- Demostrar que si $P[A] = 2x$ y $P[B] = x$, donde $x > 0$, entonces, $P[B|A] \geq \frac{3x-1}{2x}$.

29.- ¿Es cierto $\frac{P[A]+P[B]}{2} \geq P[A \cap B]$? En caso afirmativo, probarlo, en caso negativo, dar un contraejemplo.

30.- Contestar las preguntas del ejemplo 1.36, pero ahora permitiendo la repetición de los dígitos.

31.- Con los dígitos $0, 1, 2, 3, 4, 5, 6$ se forman números de 4 cifras. De todos los números posibles, se escoge al azar un número. Calcular la probabilidad de que resulte

a) Un número par.
b) Un número par y menor a 3,332.

32.- ¿Es cierto que $P[A|B] + P[A|B^c] = 1$? En caso afirmativo, probarlo, en caso negativo, dar un contraejemplo.

33.- ¿Es cierto que si A y B son eventos independientes, entonces, $P[A|B] + P[A^c|B^c] = 1$? En caso afirmativo, probarlo, en caso negativo, dar un contraejemplo.

34.- Demostrar el teorema 1.8.

35.- Para un seguro de gastos médicos mayores, la salud de sus asegurados es muy importante. Un actuario, para una población, realizó una investigación y llegó a los siguientes resultados: el 25 % de las personas tiene una adicción severa. De la gente con problemas de adicción severa, el 60 % tiene hipertensión, por otro lado, el 40 % de las personas que no tiene problemas de adicción severa, tienen problemas de hipertensión. ¿Qué porcentaje de personas que tiene hipertensión, tiene problemas de adicción severa?

36.- Consideremos dos urnas: la urna 1 tiene 15 bolas, de las cuales 6 son bolas rojas. La urna 2 tiene 6 bolas numeradas con los dígitos $3, 4, 5, 6, 7, 8$. Se saca una bola al azqr de la urna 2 y según el número obtenido, digamos n, se obtiene una muestra aleatoria de tamaño n de la urna 1. Encontrar la probabilidad de que todas las bolas seleccionadas de la urna 1 sean rojas.

37.- En una caja hay 4 monedas. Dos de ellas son monedas normales y las otras dos son anormales. De las monedas anormales, una tiene la característica de que al ser lanzada resultará sol con una probabilidad de 0.6, y la otra resultará sol con una probabilidad de 0.8. Se selecciona una moneda de la caja al azar y se lanza 3 veces.

a) Calcular la probabilidad de que resulten 3 soles.
b) Si de los 3 lanzamientos resultaron dos águilas y un sol. Calcular la probabilidad de que la moneda seleccionada sea una de las monedas normales.

38.- Si A y B son eventos independientes, tal que $P[A] = \frac{1}{2}$ y $P[B^c] = \frac{1}{3}$, encontrar $P[(A \cup B^c)^c]$.

39.- Se tiene una baraja de 24 cartas, con 4 palos: $\diamondsuit$, $\heartsuit$, $\clubsuit$, $\spadesuit$ y 6 números: as, 2, 3, 4, 5, 6. Se sacan 5 cartas aleatoriamente de la baraja, calcular la probabilidad de que resulten

a) Cuatro cartas del mismo valor.
b) Cinco cartas del mismo palo.
c) Una tercia (ningún par).
d) Dos cartas exactamente de un palo y otras dos cartas exactamente de otro palo diferente.

40.- Un hámster es colocado en un tubo, donde solo tiene dos opciones: ir a arriba o ir a abajo, si elige arriba recibirá comida, de lo contrario recibirá un estímulo negativo. Cuando un hámster es puesto por primera vez en el tubo, este elige con la misma probabilidad cualquiera de las dos opciones. Después de haber recibido alimento en una prueba, para la siguiente ocasión, el hámster elegirá el alimento con una probabilidad de 0.7. Y en caso de haber recibido el estímulo negativo, para la siguiente prueba, elegirá arriba con una probabilidad de 0.85. Calcular la probabilidad de que el hámster seleccione

a) Arriba en la segunda prueba.
b) Abajo en la tercera prueba.

41.- Si $P[A] = P[B] = P[B|A] = \frac{1}{3}$, ¿son A y B eventos independientes?

42.- Si A y B son eventos independientes tales que $P[A] = P[B] = \frac{1}{5}$, calcular $P[(A \cup B^c) \cap (A^c \cup B)]$.

43.- Para un grupo de asegurados, se tiene la siguiente información, el 60 % son del sexo masculino y el 40 % son del sexo femenino. Además, se sabe que el 30 % de los hombres y el 70 % de las mujeres tienen la adicción de fumar tabaco. Se selecciona un asegurado al azar de este grupo, calcular la probabilidad de que sea hombre, dado que se sabe que tiene la adicción de fumar.

44.- Si $P[C] = P[A|C] = P[B|A \cap C] = \frac{2}{3}$, ¿cuál es $P[A \cap B \cap C]$?

45.- Un dado bien balanceado es arrojado 5 veces. Si se sabe que el número 2 aparece al menos dos veces, ¿cuál es la probabilidad de que este aparezca exactamente 3 veces?

46.- Sean A, B, C y D eventos tales que cumplen con $B = A^c$, $C \cap D = \emptyset$, $P[A] = \frac{1}{4}$, $P[C|A] = \frac{1}{2}$, $P[C|B] = \frac{1}{8}$, $P[D|A] = \frac{1}{2}$ y $P[D|B] = \frac{1}{5}$. Calcular $P[C \cup D]$.

47.- Considerando la definición 1.21 de dos eventos independientes, demostrar que las igualdades: (I),(II) y (III) de esta definición son afirmaciones equivalentes.

48.- Un estudiante realiza un examen de opción múltiple, en el cual cada pregunta tiene 4 respuestas posibles, de las cuales una es la respuesta correcta. Si el estudiante conoce la respuesta, entonces, seleccionará la respuesta correcta. De otra manera, él contestará aleatoriamente de las 4 posibles respuestas. Supongamos que él no conoce la respuesta en un 30 % de las preguntas.

a) ¿Cuál es la probabilidad de que el estudiante conteste correctamente una pregunta?
b) Si el estudiante contesta correctamente una pregunta, ¿cuál es la probabilidad de que él realmente conociera la respuesta?

49.- Sean A, B y C eventos tales que cumplen con $P[A \cap B \cap C] > 0$. Justificando bien las respuestas, verificar si las siguientes afirmaciones son ciertas.

a) Si $P[C|A \cap B] = P[C|B]$, entonces, $P[A|B \cap C] = P[A|B]$.
b) $P[A \cap B|A \cup B] = 1$.
c) $P[A \cup B|A \cap B] = 1$.

50.- Si ocho dados son arrojados, ¿cuál es la probabilidad de obtener cuatro números pares idénticos y cuatro números impares idénticos?

51.- Un grupo de estudiantes de primaria está conformado por 10 niños y 9 niñas. El maestro selecciona 5 estudiantes aleatoriamente con reemplazo. Calcular la probabilidad de que el número de niños seleccionados exceda al número de niñas seleccionadas.

52.- Sean A, B y C tres eventos tales que los eventos A y B son independientes, además, se cumple $B \cap C = \emptyset$, $P[A] = \frac{1}{6}$, $P[B] = \frac{1}{5}$ y $P[C] = \frac{1}{2}$, calcular $P[(A \cap B)^c \cup C]$.

53.- Una carta es seleccionada aleatoriamente de una baraja ordinaria, y su resultado es anotado, entonces la carta es regresada. Este procedimiento se realiza 10 veces. Calcular la probabilidad de que resulten exactamente 2 reinas en las 10 cartas seleccionadas, si se sabe que resultó al menos una reina.

54.- Supongamos que los eventos A, B y C son mutuamente independientes y $P[A \cap B] \neq 0$. Demostrar que $P[C|A \cup B] = P[C]$.

55.- Una compañía adquiere un seguro contra accidentes laborales. Los accidentes de esta compañía son clasificados en tres categorías diferentes: pequeños, moderados y severos. De acuerdo a la experiencia, se ha encontrado las siguientes características: la probabilidad de que un accidente sea pequeño es de 0.3, que sea moderado es de 0.1 y que sea severo es de 0.05. Dos accidentes ocurren independientemente en un mes. Calcular la probabilidad de que ninguno de los accidentes sea severo y a lo más uno sea moderado.

56.- Una caja tiene 12 bolas numeradas del 1 al 12. Dos bolas son seleccionadas aleatoriamente, una tras otra, sin reemplazo. Calcular la probabilidad de que los números de las bolas seleccionadas difieran en sus números por dos o más.

57.- Si $P[A] = 0.4$ y $P[A \cup B] = 0.6$, encuentra $P[B]$ si

a) A y B son eventos independientes.
b) $P[A|B] = 0.2$.

58.- Para las probabilidades condicionales, demostrar que se cumplen las propiedades de probabilidad, esto es, comprobar que se cumplen las siguientes igualdades:

a) $P[\emptyset|B] = 0$.
b) $P[\bigcup_{i=1}^{n} A_i|B] = \sum_{i=1}^{n} P[A_i|B]$, si $A_i \cap A_j = \emptyset$, para $i \neq j$.
c) $P[A^c|B] = 1 - P[A|B]$.
d) $P[A_1 \cup A_2|B] = P[A_1|B] + P[A_2|B] - P[A_1 \cap A_2|B]$.
e) Si $A_1 \subset A_2$, entonces, $P[A_1|B] \leq P[A_2|B]$.
f) $P[\bigcup_{i=1}^{n} A_i|B] \leq \sum_{i=1}^{n} P[A_i|B]$.

59.- ¿Es cierto que $P[A \cap B^c|A] + P[A|B] = 1$? En caso afirmativo, probarlo, en caso negativo, dar un contraejemplo.

60.- Supongamos que 10 bolas son arrojadas al azar, una por una, a 10 cajas. Cada bola caerá en una y solo una caja. Calcular la probabilidad de que exactamente una caja quede vacía.

61.- Se sabe que la probabilidad de que una máquina se descomponga en un día dado es de 0.05, además, es independiente de que se descomponga o no, en cualquier otro día. Si la máquina solo se puede descomponer una sola vez al día, calcular la probabilidad de que la máquina se descomponga

a) Dos o más veces en 12 días.
b) Por primera vez, en el quinto día.

62.- En una población, se sabe que el 12 % de las personas padecen de una enfermedad, además, que la probabilidad de que una persona siga una dieta saludable, dado que tiene la enfermedad, es tres veces mayor que la probabilidad de que una persona que no tiene la enfermedad siga una dieta saludable. Se selecciona al azar una persona de esta población; calcular la probabilidad de que esta persona tenga la enfermedad, dado que se sabe que sigue una dieta saludable.

63.- Un actuario que estudia las preferencias sobre la compra de un seguro de automóvil y la compra de un seguro médico llega a las siguientes conclusiones: es tres veces más probable que una persona compre un seguro de auto que un seguro médico, la compra de un seguro de auto es independiente a la compra de un seguro médico, además, la probabilidad de que una persona compre los dos seguros es de 0.25. Calcular la probabilidad de que una persona no adquiera ninguno de los dos seguros.

64.- En un centro de servicios médicos, se sabe que el 80 % de los pacientes va a consulta médica o va a una curación. También, se sabe que el 30 % de los pacientes que asisten al centro no va a consulta médica. Por otro lado, el número de pacientes que va a consulta médica es independiente al número de pacientes que va a una curación. Calcular la probabilidad de que un paciente que asiste al centro médico vaya a una curación.

65.- Los habitantes de una población son clasificados de acuerdo con el ejercicio físico que realizan. Se sabe que el 30 % de los habitantes practican ejercicio físico a un nivel alto, el 20 % lo practican a un nivel medio y el resto lo practican a un nivel bajo o no practican ejercicio. Las personas que hacen ejercicio a un nivel medio tienen la mitad de probabilidad de desarrollar problemas de salud cardiovasculares que las personas que practican ejercicio a un nivel bajo o nulo, además, se sabe que las personas que hacen ejercicio a un nivel bajo o no lo practican tienen una probabilidad de generar problemas de salud cardiovasculares, cuatro veces más que una persona con un nivel de ejercicio alto. Se selecciona al azar un habitante de esta población y se observa que tiene problemas de salud cardiovasculares. Calcular la probabilidad de que la persona realice el ejercicio a un nivel alto.

66.- Se sabe que los profesores de una universidad de los departamentos de ciencias sociales, ingeniería y matemáticas cumplen las siguientes características: los profesores de ciencias sociales e ingeniería son eventos excluyentes, el 48 % de los profesores son de matemáticas, el 28 % no pertenece a alguna de las 3 áreas antes mencionadas, el número de profesores de ingeniería es tres veces más que el número de profesores de ciencias sociales, por otro lado, el 11 % de profesores son del área de ciencias sociales pero no son de matemáticas y el 40 % de los profesores están en dos de las tres áreas. Calcular el porcentaje de profesores que son de matemáticas e ingeniería.

67.- En una caja se encuentran 5 dados con las siguientes características: uno de ellos tiene 4 caras, dos tienen 6 caras, otro 10 caras y el quinto dado tiene 50 caras, todos numerados del 1 al número de caras que cada dado tiene. Se selecciona un dado aleatoriamente y se lanza 4 veces, calcular la probabilidad

a) De que resulte todas las veces un número múltiplo de 3.
b) De que el dado seleccionado sea el de 50 caras, dado que se sabe que resultaron 4 veces un número múltiplo de 3.

Variables
aleatorias

Capítulo

Introducción

En este capítulo se tratará con la construcción de los primeros modelos matemáticos para fenómenos donde está presente la aleatoriedad. Estos modelos como se verá son de gran utilidad, por una parte, para continuar con el cálculo de probabilidades, y por otra, para obtener diferentes tipos de promedios (valores esperados).

Se presentará primero la definición de variable aleatoria y sus principales propiedades. Posteriormente, se dará la definición de función de distribución y las propiedades que una función debe cumplir para que sea de distribución, también se explicará la importancia que tiene esta función para la clasificación de variables aleatorias.

Posteriormente, para cada uno de los tipos de variables, se presentará la definición de función de densidad, sus propiedades, así como varios ejemplos. También se tratará el concepto de valor esperado, en particular se darán las definiciones de media, varianza y momentos de una variable aleatoria. Se presentará la definición de función generadora de momentos y se explicará la importancia que tiene esta función en la contribución de resultados de una distribución.

Se enseñarán los conceptos de moda y cuantil de una variable aleatoria, en particular se tratará la definición de mediana, considerando algunos ejemplos interesantes, y explicando la importancia que tienen estos conceptos tanto en temas de probabilidad como de estadística.

Después, se presentarán dos de las desigualdades más importantes de la probabilidad, la desigualdad de Chebyshev y la de Jensen, además, se explicarán algunas posibles aplicaciones.

Finalmente, se darán las definiciones de mezclas de distribuciones, distribuciones de cola pesada y distribuciones truncadas. En cada una de estas, se ilustrará, con ejemplos, lo importante que son estos modelos en diferentes temas de actuaría.

2.1. Definición de variable aleatoria

Uno de los propósitos en este tema es construir modelos matemáticos para los fenómenos donde está presente la aleatoriedad, mas un inconveniente en algunos casos es que el correspondiente espacio muestral es un conjunto cualitativo; por ejemplo, al seleccionar al azar una carta de una baraja ordinaria, al arrojar al aire una o varias monedas, o también, al sacar aleatoriamente una bola de una urna que contiene bolas de varios colores. En este tipo de fenómenos es conveniente que, para obtener un modelo matemático o fórmula adecuada, el espacio muestral sea cuantitativo.

El objetivo principal de una variable aleatoria es el de transformar el espacio muestral en un conjunto cuantitativo. Esto es, si el espacio muestral es cualitativo, este será transformado en un conjunto cuantitativo. En el caso de que el espacio muestral sea cuantitativo, también se transformará en un conjunto que seguirá siendo cuantitativo. En ambos casos la transformación facilitará la modelación del fenómeno.

Definición 2.1. Sea Ω el espacio muestral y $\mathcal{A}$ el espacio de eventos correspondientes a un experimento. Una *variable aleatoria*, denotada por X (o por Y, Z o X_i, etc.), es una función con dominio Ω y contradominio, los números reales $\mathbb{R}$, donde para toda $r \in \mathbb{R}$, el conjunto $A_r = \{\omega \in \Omega | X(\omega) \leq r\}$ pertenece al espacio de eventos $\mathcal{A}$.

Observación. La definición de variable aleatoria equivale a la definición de función medible, véase definición D.15 del apéndice D, también se puede consultar literatura de la teoría de la medida, en particular el capítulo 3 de Capinski y Kopp [5].

A continuación, se presentará un ejemplo de variable aleatoria.

Ejemplo 2.1. Se lanza una moneda. Se define la siguiente transformación: X = Número de águilas en el lanzamiento. Demostrar que esta transformación es una variable aleatoria.

Esta transformación se puede apreciar en el siguiente dibujo:

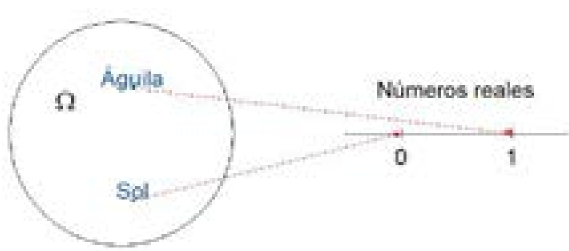


Figura 2.1. *Variable aleatoria del ejemplo 2.1*

Esto es, la regla de correspondencia es

$$X(\text{Águila}) = 1 \text{ y } X(\text{ Sol }) = 0.$$

El espacio de eventos en este caso es

$$\mathcal{A} = \{\emptyset, \Omega, \{\text{Águila }\}, \{\text{ Sol }\}\}.$$

Se comprobará que esta transformación es una variable aleatoria.
Si $r < 0$,

$$A_r = \{\omega \in \Omega | X(\omega) \leq r\} = \emptyset, \text{ esto es, } A_r \in \mathcal{A}.$$

Si $0 \leq r < 1$,

$$A_r = \{\omega \in \Omega | X(\omega) \leq r\} = \{\text{ Sol }\}, \text{ de este modo, } A_r \in \mathcal{A}.$$

Y si $r \geq 1$,

$$A_r = \{\omega \in \Omega | X(\omega) \leq r\} = \Omega, \text{ de esta manera, } A_r \in \mathcal{A}.$$

Esto es, para todo $r \in \mathbb{R}$, el conjunto A_r pertenece al espacio de eventos, por lo tanto, la función X cumple con la definición de variable aleatoria.

No es nuestro propósito estar comprobando que las funciones de interés cumplen la definición de variable aleatoria. Dichas comprobaciones son para un curso más avanzado de probabilidad. Afortunadamente son raras y complicadas de definir las funciones que no cumplen con la definición de variable aleatoria.

Tampoco nos interesa visualizar las preimágenes de las variables aleatorias, por lo que en lugar de usar la notación $X(\omega)$, solo usaremos la notación X.

La X mayúscula denotará en general a la variable aleatoria, la x minúscula denotará un valor específico de la variable. Esta notación es universal y es usada tanto en la literatura de probabilidad como de estadística.

Como se mencionó, la finalidad de utilizar variables aleatorias es para facilitar la modelación matemática de los fenómenos, tanto en materia de probabilidad, procesos estocásticos, estadística y aplicaciones en otras áreas.

Definición 2.2. El conjunto de todos los valores que puede tomar una variable aleatoria es llamado *recorrido, rango* o *soporte*.

En el ejemplo anterior el recorrido es $x = 0, 1$.

Ejemplo 2.2. Se lanzan dos dados bien equilibrados al aire. Definimos las siguientes dos variables aleatorias: X = La diferencia absoluta de los dos dados y Y = La suma de los dos dados. ¿Cuál es el recorrido de cada una de ellas?

Los recorridos de estas variables son respectivamente

$$x = 0, 1, 2, \ldots, 5, \quad \text{y} \quad y = 2, 3, 4, \ldots, 12.$$

En un problema se pueden definir varias variables aleatorias, pero dependiendo del interés se considera la definición adecuada.

2.2. Función de distribución

Se presentará primero la definición de función de distribución de una variable aleatoria, posteriormente, sus propiedades más importantes y finalmente se presentarán algunos ejemplos.

Definición 2.3. Sea X una variable aleatoria (asociada a un experimento), entonces, la *función de distribución* de X (o función acumulativa o función de distribución acumulativa) es denotada por $F_X(x)$ y se define como

$$F_X(x) = P[X \leq x] \quad \text{para toda } x \in \mathbb{R}.$$

Ejemplo 2.3. Encontrar la función de distribución de la diferencia absoluta del ejemplo 2.2

$X =$ La diferencia absoluta de ambos dados.
$x = 0, 1, 2, 3, 4, 5.$
Se calculará la función de distribución para algunos valores particulares

$$F_X(-2) = P[X \leq -2] = 0,$$

$$\begin{aligned}
F_X(0) &= P[X \leq 0] = P[X = 0] = \frac{1}{6}, \\
F_X(1) &= P[X = 0] + P[X = 1] = \frac{1}{6} + \frac{10}{36} = \frac{4}{9}, \\
F_X(2) &= P[X \leq 1] + P[X = 2] = \frac{4}{9} + \frac{8}{36} = \frac{6}{9}, \\
F_X(3) &= P[X \leq 2] + P[X = 3] = \frac{6}{9} + \frac{6}{36} = \frac{5}{6}, \\
F_X(4) &= P[X \leq 3] + P[X = 4] = \frac{5}{6} + \frac{4}{36} = \frac{17}{18}, \\
F_X(5) &= P[X \leq 4] + P[X = 5] = \frac{17}{18} + \frac{2}{36} = 1, \\
F_X(7) &= P[X \leq 5] = 1.
\end{aligned}$$

Obsérvese que $F_X(a) = 0$ para todo $a < 0$, por otro lado, $F_X(a) = 1$, para todo $a > 5$, además, $F_X(a) = F_X([a])$, donde $[a]$ es la parte entera de a, para $0 < a < 5$. En conclusión, la función de distribución expresada con funciones indicadoras se da a continuación:

$$\begin{aligned}
F_X(x) = \frac{1}{6} I_{[0,1)}(x) &+ \frac{4}{9} I_{[1,2)}(x) + \frac{6}{9} I_{[2,3)}(x) \\
&+ \frac{5}{6} I_{[3,4)}(x) + \frac{17}{18} I_{[4,5)}(x) + I_{[5,\infty)}(x).
\end{aligned}$$

La gráfica de la función de distribución es

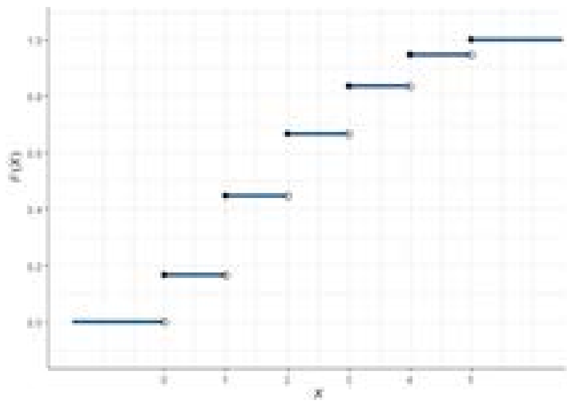


Figura 2.2. *Función de distribución, ejemplo 2.3*

Observación. Una función de distribución tiene las siguientes características:

I) Comienza en cero o tiende a comenzar en cero.
II) Nunca decrece.
III) Termina en uno o tiende a terminar en uno.
IV) Es continua por la derecha para todo número real.

Estas características se presentarán formalmente y son conocidas como las propiedades de una función de distribución.

Propiedades de una función de distribución

Sea X una variable aleatoria con función de distribución $F_X(x)$, entonces, se cumple:

(I) $\lim\limits_{x\to+\infty} F_X(x) = 1.$
(II) $\lim\limits_{x\to-\infty} F_X(x) = 0.$
(III) $F_X(x)$ es una función monótona no decreciente, esto es, si $a < b$, entonces, $F_X(a) \leq F_X(b)$.
(IV) $F_X(x)$ es una función continua por la derecha, esto es, $\lim\limits_{x\to a^+} F_X(x) = F_X(a)$.

Toda función que va de los reales a los reales que cumple las cuatro propiedades anteriores es una función de distribución, es otra manera de definir este tipo de funciones.

En el ejemplo pasado, cada salto representa la probabilidad en el valor donde se presenta este. Estos saltos no siempre se observarán, no es una generalidad de las funciones de distribución, depende del tipo de variable aleatoria que se esté tratando. Como se verá, las funciones de distribución de algunas variables aleatorias, no presentarán discontinuidades y por lo mismo tampoco saltos.

Es importante aclarar que la palabra *distribución* no se refiere necesariamente a la función de distribución, sino más bien al modelo, y el modelo puede ser determinado por medio de la función de distribución, o por medio de la función de densidad, o por medio de la función generadora de momentos, estas dos últimas se van a definir más adelante. En el mismo sentido, cuando una variable aleatoria tiene algún modelo específico, se dirá que la variable se *distribuye* con este modelo.

Observación. Para una variable aleatoria X, su función de distribución valuada en x es una medida (de probabilidad), véase ejemplos que se presentan en el apéndice D.

Las variables aleatorias o las distribuciones de las variables aleatorias se clasifican en: *discretas, continuas y mezclas* de discretas con continuas. Las definiciones de las variables discretas y continuas se presentarán a continuación, así como sus propiedades y algunos ejemplos. En las sección 2.14 de este capítulo, se dará la definición de mezcla de dos o más distribuciones, sus principales propiedades y algunos ejemplos.

2.3. Variables aleatorias discretas

En esta sección se tratarán las variables aleatorias discretas, así como las definiciones y propiedades más importantes. Se presentarán los primeros ejemplos de este tipo de variables y se comprenderá su importancia en las aplicaciones en la solución de algunos problemas reales.

Definición 2.4. Una variable aleatoria X es *discreta* si su rango es finito o infinito numerable.

Definición 2.5. Sea X una variable aleatoria discreta, la *función de densidad o función de densidad discreta* (función de probabilidad o función de probabilidad masa) es denotada como $f_X(x)$ y se define como:

$$f_X(x) = P[X = x], \quad \text{para toda } x \in \mathbb{R}.$$

Ejemplo 2.4. Se lanzan tres monedas bien balanceadas. Encontrar la función de densidad de la variable aleatoria número de soles.

$X =$ Número de soles.
$x = 0, 1, 2, 3.$
Obsérvese para valores del recorrido

$$f_X(0) = \frac{1}{8};\ f_X(1) = \frac{3}{8};\ f_X(2) = \frac{3}{8};\ f_X(3) = \frac{1}{8}.$$

Y para toda x que no está en el recorrido $f_X(x) = 0$.

A continuación, se presentarán tres formas diferentes de expresar la función de densidad.

En forma de tabla

x	$f_X(x)$
0	0.125
1	0.375
2	0.375
3	0.125

Tabla 2.1. *Función de densidad, ejemplo 2.4*

Con funciones indicadoras

$$f_X(x) = \frac{1}{8} I_{\{0,3\}}(x) + \frac{3}{8} I_{\{1,2\}}(x).$$

En forma gráfica, ya sea en forma de histograma

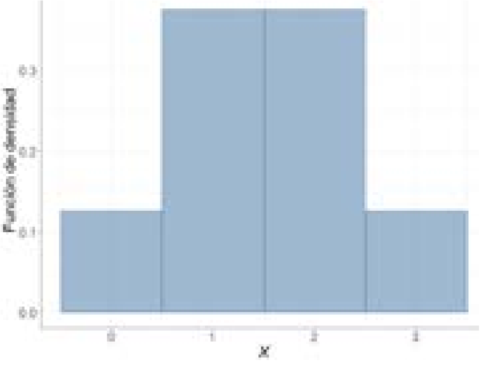


Figura 2.3. *Función de densidad como histograma, ejemplo 2.4*

O en forma de líneas y puntos

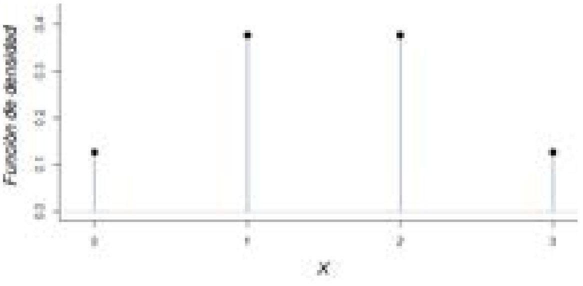


Figura 2.4. *Función de densidad, ejemplo 2.4*

La ventaja de representar la función de densidad por medio de un histograma es que la probabilidad en un valor específico del recorrido queda representada por el área del rectángulo correspondiente a ese valor. Por ejemplo, para $x = 2$, la probabilidad que le corresponde es 0.375 y el área correspondiente al rectángulo en $x = 2$ es igual a $1 \times (0.375) = 0.375$.

También se puede, en ocasiones, representar la función de densidad por medio de una fórmula, para este ejemplo,

$$\begin{aligned} f_X(x) &= \binom{3}{x} \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{3-x} I_{\{0,1,2,3\}}(x) \\ &= \binom{3}{x} \left(\frac{1}{8}\right) I_{\{0,1,2,3\}}(x). \end{aligned}$$

Ejemplo 2.5. En una prueba aplicada a niños pequeños se les pide que hagan corresponder cada uno de los siguientes cuatro dibujos: león, casa, auto, flor, con la palabra que identifica a ese objeto. Si un niño asigna al azar las cuatro palabras a los cuatro dibujos, encuentre la función de densidad para el número de correspondencias correctas.

X = Número de correspondencias correctas.

$x = 0, 1, 2, 4$.

El número de correspondencias diferentes que puede realizar el niño son $4! = 24$.

En la tabla siguiente se observarán las diferentes correspondencias posibles, en cada caso, se indica el número de aciertos.

Figuras	1	2	3	4	5	6	7	8	9	10	11	12	13
León	L	L	L	L	L	L	C	C	C	C	C	C	A
Casa	C	C	A	A	F	F	L	L	A	A	F	F	L
Auto	A	F	C	F	C	A	A	F	L	F	L	A	C
Flor	F	A	F	C	A	C	F	A	F	L	A	L	F
Aciertos	**4**	**2**	**2**	**1**	**1**	**2**	**2**	**0**	**1**	**0**	**0**	**1**	**1**

Figuras	14	15	16	17	18	19	20	21	22	23	24
León	A	A	A	A	A	F	F	F	F	F	F
Casa	L	C	C	F	F	L	L	C	C	A	A
Auto	F	L	F	L	C	C	A	L	A	L	C
Flor	C	F	L	C	L	A	C	A	L	C	L
Aciertos	**0**	**2**	**1**	**0**	**0**	**0**	**1**	**1**	**2**	**0**	**0**

Tabla 2.2. *Posibilidades de correspondencias, ejemplo 2.5*

De esta manera se puede calcular las probabilidades para cada valor del recorrido,

$$f_X(0) = \frac{9}{24} = \frac{3}{8};\ f_X(1) = \frac{8}{24} = \frac{1}{3};\ f_X(2) = \frac{6}{24} = \frac{1}{4} \text{ y } f_X(4) = \frac{1}{24}.$$

Obsérvese que la obtención de la función de densidad para este ejemplo fue por el cálculo individual de probabilidades y no por medio de una fórmula. La función de densidad expresada con funciones indicadoras es

$$f_X(x) = \frac{3}{8} I_{\{0\}}(x) + \frac{1}{3} I_{\{1\}}(x) + \frac{1}{4} I_{\{2\}}(x) + \frac{1}{24} I_{\{4\}}(x).$$

A continuación, se presentarán las propiedades que una función de densidad discreta debe de cumplir.

Propiedades de una función de densidad discreta

Una función de densidad discreta $f_X(x)$ cumple con las siguientes propiedades:

(I) $f_X(x) \geq 0$ para todo $x \in \mathbb{R}$.

(II) $\sum_x f_X(x) = 1$.

Se puede observar en cada uno de los anteriores ejemplos que la función de densidad cumple con estas propiedades.

Es importante aclarar cómo es la forma de una función de distribución discreta, esta es una función escalonada como la figura 2.2, que cumple las cuatro propiedades de una función de distribución. En cada valor del recorrido, la gráfica presenta un salto (discontinuidad), que es igual a la probabilidad o valor de la función de densidad en ese punto. Es muy importante para el estudiante tener siempre presente la forma y las propiedades de una función de distribución discreta.

Teorema 2.1. *Sea X una variable aleatoria discreta, la función de densidad puede obtenerse a partir de la función de distribución y viceversa.*

Demostración. Supongamos que conocemos la función de densidad $f_X(x)$ y que el recorrido es $x_1, x_2, \ldots, x_n, \ldots$, entonces, la función de distribución se puede obtener a partir de la función de densidad, de la siguiente manera:

$$F_X(x_0) = \sum_{\{x_i | x_i \leq x_0\}} f_X(x_i).$$

Ahora suponemos que conocemos la función de distribución $F_X(x)$, entonces, la función de densidad se puede calcular por medio de la función de distribución de la siguiente manera:

$$f_X(x) = F_X(x) - \lim_{0<h \to 0} F_X(x-h). \qquad \square$$

Es importante aclarar que en la práctica para obtener la función de densidad, a partir de la función de distribución, no es necesario hacerlo por medio de la igualdad anterior, esto es, por medio de límites. El cálculo de la función de densidad en un valor específico del recorrido será igual a la diferencia de la función de distribución, valuada en este valor específico, y la función de distribución, valuada en el valor del recorrido anterior al valor específico, que corresponde a uno de los saltos en la función de distribución. Por último, el valor de la función de densidad para valores que no es del recorrido es cero.

2.4. Variables aleatorias continuas

En esta sección trataremos la definición, algunas propiedades y varios ejemplos de las variables aleatorias continuas. Por medio de los ejemplos ilustraremos la importancia que tienen las variables aleatorias continuas en la práctica.

Es importante aclarar que hay una gran diferencia en la forma de hacer los cálculos correspondientes entre las variables aleatorias continuas y las variables discretas. Entre los cálculos más importantes, se encuentran las probabilidades y diversos valores esperados.

Definición 2.6. Una variable aleatoria X es *continua* si su función de distribución correspondiente es una función continua.

Definición 2.7. Una variable aleatoria X es *absolutamente continua* si es una variable continua y además existe una función $f_X(x)$ tal que

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt,$$

donde la función $f_X(x)$ es conocida como *función de densidad o función de densidad continua* de X.

Las únicas variables aleatorias continuas que se tratarán son las variables que son absolutamente continuas, por esta razón, cuando se mencione en una definición, teorema o ejemplo que se trata de una variable continua es porque esta variable también es absolutamente continua.

La forma de la gráfica de una función de distribución para una variable aleatoria continua, en términos generales, se presenta en la siguiente gráfica

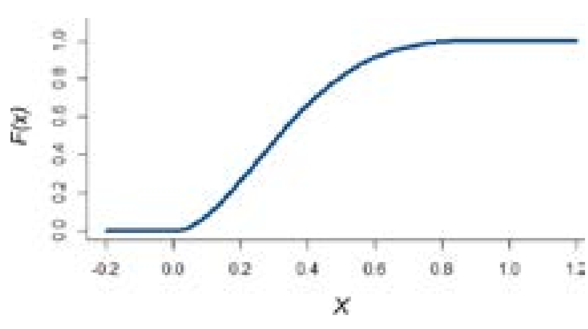


Figura 2.5. *Función de distribución continua*

Obsérvese que esta última gráfica no tiene discontinuidades (saltos), en todos los reales es continua.

La forma de la gráfica de una función de densidad para una variable continua, en términos generales, es

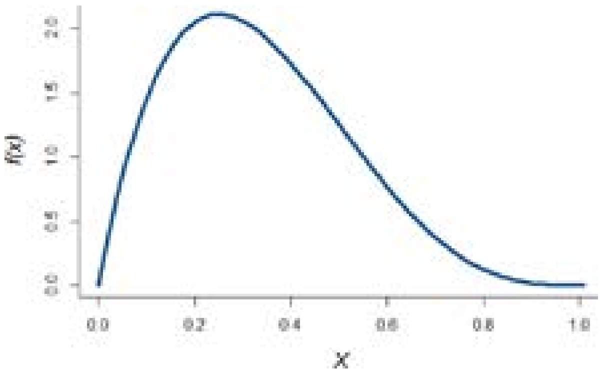


Figura 2.6. *Función de densidad continua*

Obsérvese, para el caso particular de las anteriores gráficas, que el recorrido de esta variable aleatoria es el intervalo $[0, 1]$.

Tanto las funciones de distribución y de densidad de una variable aleatoria continua, serán utilizadas para el cálculo de probabilidades, pero, además, la función de densidad también se usará para calcular valores esperados de la variable aleatoria, entre otros cálculos.

Obsérvese, si X es una variable aleatoria continua, entonces,

$$\begin{aligned} P[a < X \leq b] &= F_X(b) - F_X(a) \\ &= \int_{-\infty}^{b} f_X(t)dt - \int_{-\infty}^{a} f_X(t)dt \\ &= \int_{a}^{b} f_X(t)dt. \end{aligned}$$

Dicho de otra manera, la probabilidad de que la variable aleatoria continua tenga un valor en el intervalo $(a, b]$, geométricamente, es el área por debajo de la curva de la función de densidad comprendida entre los valores de a y b.

También obsérvese

$$\begin{aligned} P[a \leq X \leq b] &= P[a < X \leq b] \\ &= P[a \leq X < b] \\ &= P[a < X < b]. \end{aligned}$$

Esto es, para calcular la probabilidad de que la variable tome un valor entre a y b, da lo mismo si se consideran los extremos o no del intervalo.

Se puede apreciar también

$$\begin{aligned} P[X = a] &= P[a \leq X \leq a] \\ &= \int_a^a f_X(t)dt = 0. \end{aligned}$$

Esto es, la probabilidad de que una variable aleatoria continua tome un valor particular a es cero para cualquier valor de a.

A continuación, se presentan las propiedades de una función de densidad continua.

Propiedades de una función de densidad continua

La función de densidad de una variable aleatoria continua $f_X(x)$ cumple con las siguientes propiedades:

(I) $f_X(x) \geq 0$ para toda $x \in \mathbb{R}$.

(II) $\int_{-\infty}^{\infty} f_X(x)dx = 1$.

Teorema 2.2. *Sea X una variable aleatoria continua con función de densidad $f_X(x)$ y con función de distribución $F_X(x)$. Entonces, la función de distribución es posible obtenerla a partir de la función de densidad y viceversa.*

Demostración. Supongamos que conocemos la función de densidad $f_X(x)$, entonces, por definición (de variable aleatoria absolutamente continua), la función de distribución es igual a

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt.$$

Ahora, supongamos que conocemos la función de distribución $F_X(x)$, entonces, por el teorema fundamental del cálculo, $f_X(x)$ se puede obtener de la siguiente manera:

$$f_X(x) = \frac{dF_X(x)}{dx}. \qquad \square$$

Ejemplo 2.6. Sea X una variable aleatoria con función de distribución $F_X(x) = xI_{[0,1]}(x) + I_{(1,\infty)}(x)$.

a) Demostrar que $F_X(x)$ es función de distribución.

b) Graficar $F_X(x)$.

c) Encontrar la función de densidad $f_X(x)$ y graficarla.

d) Calcular $P[\frac{1}{2} < X < \frac{3}{4}]$.

e) Calcular $P[\frac{1}{2} < X | X < \frac{3}{4}]$.

a) Se puede observar

$$\lim_{x \to +\infty} F_X(x) = 1.$$

Además,

$$\lim_{x \to -\infty} F_X(x) = 0.$$

Ahora supongamos que $a < b$, entonces,

$$F_X(a) = F_X(b), \text{ si } a < b < 0, \text{ ó } 1 < a < b.$$

De otra manera,

$$F_X(a) < F_X(b), \text{ en cualquier otro caso,}$$

donde cualquier otro caso puede ser si

$$a < 0 < b \leq 1; \quad 0 \leq a < b \leq 1, \text{ o } 0 \leq a < 1 < b.$$

Por lo tanto, $F_X(x)$ es una función monótona no decreciente.

Finalmente, se puede notar que la función $F_X(x)$ es una función continua en todos los reales, esto es, se cumple

$$\lim_{x \to x_0} F_X(x) = F_X(x_0).$$

En conclusión, $F_X(x)$ es función de distribución.

b) La gráfica de la función de distribución se da a continuación:

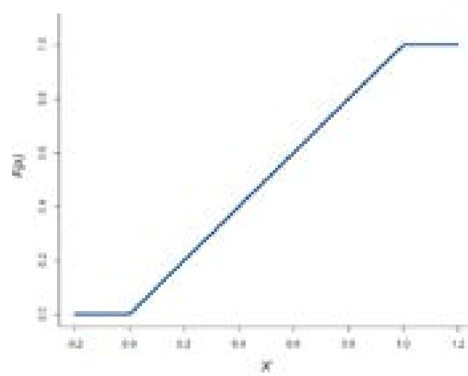


Figura 2.7. *Función de distribución, ejemplo 2.6*

c) Obsérvese

$$f_X(x) = \frac{dF_X(x)}{dx} = 1, \quad \text{para toda } x \in (0, 1).$$

Para valores de $x < 0$ o $x > 1$, la derivada vale 0. Pero la derivada se indefine para los valores de $x = 0$ y $x = 1$. Lo que significa que la derivada no existe en estos valores.

Como la probabilidad de una variable aleatoria continua en un punto es cero, entonces, podemos redefinir $f_X(x)$ para los valores donde no existe la derivada.

Entonces, la función de densidad puede ser expresada como

$$f_X(x) = I_{[0,1]}(x).$$

La gráfica de la función de densidad es

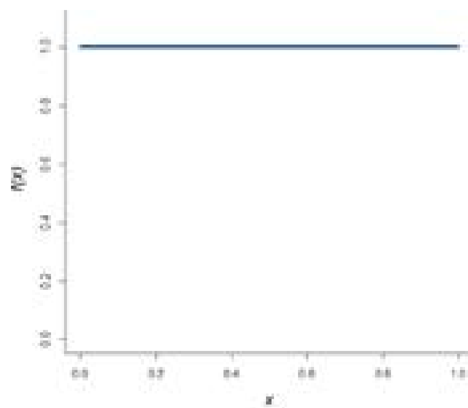


Figura 2.8. *Función de densidad, ejemplo 2.6*

d) La probabilidad pedida es

$$P\left[\frac{1}{2} < X < \frac{3}{4}\right] = F_X(0.75) - F_X(0.5) = 0.25.$$

e) La probabilidad condicional se calcula como

$$\begin{aligned} P\left[\frac{1}{2} < X \middle| X < \frac{3}{4}\right] &= \frac{P\left[\left(\frac{1}{2} < X\right) \cap \left(X < \frac{3}{4}\right)\right]}{P\left[X < \frac{3}{4}\right]} \\ &= \frac{P\left[\frac{1}{2} < X < \frac{3}{4}\right]}{P\left[X < \frac{3}{4}\right]} \\ &= \frac{0.25}{0.75} = 0.33333. \end{aligned}$$

Ejemplo 2.7. El tiempo requerido por un estudiante para terminar un examen con tiempo máximo de una hora es una variable aleatoria con función de densidad dada por $f_X(x) = (cx^2 + x)I_{[0,1]}(x)$. Encontrar

a) La constante c para que $f_X(x)$ sea función de densidad.
b) La función de distribución $F_X(x)$.
c) La probabilidad de que un estudiante termine en menos de media hora.
d) La probabilidad de que un estudiante termine en por lo menos media hora, dado que se sabe que no terminó en los primeros 15 minutos.

a) Se sabe

$$\begin{aligned} \int_{-\infty}^{\infty} f_X(x)dx &= \int_0^1 (cx^2 + x)dx \\ &= \frac{c}{3} + \frac{1}{2} = 1. \end{aligned}$$

Por lo tanto, $c = \frac{3}{2}$.
De esta manera,

$$f_X(x) = \left(\frac{3}{2}x^2 + x\right) I_{[0,1]}(x).$$

b) Para valores $0 \leq x \leq 1$, la función de distribución es

$$\begin{aligned} F_X(x) &= \int_{-\infty}^{x} f_X(t)dt \\ &= \int_{0}^{x} \left(\frac{3}{2}t^2 + t\right) dt \\ &= \frac{x^3}{2} + \frac{x^2}{2}. \end{aligned}$$

Obsérvese que para $x < 0$, la función de distribución es igual a 0, y para $x > 1$, la función de distribución es igual a 1. Por lo tanto, la función de distribución queda de la siguiente manera:

$$F_X(x) = \left(\frac{x^3}{2} + \frac{x^2}{2}\right) I_{[0,1]}(x) + I_{(1,\infty)}(x).$$

c) La probabilidad pedida es

$$F_X\left(\frac{1}{2}\right) = \frac{1}{2}\left(\frac{1}{8} + \frac{1}{4}\right) = 0.1875.$$

d) La probabilidad condicional se calcula como

$$\begin{aligned} P\left[X \geq \frac{1}{2} \middle| X \geq \frac{1}{4}\right] &= \frac{P\left[X \geq \frac{1}{2}\right]}{P\left[X \geq \frac{1}{4}\right]} \\ &= \frac{1 - F_X\left(\frac{1}{2}\right)}{1 - F_X\left(\frac{1}{4}\right)} = 0.8455. \end{aligned}$$

Ejemplo 2.8. Cuando un asegurado tiene un accidente, la cantidad reclamada por él, a la compañía de seguros por los daños ocasionados debido al siniestro, es una variable aleatoria X con función de densidad dada por $f_X(x) = c\left[xI_{(0,1.5)}(x) + (3-x)I_{[1.5,3)}(x)\right]$.
a) Encontrar c.
b) Encontrar la función de distribución $F_X(x)$.
c) Graficar las funciones de densidad y distribución.
d) Calcular $P[1.2 < X < 1.8]$.
e) Calcular $P[X > 2|X > 1.5]$.

a) La constante c se calcula de la siguiente manera:

$$\begin{aligned}
\int_{-\infty}^{\infty} f_X(x)dx &= c\left(\int_0^{1.5} xdx + \int_{1.5}^{3}(3-x)dx\right) \\
&= c\left(\left[\frac{x^2}{2}\right]_0^{1.5} + \left[3x - \frac{x^2}{2}\right]_{1.5}^{3}\right) \\
&= c\left(\frac{1.5^2}{2} + 3(3) - \frac{9}{2} - 4.5 + \frac{1.5^2}{2}\right) = 1.
\end{aligned}$$

De esta manera, $c = \dfrac{4}{9}$.

Entonces, la función de densidad se puede expresar de la siguiente manera:

$$f_X(x) = \frac{4}{9}\left[xI_{(0,1.5)}(x) + (3-x)I_{[1.5,3)}(x)\right].$$

b) La función de distribución se encontrará considerando 4 casos diferentes.

Caso 1. Cuando $x \leq 0$.

$$F_X(x) = 0.$$

Caso 2. Cuando $0 < X < 1.5$.

$$\begin{aligned}
F_X(x) &= \int_{-\infty}^{x} f_X(t)dt \\
&= \int_0^x \frac{4}{9}tdt \\
&= \frac{2}{9}x^2.
\end{aligned}$$

Caso 3. Cuando $1.5 \leq X < 3$.

$$\begin{aligned}
F_X(x) &= \int_{-\infty}^{x} f_X(t)dt \\
&= \int_0^{1.5} f_X(t)dt + \int_{1.5}^{x} f_X(t)dt \\
&= F_X(1.5) + \int_{1.5}^{x} \frac{4}{9}(3-t)dt \\
&= 0.5 + \left[\frac{4}{3}t - \frac{2}{9}t^2\right]_{1.5}^{x} \\
&= \frac{4}{3}x - \frac{2}{9}x^2 - 1.
\end{aligned}$$

Caso 4. Cuando $x > 3$.

$$F_X(x) = 1.$$

La función de distribución expresada con funciones indicadoras es

$$F_X(x) = \frac{2}{9}x^2 I_{(0,1.5)}(x) + \left(\frac{4}{3}x - \frac{2}{9}x^2 - 1\right) I_{[1.5,3)}(x) + I_{[3,\infty)}(x).$$

c) La gráfica de la función de densidad es

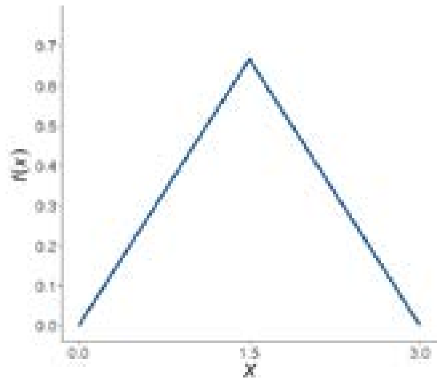


Figura 2.9. *Función de densidad, ejemplo 2.8*

La gráfica de la función de distribución es

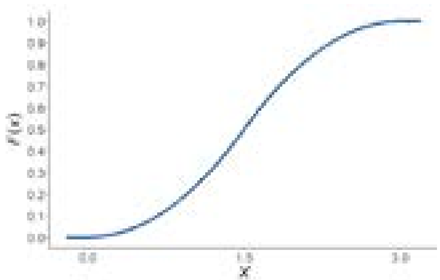


Figura 2.10. *Función de distribución, ejemplo 2.8*

d) La probabilidad pedida es

$$\begin{aligned} P[1.2 < X < 1.8] &= F_X(1.8) - F_X(1.2) \\ &= \frac{4}{3}(1.8) - \frac{2}{9}(1.8^2) - 1 - \frac{2}{9}(1.2^2) = 0.36. \end{aligned}$$

e) La probabilidad condicional se calcula de la siguiente manera:

$$\begin{aligned} P[X > 2 | X > 1.5] &= \frac{P[X > 2]}{P[X > 1.5]} \\ &= \frac{1 - F_X(2)}{1 - F_X(1.5)} \\ &= \frac{1 - \frac{4}{3}(2) + \frac{2}{9}(4) + 1}{\frac{1}{2}} = 0.44444. \end{aligned}$$

Las últimas dos probabilidades también pueden ser calculadas en forma geométrica usando la gráfica de la función de densidad. Esto en general es posible cuando las probabilidades son representadas por áreas de una figura geométrica conocida, como la de un triángulo o un rectángulo, entre otras. La función de distribución también puede ser obtenida en forma geométrica, por medio de áreas debajo de la gráfica de la función de densidad.

Las siguientes definiciones tienen aplicaciones en temas de confiabilidad y en temas de actuaría, como seguros, finanzas y teoría del riesgo, entre otras.

Definición 2.8. Sea X una variable aleatoria (asociada a un experimento), la *función de supervivencia o función de confiabilidad* es denotada como $S_X(x)$ y se define como el complemento de la función de distribución,

$$S_X(x) = 1 - F_X(x) = P[X > x].$$

La función de supervivencia se puede interpretar como la cola derecha de la distribución a partir de cierto valor x de la variable. Supongamos que X es la vida o duración de un artículo (o ser vivo), entonces la función de supervivencia evaluada en x, es la probabilidad de que el artículo (o ser vivo) sobreviva más de x tiempo.

Definición 2.9. Sea X una variable aleatoria continua la *función de tasa de fallos o función de riesgo* se denota como $h_X(x)$ y se define como

$$h_X(x) = \frac{f_X(x)}{S_X(x)}.$$

La función de tasa de fallos es muy útil en temas de análisis de supervivencia o confiabilidad. Como su nombre lo dice, nos indicará la fuerza con la que un artículo (o ser vivo) falla en un momento dado.

Suponemos de nuevo que X es la vida de un componente y $h_X(x)$ su función de riesgo. Es común observar que el comportamiento de la función de riesgo sea como la *curva de la bañera*, esto es, se espera que los primeros momentos de vida, la tasa de fallos sea alta, debido a errores de fabricación, posteriormente la tasa decrecerá conforme pase el tiempo, hasta llegar a un momento de estabilidad el cual se llama vida útil (del componente). En este periodo la tasa será constante o crecerá suavemente, hasta llegar al periodo de envejecimiento, donde la tasa de fallos crecerá notablemente. Es muy común que este comportamiente también se presente en un ser vivo.

Observación. En el tema de valores extremos, la función de supervivencia se denota como $\bar{F}_X(x)$ y se define como la *cola derecha de la distribución.* Es común que los análisis de valores extremos sean usados en temas de actuaría.

2.5. Valor esperado

El concepto de valor esperado es sumamente importante en temas de probabilidad y de estadística, y en diferentes áreas de aplicación, como las ciencias actuariales, confiabilidad, mejora continua y, en general, en todas las áreas donde se aplique la probabilidad y la estadística.

El concepto de valor esperado, media o promedio de una variable aleatoria será motivado por medio del siguiente ejemplo.

Ejemplo 2.9. Se tienen las calificaciones de 10 estudiantes: 9.8, 9.9, 9.7, 9.8, 10, 10, 9.8, 9.3, 5.0, 9.9. Calcular el promedio (muestral) de estas notas.

El promedio es

$$
\begin{aligned}
\bar{x} &= \frac{3(9.8) + 2(9.9) + 9.7 + 9.3 + 5 + 2(10)}{10} \\
&= 9.8\left(\frac{3}{10}\right) + 9.9\left(\frac{2}{10}\right) + 9.7\left(\frac{1}{10}\right) \\
&+ 9.3\left(\frac{1}{10}\right) + 5\left(\frac{1}{10}\right) + 10\left(\frac{2}{10}\right) = 9.32.
\end{aligned}
$$

Observemos que las cantidades que se encuentran dentro de los paréntesis son las frecuencias relativas (proporciones) de las calificaciones que se encuentran multiplicadas por estas. El principio de regularidad estadística (véase capítulo anterior, sección 1.4) dice que, bajo ciertas condiciones, la frecuencia relativa de un evento se aproximará a la probabilidad del evento, cuando el experimento se repite un gran número de veces.

El cálculo de la media o promedio de una variable aleatoria, en lugar de considerar las frecuencias relativas, considera las probabilidades de los eventos, véase la siguiente definición.

Definición 2.10. Sea X una variable aleatoria, el *valor esperado* (la *media*, *media poblacional*, *promedio, esperanza, esperanza matemática*) de X es denotado por $E[X]$, μ_X o simplemente μ y se define como

$$E[X] = \sum_x x f_X(x), \text{ si } X \text{ es una variable discreta,}$$

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx, \text{ si } X \text{ es una variable continua,}$$

siempre y cuando la suma o la integral sea finita.

Ejemplo 2.10. Considerando el ejemplo 2.4, del lanzamiento de tres monedas, encontrar el valor esperado del número de soles.

La función de densidad en este caso es

$$f_X(x) = \binom{3}{x}\left(\frac{1}{8}\right) I_{\{0,1,2,3\}}(x).$$

La media se calcula de la siguiente manera:

$$\begin{aligned} E[X] &= \sum_{x=0}^{3} x f_X(x) \\ &= 0 + \frac{3}{8} + \frac{6}{8} + \frac{3}{8} = 1.5. \end{aligned}$$

El significado de este valor es, si las tres monedas se arrojan un gran número de veces, entonces, el promedio del número de soles será alrededor de 1.5. Es importante aclarar que el promedio no necesariamente deberá de coincidir con un valor del recorrido.

Ejemplo 2.11. Considerando de nuevo el problema del seguro de automóvil, el ejemplo 2.8, calcular la cantidad promedio de reclamo por un asegurado.

En este caso, la función de densidad es

$$f_X(x) = \frac{4}{9}\left[xI_{(0,1.5)}(x) + (3-x)I_{[1.5,3)}(x)\right].$$

Entonces, el valor esperado de X se calcula como

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} x f_X(x)dx \\ &= \frac{4}{9}\left(\int_0^{1.5} x(x)dx + \int_{1.5}^{3} x(3-x)dx\right) \\ &= \frac{4}{9}\left(\left[\frac{x^3}{3}\right]_0^{1.5} + \left[3\frac{x^2}{2} - \frac{x^3}{3}\right]_{1.5}^{3}\right) = 1.5. \end{aligned}$$

A continuación, se generalizará la definición de valor esperado.

Definición 2.11. Sea X una variable aleatoria y $g(X)$ una función de X. Entonces, el *valor esperado* (la *media, media poblacional, promedio, esperanza, esperanza matemática)* de $g(X)$ se define como

$$E[g(X)] = \sum_x g(x)f_X(x), \text{ si } X \text{ es una variable discreta,}$$

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)f_X(x)dx, \text{ si } X \text{ es una variable continua,}$$

siempre y cuando la suma o la integral sea finita.

2.6. Propiedades de un valor esperado

El siguiente teorema considera las principales propiedades de un valor esperado.

Teorema 2.3. Propiedades de un valor esperado. *Sea X una variable aleatoria, $g_1(X)$ y $g_2(X)$ funciones de X, sean c, c_1 y c_2 constantes, entonces,*

(I) $E[c] = c$.

(II) $E[c_1 g_1(X) + c_2 g_2(X)] = c_1 E[g_1(X)] + c_2 E[g_2(X)]$.

(III) *Si* $g_1(x) \leq g_2(x)$ *para toda* $x \in \mathbb{R}$, *entonces,* $E[g_1(X)] \leq E[g_2(X)]$.

Demostración de (III) (caso continuo). Definimos la siguiente función

$$h(x) = g_2(x) - g_1(x).$$

Por hipótesis

$$h(x) \geq 0.$$

Por lo tanto,

$$h(x) f_X(x) \geq 0.$$

Entonces,

$$E[h(X)] = \int_{-\infty}^{\infty} h(x) f_X(x) dx \geq 0.$$

Ahora

$$\begin{aligned} E[h(X)] &= \int_{-\infty}^{\infty} h(x) f_X(x) dx \\ &= \int_{-\infty}^{\infty} g_2(x) f_X(x) dx - \int_{-\infty}^{\infty} g_1(x) f_X(x) dx \\ &= E[g_2(X)] - E[g_1(X)]. \end{aligned}$$

De este modo,

$$E[g_2(X)] - E[g_1(X)] \geq 0.$$

Esto es,

$$E[g_2(X)] \geq E[g_1(X)].$$

□

La demostración donde X es una variable discreta es análoga, en lugar de considerar integrales se consideran sumas sobre los valores del recorrido. Las demostraciones de los incisos I y II se quedan como ejercicio para el estudiante.

Observación. Si X es una variable aleatoria con función de densidad $f_X(x)$, la cual es simétrica con respecto c, esto es, $f_X(c - x) = f_X(c + x)$, entonces, $E[(X - c)^k] = 0$, para todo k entero impar. En particular, $E[X] = c$.

La demostración de la anterior afirmación se queda como ejercicio para el estudiante.

2.7. Varianza y desviación estándar

Se procederá a dar las definiciones de varianza y desviación estándar de una variable aleatoria, se considerarán algunas propiedades y se explicará la gran importancia que tienen estos conceptos, tanto en la probabilidad como en la estadística y en diversas áreas de aplicación.

Definición 2.12. Sea X una variable aleatoria, la *varianza* de X es denotada como $V[X]$ o σ_X^2 o simplemente σ^2, y se define como $V[X] = E[(X-\mu)^2]$, siempre y cuando el valor esperado sea finito. La *desviación estándar* se define como la raíz cuadrada positiva de la varianza, la cual se denota como σ_X o σ.

Obsérvese

$$V[X] = \sum_x (x - E[X])^2 f_X(x), \text{ si } X \text{ es discreta.}$$

$$V[X] = \int_{-\infty}^{\infty} (x - E[X])^2 f_X(x)dx, \text{ si } X \text{ es continua.}$$

Tanto la varianza como la desviación estándar son medidas de dispersión de los valores que toma una variable aleatoria. Esto es, si la varianza (desviación estándar) es «grande», entonces los valores que toma la variable, se encontrarán dispersos entre sí. Al contrario, si la varianza (desviación estándar) es «pequeña», entonces, los valores que toma la variable se parecerán entre sí.

Es importante aclarar, las unidades de medida de la varianza son las unidades de la variable aleatoria al cuadrado. La desviación estándar tiene las mismas unidades que la variable aleatoria.

Teorema 2.4. *Sea X una variable aleatoria, entonces,*

$$V[X] = E[X^2] - E^2[X].$$

Demostración.

$$\begin{aligned} V[X] &= E[(X-\mu)^2] \\ &= E[X^2] - 2\mu E[X] + \mu^2 \\ &= E[X^2] - \mu^2. \end{aligned}$$

□

Teorema 2.5. *Sea X una variable aleatoria, entonces,*

a) $V[X] \geq 0$, *y* $V[X] = 0$ *si y solo si X es una constante.*

b) $V[aX + b] = a^2 V[X]$, *donde a y b son constantes.*

La demostración del anterior teorema se deja como ejercicio para el estudiante.

Ejemplo 2.12. Considerando el ejemplo 2.4 donde se lanzan 3 monedas, calcular la varianza y desviación estándar de la variable número de soles.

El valor esperado de X^2 es

$$\begin{aligned} E[X^2] &= \sum_{x=0}^{3} x^2 \binom{3}{x} \frac{1}{8} \\ &= 0 + \frac{3}{8} + \frac{12}{8} + \frac{9}{8} = 3. \end{aligned}$$

Por lo tanto,

$$V[X] = 3 - (1.5)^2 = 0.75,$$

y la desviación estándar

$$\sigma = 0.86603.$$

Ejemplo 2.13. Considerando el ejemplo 2.8, del seguro de auto, calcular la varianza y desviación estándar de la cantidad reclamada por un asegurado.

El cálculo de $E[X^2]$

$$\begin{aligned} E[X^2] &= \int_{-\infty}^{\infty} x^2 f_X(x)dx \\ &= \frac{4}{9}\left(\int_0^{1.5} x^2(x)dx + \int_{1.5}^{3} x^2(3-x)dx\right) \\ &= \frac{4}{9}\left(\left[\frac{x^4}{4}\right]_0^{1.5} + \left[x^3 - \frac{x^4}{4}\right]_{1.5}^{3}\right) = \frac{21}{8}. \end{aligned}$$

Por lo tanto,

$$V[X] = \frac{21}{8} - (1.5)^2 = \frac{3}{8},$$

y la desviación estándar

$$\sigma = 0.61237.$$

La esperanza de una variable aleatoria no siempre existe. Si la suma o la integral correspondiente no es finita, entonces, el valor esperado no existe, véase los siguientes ejercicios.

Ejemplo 2.14. Sea X una variable aleatoria con función de densidad dada por $f_X(x) = \frac{1}{x^2} I_{(1,\infty)}(x)$.
a) Demostrar que $f_X(x)$ es función de densidad.
b) Verificar si existe el valor esperado de X.

a) Obsérvese

$$\frac{1}{x^2} > 0 \text{ para todo } x > 1, \text{ y } I_{(1,\infty)}(x) \geq 0 \text{ para toda } x \in \mathbb{R}.$$

Por lo tanto,

$$f_x(x) = \frac{1}{x^2} I_{(1,\infty)}(x) \geq 0.$$

Por otro lado,

$$\int_1^\infty \left(\frac{1}{x^2}\right) dx = \left[-\left(\frac{1}{x}\right)\right]_1^\infty = 1.$$

De este modo, $f_X(x)$ es función de densidad.
b) Ahora,

$$\begin{aligned} \int_1^\infty x\left(\frac{1}{x^2}\right) dx &= \int_1^\infty \frac{1}{x} dx \\ &= [ln(x)]_1^\infty \\ &= \lim_{x\to\infty} ln(x) - 0 \to \infty. \end{aligned}$$

Esto es, el valor esperado de X no existe.

Ejemplo 2.15. Sea X una variable aleatoria con función de densidad dada por $f_X(x) = \frac{6}{\pi^2 x^2} I_{\{1,2,3,\ldots\}}(x)$.
a) Demostrar que $f_X(x)$ es función de densidad.
b) Verificar si existe el valor esperado de X.

a) Obsérvese $\dfrac{6}{\pi^2 x^2} > 0$ para $x \in \{1, 2, \ldots\}$, y $I_{\{1,2,3,\ldots\}}(x) \geq 0$.
De este modo,

$$f_X(x) = \frac{6}{\pi^2 x^2} I_{\{1,2,3,\ldots\}}(x) \geq 0.$$

Por otro lado,

$$\begin{aligned} \sum_{x=1}^\infty \frac{6}{\pi^2 x^2} &= \frac{6}{\pi^2} \sum_{x=1}^\infty \frac{1}{x^2} \\ &= \frac{6}{\pi^2}\frac{\pi^2}{6} = 1, \end{aligned}$$

aplicando el resultado de la serie C.10 del apéndice.
Por lo tanto, $f_X(x)$ es función de densidad.
b) Ahora obsérvese,

$$\begin{aligned} \sum_{x=1}^\infty x\frac{6}{\pi^2 x^2} &= \sum_{x=1}^\infty \frac{6}{\pi^2 x} \\ &= \frac{6}{\pi^2} \sum_{x=1}^\infty \frac{1}{x} \to \infty. \end{aligned}$$

Concluimos que el valor esperado de X no existe.

2.8. Función generadora de momentos

En esta sección se tratarán los conceptos de momentos con respecto al origen, los momentos centrales y la definición de función generadora de momentos.

Es importante mencionar que los momentos de una variable aleatoria son importantes para interpretar algunas características de la distribución de la variable aleatoria. Cabe señalar que los momentos también son utilizados en métodos de inferencia estadística, en particular en estimación puntual.

Por otro lado, la función generadora de momentos, como su nombre lo dice, genera momentos con respecto al origen, pero además tiene más aplicaciones, por ejemplo, para encontrar la distribución de una función de variables aleatorias (véase capítulo 6), o también para encontrar la convergencia en distribución de sucesiones de variables aleatorias (véase capítulo 7), entre otras aplicaciones.

Definición 2.13. Sea X una variable aleatoria, el *r-ésimo momento con respecto al origen* de X se denota como ${\mu'}_r$ y se define como ${\mu'}_r = E[X^r]$, y el *r-ésimo momento centra*l de X se denota como μ_r y se define como $\mu_r = E[(X-\mu)^r]$.

Observación. El primer momento de X con respecto al origen es la media de X. El primer momento central de X es 0 y el segundo momento central de X es la varianza de X.

A continuación, se dará la definición de la función generadora de momentos de una variable aleatoria X.

Definición 2.14. Sea X una variable aleatoria, la *función generadora de momentos* (o *función generatriz de momentos*) de X es denotada como $m_X(t)$, y se define como $m_X(t) = E[e^{Xt}]$.

Obsérvese que en general se cumple $m_X(0) = 1$.

Por otro lado, si $P[X \leq 0] = 1$, entonces, la función generadora de momentos $m_X(t)$ es finita, para toda $t > 0$. Esto es, si $x \leq 0$, cuando $t > 0$, se cumple $e^{tx} \leq 1$, y por lo tanto, $E[e^{tX}] \leq 1$.

De manera similar, si $P[X \geq 0] = 1$, entonces, $m_X(t)$ también es finita, cuando $t < 0$. Esto es, si $x \geq 0$, cuando $t < 0$, se puede ver $e^{tx} \leq 1$, y por lo tanto, $E[e^{tX}] \leq 1$.

La primera utilidad importante de la función generadora de momentos es, como su nombre lo dice, generar los momentos con respecto al origen.

Esto es, si X es una variable aleatoria con función generadora de momentos $m_X(t)$, la cual existe para toda $t \in (-h, h)$, para alguna $h > 0$, entonces,

$$\mu'_r = E[X^r] = \left.\frac{d^r m_X(t)}{dt^r}\right|_{t=0}.$$

Obsérvese que es posible desarrollar e^{tX} en serie de Maclaurin (véase ecuación C.13 del apéndice)

$$\begin{aligned} e^{tX} &= \sum_{k=0}^{\infty} \frac{(tX)^k}{k!} \\ &= 1 + tX + \frac{(tX)^2}{2!} + \frac{(tX)^3}{3!} + \cdots \end{aligned}$$

De esta manera,

$$\begin{aligned} m_X(t) &= E[e^{tX}] \\ &= E\left[\sum_{k=0}^{\infty} \frac{(tX)^k}{k!}\right] \\ &= 1 + tE[X] + \frac{t^2}{2!}E[X^2] + \frac{t^3}{3!}E[X^3] + \cdots \end{aligned}$$

Como el valor esperado de e^{tX} existe, cada valor esperado $E[X^k]$ también existe.

Ahora obsérvese

$$\frac{dm_X(t)}{dt} = E[X] + \frac{2t}{2!}E[X^2] + \frac{3t^2}{3!}E[X^3] + \frac{4t^3}{4!}E[X^4] + \cdots$$

Por lo tanto,

$$\left.\frac{dm_X(t)}{dt}\right|_{t=0} = E[X].$$

La segunda derivada es

$$\begin{aligned} \frac{d^2m_X(t)}{dt^2} &= E[X^2] + \frac{3(2)t}{3!}E[X^3] + \frac{4(3)t^2}{4!}E[X^4] + \frac{5(4)t^3}{5!}E[X^5] + \cdots \\ &= E[X^2] + tE[X^3] + \frac{t^2}{2!}E[X^4] + \frac{t^3}{3!}E[X^5] + \cdots \end{aligned}$$

De esta manera,

$$\left.\frac{d^2m_X(t)}{dt^2}\right|_{t=0} = E[X^2].$$

Es claro que la r-ésima derivada se puede expresar como

$$\frac{d^rm_X(t)}{dt^r} = E[X^r] + tE[X^{r+1}] + \frac{t^2}{2!}E[X^{r+2}] + \frac{t^3}{3!}E[X^{r+3}] + \cdots$$

De este modo,

$$\left.\frac{d^rm_X(t)}{dt^r}\right|_{t=0} = E[X^r] = \mu'_r.$$

Observación. De las anteriores expresiones, vamos a resaltar, que al existir la función generadora de momentos, en un intervalo para t que contenga al

0, esta se puede desarrollar como una serie, siendo esta la serie de Maclaurin para $m_X(t)$, esto es,

$$m_X(t) = 1 + t\mu_1' + \frac{t^2}{2!}\mu_2' + \frac{t^3}{3!}\mu_3' + \cdots \tag{2.1}$$

Observación. Para una variable aleatoria X con función generadora de momentos $m_X(t)$ que existe para $t \in (-h, h)$, se cumple

$$\left.\frac{d}{dt} ln[m_X(t)]\right|_{t=0} = E[X],$$

$$\left.\frac{d^2}{dt^2} ln[m_X(t)]\right|_{t=0} = V[X].$$

La demostración se deja como ejercicio para el estudiante.

Ejemplo 2.16. El tiempo en horas para que un ajustador de seguros llegue al lugar de los hechos de un accidente automovilístico, reportado por un asegurado en una ciudad, es una variable aleatoria X con función de densidad dada por $f_X(x) = e^{-x} I_{(0,\infty)}(x)$. Encontrar la función generadora de momentos, la media y la varianza de X.

Obsérvese

$$\begin{aligned} m_X(t) &= E\left[e^{tX}\right] \\ &= \int_0^\infty e^{tx} e^{-x} dx \\ &= \int_0^\infty e^{x(t-1)} dx \\ &= (t-1)^{-1} e^{x(t-1)}\Big|_0^\infty \\ &= (1-t)^{-1}, \quad \text{si} \quad t-1 < 0. \end{aligned}$$

Por lo tanto,

$$m_X(t) = (1-t)^{-1}, \quad \text{si} \quad t < 1.$$

Ahora calcularemos los primeros dos momentos con respecto al origen

$$\begin{aligned} E[X] &= \left.\frac{dm_X(t)}{dt}\right|_{t=0} \\ &= (1-t)^{-2}\Big|_{t=0} = 1, \end{aligned}$$

$$\begin{aligned} E[X^2] &= \left.\frac{d^2 m_X(t)}{dt^2}\right|_{t=0} \\ &= 2(1-t)^{-3}\Big|_{t=0} = 2. \end{aligned}$$

Por lo tanto,

$$V[X] = E[X^2] - E^2[X] = 1.$$

Ejemplo 2.17. Encontrar la función generadora de momentos, la media y la varianza de la variable aleatoria número de soles, del ejemplo 2.4.

La función de densidad en este caso es

$$f_X(x) = \binom{3}{x}\frac{1}{8}I_{\{0,1,2,3\}}(x).$$

Entonces,

$$\begin{aligned} m_X(t) &= E[e^{tX}] \\ &= \sum_{x=0}^{3} e^{tx}\binom{3}{x}\frac{1}{8} \\ &= \frac{1}{8}\left(1 + 3e^t + 3e^{2t} + e^{3t}\right). \end{aligned}$$

De esta manera,

$$\begin{aligned} E[X] &= \left.\frac{dm_X(t)}{dt}\right|_{t=0} \\ &= \frac{1}{8}\left(0 + 3e^t + 6e^{2t} + 3e^{3t}\right)\Big|_{t=0} = 1.5, \end{aligned}$$

$$\begin{aligned} E[X^2] &= \left.\frac{d^2m_X(t)}{dt^2}\right|_{t=0} \\ &= \frac{1}{8}\left(3e^t + 12e^{2t} + 9e^{3t}\right)\Big|_{t=0} = 3. \end{aligned}$$

Por lo tanto,

$$V[X] = E[X^2] - E^2[X] = 0.75.$$

Otra manera de obtener la función generadora de momentos es aplicando el teorema del binomio (véase apéndice C), esto es,

$$\begin{aligned} m_X(t) &= E[e^{tX}] \\ &= \sum_{x=0}^{3} e^{tx}\binom{3}{x}\frac{1}{8} \\ &= \sum_{x=0}^{3} e^{tx}\binom{3}{x}\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{3-x} \\ &= \sum_{x=0}^{3}\binom{3}{x}\left(\frac{1}{2}e^t\right)^x\left(\frac{1}{2}\right)^{3-x} \\ &= \left(\frac{1}{2}e^t + \frac{1}{2}\right)^3. \end{aligned}$$

Ejemplo 2.18. El número de llamadas en una hora a una compañía de seguros, para reportar un siniestro de automóvil, es una variable aleatoria X con función de densidad dada por $f_X(x) = c\left(\frac{5^x}{x!}\right) I_{\{0,1,2,\ldots\}}(x)$.
a) Encontrar c.
b) Encontrar la función generadora de momentos de X.
c) Encontrar la media y la varianza de X.

a) Obsérvese

$$\begin{aligned}\sum_{x=0}^{\infty} c\left(\frac{5^x}{x!}\right) &= c\sum_{x=0}^{\infty}\left(\frac{5^x}{x!}\right)\\ &= ce^5 = 1.\end{aligned}$$

Por lo tanto,

$$c = e^{-5}.$$

De este modo,

$$f_X(x) = \frac{5^x}{x!}e^{-5}I_{\{0,1,2,\ldots\}}(x).$$

b) La función generadora de momentos es

$$\begin{aligned}m_X(t) &= E[e^{Xt}]\\ &= \sum_{x=0}^{\infty} e^{xt}\frac{5^x}{x!}e^{-5}\\ &= e^{-5}\sum_{x=0}^{\infty}\frac{(5e^t)^x}{x!}\\ &= e^{-5}e^{5e^t}\\ &= e^{5(e^t-1)}.\end{aligned}$$

Obsérvese que se aplicó el desarrollo de e^{5e^t} en series de Maclaurin (véase apéndice C).

c) La media y la varianza

$$\begin{aligned}E[X] &= \left.\frac{dm_X(t)}{dt}\right|_{t=0}\\ &= \left.e^{5(e^t-1)}5e^t\right|_{t=0} = 5,\end{aligned}$$

$$\begin{aligned}E[X^2] &= \left.\frac{d^2m_X(t)}{dt^2}\right|_{t=0}\\ &= \left.e^{5(e^t-1)}5e^t + 5e^te^{5(e^t-1)}5e^t\right|_{t=0} = 30.\end{aligned}$$

Por lo tanto,

$$V[X] = E[X^2] - E^2[X] = 5.$$

El siguiente teorema asegura una relación uno a uno, entre las distribuciones y las funciones generadoras de momentos, y se le conoce como el *teorema de unicidad.*

Teorema 2.6. Teorema de unicidad. *Sean X y Y dos variables aleatorias con funciones de densidad $f_X(x)$ y $f_Y(y)$, respectivamente, y funciones generadoras de momentos $m_X(t)$ y $m_Y(t)$, respectivamente. Si $m_X(t) = m_Y(t)$ para toda t en el intervalo $(-h, h)$, donde $h > 0$, entonces, $f_X(x) = f_Y(x)$.*

Este teorema será de gran utilidad para el tema de distribuciones de funciones de variables aleatorias y su demostración no se realizará.

A continuación, se presentará la definición de función característica, a partir de esta, también se pueden generar momentos con respecto al origen.

Definición 2.15. Sea X una variable aleatoria, la *función característica* de X se denota como $\varphi_X(t)$ y se define como

$$\varphi_X(t) = E[e^{Xti}], \text{ donde } i = \sqrt{-1}.$$

Se puede demostrar que la función característica siempre existe para toda variable aleatoria. La demostración para el caso continuo se presenta a continuación:

$$\begin{aligned}
|\varphi_X(t)| &= \left|E\left[e^{Xti}\right]\right| \\
&= \left|\int_{-\infty}^{\infty} e^{Xti} f_X(x)dx\right| \\
&\leq \int_{-\infty}^{\infty} \left|e^{Xit}\right| f_X(x)dx \\
&= \int_{-\infty}^{\infty} |cos(xt) + isen(xt)|\, f_X(x)dx \\
&= \int_{-\infty}^{\infty} \sqrt{cos^2(xt) + sen^2(xt)} f_X(x)dx = 1.
\end{aligned}$$

Por otro lado, se puede verificar que

$$\left.\frac{d^r\varphi_X(t)}{dt^r}\right|_{t=0} = i^r E[X^r].$$

Cabe aclarar que no es propósito de este material utilizar la función característica.

2.9. Función generadora de momentos factoriales

En esta sección se tratarán los momentos factoriales y la función generadora de momentos factoriales. Cabe señalar que otra forma de calcular la media y la varianza de una variable aleatoria es por medio de sus dos primeros momentos factoriales.

Definición 2.16. Sea X una variable aleatoria, el *r-ésimo momento factorial* de X se define como $E[X(X-1)(X-2)\cdots(X-(r-1))]$.

A continuación, se presenta la definición de función generadora de momentos factoriales.

Definición 2.17. Sea X una variable aleatoria, la *función generadora de momentos factoriales* de X se denota como $M_X(t)$ y se define como $M_X(t) = E\left[t^X\right]$ si el valor esperado existe.

La función generadora de momentos factoriales sirve para obtener los momentos factoriales de una variable aleatoria. Esto es, si X es una variable aleatoria con función generadora de momentos factoriales $M_X(t)$, entonces,

$$E[X(X-1)(X-2)\cdots(X-(r-1))] = \left.\frac{d^r M_X(t)}{dt^r}\right|_{t=1}.$$

Supongamos que X es una variable aleatoria discreta, obsérvese

$$\begin{aligned} M_X(t) &= E\left[t^X\right] \\ &= \sum_x t^x f_X(x). \end{aligned}$$

Por lo tanto,

$$\begin{aligned} \frac{dM_X(t)}{dt} &= \frac{d}{dt}\sum_x t^x f_X(x) \\ &= \sum_x \frac{d}{dt} t^x f_X(x) \\ &= \sum_x x t^{x-1} f_x(x). \end{aligned}$$

La segunda derivada es

$$\frac{d^2}{dt^2} M_X(t) = \sum_x x(x-1)t^{x-2} f_x(x).$$

De este modo, podemos deducir que la r-ésima derivada es

$$\frac{d^r}{dt^r} M_X(t) = \sum_x x(x-1)\cdots(x-(r-1))t^{x-r} f_x(x).$$

De esta manera, concluimos

$$\frac{d^r}{dt^r}M_X(t)\bigg|_{t=1} = E[X(X-1)(X-2)\cdots(X-(r-1))].$$

Para una variable aleatoria continua X, el desarrollo anterior es análogo, en lugar de usar sumas, se considera integrales.

Ejemplo 2.19. Encontrar la función generadora de momentos factoriales de la variable aleatoria X con función de densidad dada por $f_X(x) = e^{-x}I_{(0,\infty)}(x)$, y a partir de esta, calcular la media y la varianza.

Obsérvese

$$\begin{aligned}
M_X(t) &= E[t^X] \\
&= \int_0^\infty t^x e^{-x}dx \\
&= \int_0^\infty \left(\frac{t}{e}\right)^x dx \\
&= \left(ln\left(\frac{t}{e}\right)\right)^{-1}\left(\frac{t}{e}\right)^x\bigg|_0^\infty \\
&= 0 - \left(ln\left(\frac{t}{e}\right)\right)^{-1}, \text{ si } \quad \frac{t}{e} < 1.
\end{aligned}$$

Por lo tanto,

$$M_X(t) = (1 - ln(t))^{-1}, \quad \text{si} \quad t < e.$$

El primer y segundo momento factorial se calculan, respectivamente,

$$\begin{aligned}
E[X] &= \frac{d}{dt}M_X(t)\bigg|_{t=1} \\
&= (t(1-ln(t))^2)^{-1}\bigg|_{t=1} = 1,
\end{aligned}$$

$$\begin{aligned}
E[X(X-1)] &= \frac{d^2}{dt^2}M_X(t) \\
&= 2t^{-2}(1-ln(t))^{-3} - t^{-2}(1-ln(t))^{-2}\bigg|_{t=1} = 1.
\end{aligned}$$

Concluimos

$$V[X] = E[X(X-1)] + E[X] - E^2[X] = 1.$$

Se deja al estudiante encontrar la función generadora de momentos factoriales, la media y la varianza, para la variable aleatoria del ejemplo 2.4.

Ejemplo 2.20. Consideremos el ejemplo 2.18 donde la variable aleatoria X es el número de llamadas en una hora para reportar siniestros automovilísticos. Encontrar la función generadora de momentos factoriales de X y a partir de esta, calcular la media y la varianza de X.

La función de densidad en este caso es $f_X(x) = \frac{5^x}{x!}e^{-5}I_{\{0,1,2,\cdots\}}(x)$. Primero obtendremos la función generadora de momentos factoriales

$$\begin{aligned} M_X(t) &= E[t^X] \\ &= \sum_{x=0}^{\infty} t^x \frac{5^x}{x!} e^{-5} \\ &= e^{-5} \sum_{x=0}^{\infty} \frac{(5t)^x}{x!} \\ &= e^{5(t-1)}. \end{aligned}$$

El primer y segundo momentos factoriales se calculan, respectivamente,

$$\begin{aligned} E[X] &= \frac{d}{dt} M_X(t) \bigg|_{t=1} \\ &= e^{5(t-1)} 5 \bigg|_{t=1} = 5, \end{aligned}$$

$$\begin{aligned} E[X(X-1)] &= \frac{d^2}{dt^2} M_X(t) \\ &= e^{5(t-1)} 5(5) \bigg|_{t=1} = 25. \end{aligned}$$

Por lo tanto,

$$V[X] = E[X(X-1)] + E[X] - E^2[X] = 5.$$

2.10. Función generadora de probabilidades

En esta parte, se definirá la función generadora de probabilidades, la cual es válida para variables aleatorias discretas con valores enteros no negativos y, como su nombre lo dice, con esta función se podrán generar probabilidades para este tipo de variables aleatorias.

Definición 2.18. Sea X una variable aleatoria discreta con valores enteros no negativos, entonces, a la función generadora de momentos factoriales de X, $M_X(t)$, se le conoce también como *función generadora de probabilidades* de X, si esta existe.

Supongamos que X es una variable aleatoria discreta con valores enteros no negativos y con función generadora de probabilidades $M_X(t)$, entonces,

$$f_X(r) = \frac{1}{r!}\left(\frac{d^r M_X(t)}{dt^r}\bigg|_{t=0}\right), \text{ donde } r = 0, 1, \ldots, n,$$

entendiendo que la derivada de orden cero es la misma función generadora de probabilidades.

Obsérvese que la función generadora de probabilidades se puede expresar de la siguiente manera:

$$\begin{aligned} M_X(t) &= E\left[t^X\right] \\ &= \sum_{x=0}^{\infty} t^x f_X(x) \\ &= f_X(0) + t f_X(1) + t^2 f_X(2) + \cdots + t^n f_X(n) + \cdots \end{aligned}$$

Entonces, la primera derivada es

$$\frac{d}{dt} M_X(t) = f_X(1) + 2t f_X(2) + 3t^2 f_X(3) + \cdots + n t^{n-1} f_X(n) + \cdots$$

La segunda derivada es

$$\frac{d^2}{dt^2} M_X(t) = 2(1) f_X(2) + 3(2) t f_X(3) + \cdots + n(n-1) t^{n-2} f_X(n) + \cdots$$

De este modo, se puede deducir que la r-ésima derivada es

$$\begin{aligned} \frac{d^r}{dt^r} M_X(t) = r(r-1)\cdots(2)(1) f_X(r) + (r+1)(r)\cdots(2) t f_X(r+1)\cdots \\ + n(n-1)\cdots(n-(r-1)) t^{n-r} f_X(n) + \cdots \end{aligned}$$

Entonces, las anteriores derivadas valuadas en cero son

$$M_X(t)\bigg|_{t=0} = f_X(0),$$

$$\frac{d}{dt} M_X(t)\bigg|_{t=0} = f_X(1),$$

$$\frac{d^2}{dt^2} M_X(t)\bigg|_{t=0} = 2(1) f_X(2) = 2! f_X(x),$$

$$\begin{aligned} \frac{d^r}{dt^r} M_X(t) &= r(r-1)\cdots(2)(1) f_X(r)\bigg|_{t=0} \\ &= r(r-1)\cdots(2)(1) f_X(r) \\ &= r! f_X(r). \end{aligned}$$

Por lo tanto,

$$f_X(r) = \frac{1}{r!}\left(\frac{d^r M_X(t)}{dt^r}\bigg|_{t=0}\right).$$

Ejemplo 2.21. Si X es una variable aleatoria discreta que toma valores enteros no negativos y con función generadora de probabilidades $M_X(t) = \frac{1}{8}\left(1+3t+3t^2+t^3\right)$. Encontrar la función de densidad de X.

Obsérvese los siguientes cálculos

$$f_X(0) = M_X(t)\Big|_{t=0} = \frac{1}{8},$$

$$\begin{aligned} f_X(1) &= \frac{d}{dt}M_X(t)\Big|_{t=0} \\ &= \frac{1}{8}\left(3+6t+3t^2\right)\Big|_{t=0} = \frac{3}{8}, \end{aligned}$$

$$\begin{aligned} 2!f_X(2) &= \frac{d^2}{dt^2}M_X(t)\Big|_{t=0} \\ &= \frac{1}{8}\left(6+6t\right)\Big|_{t=0} = \frac{6}{8}, \end{aligned}$$

$$\begin{aligned} 3!f_X(3) &= \frac{d^3}{dt^3}M_X(t)\Big|_{t=0} \\ &= \frac{1}{8}\left(6\right)\Big|_{t=0} = \frac{6}{8}. \end{aligned}$$

Para cualquier otro caso, la derivada vale cero, de este modo,

$$f_X(0) = \frac{1}{8}; f_X(1) = \frac{3}{8}; f_X(2) = \frac{3}{8} \text{ y } f_X(3) = \frac{1}{8}.$$

Concluimos,

$$f_X(x) = \frac{1}{8}I_{\{0,3\}}(x) + \frac{3}{8}I_{\{1,2\}}(x).$$

Ejemplo 2.22. Sea X una variable aleatoria discreta con valores enteros no negativos, con función generadora de probabilidades dada por $M_X(t) = e^{5(t-1)}$, obtener la función de densidad.

Calculemos la primera y segunda derivada, respectivamente,

$$\frac{d}{dt}M_X(t) = e^{5(t-1)}(5),$$

$$\frac{d^2}{dt^2}M_X(t) = 5e^{5(t-1)}(5) = 5^2e^{5(t-1)}.$$

Se puede observar que la r-ésima derivada es

$$\frac{d^r}{dt^r}M_X(t) = 5^{r-1}e^{5(t-1)}(5) = 5^re^{5(t-1)}.$$

Por lo tanto,

$$\frac{d^r}{dt^k} M_X(t)\bigg|_{t=0} = 5^r e^{-5}.$$

De esta manera,

$$f_X(r) = \frac{1}{r!} 5^r e^{-5} I_{\{0,1,2,\ldots\}}(r).$$

2.11. Moda, cuantiles y mediana

En esta sección se darán las definiciones de moda, cuantiles y mediana de una variable aleatoria. Se tratarán varios ejemplos, tanto para variables discretas como para variables continuas. Cabe señalar que estos conceptos son muy importantes en diversos temas de análisis estadístico, también en diferentes problemas de ciencias actuariales, entre otras áreas de aplicación.

Definición 2.19. Sea X una variable aleatoria, la *moda* de X o de la distribución de X es el valor o los valores del recorrido de la variable donde la función de densidad alcanza su máximo.

Observaciones.

1. Para una determinada variable aleatoria, puede existir una única moda. Otra posibilidad es que exista más de una moda, o puede suceder que no existan modas. En este capítulo y los siguientes dos, se tratarán algunos ejemplos.
2. Si la función de densidad tiene diferentes máximos de diferentes valores, entonces, podemos decir que existen la moda primaria, la moda secundaria, etc.
3. Los criterios de cálculo diferencial, para encontrar mínimos y/o máximos, puede ser de utilidad para encontrar la moda o modas de una variable aleatoria, tanto en el caso discreto como en el caso continuo, mas no es una metodología matemática completamente determinante para encontrar modas. En general, el cálculo diferencial será una herramienta útil para el análisis de la gráfica de una función de densidad y de esta manera encontrar lo que se pretende.
4. Sobre todo para variables aleatorias discretas, no existe una metodología bien definida para encontrar modas. En el siguiente capítulo, para algunos modelos discretos, se presentarán sus modas, y no necesariamente se encontrarán por medio de metodologías de cálculo diferencial. En los siguientes ejemplos se tratarán algunos casos discretos, y estos se resolverán observando y analizando la gráfica de la función de densidad.

Ejemplo 2.23. Sea X una variable aleatoria con función de densidad dada por $f_X(x) = \binom{6}{x}(0.5)^6 I_{\{0,1,2,3,4,5,6\}}(x)$. Encontrar la moda.

Una forma simple de encontrar la moda en un caso discreto es realizando la gráfica de la función de densidad. Es recomendable apoyarse con un paquete computacional adecuado. La gráfica de esta función de densidad se realizó por medio de R.

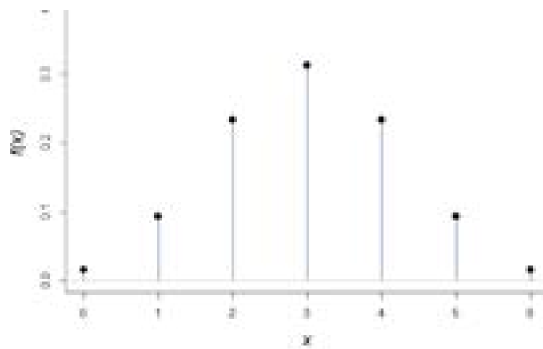


Figura 2.11. *Función de densidad, ejemplo 2.23*

Se puede observar que el máximo de la función de densidad es 0.3125 y este es alcanzado en $x = 3$, por lo tanto, la moda es igual a 3.

Ejemplo 2.24. Sea X una variable aleatoria con función de densidad dada por $f_X(x) = \binom{5}{x}(0.5)^5 I_{\{0,1,2,3,4,5\}}(x)$. Encontrar la moda.

Realizaremos la gráfica de la función de densidad.

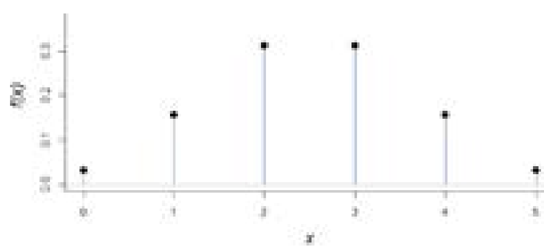


Figura 2.12. *Función de densidad, ejemplo 2.24*

De esta manera, se observan dos modas, $x = 2$ y $x = 3$.

En el siguiente capítulo, para algunos modelos discretos se presentarán criterios más específicos y formales para encontrar la moda o las modas, según sea el caso.

Ejemplo 2.25. Sea X una variable aleatoria con función de densidad dada por $f_X(x) = \left(\frac{3}{2}x^2 + x\right) I_{[0,1]}(x)$. Encontrar la moda.

Usaremos los criterios de cálculo diferencial para encontrar el máximo o mínimo relativo

$$\frac{df_X(x)}{dx} = 3x + 1 = 0.$$

Despejando, nos queda el siguiente punto crítico $x = -\frac{1}{3}$.

Realizaremos la segunda derivada.

$$\frac{d^2 f_X(x)}{dx^2} = 3.$$

Lo que significa que en $x = -\dfrac{1}{3}$, hay un mínimo.

Con los criterios de cálculo diferencial, no se encontraron puntos críticos dentro del recorrido de X. Ahora, si en $x = -\frac{1}{3}$ hay un mínimo y la gráfica de la función de densidad es una parábola, entonces, a partir de $x = -\frac{1}{3}$ la función es creciente. En particular, la gráfica de la función de densidad en el intervalo $[0, 1]$ es creciente, por lo que se concluye que la función de densidad alcanza su máximo en $x = 1$, considerando su recorrido. En conclusión, la moda de X es $x = 1$.

Los criterios de cálculo diferencial para encontrar máximos y mínimos son útiles, pero no son definitivos para encontrar una moda, fue necesario realizar un análisis más detallado de la gráfica de la función de densidad, en específico en valores del recorrido. Esto se debe a que muchas veces las funciones de densidad están restringidas en un intervalo y los métodos de optimización de cálculo consideran a todos los números reales.

A continuación, se presentará la definición del r-ésimo cuantil de una variable aleatoria.

Definición 2.20. Sea X una variable aleatoria y $0 \leq r \leq 1$, el *r-ésimo cuantil* de X se denota como ζ_r y se define como el número más pequeño del recorrido de X que cumple con la siguiente desigualdad: $P[X \leq \zeta_r] \geq r$.

Obsérvese que si X es una variable aleatoria continua, el r-ésimo cuantil de X se define como el número ζ_r que cumple con la siguiente igualdad $P[X \leq \zeta_r] = F_X(\zeta_r) = r$.

Para el caso discreto, es el primer número del recorrido ζ_r que cumple con $P[X \leq \zeta_r] > r$ o $P[X \leq \zeta_r] = F_X(\zeta_r) = r$.

Definición 2.21. Sea X una variable aleatoria, la *mediana* de X se denota como m y se define como $m = 0.5$-ésimo cuantil.

Observación. La definición de percentil de una variable aleatoria es prácticamente la misma que la de cuantil, la diferencia es que la definición de percentil está dada en términos de porcentaje. Por ejemplo, el 0.3-ésimo cuantil equivale al percentil 30.

Observación. Si una variable aleatoria X tiene una función de densidad $f_X(x)$ simétrica con respecto a c, entonces, la mediana de X es c.

La demostración se deja como ejercicio para el estudiante.

A continuación, se presentarán algunos ejemplos de cálculo de cuantiles, tanto de variables aleatorias discretas como continuas.

Ejemplo 2.26. Sea X una variable aleatoria con función de densidad dada por $f_X(x) = \binom{5}{x}(0.5)^5 I_{\{0,1,2,3,4,5\}}(x)$. Encontrar el 0.35-ésimo cuantil y la mediana de X.

Calcularemos los valores de la función de densidad y de la función de distribución para los primeros valores del recorrido de la variable, hasta que la probabilidad acumulada llegue o pase por primera vez de 0.35, y posteriormente de 0.5

x	$f_X(x)$	$F_X(x)$
0	0.03125	0.03125
1	0.15625	0.1875
2	0.3125	0.5

Tabla 2.3. *Probabilidades, ejemplo 2.26*

Con estos valores es suficiente para poder concluir $\zeta_{0.35} = m = 2$. En este caso, de acuerdo a la definición de cuantil, $\zeta_{0.35} = m$.

Ejemplo 2.27. Encontrar el 0.35-ésimo cuantil y la mediana de la variable aleatoria, cantidad reclamada por un asegurado, del ejemplo 2.8.

La función de distribución, en este ejemplo es

$$F_X(x) = \frac{2}{9}x^2 I_{(0,1.5)}(x) + \left(\frac{4}{3}x - \frac{2}{9}x^2 - 1\right) I_{[1.5,3)}(x) + I_{[3,\infty)}(x).$$

El r-ésimo cuantil de una variable aleatoria continua se encuentra como el valor del recorrido ζ_r que cumple con la ecuación $F_X(\zeta_r) = r$.

En este ejemplo, la función de distribución está dada en partes para valores del recorrido. Primero se sugiere evaluar la función de distribución en 1.5 (valor del recorrido donde cambia la función) para saber la probabilidad acumulada en ese valor, y saber qué parte de la función de distribución será utilizada para encontrar el cuantil solicitado.

Obsérvese que si $F_X(1.5) = 0.5$, entonces, $m = 1.5$.

El 0.35-ésimo cuantil se buscará con la primera parte de la función de distribución, esto es,

$$\frac{2}{9}(\zeta_{0.35}^2) = 0.35.$$

Las soluciones son

$$\zeta_{0.35} = \pm\sqrt{0.35\left(\frac{9}{2}\right)}.$$

La solución negativa se descarta, ya que el recorrido de la variable es el intervalo $(0, 3)$. La solución positiva sí se encuentra en el intervalo, por lo tanto,

$$\zeta_{0.35} = \sqrt{0.35\left(\frac{9}{2}\right)} = 1.25499.$$

2.12. Desigualdad de Chebyshev

En esta sección se tratará con la desigualdad de Chebyshev que es útil para acotar algunas probabilidades. Esta desigualdad tiene varias aplicaciones, por ejemplo, en temas de convergencia estocástica, también en muestreo y en estadística.

La desigualdad de Chebyshev es un caso particular de la desigualdad de Markov, la cual se presenta en el siguiente resultado.

Teorema 2.7. Desigualdad de Markov. *Sea X una variable aleatoria y $g(X)$ una función real no negativa de X, entonces,*

$$P[g(X) \geq r] \leq r^{-1}E[g(X)], \ \textit{para} \quad r > 0.$$

Demostración (caso continuo). El valor esperado de $g(X)$ se calcula como

$$\begin{aligned} E[g(X)] &= \int_{-\infty}^{\infty} g(x)f_X(x)dx \\ &= \int_{\{x|g(x)\geq r\}} g(x)f_X(x)dx + \int_{\{x|g(x)<r\}} g(x)f_X(x)dx. \end{aligned}$$

Obsérvese

$$g(x)f_X(x) \geq 0 \text{ para todo } x \in \mathbb{R}.$$

Entonces,

$$\int_{\{x|g(x)<r\}} g(x)f_X(x)dx \geq 0.$$

Por lo tanto,

$$\begin{aligned} E[g(X)] &\geq \int_{\{x|g(x)\geq r\}} g(x)f_X(x)dx \\ &\geq \int_{\{x|g(x)\geq r\}} rf_X(x)dx \\ &= r\int_{\{x|g(x)\geq r\}} f_X(x)dx \\ &= rP[g(X) \geq r]. \end{aligned}$$

De esta manera,

$$r^{-1}E[g(X)] \geq P[g(X) \geq r].$$ □

A continuación, se tratará la desigualdad de Chebyshev.

Teorema 2.8. Desigualdad de Chebyshev. *Sea X una variable aleatoria con media μ y varianza σ^2, entonces,*

$$P[|X-\mu| < k\sigma] \geq 1 - k^{-2}, \ \textit{donde} \quad k > 0,$$

o equivalentemente,

$$P[|X-\mu| \geq k\sigma] \leq k^{-2}, \ \textit{donde} \quad k > 0.$$

Demostración. Obsérvese

$$P[|X-\mu| \geq k\sigma] = P[(X-\mu)^2 \geq k^2\sigma^2].$$

Se aplicará el teorema anterior, considerando como $g(X) = (X-\mu)^2$ y $r = k^2\sigma^2$, de esta manera,

$$\begin{aligned} P[|X-\mu| \geq k\sigma] &= P[(X-\mu)^2 \geq k^2\sigma^2] \\ &\leq k^{-2}\sigma^{-2}E[(X-\mu)^2] \\ &= k^{-2}. \end{aligned}$$

□

Ejemplo 2.28. El departamento de producción de una empresa se conforma por dos personas que tienen alguna profesión y tres que no tienen profesión, pero con mucha experiencia. Se sacan al azar 3 personas para un proyecto. Para la variable aleatoria X, número de trabajadores sin profesión, verificar la desigualdad de Chebyshev para $k = 2$.

Obsérvese
$X =$ Número de trabajadores sin profesión.
$x = 1, 2, 3.$
Por lo tanto,

$$f_X(x) = \frac{\binom{3}{x}\binom{2}{3-x}}{\binom{5}{3}} I_{\{1,2,3\}}(x).$$

Se procede a calcular $E[X]$ y σ.

$$E[X] = 1\left(\frac{3}{10}\right) + 2\left(\frac{6}{10}\right) + 3\left(\frac{1}{10}\right) = \frac{9}{5},$$

$$E[X^2] = 1^2\left(\frac{3}{10}\right) + 2^2\left(\frac{6}{10}\right) + 3^2\left(\frac{1}{10}\right) = 3.6,$$

$$\sigma = 0.6 = \sqrt{3.6-(1.8)^2} = 0.6.$$

Obsérvese

$$\begin{aligned} P[|X-1.8| < 2(0.6)] &= P[-1.2 < X-1.8 < 1.2] \\ &= P[0.6 < X < 3] \\ &= P[X=1] + P[X=2] = 0.9, \end{aligned}$$

además,

$$1-k^{-2} = 1-2^{-2} = 0.75 < 0.90.$$

Por lo tanto, sí se cumple la desigualdad de Chebyshev, para $k = 2$.

El anterior ejemplo tiene como objetivo que el estudiante se dé cuenta de cómo se verifica el cumplimiento de la desigualdad de Chebyshev para un ejemplo particular. De antemano se sabía que se tenía que cumplir, ya que es un resultado que fue ya demostrado en general.

Ejemplo 2.29. Sea X una variable aleatoria con media 3 y segundo momento con respecto al origen 13, encontrar una cota inferior para la siguiente probabilidad $P[-2 < X < 8]$.

Se puede calcular la desviación estándar, esto es, $\sigma = \sqrt{13-9} = 2$.

Consideremos la desigualdad de Chebyshev,

$$P[\mu - k\sigma < X < \mu + k\sigma] \geq 1 - k^{-2}.$$

Nos podemos percatar que $\mu - k\sigma = -2 \quad \text{y} \quad \mu + k\sigma = 8$, de esta manera, se deduce que $k = 2.5$. Por lo tanto, la cota inferior para la probabilidad $P[-2 < X < 8]$, usando la parte derecha de la desigualdad, es $1 - \frac{4}{25} = \frac{21}{25}$.

2.13. Desigualdad de Jensen

La desigualdad de Jensen tiene que ver con funciones convexas de una variable aleatoria. Primero daremos la definición de una función convexa y luego se presentará el teorema, donde se expone esta desigualdad. Como se explicará, este resultado tiene aplicaciones, tanto en probabilidad como en otras áreas de aplicación.

Definición 2.22. Una función $g(X)$ con dominio y contra dominio de los números reales es *convexa*, si es una función continua y además cumple la siguiente propiedad: para todo $x_0 \in \mathbb{R}$, existe una línea recta $l(x)$ tal que $l(x_0)$ $= g(x_0)$ y además $l(x) \leq g(x)$ para todo $x \in \mathbb{R}$, donde $x \neq x_0$.

La desigualdad de Jensen es válida para funciones convexas y es de gran utilidad para algunos resultados importantes, tanto en probabilidad, en el área de inferencia estadística, como también en matemáticas actuariales.

Teorema 2.9. Desigualdad de Jensen. *Sea X una variable aleatoria y $g(X)$ una función convexa, entonces,*

$$E[g(X)] \geq g(E[X]).$$

Demostración. Supongamos que $g(X)$ es una función convexa.

Entonces, para un $x_0 \in \mathbb{R}$, existe una línea recta $l(X) = a + bX$ tal que $l(x_0) = g(x_0)$ y $l(x) \leq g(x)$ para todo $x \neq x_0$.

En particular, se puede considerar $x_0 = E[X]$, por lo que existe $l(X)$, tal que $l(E[X]) = g(E[X])$, y además $l(x) \leq g(x)$ para todo $x \neq E[X]$.

Por otro lado, $g(E[X]) = l(E[X]) = aE[X] + b = E[aX + b] = E[l(X)]$. Además, si $l(x) \leq g(x)$, entonces, $E[l(X)] \leq E[g(X)]$.

Por lo tanto,

$$g(E[X]) \leq E[g(X)].$$

□

Observación. Si la función $g(x)$ y su derivada $\frac{d}{dx}g(x)$ son diferenciables, entonces, $g(x)$ es una función convexa, si y solo si, la segunda derivada $\frac{d^2}{dx}g(x) \geq 0$.

Ejemplo 2.30. Para la variable aleatoria X.
a) Aplicar la desigualdad de Jensen para las funciones e^X, X^2.
b) Comprobar que $V[X] \geq 0$.

a) Las funciones e^X, X^2 son convexas, entonces, se cumple $E[e^X] \geq e^{E[X]}$ y $E[X^2] \geq E^2[X]$.

b) Obsérvese que de la última desigualdad del inciso anterior, se puede afirmar que $V[X] \geq 0$.

Observación. Una función $g(X)$ es *cóncava*, si es una función continua, con dominio y contra dominio los reales y además cumple la siguiente propiedad: para todo $x_0 \in \mathbb{R}$, existe una línea recta $l(x)$ tal que $l(x_0) = g(x_0)$ y $l(x) \geq g(x)$ para todo $x \in \mathbb{R}$ donde $x \neq x_0$.

Existe una versión de la desigualdad de Jensen para funciones cóncavas, esto es, si X es una variable aleatoria y $g(X)$ es una función cóncava de X, entonces, se cumple la siguiente desigualdad

$$E[g(X)] \leq g(E[X]).$$

Para comprobar la anterior desigualdad, se puede verificar fácilmente que una función $g(X)$ es cóncava si y solo si $-g(X)$ es convexa, entonces se puede aplicar el teorema anterior.

2.14. Mezcla de distribuciones

En esta sección se tratarán las mezclas, combinaciones o promedios ponderados de dos o más distribuciones. El promedio ponderado de varias distribuciones es una forma de ampliar los modelos de probabilidad, en particular, es una opción para modelos de variables aleatorias como el reclamo de un asegurado a una compañía de seguros, entre otros ejemplos.

Teorema 2.10. *Sean $f_1(x), f_2(x), \ldots, f_k(x)$ funciones de densidad, y $p_1, p_2, \ldots, p_k$ constantes, donde $p_i \geq 0$ para $i = 1, 2, \ldots, k$ y $\sum_{i=1}^{k} p_i = 1$, entonces,*

$$f_X(x) = \sum_{i=1}^{k} p_i f_i(x)$$

es una función de densidad.

La función de densidad $f_X(x)$, considerada en el anterior teorema, es conocida como la *mezcla de distribuciones* con funciones de densidad dadas por $f_1(x), f_2(x), \ldots, f_k(x)$.

Se sugiere al lector, realizar la demostración del anterior teorema.

Las funciones $f_i(x)$ pueden ser discretas o continuas, o una combinación de discretas y continuas.

Obsérvese que la suma

$$f_X(x) = \sum_{i=1}^{k} p_i f_i(x)$$

es un promedio ponderado de las funciones de densidad $f_i(x)$, $i = 1, 2, \ldots, k$, donde los p_i son considerados los pesos de este promedio.

Por ejemplo, las estaturas de una población puede considerarse como la mezcla de dos distribuciones continuas, donde una de ellas modela las estaturas de las mujeres y la otra la de los hombres.

Un ejemplo de la gráfica de una función de densidad de una mezcla de dos distribuciones continuas se da a continuación:

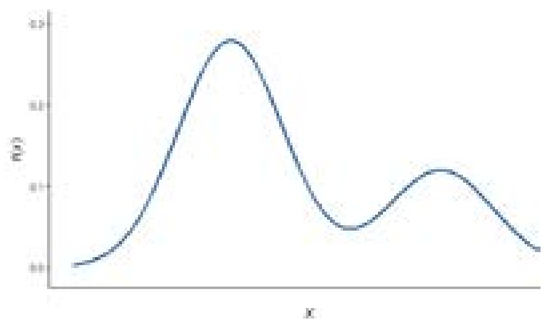


Figura 2.13. *Función de densidad, ejemplo 2.24*

Otro ejemplo, el reclamo de un asegurado de acuerdo a cómo está estipulada su póliza de seguros, puede considerarse como la mezcla de dos distribuciones, una parte discreta, donde la variable aleatoria (el reclamo del asegurado) tiene la posibilidad de tomar valores específicos con probabilidades mayor que cero, como el reclamo nulo o el reclamo de toda la suma asegurada, y en cualquier otro caso, la variable es continua, dando la posibilidad de que el asegurado reclame una cantidad mayor que cero y menor que la suma asegurada.

Teorema 2.11. *Sea X una variable aleatoria con función de densidad $f_X(x)$ la cual puede ser expresada como una mezcla de distribuciones con funciones de densidad $f_1(x), f_2(x), \ldots, f_k(x)$ y pesos p_i, $i = 1, 2, \ldots, k$, entonces,*

$$E[X^n] = \sum_{i=1}^{k} p_i E_i[X^n],$$

$$F_X(x) = \sum_{i=1}^{k} p_i F_i(x),$$

$$m_X(t) = \sum_{i=1}^{k} p_i m_i(t),$$

donde $E_i[X^n]$, $F_i(x)$ y $m_i(t)$ son, respectivamente, el n-ésimo momento con respecto al origen, la función de distribución y la función generadora de momentos, de la distribución correspondiente, aquella que tiene como función de densidad a $f_i(x)$, $i = 1, \ldots, n$.

La demostración de este teorema se deja como ejercicio para el estudiante.

Ejemplo 2.31. Una persona tiene 4 dados bien balanceados con diferentes números en las caras: el dado 1 tiene 6 caras, el dado 2 tiene 20, el dado 3 está formado con 30 números diferentes, y el dado 4 es de 100 números. La persona toma el dado 1 y lo arroja, si cae el número 1, 2 o 3, entonces, lanza el dado 2 una vez, si cae el número 4, entonces, lanza el 3, y de otra manera arroja el dado 4. Sea X el número de caras menores a 11 en el segundo dado que se lanza, para esta variable aleatoria encontrar la
a) Función de densidad.
b) Media.
c) Varianza.

a) Obsérvese que el recorrido de X es $x = 0, 1$.

La distribución de X es una mezcla de tres distribuciones discretas, donde sus funciones de densidad se dan a continuación: $f_1(x) = 0.5I_{\{0,1\}}(x)$; $f_2(x) = \frac{2}{3}I_{\{0\}} + \frac{1}{3}I_{\{1\}}(x)$ y $f_3(x) = \frac{9}{10}I_{\{0\}}(x) + \frac{1}{10}I_{\{1\}}(x)$. Las ponderaciones que se deben considerar para la función de densidad de X son respectivamente, $p_1 = 0.5$; $p_2 = \frac{1}{6}$ y $p_3 = \frac{1}{3}$.

Por lo tanto, la función de densidad de X es

$$f_X(x) =$$

$$(0.5)0.5I_{\{0,1\}}(x) + \frac{1}{6}\left[\frac{2}{3}I_{\{0\}(x)} + \frac{1}{3}I_{\{1\}}(x)\right] + \frac{1}{3}\left[\frac{9}{10}I_{\{0\}}(x) + \frac{1}{10}I_{\{1\}}(x)\right].$$

Escrita de otra manera,

$$f_X(x) = 0.66111I_{\{0\}}(x) + 0.33889I_{\{1\}}(x).$$

b) Las medias de las 3 distribuciones son, respectivamente: $E_1[X] = 0.5$; $E_2[X] = \frac{1}{3}$ y $E_3[X] = \frac{1}{10}$.

Por lo tanto, la media de X, considerando las medias anteriores, se calcula de la siguiente manera:

$$E[X] = 0.5(0.5) + \left(\frac{1}{6}\right)\left(\frac{1}{3}\right) + \left(\frac{1}{3}\right)\left(\frac{1}{10}\right) = 0.33889.$$

c) Obsérvese que los segundos momentos de las 3 distribuciones son iguales que los primeros, esto es, $E_1[X^2] = 0.5$; $E_2[X^2] = \frac{1}{3}$ y $E_3[X^2] = \frac{1}{10}$.

De esta manera, el segundo momento de X es también $E[X^2] = 0.33889$, por lo tanto, $V[X] = 0.33889 - (0.33889)^2 = 0.22404$.

Es importante aclarar que la varianza de X no se puede obtener como el promedio ponderado de las tres varianzas $V_1[X]$, $V_2[X]$ y $V_3[X]$.

Ejemplo 2.32. Los daños de una máquina asegurada reportados en un mes, es una variable aleatoria X que cumple con las siguientes características: $P[X \leq x] = 1 - \frac{2}{5}e^{-\frac{x}{5}}$ para $x \geq 0$ y $P[X \leq x] = 0$ para $x < 0$. Encontrar para la variable X, la

a) Función de densidad escrita como mezcla de distribuciones.
b) Media mensual.
c) Función generadora de momentos.
d) Varianza mensual.

a) Se puede observar que la función de distribución de la variable X es continua para todo valor de x, excepto en $x = 0$, donde se observa un salto. Lo que significa que la distribución de X es una mezcla de dos distribuciones, una distribución discreta con función de densidad $f_1(x)$ en $x = 0$ y otra continua con función de densidad $f_2(x)$ para valores de $x > 0$.

Obsérvese que la probabilidad de tener daños reportados de 0 en un mes es igual a $\frac{3}{5}$, y como la parte discreta de esta distribución se concentra en solo un punto, esto es, en $x = 0$ (solo hay una discontinuidad), entonces, la función de densidad discreta es $f_1(x) = I_{\{0\}}(x)$, de aquí se deduce que el peso de esta función de densidad es igual a $p_1 = \frac{3}{5}$. Por lo tanto, el peso de función de densidad continua p_2 se obtiene del complemento de p_1, esto es, $p_2 = \frac{2}{5}$.

Ahora, la función de densidad $f_2(x)$ se obtiene derivando la función de distribución para valores de $x > 0$, y después el resultado se ajusta con la constante adecuada, para que sea función de densidad.

De este modo, la derivada de la función de distribución es $\frac{2}{25}e^{-\frac{x}{5}}$. Obsérvese, $\int_0^\infty \frac{2}{25}e^{-\frac{x}{5}} = P[X > 0] = \frac{2}{5}$, siendo otra forma de calcular p_2.

El recíproco de este último resultado es la constante para ajustar la derivada de la función de distribución y de esta manera obtener la función de densidad $f_2(x)$.

Esto es, $f_2(x) = \frac{1}{\frac{2}{5}}\left(\frac{2}{25}\right)e^{-\frac{x}{5}}I_{(0,\infty)} = \frac{1}{5}e^{-\frac{x}{5}}I_{(0,\infty)}$.

Por lo tanto, la función de densidad $f_X(x)$ expresada como mezcla de dos distribuciones queda de la siguiente manera:

$$f_X(x) = \left(\frac{3}{5}\right)I_{\{0\}}(x) + \left(\frac{2}{5}\right)\frac{1}{5}e^{-\frac{x}{5}}I_{(0,\infty)}(x).$$

b) Ahora obsérvese que $E_1[X] = 0$ y $E_2[X] = 5$.

Por lo tanto,

$$\begin{aligned} E[X] &= p_1E_1[X] + p_2E_2[X] \\ &= \left(\frac{3}{5}\right)(0) + \left(\frac{2}{3}\right)(5) = 3.33333. \end{aligned}$$

c) Además, $m_1(t) = e^{0t} = 1$ y $m_2(t) = (1 - 5t)^{-1}$, si $t < 1/5$.
Entonces,

$$\begin{aligned} m_X(t) &= p_1 m_1(t) + p_2 m_2(t) \\ &= \tfrac{3}{5} + \tfrac{2}{5}(1 - 5t)^{-1}, \text{si } t < 1/5. \end{aligned}$$

El inciso d se deja como ejercicio para el estudiante.

2.15. Distribuciones de cola pesada

En la vida real es muy común que se trate con variables aleatorias continuas con recorrido positivo, por ejemplo, el tiempo necesario para observar un acontecimiento, las ganancias (o pérdidas) de una compañía, el tamaño de reclamos mensual a una compañía de seguros. Para este tipo de variables, las distribuciones de cola pesada pueden ser una opción de mucha utilidad.

En particular, las distribuciones de cola pesada son muy útiles tanto en problemas de seguros, como de finanzas o también de economía, por mencionar algunas áreas de aplicación. Estos modelos pueden ser usados, en general, para variables aleatorias con probabilidad no «tan pequeña», donde la variable tome valores grandes o extremos, se puede consultar [24].

En seguros, es común que se cubran siniestros que, de momento, puedan ser catastróficos, como un huracán, un terremoto, una nevada o un incendio extremo, y en estos casos es necesario contar con distribuciones de cola pesada, esto es, modelos donde la probabilidad de tener un reclamo muy grande o extremo no sea «tan pequeña», o dicho de otra manera, la probabilidad de que no sea prácticamente nula. El tener un modelo adecuado para el tamaño de reclamos a una compañía de seguros contribuirá a lograr una estimación más realista de la probabilidad de ruina, por reclamos excesivos.

Pensemos ahora en los créditos que un banco ofrece a sus clientes, por lo general siempre existe lo que se le llama cartera vencida, la cantidad que no es pagada cada mes por algunos de sus clientes. Para un mes determinado, la cantidad no pagada al banco por sus clientes puede tener un valor extremo, debido posiblemente a alguna contingencia, por ejemplo, una catástrofe natural que afecta a todo el país, o también, alguna devaluación, etc. Una distribución de cola pesada sería un modelo adecuado para este tipo de circunstancias.

Los ejemplos previamente expuestos son casos de variables aleatorias con recorrido $(0, \infty)$, que, cuando la variable toma valores extremos, serían valores grandes positivos, por esta razón solo nos interesa la definición de distribución de cola pesada considerando solo la cola derecha. Cabe mencionar que también se puede definir este tipo de distribuciones, considerando la cola izquierda, o también tomando en cuenta las dos colas de la distribución.

Se ha tratado de definir o explicar de varias maneras lo que es una distribución de cola pesada, a continuación, se expondrán algunas explicaciones o definiciones que algunos autores han expuesto. Estas, principalmente, se pueden encontrar en Asmussen [1].

Una primera manera de definir una distribución de cola pesada es aquella que va tendiendo por la derecha a cero, más lentamente que una distribución exponencial, véase la definición 4.8 de distribución exponencial. De esta manera, una variable aleatoria con distribución de cola pesada podrá tomar con mayor facilidad un valor extremo.

Una segunda definición es que una distribución se considera de cola pesada si su varianza es infinita.

Otra manera es explicada en términos actuariales; una distribución es de cola pesada, si el 20 % del conteo de los reclamos son al menos el 80 % del total reclamado. Esto es, consideremos a la variable X como el total reclamado a una compañía de seguros en un periodo dado, sea $f_X(x)$ su función de densidad, μ su media, y $\phi_{0.80}$ es el 0.80-cuantil. Se dice que la distribución de X es de cola pesada si cumple:

$$\frac{1}{\mu}\int_{\phi_{0.8}}^{\infty} x f_X(x)dx \geq 0.80.$$

A continuación, se dará la definición más común de distribución de cola pesada encontrada en la literatura, también se presentará la de distribución de cola ligera. Estas definiciones son las que consideraremos, de aquí en adelante, para indicar si una distribución particular con recorrido $(0, \infty)$ es de cola ligera o pesada.

Definición 2.23. Sea X una variable aleatoria con función de distribución $F_X(x)$ y función generadora de momentos $m_X(t)$. Se dice que X tiene *distribución de cola pesada* si

$$m_X(t) = \infty, \quad \text{para todo} \quad t > 0.$$

Por otro lado, se dice que X tiene *distribución de cola ligera* si

$$m_X(t) < \infty, \quad \text{para algún} \quad t > 0.$$

Cabe señalar que al considerar distribuciones con recorrido $(0, \infty)$, la función generadora de momentos siempre va a existir para $t \leq 0$, véase sección 2.10 de este capítulo.

En el caso de una distribución de cola ligera, la función generadora de momentos va existir para valores de t, en un entorno que contenga al cero. De esta manera, para estas distribuciones también van a existir todos los momentos con respecto al origen, véase el desarrollo en serie de Maclaurin de la función generadora de momentos (ecuación 2.1).

Pero para una distribución de cola pesada, la función generadora de momentos no va existir para valores de t en un intervalo $(-h, h)$, donde $h > 0$. No obstante, pueden existir todos los momentos con respecto al origen, pero la suma de la parte derecha de la expresión 2.1 puede no converger, tal es el caso de la distribución lognormal, véase definición y propiedades de esta distribución en el capítulo 4. Lo que si podemos afirmar es que si algún momento $E[X^r]$ no existe, entonces, la función generadora tampoco, lo cual será suficiente para afirmar que se trata de una distribución de cola pesada.

Un primer ejemplo de distribución de cola ligera es la distribución del ejemplo 2.16, en este caso la función generadora de momentos existe para alguna $t > 0$.

A continuación, se dará un ejemplo de una distribución de cola pesada.

Ejemplo 2.33. Sea X una variable aleatoria con función de densidad $f_X(x) = \frac{1}{(1+x)^2} I_{(0,\infty)}(x)$. Para esta función

a) Verificar que sea función de densidad.

b) Verificar que sea de cola pesada.

c) Realizar la gráfica de esta función junto con la función de densidad del ejemplo 2.16.

a) Obsérvese

$$(1+x)^{-2} > 0 \quad \text{e} \quad I_{(0,\infty)}(x) \geq 0.$$

De esta manera,

$$f_X(x) \geq 0$$

.

Por otro lado,

$$\int_0^\infty (1+x)^{-2} dx = -(1+x)^{-1}\Big|_0^\infty = 1.$$

Por lo tanto, $f_X(x)$ es función de densidad.

b) El valor esperado se calcula de la siguiente manera:

$$E[X] = \int_0^\infty x(1+x)^{-2} dx,$$

donde se considera el siguiente cambio de variable $u = 1 + x$, de esta manera $du = dx$.

De este modo,

$$\begin{aligned}
\int_0^\infty x(1+x)^{-2} dx &= \int_1^\infty (u-1)u^{-2} du \\
&= \int_1^\infty u^{-1} du - \int_1^\infty u^{-2} du \\
&= [ln(u)]_{t=1}^\infty - \left[u^{-1}\right]_1^\infty \\
&= \lim_{u\to\infty} ln(u) - 1 = \infty.
\end{aligned}$$

Por lo tanto, al no existir el valor esperado de X (función generadora tampoco existe), $f_X(x)$ es la función de densidad de una distribución de cola pesada.

c) A continuación, se presentan las gráficas de las funciones de densidad de los ejemplos 2.16 (cola ligera) y 2.33 (cola pesada).

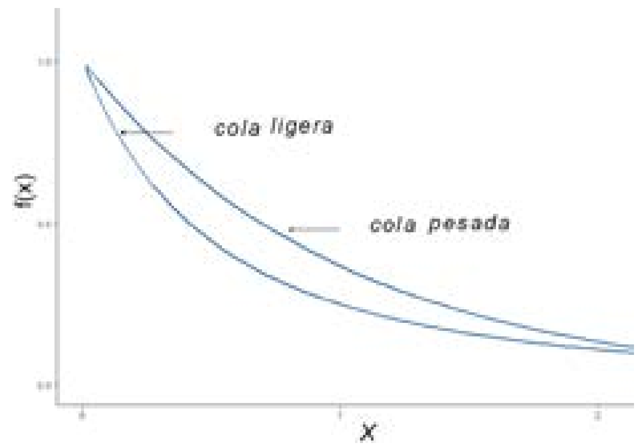


Figura 2.14. *Funciones de densidad de los ejemplos 2.16 y 2.33.*

2.16. Distribuciones truncadas

En los siguientes dos capítulos se tratarán las familias de distribuciones, tanto discretas como continuas, más importantes que existen en la literatura. Estas familias toman mucha importancia, sobre todo, por las aplicaciones en el mundo real.

Una manera de ampliar los modelos es modificándolos, existen varias formas de hacerlo. En esta sección se presentará un procedimiento llamado *truncación.* Para entender en qué consiste este procedimiento, se presentarán, a continuación, algunos ejemplos.

Ejemplo 2.34. El número de personas que llegan a una conferencia que se realiza cada año es una variable aleatoria, la cual tiene función de densidad dada por $f_X(x) = \frac{10^x}{x!}e^{-10}I_{\{0,1,2,\ldots\}}(x)$. Se sabe que al menos 2 personas asistirán a la conferencia. Considerando este hecho, encontrar la nueva función de densidad.

El nuevo modelo discreto calculará probabilidades en cada valor del recorrido con la condición de que $X > 1$. Sea $g_X(x)$ la nueva función de densidad. Esta función se obtendrá de la siguiente manera:

$$g_X(x) = P[X = x|X > 1],$$

donde el nuevo recorrido es $\{0, 1, \ldots\} \cap \{x|x > 1\} = \{2, \ldots\}$.

De esta manera,

$$g_X(x) = \frac{P[(X = x) \cap (X > 1)]}{P[X > 1]} = \frac{10^x e^{-10}}{x!P[X > 1]}.$$

Obsérvese

$$P[X > 1] = 1 - P[X = 0] - P[X = 1] = 1 - e^{-10} - 10e^{-10}.$$

Por lo tanto,

$$g_X(x) = \frac{10^x e^{-10}}{x!(1 - e^{-10} - 10e^{-10})} I_{\{2,\ldots\}}(x).$$

Esta última es la función de densidad para la variable aleatoria X considerando la restricción $X > 1$. Se dice que esta nueva distribución para X está *truncada* en el conjunto $\{0, 1\}$.

En este ejemplo se truncó la función de densidad en $\{0, 1\}$, pero también se puede truncar en otro conjunto, por ejemplo, se puede truncar en $x > 7$, o en $x < 7$, o en $3 < X < 8$, etc. Esto depende del problema donde se quiera aplicar el nuevo modelo.

A continuación, se presenta un ejemplo de distribución continua truncada en algún conjunto.

Ejemplo 2.35. Considerando la variable aleatoria del ejemplo 2.16, tiempo en horas para que un ajustador de seguros llegue al lugar de los hechos por el reporte de un accidente automovilístico, se sabe que este traslado por parte del ajustador no puede ser menor o igual a 30 minutos. Modificar la función de densidad para esta variable, considerando esta restricción.

La nueva función de densidad servirá para calcular probabilidades del tiempo de traslado por el ajustado al lugar de los hechos, con la condición de que $X > 0.5$. Sea $g_X(x)$ la nueva función de densidad, esta se obtendrá de la siguiente manera:

$$g_X(x) = \frac{f_x(x)}{P[X > 0.5]},$$

donde el nuevo recorrido es $(0, \infty) \cap (0.5, \infty)$.

Obsérvese

$$P[X > 0.5] = \int_{0.5}^{\infty} e^{-x} dx = e^{-0.5}.$$

Por lo tanto,

$$g_X(x) = e^{-(x-0.5)} I_{(0.5,\infty)}(x).$$

Se dice que la nueva distribución está *truncada* en el intervalo $(0, 0.5]$ o que está *truncada* a la izquierda de 0.5, pero también se puede truncar en otros intervalos o conjuntos, por ejemplo, en $(0.5, \infty)$ o en $(0.5, 1)$, o en $(0, 0.5) \cup (1, \infty)$, o en algún otro conjunto, dependiendo de la naturaleza de la variable.

2.17. Ejemplos diversos

En esta sección se tratarán diferentes ejemplos vinculados con los temas vistos durante este capítulo.

Ejemplo 2.36. Un seguro de vida por un año tiene las siguientes características: se pagará 200 en caso de que el asegurado tenga una muerte natural, en caso de muerte accidental se pagará el doble. La probabilidad de sufrir una muerte accidental es de 0.001 y de morir en forma natural es de 0.01. Si la compañía de seguros pretende tener una ganancia promedio de 2 por cada seguro que venda, ¿en cuánto deberá vender cada seguro de vida?

X = Pago que la compañía realizará al beneficiario del asegurado al cabo de un año.
$x = 0, 200, 400.$
Se encontrará la función de densidad para cada valor de la variable,

$$\begin{aligned} f_X(0) &= P[X = 0] \\ &= 1 - 0.01 - 0.001 = 0.989, \end{aligned}$$

$$f_X(200) = P[X = 200] = 0.01,$$

$$f_X(400) = P[X = 400] = 0.001.$$

Se calculará el valor esperado de lo que la compañía pagará por seguro:

$$E[X] = 0(0.989) + 200(0.01) + 400(0.001) = 2.4.$$

Si en promedio quiere ganar 2 por cada seguro, entonces, tendrá que cobrar por cada seguro $2.4 + 2 = 4.4$.

Ejemplo 2.37. Sea X una variable aleatoria con función de densidad $f_X(x) = \frac{6}{\pi^2 x^2} I_{\{1,2,3,\ldots\}}(x)$, ¿existe la función generadora de momentos para $t > 0$? La distribución, ¿de qué tipo de cola es?

Obsérvese

$$\begin{aligned} E[e^{tX}] &= \sum_{x=1}^{\infty} e^{tx} \frac{6}{\pi^2 x^2} \\ &= \left(\frac{6}{\pi^2}\right) \sum_{x=1}^{\infty} x^{-2} e^{tx}. \end{aligned}$$

Analizaremos la convergencia de la sucesión de la serie anterior,

$$\begin{aligned}\lim_{x\to\infty} x^{-2}e^{tx} &= \lim_{x\to\infty}\frac{te^{tx}}{2x}\\ &= \lim_{x\to\infty}\frac{t^2e^{tx}}{2}\\ &= \frac{t^2}{2}\lim_{x\to\infty} e^{tx}.\end{aligned}$$

El último límite no existe para valores de $t > 0$. Lo que significa, de acuerdo a la definición 2.23, que esta distribución es de cola pesada.

Ejemplo 2.38. Sea X una variable aleatoria discreta con función generadora de momentos $m_X(t) = \frac{1}{6}e^t + \frac{2}{6}e^{2t} + \frac{1}{2}e^{3t}$. Encontrar la función de densidad de X.

La función generadora de momentos puede expresarse como

$$m_X(t) = \sum_{x=1}^{3} e^{tx}\left(\frac{x}{6}\right).$$

Por otro lado, se sabe

$$m_X(t) = \sum_{x} e^{tx} f_X(x).$$

Por lo tanto,

$$f_X(x) = \left(\frac{x}{6}\right) I_{\{1,2,3\}}(x).$$

Ejemplo 2.39. Sea X una variable aleatoria con función de densidad $f_X(x) = \left(\frac{1}{2}\right) I_{(1,3)}(x)$.
a) Encontrar la función generadora de momentos de X.
b) Encontrar la media y la varianza de X a partir de a).
c) Verificar si se cumple el teorema de Chebyshev para $k = 2$

a) Obsérvese

$$\begin{aligned}E[e^{tX}] &= \int_1^3 \left(\frac{1}{2}\right) e^{tx}dx\\ &= \left.\frac{e^{tx}}{2t}\right|_1^3, \quad \text{si} \quad t \neq 0.\end{aligned}$$

Entonces, para $t \neq 0$,

$$E[e^{tX}] = \frac{1}{2}\left(t^{-1}(e^{3t} - e^t)\right).$$

La función $E[e^{tX}]$ es discontinua en $t = 0$.

Se obtendrá el límite de $E[e^{tX}]$ cuando t tiende a 0.

$$\begin{aligned}\lim_{t\to 0} E[e^{tX}] &= \frac{1}{2}\lim_{t\to 0}\frac{e^{3t}-e^{t}}{t}\\ &= \frac{1}{2}\lim_{t\to 0}\frac{3e^{3t}-e^{t}}{1} = 1.\end{aligned}$$

Entonces, la función generadora de momentos de X es

$$m_X(t) = \frac{1}{2}\left(t^{-1}(e^{3t}-e^{t})\right) I_{\{t\neq 0\}}(t) + I_{\{0\}}(t).$$

Obsérvese que esta última función es continua para todo $t \in \mathbb{R}$.
b) La primera derivada de la función generadora de momentos es

$$\frac{d}{dt}m_X(t) = \frac{1}{2}\left(\frac{t(3e^{3t}-e^{t})-e^{3t}+e^{t}}{t^2}\right).$$

Observemos que la primera derivada no existe en $t=0$. Por lo que $E[X]$ se obtendrá sacando el límite de la primera derivada cuando t tiende a 0.

$$\begin{aligned}E[X] &= \lim_{t\to 0}\frac{d}{dt}m_X(t)\\ &= \lim_{t\to 0}\frac{1}{2}\left(\frac{t(3e^{3t}-e^{t})-e^{3t}+e^{t}}{t^2}\right)\\ &= \lim_{t\to 0}\frac{1}{2}\left(\frac{t(9e^{3t}-e^{t})+3e^{3t}-e^{-t}-3e^{3t}+e^{t}}{2t}\right)\\ &= \lim_{t\to 0}\frac{1}{2}\left(\frac{t(9e^{3t}-e^{t})}{2t}\right)\\ &= \lim_{t\to 0}\frac{1}{2}\left(\frac{t(27e^{3t}-e^{t})+9e^{3t}-e^{t}}{2}\right) = 2.\end{aligned}$$

Para el segundo momento, $E[X^2]$, se obtiene el límite de la segunda derivada de la función generadora de momento cuando t tiende a cero. El segundo momento es $E[X^2] = \frac{13}{3}$. Se le pide al estudiante confirmar este cálculo.

Por lo tanto,

$$\sigma = \sqrt{\frac{13}{3}-4} = \sqrt{\frac{1}{3}}.$$

c) Entonces,

$$P\left[|X-2| < 2\sqrt{\frac{1}{3}}\right] = P[0.84 < X < 3.15] = 1 \geq 0.75.$$

De este modo, se cumple la desigualdad de Chebyshev.

Ejemplo 2.40. Sea X una variable aleatoria con función generadora de momentos $m_X(t) = e^{\frac{t^2}{2}}$. Encontrar el r-ésimo momento $E[X^r]$.

Se desarrollará la función generadora de momentos en series de Maclaurin (véase apéndice C).

$$\begin{aligned} m_X(t) &= e^{\frac{t^2}{2}} \\ &= 1 + \frac{t^2}{2} + \frac{t^4}{2^2 2!} + \frac{t^6}{2^3 3!} + \cdots + \frac{t^{2k}}{2^k k!} + \cdots \end{aligned}$$

Ahora, el desarrollo en general de la función generadora de momentos en series de Maclaurin es (véase ecuación 2.1),

$$m_X(t) = 1 + tE[X] + \frac{t^2}{2!}E[X^2] + \frac{t^3}{3!}E[X^3] + \frac{t^4}{4!}E[X^4] + \cdots$$

La función generadora de este ejemplo se puede expresar también de la siguiente manera:

$$\begin{aligned} m_X(t) = 1 + (0)t + \frac{t^2}{2} + (0)\frac{t^3}{3!} + \left(\frac{4!}{2^2 2!}\right)\left(\frac{t^4}{4!}\right) + (0)\frac{t^5}{5!} \\ + \left(\frac{6!}{2^3 3!}\right)\left(\frac{t^6}{6!}\right) + (0)\frac{t^7}{7!} + \cdots + \left(\frac{(2k)!}{2^k k!}\right)\left(\frac{t^{2k}}{(2k)!}\right) + \cdots \end{aligned}$$

De este modo, se concluye

$$E[X^r] = \begin{cases} 0, & \text{cuando} \quad r \text{ es impar}, \\ \dfrac{r!}{2^{r/2}(r/2)!}, & \text{cuando} \quad r \text{ es par}. \end{cases}$$

Ejemplo 2.41. Sea X una variable aleatoria con función de densidad dada por $f_X(x) = c|x-4|I_{[2,6]}(x)$. Encontrar

a) La constante c.

b) La función de distribución de X.

c) El valor esperado de X.

d) La mediana.

e) La moda.

f) El 0.8-ésimo cuantil.

a) $f_X(x)$ se puede expresar como

$$f_X(x) = c[(4-x)I_{[2,4]}(x) + (x-4)I_{(4,6]}(x)].$$

Por lo tanto,

$$c\left[\int_2^4 (4-x)dx + \int_4^6 (x-4)dx\right] = 1.$$

De esta manera,

$$c\left[\left[4x-\frac{x^2}{2}\right]_2^4+\left[\frac{x^2}{2}-4x\right]_4^6\right]=1.$$

Evaluando y despejando, $c=\frac{1}{4}$.
Entonces, la función de densidad es

$$f_X(x)=\frac{1}{4}\left[(4-x)I_{[2,4]}(x)+(x-4)I_{(4,6]}(x)\right].$$

b) Se calculará la función de distribución por casos
Caso 1. Cuando $x<2$.

$$F_X(x)=0.$$

Caso 2. Cuando $2\leq X\leq 4$.

$$\begin{aligned}F_X(x)&=\int_2^x\frac{1}{4}(4-t)dt\\&=\frac{1}{4}\left[4t-\frac{t^2}{2}\right]_2^x\\&=x-\frac{x^2}{8}-\frac{3}{2}.\end{aligned}$$

Caso 3. Cuando $4<X\leq 6$.

$$\begin{aligned}F_X(x)&=\int_2^x f_X(t)dt\\&=F_X(4)+\frac{1}{4}\int_4^x(t-4)dt\end{aligned}$$

$$\begin{aligned}&=\frac{1}{2}+\frac{1}{4}\left[\frac{t^2}{2}-4t\right]_4^x\\&=\frac{1}{2}+\frac{1}{4}\left[\frac{x^2}{2}-4x-8+16\right]\\&=\frac{x^2}{8}-x+\frac{5}{2}.\end{aligned}$$

Caso 4. Cuando $x>6$.

$$F_X(x)=1.$$

Se concluye que

$$F_X(x)=\left(x-\frac{x^2}{8}-\frac{3}{2}\right)I_{(2,4]}(x)+\left(\frac{x^2}{8}-x+\frac{5}{2}\right)I_{(4,6]}(x)+I_{(6,\infty)}(x).$$

c) Ahora, el valor esperado se calcula como

$$\begin{aligned}E[X] &= \int_2^4 x\left(\frac{1}{4}(4-x)\right)dx + \int_4^6 x\left(\frac{1}{4}(x-4)\right)dx\\ &= \left[\frac{x^2}{2}-\frac{x^3}{12}\right]_2^4 + \left[\frac{x^3}{12}-\frac{x^2}{2}\right]_4^6 = 4.\end{aligned}$$

d) Obsérvese que $F_X(4) = 0.5$, entonces, la mediana es $m = 4$.

e) La función de densidad está compuesta por dos rectas, la primera con pendiente negativa, $1-0.25x$, en el intervalo $[2,4]$, y con un valor máximo de 0.5, en $x = 2$, y la segunda $0.25x-1$, con pendiente positiva, en el intervalo $(4,6]$, y con un valor máximo de 0.5, en $x = 6$. De esta manera, se concluye que las modas son 2 y 6.

f) El 0.8-ésimo cuantil es el valor del recorrido $\zeta_{0.8}$ que cumple con

$$F_X(\zeta_{0.8}) = 0.8.$$

De esta manera, usando la segunda parte de la función de distribución

$$\frac{\zeta_{0.8}^2}{8} - \zeta_{0.8} + \frac{5}{2} = 0.8.$$

Esto es,

$$\begin{aligned}\zeta_{0.8}^2 - 8\zeta_{0.8} + 13.6 &= 0,\\ \zeta_{0.8} &= \frac{8 \pm \sqrt{64-4(13.6)}}{2},\\ \zeta_{0.8} = 5.54919, \quad &\text{o} \quad \zeta_{0.8} = 2.45081.\end{aligned}$$

De estas dos soluciones, descartamos aquella que no se encuentra en el recorrido X, por lo tanto, $\zeta_{0.8} = 5.54919$.

Ejemplo 2.42. La cantidad reclamada mensualmente por los asegurados a una compañía por un tipo de seguro, tiene la siguiente función de distribución:

$$F_X(x) = \begin{cases} 0 & \text{si} \quad x < 0,\\ 0.2x & \text{si} \quad 0 \le x < 2.5,\\ 0.05(x+13.5) & \text{si} \quad 2.5 \le x < 6.5,\\ 1 & \text{si} \quad x \ge 6.5.\end{cases}$$

Encontrar
a) $P[2.5 < X \le 3.5]$.
b) $P[2.5 \le X \le 3.5]$.
c) La función de densidad.

d) El valor esperado de X.
e) La varianza.
f) El 0.9-ésimo cuantil.

Se puede observar que la función de distribución es continua, excepto en el punto $x = 2.5$, lo que significa que esta distribución es una mezcla de dos distribuciones, de una continua y una discreta. La variable cantidad reclamada es discreta en el punto $x = 2.5$ y continua en los intervalos $[0, 2.5)$ y $(2.5, 6.5]$.

a) $P[2.5 < X \leq 3.5] = F_X(3.5) - F_X(2.5) = 0.85 - 0.80 = 0.05$.

b) $P[2.5 \leq X \leq 3.5] = F_X(3.5) - F_X(2.5^-) = 0.85 - 0.5 = 0.35$.

c) El salto en $x = 2.5$ es de 0.3, lo que significa que $P[X = 2.5] = 0.3$. Dicho de otra manera, la función de densidad de la distribución discreta es $f_1(x) = I_{\{2.5\}}(x)$ con un peso $p_1 = 0.3$. Entonces, el peso de la función de densidad continua es $p_2 = 0.7$.

Derivando la función de distribución en los intervalos $[0, 2.5)$ y $(2.5, 6.5]$, queda respectivamente 0.2 y 0.05.

Obsérvese que el área total de la parte continua es precisamente 0.7. Por lo tanto, la función de densidad para la parte continua es

$$\begin{aligned} f_2(x) &= \frac{0.2}{0.7} I_{[0,2.5)}(x) + \frac{0.05}{0.7} I_{(2.5,6.5]}(x) \\ &= \frac{2}{7} I_{[0,2.5)}(x) + \frac{1}{14} I_{(2.5,6.5]}(x). \end{aligned}$$

De esta manera, la función de densidad de la cantidad reclamada mensualmente escrita en forma de mezcla de distribuciones es

$$f_X(x) = 0.3 I_{\{2.5\}}(x) + 0.7 \left\{ \frac{2}{7} I_{[0,2.5)}(x) + \frac{1}{14} I_{(2.5,6.5]}(x) \right\}.$$

d) Los valores esperados de $f_1(x)$ y $f_2(x)$, respectivamente, son

$$\begin{aligned} E_1[X] &= 2.5 \times 1 = 2.5, \\ E_2[X] &= \int_0^{2.5} x(\frac{2}{7})dx + \int_{2.5}^{6.5} x(\frac{1}{14})dx \\ &= \frac{2}{7} \left[\frac{x^2}{2} \right]_0^{2.5} + \frac{1}{14} \left[\frac{x^2}{2} \right]_{2.5}^{6.5} = 2.17857. \end{aligned}$$

De esta manera, el valor esperado de la cantidad reclamada es

$$\begin{aligned} E[X] &= 0.3 E_1[X] + 0.7 E_2[X] \\ &= 0.3(2.5) + 0.7(2.17857) = 2.275. \end{aligned}$$

e) Los segundos momentos de $f_1(x)$ y $f_2(x)$ son, respectivamente,

$$E_1[X^2] = (2.5)^2 \times 1 = 6.25,$$

$$
\begin{aligned}
E_2[X^2] &= \int_0^{2.5} x^2(\frac{2}{7})dx + \int_{2.5}^{6.5} x^2(\frac{1}{14})dx \\
&= \frac{2}{7}\left[\frac{x^3}{3}\right]_0^{2.5} + \frac{1}{14}\left[\frac{x^3}{3}\right]_{2.5}^{6.5} = 7.65476.
\end{aligned}
$$

Por lo tanto, el segundo momento de la cantidad reclamada es

$$
\begin{aligned}
E[X^2] &= 0.3E_1[X^2] + 0.7E_2[X^2] \\
&= 0.3(6.25) + 0.7(7.65476) = 7.23333.
\end{aligned}
$$

De esta manera, la varianza de la cantidad reclamada es

$$V[X] = 7.23333 - (2.275)^2 = 2.05771.$$

f) Para obtener el cuantil $\zeta_{0.9}$, se usará la tercera parte de la función de distribución (véase $F_X(x)$), esto se debe a que $F_X(2.5) = 0.8$. Entonces,

$$F_X(\zeta_{0.9}) = 0.9.$$

De este modo,

$$0.05(\zeta_{0.9} + 13.5) = 0.9.$$

De esta última igualdad, se obtiene $\zeta_{0.9} = 4.5$.

Ejercicios del capítulo 2

1.- El departamento de finanzas de una empresa motiva a sus trabajadores a invertir su dinero en algunos instrumentos de inversión. Se sabe que el 45 % de los empleados tienen dinero invertido en el instrumento A y el 60 % en el instrumento B, también se sabe que el 10 % de los trabajadores no tienen dinero invertido en algún instrumento. Se selecciona aleatoriamente a un empleado de esta empresa, encontrar la función de densidad para el número de instrumentos de inversión del empleado.

2.- En una caja hay 5 artículos, de los cuales 3 son de buena calidad. Una persona saca artículos al azar de la caja, uno por uno hasta tener fuera los 3 artículos buenos. Encontrar la función de densidad para el número de artículos sacados hasta lograr el tercer artículo bueno.

3.- En una caja donde hay 11 artículos, 6 son defectuosos. Se seleccionan al azar 7 artículos de la caja. Encontrar la función de densidad para el número de artículos buenos en la muestra.

4.- Seis pelotas numeradas del 1 al 6 se encuentran en una urna. Se sacan al azar dos pelotas de la urna. Encontrar la función de densidad para las siguientes variables aleatorias:

a) El mayor de los dos números seleccionados.

b) La suma de los dos números seleccionados.

5.- Un trabajador logra realizar su trabajo exitosamente en un día dado, con probabilidad del 40 %. El trabajo realizado por el trabajador en un día es independiente del trabajo realizado en cualquier otro día. Encontrar la función de densidad para el número de días entre 2 días exitosos.

6.- Encontrar la función de distribución para valores del recorrido de la variable aleatoria del ejercicio 5.

7.- Un juego consiste en lanzar 2 dados simultáneamente, si resultan ambos dados impares, entonces, se gana. Una persona juega 20 veces. Calcular la función de densidad para el número de veces que la persona gana.

8.- Sacar la media y la varianza del ejercicio 3.

9.- Un juego consiste en sacar 5 cartas de una baraja ordinaria de 52 cartas. Si un jugador obtiene exactamente un par, entonces, gana 5, si obtiene dos pares (distintos) gana 10 y si saca una tercia (no par) gana 30, y en cualquier otro caso paga 15. Encontrar la ganancia esperada del jugador.

10.- Demostrar el teorema 2.5.

11.- Para la variable aleatoria del ejercicio 3 verificar que se cumple la desigualdad de Chebyshev para $k = 2$.

12.- Verificar si las siguientes funciones son de densidad,

a) $f_X(x) = 5e^{-5x} I_{(0,\infty)}(x)$.
b) $f_X(x) = \theta e^{-\theta x} I_{(0,\infty)}(x)$.
c) $f_X(x) = \frac{a}{x^2} I_{(0,\infty)}(x)$, donde $a > 0$.
d) $f_X(x) = e^{-(x-4)} I_{(4,\infty)}(x)$.
e) $f_X(x) = \frac{5^x}{x!} e^{-5} I_{\{0,1,2,\ldots\}}(x)$.
f) $f_X(x) = x I_{(0,1)}(x) + (2 - x) I_{[1,2)}(x)$.
g) $f_X(x) = (\frac{1}{3})^x I_{\{0,1,2,\ldots\}}(x)$.
h) $f_X(x) = x I_{(0,1)}(x) + (2-x) I_{[1,3)}(x) + (x-4) I_{[3,5)}(x) + (6-x) I_{[5,6)}(x)$.
i) $f_X(x) = \frac{6}{\pi^2 x^2} I_{\{1,2,3,\ldots\}}(x)$.

13.- Sea X una variable aleatoria con función de densidad dada por $f_X(x) = \frac{c}{x^2} I_{(1,\infty)}(x)$. Encontrar c.

14.- Supongamos que $f_1(x)$ y $f_2(x)$ son funciones de densidad continuas (discretas), definimos la función $f_X(x) = \lambda_1 f_1(x) + \lambda_2 f_2(x)$, donde λ_1 y λ_2 son constantes. Para los siguientes casos, justificando bien la respuesta, ¿$f_X(x)$ es función de densidad?

a) Si $\lambda_1 + \lambda_2 = 1$.
b) Si $\lambda_1 + \lambda_2 = 1$ y $\lambda_i \geq 0, i = 1, 2$.

15.- Encontrar la constante c para que $f_X(x)$ sea función de densidad, donde $f_X(x) = cx^2 I_{(-c,c)}(x)$.

16.- Encontrar media, mediana y moda del ejercicio 15.

17.- Demostrar los incisos I y II del teorema 2.3.

18.- Una persona lanza 4 pelotas en siete cajas, se sabe que cada pelota tiene la misma probabilidad de caer en cualquiera de las 7 cajas. Sea X el número de pelotas que caen en la última caja. Encontrar

a) Para valores del recorrido de X, la función de distribución.
b) La función de densidad de X.
c) La media y la varianza de X.
d) El 0.4-ésimo cuantil.
e) La moda.

19.- Considerando la variable aleatoria definida como número de soles, en el ejemplo 2.4, calcular la función generadora de probabilidades, y a partir de esta, encontrar la media y la varianza.

20.- Un dado es lanzado varias veces hasta que aparece un número menor que 3. Sea X el número de lanzamientos requeridos. Para X, encontrar

a) La función de densidad.
b) La función generadora de momentos.
c) La función generadora de probabilidades.
d) La media y la varianza.
e) La moda.

21.- Sea X una variable aleatoria con función de densidad dada por $f_X(x) = \frac{1}{2} \mid 2 - x \mid I_{[0,c]}(x)$. Encontrar

a) c.
b) La función de distribución de X.
c) La mediana.
d) La moda.

22.- Sea X una variable aleatoria con función de densidad dada por $f_X(x) = \frac{1}{2}[1 - \mid \frac{x-a}{2} \mid] I_{(a-2,a+2)}(x)$.

a) Demostrar que $f_X(x)$ es función de densidad.
b) Encontrar la función de distribución de X.
c) Encontrar la media.
d) Encontrar la mediana.
e) Encontrar el 0.8-ésimo cuantil.

23.- Sea X una variable aleatoria con función de densidad dada por $f_X(x) = c[1 - (x-a)^2] I_{(a-1,a+1)}(x)$.

a) Encontrar c.
b) Encontrar la media.
c) Encontrar la mediana.

24.- Sea $f_X(x) = c[\lambda I_{(0,1)}(x) + I_{[1,2]}(x) + (1-\lambda)I_{(2,3)}(x)]$, donde λ es una constante.

a) Si $\lambda = 0.5$, encontrar
   I) La constante c para que $f_X(x)$ sea función de densidad.
   II) La función de distribución.
   III) La mediana.

b) Para cualquier valor de λ que satisface $0 \leq \lambda \leq 1$, encontrar
   I) La constante c para que $f_X(x)$ sea función de densidad.
   II) La función de distribución.
   III) La mediana.

25.- Una persona lanza una moneda hacia una raya (juego de la rayuela). Si la moneda cae a lo más 50 mm de la raya (distancia perpendicular a la raya), la persona gana. Sea X la distancia perpendicular en mm de la raya al punto donde cayó la moneda. La función de densidad de X está dada por $f_X(x) = \frac{120-x}{7200} I_{(0,120)}(x)$.

a) Si la persona lanza 3 monedas, calcular la probabilidad de que al menos gane una vez.
b) Si la persona lanza 8 monedas, calcular la probabilidad de que no gane ninguna de las veces.

26.- Sea X una variable aleatoria con media μ. Demostrar que $E[(X-c)^2]$, tiene su valor mínimo en $c = \mu$.

27.- Sea X una variable aleatoria con primer y segundo momentos respectivamente, $E[X] = 5$ y $E[X^2] = 34$. Determinar el límite inferior de la siguiente probabilidad $P[-4 < X < 14]$.

28.- Sea X una variable aleatoria con media μ y que cumple $P[X \leq 0] = 0$, demostrar que $P[X > 2\mu] \leq \frac{1}{2}$. *Sugerencia: aplicar el teorema 2.7.*

29.- Una urna contiene pelotas numeradas del 1 al 4. Primero se saca una pelota de la urna, y después se lanza una moneda el número de veces observado en la pelota seleccionada. Encontrar el valor esperado del número de águilas obtenidas.

30.- Sea X una variable aleatoria tal que tiene una distribución dada por $P[X = 0] = 1 - p$ y $P[X = 1] = p$, donde $0 \leq p \leq 1$. ¿Para qué valor de p, la varianza de X es máxima?

31.- Sea $f_X(x) = (cx + \frac{1}{4})I_{(-2,2)}(x)$ donde c es una constante. Encontrar

a) El rango de valores de c para los cuales $f_X(x)$ es una función de densidad.
b) $E[X]$.
c) La función de distribución.
d) La moda.

32.- Sea X una variable aleatoria con media μ, desviación estándar σ, y función generadora de momentos $m_X(t), -h < t < h$. Sea $Z = \frac{X-\mu}{\sigma}$. Al procedimiento de restar la media a una variable aleatoria y luego dividir entre su desviación estándar, se le conoce como procedimiento de estandarización. Demostrar

a) $E[Z] = 0$.
b) $V[Z] = 1$.
c) $m_Z(t) = e^{-\frac{\mu}{\sigma}t} m_X(\frac{t}{\sigma}),\ -h\sigma < t < h\sigma$.

33.- Sea X una variable aleatoria tal que $P[X = c] = 1$. A este tipo de distribución se le llama *distribución degenerada en c*. Para una variable aleatoria degenerada en c. Encontrar

a) La función de densidad.
b) La función de distribución.
c) $E[X]$ y $V[X]$.
d) La función generadora de momentos.

34.- Sea X una variable aleatoria continua con funciones de densidad y de distribución $f_X(x)$ y $F_X(x)$, respectivamente, y con el siguiente recorrido $0 < x < c < \infty$. Demostrar que $E[X] = \int_0^c [1 - F_X(x)]dx$.

35.- Sea $f_X(x) = c3^{-x} I_{\{1,2,3,..\}}(x)$ una función de densidad. Encontrar c.

36.- En una caja hay 12 pelotas numeradas del 1 al 12. Se seleccionan al azar dos pelotas, una tras otra con reemplazo. Sea X el número más grande de las pelotas seleccionadas. Obtener la función de densidad de X.

37.- Considerando el ejercicio 36, supongamos que ahora las dos pelotas son seleccionadas sin reemplazo. Calcular la función de densidad para X.

38.- En una caja hay r bolas numeradas del 1 al r. Se sacan al azar n bolas sin reemplazo. Sea X el número más grande de la muestra. Calcular $P[X \leq x]$, para x un valor del recorrido.

39.- Sea X una variable aleatoria discreta con valores enteros no negativos, y con función generadora de probabilidades $M_X(t) = e^{c(t^2-1)}$, donde $c > 0$. Encontrar la función de densidad de X. *Sugerencia: Desarrollar* e^{ct^2} *en serie de Maclaurin y multiplicar cada sumando de la serie por* e^{-c}.

40.- Considerando el ejemplo 2.39, calcular la varianza de X por medio de la función generadora de momentos.

41.- Sea X una variable aleatoria con función de densidad dada por $f_X(x) = cxI_{\{1,2,3,\ldots,n\}}(x)$, donde n es un número entero positivo. Encontrar

a) c.
b) La media de X.
c) La moda.

42.- El tiempo que tarda un asesor de un banco en realizar el trámite de un crédito es una variable aleatoria X con función de densidad $f_X(x) = ce^{-cx}I_{[0,\infty)}(x)$, tal que cumple $P[X \leq c] = P[X \geq c]$. Calcular

a) La constante c.
b) El tiempo promedio para hacer el trámite.
c) La mediana de la variable X.
d) La moda de X.

43.- Sea X una variable aleatoria con densidad $f_X(x) = kx^2 I_{[0,\sqrt[3]{3/k}]}(x)$. Supongamos que el 0.75-ésimo cuantil es igual a $\sqrt[3]{6}$. Encontrar

a) k.
b) La moda.

44.- El tiempo en meses para el mantenimiento de una máquina es una variablea aleatoria X con función de densidad $f_X(x) = e^{-cx}I_{[0,\infty)}(x)$. Encontrar

a) c.
b) La función de distribución.
c) $P[0.5 \leq e^X \leq 7]$.

45.- Un triángulo rectángulo es construido con dimensiones de sus catetos de X y $2X$, donde X es una variable aleatoria con función de densidad dada por $f_X(x) = \frac{x}{2}I_{(0,2)}(x)$, ¿cuál es el área esperada del triángulo?

46.- Sea X una variable aleatoria con función generadora de momentos $m_X(t) = (1-\beta t)^{-k}$, $t < 1/\beta$. Para n entero no negativo, calcular $E[X^n]$.

47.- Dos bolas son lanzadas de tal manera que cada bola tiene igual posibilidad de caer en cualquiera de 4 hoyos existentes. Ambas bolas pueden caer en el mismo hoyo. Sea X el número de hoyos desocupados al final del experimento. Encontrar la función generadora de momentos de la variable X.

48.- Sea X una variable aleatoria continua con función de densidad $f_X(x) = \frac{1}{32}kxI_{[0,k]}(x)$, si la moda es $x = 4$, encontrar la mediana.

49.- Las ventas semanales de una empresa es una variable aleatoria X con función de densidad $f_X(x) = \frac{1}{2}x^2e^{-x}I_{(0,\infty)}(x)$. Calcular la moda de X.

50.- La probabilidad de que una acción suba en un periodo dado es de $\frac{1}{2}$. Consideremos la siguiente variable aleatoria X, número de periodos hasta observar por primera vez que la acción sube.

a) Demostrar que $f_X(x) = \frac{1}{2^x} I_{\{1,2,3,\ldots\}}(x)$.
b) Calcular $P[X \text{ es par}]$.
c) Calcular $P[X \geq 5]$.
d) Calcular $P[X \text{ es divisible por } 3]$.

51.- Considerando la función de densidad del ejercicio 50, encontrar

a) La función generadora de probabilidades.
b) La media.
c) La varianza.
d) La función de distribución para valores del recorrido.

52.- Sea X una variable aleatoria con función de densidad dada por $f_X(x) = c\{xI_{[0,1]}(x) + I_{(1,2]}(x) + (3-x)I_{(2,3]}(x)\}$.

a) Calcular c.
b) Determinar la función de distribución y graficarla.
c) Encontrar la mediana.
d) Si X_1, X_2, X_3 son tres observaciones independientes de X, calcular la probabilidad de que
   1) Exactamente una de estas tres sea mayor que 1.5.
   2) Al menos una de las tres sea menor que 1.5.

53.- La tasa de interés en una inversión, varía entre periodo y periodo. Sea X, el número de periodos hasta observar por primera vez un periodo con tasa menor al 10 %. La función de densidad de la variable X es $f_X(x) = c(1-p)^{x-1} I_{\{1,2,\ldots,\}}(x)$, $0 < p < 1$. Determinar

a) c.
b) La moda de la distribución.
c) La función de distribución para valores del recorrido.

54.- Sea X una variable aleatoria con función de distribución dada por $F_X(x) = (1 - e^{-cx})I_{[0,\infty)}(x)$. Si el 0.75-ésimo cuantil es $\frac{1}{3}$. Encontrar

a) La constante c.
b) La moda.
c) La mediana.

55.- La póliza de un seguro de automóvi, en caso de accidente, pagará una pérdida sujeta a un deducible c. La pérdida por un siniestro tiene función de densidad dada por $f_X(x) = \frac{1}{150} e^{-\frac{1}{150}x} I_{(0,\infty)}(x)$. Se sabe que la probabilidad de que la compañía pague menos de 150 es igual a 0.70. Calcular el valor del deducible c.

56.- Una tienda que vende determinado dispositivo electrónico ofrece la siguiente garantía al comprador: si el dispositivo falla dentro de los primeros 24 meses a partir de la compra, la tienda regresa al comprador el 100 % del costo. Si falla después del mes 24 y antes del mes 30, la tienda regresará el 70 % del costo total, en caso de que se descomponga a partir del mes 30 pero antes del 36, la tienda regresará el 50 % del total. En otro caso, no hay garantía. Se sabe que la vida útil en meses del dispositivo tiene una función de densidad dada por $f_X(x) = \frac{1}{50}e^{-\frac{1}{50}x}I_{(0,\infty)}(x)$. Si una empresa le compra 250 dispositivos y el costo unitario es de 100, calcular el valor esperado de lo que la tienda reembolsará por garantía de esta compra.

57.- Una póliza colectiva de seguros cubre los gastos médicos mayores a los empleados de una compañía, suponiendo que existe una suma asegurada limitada, la proporción pagada por la aseguradora a la compañía en un año es una variable con función de densidad $f_X(x) = cx(1-x)I_{[0,1]}(x)$. Encontrar

a) c.
b) La moda.
c) La función de distribución.
d) La probabilidad de que la aseguradora le pague a la compañía en un año una proporción mayor al 69 %, dado que se sabe que la proporción ya excedió el 50 %.
e) La mediana.

58.- Sea X una variable aleatoria con función generadora de momentos $m_X(t)$. Demostrar que se cumplen las igualdades: $\frac{d}{dt}ln[m_X(t)]\Big|_{t=0} = E[X]$; $\frac{d^2}{dt^2}ln[m_X(t)]\Big|_{t=0} = V[X]$.

59.- Sea X una variable aleatoria continua con función de densidad $f_X(x)$ la cual es simétrica con respecto a c, demostrar que se cumplen los siguientes incisos:

a) La media es c.
b) La mediana es c.
c) $E[(X-\mu)^k] = 0$ para todo k entero impar.

60.- Calcular la media y mediana de la función de densidad dada en el ejercicio 57.

61.- Sea X una variable aleatoria con función de densidad dada por $f_X(x) = c(2-|x|)I_{(-1,2)}(x)$. Encontrar

a) La constante c.
b) La media.

c) Función de distribución.
d) La mediana.
e) La moda.

62.- Sea X una variable aleatoria con función de distribución dada por

$$F_X(x) = \begin{cases} 0 & \text{si} \quad x < 0, \\ x/10 & \text{si} \quad 0 \leq x < 2, \\ x/4 & \text{si} \quad 2 \leq x < 3, \\ 3/4 + x/24 & \text{si} \quad 3 \leq x < 6, \\ 1 & \text{si} \quad x \geq 6. \end{cases}$$

Calcular

a) $P[X = 2]$.
b) $P[X = 2.5]$.
c) $P[2 \leq X \leq 3]$.
d) $P[2 < X < 3]$.

63.- Encontrar la función de densidad truncada a la izquierda de 0, del ejercicio 61.

64.- Considerando la función de densidad del ejemplo 50, obtener la función de densidad truncada en el conjunto $\{1, 2, 3\}$.

65.- Calcular la varianza del ejemplo 2.32.

66.- Demostrar el teorema 2.11.

67.- En una compañía aseguradora, se sabe que el reclamo individual de las mujeres por un seguro de auto es una variable aleatoria con un promedio de 45,600 y una varianza de 98,800. Además, el reclamo individual de los hombres es otra variable aleatoria, con un promedio de 47,000 y una varianza de 122,100. Se sabe que el 67 % de los conductores son hombres y el resto son mujeres. Calcular

a) El reclamo promedio de todos los asegurados.
b) La desviación estándar del reclamo de todos los asegurados.

68.- Sea X una variable aleatoria con función de densidad dada por $f_X(x) = c|3 - x|I_{[0,7]}(x)$. Encontrar

a) La constante c.
b) La media.
c) La moda.
d) La función de distribución.
e) La mediana.

69.- Sea X una variable aleatoria con recorrido el intervalo $(0, 2c)$, donde la gráfica de la función de densidad se da a continuación:

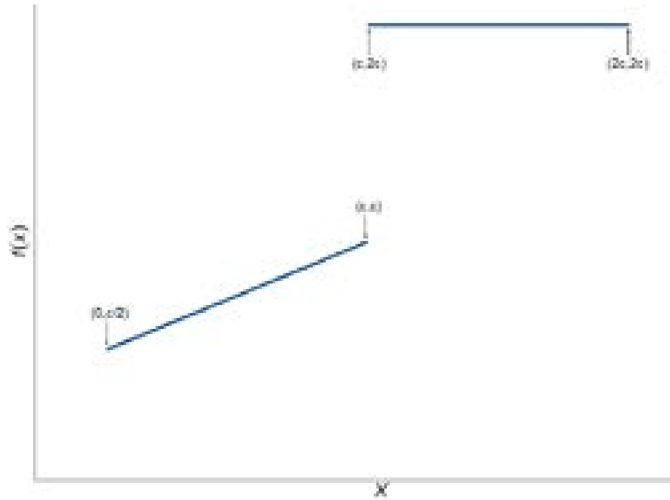
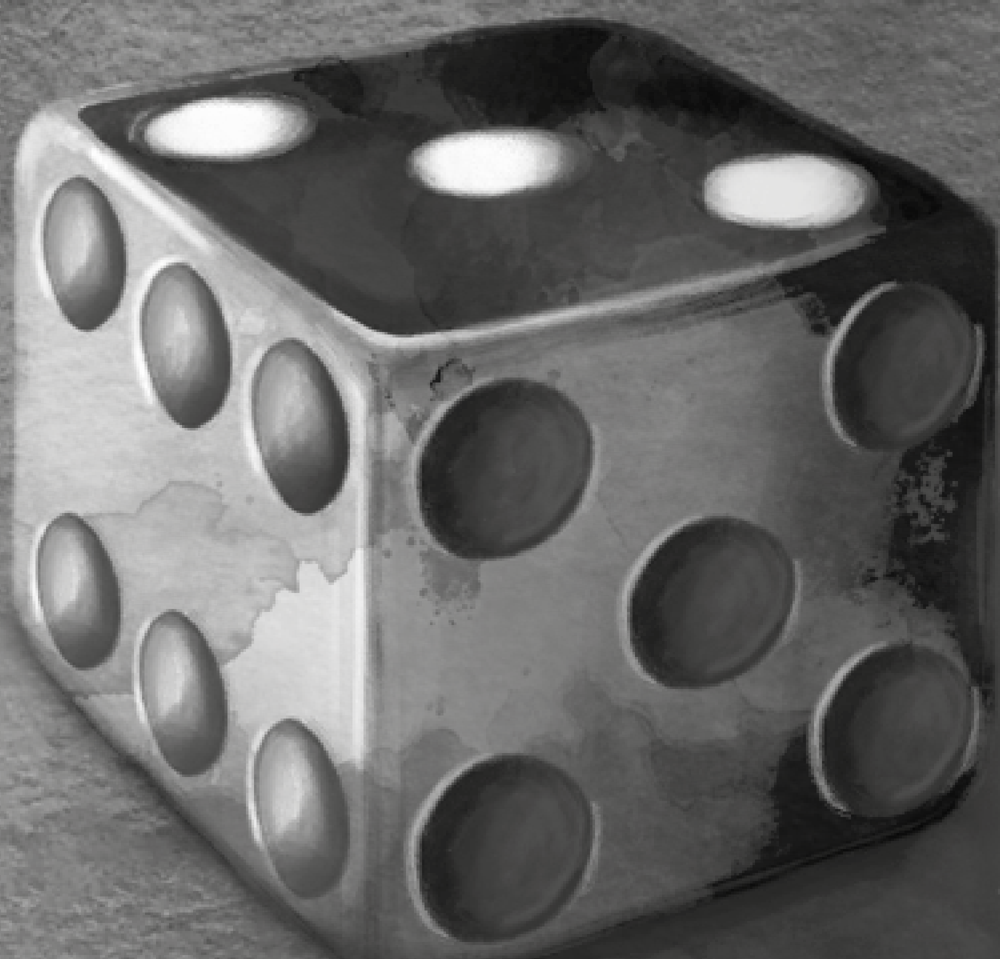


Figura 2.15. *Función de densidad del ejercicio 69*

a) Encontrar el valor de c.
b) Encontrar la función de densidad.
c) Calcular $P[1/3 < X < 1]$.
d) Calcular la media.

Distribuciones
discretas

Capítulo



Introducción

En este capítulo se estudiarán los modelos discretos más importantes y conocidos en la literatura de probabilidad y estadística. Estos cobran mucha importancia por las aplicaciones que tienen en la vida real.

Para cada modelo discreto, se explicará en qué consiste el experimento, se presentará la definición de la distribución, y se darán las propiedades más importantes, como la media, la varianza, la función generadora de momentos, entre otras. Después se considerarán algunos ejemplos que ayudarán a entender mejor cómo se puede aplicar cada modelo.

3.1. Distribución uniforme discreta

En esta sección se presentará la definición de la distribución uniforme discreta, sus principales propiedades y algunos ejemplos. Además, se explicará la importancia que tiene esta distribución en simulación estocástica.

Definición 3.1. Una variable aleatoria discreta X tiene distribución *uniforme discreta* sobre el conjunto $\{a_1, a_2, \ldots, a_N\}$ si su función de densidad está dada por

$$f_X(x) = \frac{1}{N} I_{\{a_1, a_2, \ldots, a_N\}}(x),$$

donde N es un número entero positivo.

El recorrido de esta variable es $a_1, a_2, \ldots, a_N$.

Observación. La probabilidad se reparte en forma equitativa en cada valor del recorrido.

Notación. $X \sim UD\{a_1, a_2, \ldots, a_N\}$ significa que X es una variable aleatoria con distribución uniforme discreta sobre el conjunto $\{a_1, a_2, \ldots, a_N\}$.

La gráfica de la función de densidad de una distribución uniforme discreta tiene la siguiente forma:

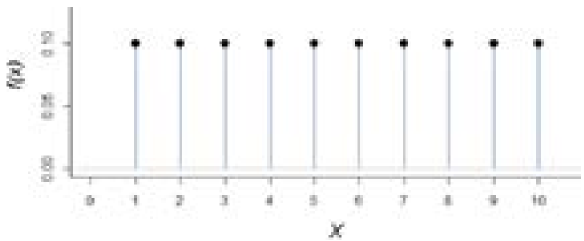


Figura 3.1. *Función de densidad. Uniforme discreta*

Se puede verificar que $f_X(x)$ es función de densidad.

Obsérvese

$$\frac{1}{N} > 0 \quad \text{e} \quad I_{\{a_1,a_2,\ldots,a_N\}}(x) \geq 0, \text{ donde } N \in \mathbb{N}.$$

Por lo tanto,

$$f_X(x) = \frac{1}{N} I_{\{a_1,a_2,\ldots,a_N\}}(x) \geq 0.$$

Por otro lado,

$$\begin{aligned} \sum_x f_X(x) &= \sum_x \frac{1}{N} \\ &= \frac{1}{N} N = 1. \end{aligned}$$

Por lo tanto, $f_X(x)$ es función de densidad.

A continuación, estudiaremos algunas propiedades que tiene la distribución $UD\{1, 2, \ldots, N\}$.

Teorema 3.1. *Sea X una variable aleatoria con distribución uniforme discreta sobre el conjunto $\{1, 2, \ldots, N\}$, entonces,*

$$E[X] = \frac{N+1}{2},$$

$$V[X] = \frac{N^2-1}{12},$$

$$m_X(t) = \frac{e^t(1-e^{tN})}{N(1-e^t)} I_{(-\infty,0)\cup(0,\infty)}(t) + I_{\{0\}}(t).$$

Demostración. Obsérvese

$$\begin{aligned} m_X(t) &= E[e^{tX}] \\ &= \sum_{x=1}^{N} \frac{1}{N} e^{tx} \\ &= \frac{1}{N} e^t [1 + e^t + e^{2t} + \cdots + e^{t(N-1)}] \\ &= \frac{e^t(1-e^{tN})}{N(1-e^t)}, \text{ si } t \neq 0. \end{aligned}$$

En principio, la suma anterior existe para todo valor de t diferente a cero. Es posible definir la función generadora de momentos en t igual a cero, obteniendo el límite de la anterior función cuando t tiende a cero.

$$\begin{aligned}\lim_{t\to 0}\frac{e^t(1-e^{tN})}{N(1-e^t)} &= \lim_{t\to 0}\frac{e^t-(N+1)e^{t(N+1)}}{-Ne^t}\\ &= \frac{1-(N+1)}{-N}=1.\end{aligned}$$

De esta manera, queda definida la función generadora de momentos para todo valor de t, esto es

$$m_X(t)=\frac{e^t(1-e^{tN})}{N(1-e^t)}I_{(-\infty,0)\cup(0,\infty)}(t)+I_{\{0\}}(t).$$

Procederemos a obtener el valor esperado y la varianza de X. En esta ocasión es más fácil la realización de estos cálculos usando la función de densidad.

$$\begin{aligned}E[X] &= \sum_{x=1}^{N} x\frac{1}{N}\\ &= \frac{1}{N}\sum_{x=1}^{N} x\\ &= \frac{1}{N}\left[\frac{N(N+1)}{2}\right]\\ &= \frac{N+1}{2}.\end{aligned}$$

Observése que se aplicó la suma de los primeros N números naturales (véase suma C.3 del apéndice C).

$$\begin{aligned}E[X^2] &= \sum_{x=1}^{N} x^2\frac{1}{N}\\ &= \frac{1}{N}\sum_{x=1}^{N} x^2\\ &= \frac{1}{N}\left[\frac{N(N+1)(2N+1)}{6}\right]\\ &= \frac{(N+1)(2N+1)}{6}.\end{aligned}$$

Aquí se aplicó la suma de los cuadrados de los primeros N números naturales (véase suma C.4 del apéndice C).

Por lo tanto,

$$\begin{aligned}
V[X] &= E[X^2] - E^2[X] \\
&= \frac{(N+1)(2N+1)}{6} - \left(\frac{N+1}{2}\right)^2 \\
&= (N+1)\left[\frac{2N+1}{6} - \frac{N+1}{4}\right] \\
&= \frac{N^2-1}{12}.
\end{aligned}$$

□

Observación. La distribución uniforme discreta no tiene moda, ya que las probabilidades son iguales para cada uno de los valores del recorrido.

Las aplicaciones de la distribución uniforme discreta serán en aquellos fenómenos donde cada valor del recorrido de la variable aleatoria considerada tiene la misma probabilidad de ocurrir. Por ejemplo, el número logrado en una rifa, el resultado observado al lanzar un dado, etc. Véase siguiente ejemplo.

Ejemplo 3.1. Se lanza un dado bien balanceado y se define la siguiente variable aleatoria X = número que resulta en el lanzamiento. Indicar la función de densidad, valor esperado, varianza y la desviación estándar de la variable X.

En este caso, el recorrido es $x = 1, 2, 3, 4, 5, 6$.

La función de densidad es $f_X(x) = \frac{1}{6} I_{\{1,2,3,4,5,6\}}(x)$.

El valor esperado de X es $E[X] = \frac{7}{2} = 3.5$.

La varianza se calcula como $V[X] = \frac{36-1}{12} = 2.917$ y, por lo tanto, la desviación estándar es $\sigma = 1.70783$.

3.2. Distribución Bernoulli

Se tratará la distribución Bernoulli, su definición y sus principales propiedades. Se darán algunos ejemplos de esta distribución.

Definición 3.2. Un *experimento Bernoulli* consiste en realizar una prueba o ensayo, la cual tiene dos opciones: «éxito» o «fracaso». La probabilidad de «éxito» es p y la de «fracaso» es $q = 1 - p$. El objetivo de realizar la prueba es observar si resulta «éxito» o no. La variable aleatoria se define como número de «éxitos» en la prueba.

En general, el recorrido de la variable aleatoria es $x = 0, 1$.

Definición 3.3. Una variable aleatoria discreta X tiene *distribución Bernoulli* con parámetro p, si su función de densidad está dada por

$$f_X(x) = p^x(1-p)^{1-x} I_{\{0,1\}}(x), \quad \text{donde} \quad 0 < p < 1.$$

Notación. $X \sim$ Bernoulli(p) significa que la variable aleatoria X se distribuye Bernoulli con parámetro p.

Se demostrará que $f_X(x)$ es función de densidad,

$$p^x > 0; \quad (1-p)^{1-x} > 0 \quad \text{e} \quad I_{\{0,1\}}(x) \geq 0.$$

Entonces,

$$f_X(x) = p^x(1-p)^{1-x} I_{\{0,1\}}(x) \geq 0.$$

Por otro lado,

$$f_X(0) + f_X(1) = 1 - p + p = 1.$$

Por lo tanto, $f_X(x)$ es función de densidad.

Teorema 3.2. *Sea X una variable aleatoria con distribución Bernoulli con parámetro p, entonces,*

$$\begin{aligned} E[X] &= p, \\ V[X] &= p(1-p), \\ m_X(t) &= 1 - p + e^t p. \end{aligned}$$

Demostración. Se calculará la función generadora de momentos

$$\begin{aligned} m_X(t) &= E[e^{tX}] \\ &= \sum_{x=0}^{1} e^{tx} p^x (1-p)^{1-x} \\ &= 1 - p + e^t p. \end{aligned}$$

El primer y segundo momentos, respectivamente,

$$\begin{aligned} E[X] &= \frac{d}{dt} m_X(t) \Big|_{t=0} \\ &= pe^t \Big|_{t=0} = p, \\ \frac{d^2}{dt^2} m_X(t) \Big|_{t=0} &= pe^t \Big|_{t=0} = p. \end{aligned}$$

Por lo tanto,

$$V[X] = p - p^2 = p(1-p).$$

□

El valor esperado y la varianza también se pueden calcular con mucha facilidad por medio de la función de densidad.

Obsérvese que si $p < 0.5$, la moda es $x = 0$, si $p > 0.5$, entonces, la moda es $x = 1$. En el caso de $p = 0.5$, no hay moda.

Las aplicaciones de la distribución Bernoulli serán en fenómenos donde solo se quiere observar la ocurrencia de un suceso o no, y la probabilidad de que ocurra el suceso es p. Algunos ejemplos son el lanzamiento de una moneda para observar águila o no, la selección de una carta de una baraja para observar si sale as o no, etc. Véase siguiente ejemplo.

Ejemplo 3.2. Se lanza una moneda y definimos la siguiente variable aleatoria, X = Número de soles. Indicar la función de densidad, la media y la varianza de la variable X.

La función de densidad es

$$f_X(x) = (0.5)^x(1-0.5)^{1-x}I_{\{0,1\}}(x) = 0.5I_{\{0,1\}}.$$

La media y la varianza se dan a continuación, $E[X] = 0.5$ y $V[X] = 0.25$.

Ejemplo 3.3. Se lanzan dos dados y definimos la siguiente variable aleatoria, X = Número de veces que ocurrieron ambos dados diferentes. Indicar la función de densidad, el valor esperado y la varianza de la variable aleatoria.

La función de densidad es

$$f_X(x) = \left(\frac{5}{6}\right)^x \left(\frac{1}{6}\right)^{1-x} I_{\{0,1\}}(x).$$

La media y la varianza de X son, respectivamente, $E[X] = \frac{5}{6}$ y $V[X] = \left(\frac{5}{6}\right)\left(\frac{1}{6}\right) = \frac{5}{36}$.

3.3. Distribución binomial

En esta parte se van a generalizar las definiciones de experimento y distribución Bernoulli, obteniendo de estas, el experimento y la distribución binomial. También, se tratarán sus principales propiedades y algunos ejemplos que nos ayudarán a entender la importancia que tiene este modelo en diversas aplicaciones.

Definición 3.4. Un *experimento binomial* consiste en realizar n pruebas o ensayos idénticos e independientes. Cada prueba tiene dos opciones: «éxito» o «fracaso». La probabilidad de «éxito» en cada prueba es p y la de «fracaso» es $q = 1 - p$. El objetivo de realizar las n pruebas es observar el número de «éxitos». La variable aleatoria se define como número de «éxitos» en las n pruebas.

Obsérvese que el recorrido de la variable aleatoria binomial en general es $x = 0, 1, 2, \ldots, n$.

Una aclaración, la realización de n pruebas o ensayos idénticos e independientes equivale a la obtención de una muestra aleatoria de tamaño n (véase definición 6.1).

Definición 3.5. Una variable aleatoria discreta X tiene *distribución binomial* con parámetros n y p, si su función de densidad está dada por

$$f_X(x) = \binom{n}{x} p^x(1-p)^{n-x}I_{\{0,1,2,\ldots,n\}}(x), \quad \text{donde} \quad 0 < p < 1.$$

Notación. $X \sim B(n, p)$ significa que la variable aleatoria X se distribuye binomial con parámetros n y p.

La gráfica de una distribución binomial puede tener 3 formas diferentes, esto depende del valor de p.

Obsérvese que la gráfica es simétrica para $p = 0.5$.

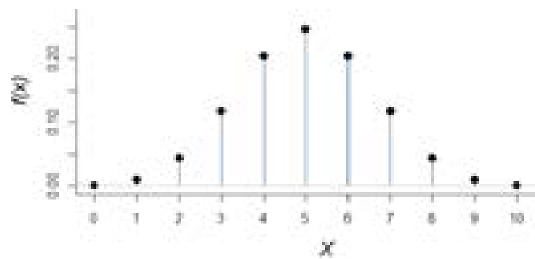


Figura 3.2. *Función de densidad. Binomial,* $p = 0.5$

Para $p < 0.5$, la gráfica es sesgada a la derecha.

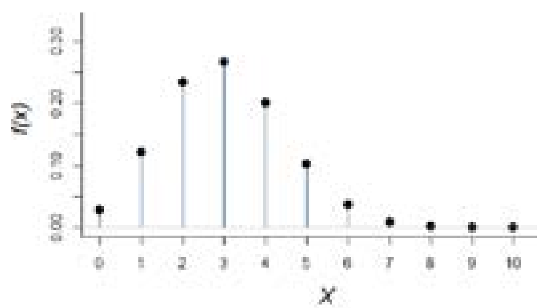


Figura 3.3. *Función de densidad. Binomial,* $p < 0.5$

Y para $p > 0.5$, la gráfica es sesgada a la izquierda.

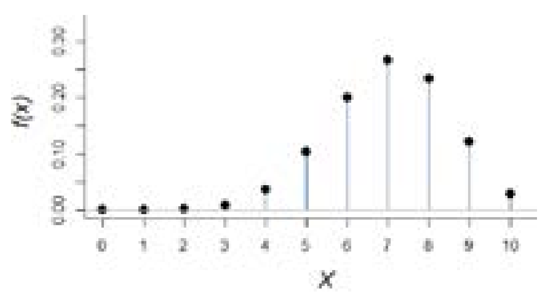


Figura 3.4. *Función de densidad. Binomial,* $p > 0.5$

Otras formas no se van a presentar. Estas gráficas ayudarán, entre otras cosas, a encontrar de una manera más fácil, algún cuantil específico, la mediana o la moda.

La función $f_X(x)$ cumple con las propiedades de una función de densidad discreta, esto es,

$$\binom{n}{x} > 0; \quad p^x > 0; \quad (1-p)^{n-x} > 0 \text{ e } I_{\{0,1,2,\ldots,n\}}(x) \geq 0.$$

Por lo tanto,

$$f_X(x) = \binom{n}{x} p^x (1-p)^{n-x} I_{\{0,1,2,\ldots,n\}}(x) \geq 0.$$

Por otro lado, aplicando el teorema del binomio (véase suma C.7 del apéndice C)

$$\begin{aligned}\sum_{x=0}^{n} f_X(x) &= \sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x} \\ &= (p+1-p)^n = 1.\end{aligned}$$

Por lo tanto, $f_X(x)$ es función de densidad.
Se puede observar que la función $f_X(x)$ representa la probabilidad de que el número de «éxitos» en las n pruebas sea igual a x, esto es, $f_X(x) = P$[número de éxitos sea igual a x].

Obsérvese que si de n pruebas, x son «éxitos», entonces, el resto, $n-x$, son «fracasos». Ahora, la probabilidad de que en x pruebas específicas resulten «éxitos» y en el resto «fracasos», es igual a $p^x(1-p)^{n-x}$.

Por otro lado, el número de maneras de seleccionar x pruebas del total de n, para que en estas ocurran «éxitos», es igual a $\binom{n}{x}$.

Por lo tanto, la probabilidad de tener x «éxitos» en n pruebas es igual a

$$f_X(x) = \binom{n}{x} p^x (1-p)^{n-x}.$$

Teorema 3.3. *Si X es una variable aleatoria con distribución binomial con parámetros n y p, entonces,*

$$E[X] = np,$$

$$V[X] = np(1-p),$$

$$m_X(t) = (1-p+pe^t)^n.$$

Demostración. La función generadora de momentos es

$$\begin{aligned}m_X(t) &= E[e^{tX}] \\ &= \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x (1-p)^{n-x} \\ &= \sum_{x=0}^{n} \binom{n}{x} (pe^t)^x (1-p)^{n-x} \\ &= (pe^t + 1 - p)^n.\end{aligned}$$

Obsérvese que se aplicó el teorema del binomio (véase apéndice C).
La media es

$$\begin{aligned}\frac{d}{dt}m_X(t)\bigg|_{t=0} &= n(pe^t+1-p)^{n-1}pe^t\bigg|_{t=0}\\ &= np.\end{aligned}$$

El segundo momento es

$$\begin{aligned}\frac{d^2}{dt^2}m_X(t)\bigg|_{t=0} &= n(pe^t+1-p)^{n-1}pe^t\\ &\quad + pe^t n(n-1)(pe^t+1-p)^{n-2}pe^t\bigg|_{t=0}\\ &= np+np(n-1)p.\end{aligned}$$

Por lo tanto, la varianza

$$\begin{aligned}V[X] &= np+np(n-1)p-n^2p^2\\ &= np(1-p).\end{aligned}$$

□

Teorema 3.4. *La moda de una distribución binomial con parámetros n y p, es un número entero M, que satisface la siguiente desigualdad*

$$np-q \leq M \leq np+p.$$

Demostración. Supongamos que x es un valor específico del recorrido, se puede observar

$$\begin{aligned}\frac{f_X(x)}{f_X(x-1)} &= \frac{\binom{n}{x}p^x(1-p)^{n-x}I_{\{0,1,2,\ldots,n\}}(x)}{\binom{n}{x-1}p^{x-1}(1-p)^{n-(x-1)}I_{\{0,1,2,\ldots,n\}}(x)}\\ &= \frac{\frac{n!}{x!(n-x)!}}{\frac{n!}{(x-1)!(n-(x-1))!}}\left(\frac{p}{1-p}\right)\\ &= \frac{(n-x+1)p}{xq}\\ &= \frac{xq}{xq}+\frac{(n-x+1)p-xq}{xq}\\ &= 1+\frac{(n+1)p-xp-x(1-p)}{xq}\\ &= 1+\frac{(n+1)p-x}{xq}.\end{aligned}$$

De esta última igualdad, podemos afirmar

$$f_X(x-1) < f_X(x), \text{ para } x < (n+1)p,$$
$$f_X(x-1) > f_X(x), \text{ para } x > (n+1)p,$$
$$f_X(x-1) = f_X(x), \ \text{ para } x = (n+1)p, \ \text{ si } (n+1)p \text{ es entero.}$$

Sea M la moda. Si $M > (n+1)p$, entonces, $f_X(M-1) > f_X(M)$, lo cual no puede ser, por lo tanto, $M \leq (n+1)p$.

Ahora, si $M+1 < (n+1)p$, entonces, $f_X(M) < f_X(M+1)$, y tampoco puede ser, por lo tanto, $M+1 \geq (n+1)p$, esto es, $M \geq (n+1)p-1 = np-q$.

En conclusión, $np - q \leq M \leq np + p$. □

Es importante aclarar que puede existir más de una moda, en este caso, las modas deberán cumplir la anterior desigualdad.

Observación. En la anterior demostración se llegó a la igualdad

$$\frac{f_X(x)}{f_X(x-1)} = 1 + \frac{(n+1)p - x}{xq}.$$

Dicho de otra manera,

$$f_X(x) = f_X(x-1)\left[\frac{(n-x+1)p}{xq}\right].$$

La igualdad anterior es una fórmula recursiva para calcular probabilidades de una distribución binomial, la cual es útil para calcular probabilidades de esta distribución fácilmente.

Ejemplo 3.4. Para una determinada acción financiera, se sabe que en un periodo, la acción sube con una probabilidad del 65 % y baja con una probabilidad del 35 %. Se supone independencia entre los periodos. Para los siguientes 25 periodos, considerar la variable aleatoria X = número de periodos a la baja, calcular

a) La probabilidad de $X = 10$.

b) La probabilidad de X sea mayor 2.

c) El promedio y desviación estándar.

d) El 0.1-ésimo cuantil.

e) La moda.

Obsérvese

X = Número de periodos a la baja de los siguientes 25 periodos.

$x = 0, 1, \ldots, 25$.

$X \sim B(25, 0.35)$.

a) La primera probabilidad es

$$P[X = 10] = \binom{25}{10}(0.35)^{10}(0.65)^{15} = 0.14085.$$

b) La siguiente probabilidad es

$$\begin{aligned} P[X > 2] &= 1 - P[X \leq 2] \\ &= 1 - \binom{25}{0}(0.35)^0(0.65)^{25} - \binom{25}{1}(0.35)^1(0.65)^{24} \\ &- \binom{25}{2}(0.35)^2(0.65)^{23} = 0.99787. \end{aligned}$$

c) La media, la varianza y la desviación estándar respectivamente son $E[X] = 25(0.35) = 8.75$; $V[X] = 8.75(0.65) = 5.6875$ y $\sigma = 2.38485$.

d) Se calcularán las probabilidades acumuladas para valores del recorrido, hasta que la probabilidad acumulada llegue o pase por primera vez de 0.10.

x	$f_X(x)$	$F_X(x)$
0	0.000021	0.000021
1	0.000283	0.000304
2	0.001829	0.002133
3	0.007551	0.009684
4	0.0223637	0.032048
5	0.050576	0.082624
6	0.090778	0.173402

Tabla 3.1. *Probabilidades, ejemplo 3.4*

Se observa que el 0.1-ésimo cuantil es 6.

e) Se sabe que la moda M cumple

$$np - q \leq M \leq np + p,$$

$$8.1 \leq M \leq 9.1.$$

Por lo tanto, la moda es $M = 9$.

Los incisos a, b y d se pueden calcular por medio de R, considerando los siguientes códigos:

a) $dbinom(10, 25, 0.35)$, dando como resultado 0.1408523.

b) $1 - pbinom(2, 25, 0.35)$, siendo el resultado 0.9978667.

c) $qbinom(0.1, 25, 0.35)$, donde el cuantil es 6.

3.4. Distribución hipergeométrica

En esta sección se explicará primero en qué consiste un experimento hipergeométrico, después se dará la definición de distribución hipergeométrica. También, se presentarán las propiedades más importantes de esta distribución. Además, se considerarán algunos ejemplos prácticos.

Definición 3.6. Considérese una población con N elementos, donde K de ellos tienen una misma característica, llamada «éxito», y el resto de los elementos $N-K$ tienen la característica opuesta a la anterior, llamada «fracaso». El *experimento hipergeométrico* consiste en sacar una muestra aleatoria de tamaño n ($n < N$) al mismo tiempo de esta población, o uno tras otro sin reemplazo, con el objetivo de observar el número de «éxitos». La variable aleatoria se define como el número de «éxitos» en la muestra seleccionada.

Observación. En este caso, el recorrido de la variable depende del tamaño de la muestra n y del tamaño de la subpoblación de «éxitos» K. En otras palabras, el recorrido no siempre empezará en cero y no siempre terminará en n, véase el siguiente ejemplo.

Ejemplo 3.5. Supongamos una población de 15 elementos, con una subpoblación de «éxitos» de K elementos, y una muestra aleatoria de tamaño $n = 5$. Sea X = número de «éxitos» en la muestra. Encontrar el recorrido de X para
a) $K = 4$.
b) $K = 11$.

a) Considerando $K = 4$, el recorrido es $x = 0, 1, 2, 3, 4$.
b) Considerando $K = 11$, el recorrido es $x = 1, 2, 3, 4, 5$.

Con la idea de que la función indicadora que es usada en la función de densidad siempre se considere sobre el conjunto $\{0, 1, 2, \ldots, n\}$, a pesar de que el recorrido no siempre sea este conjunto, se va a definir el siguiente número combinatorio

$$\binom{N}{M} = 0, \quad \text{si } M > N, \text{o } M < 0, \text{o } N < 0.$$

Definición 3.7. Una variable aleatoria discreta X tiene *distribución hipergeométrica* con parámetros N, K y n, si su función de densidad está dada por

$$f_X(x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}} I_{\{0,1,2,\ldots,n\}}(x),$$

donde N, K y n son números enteros no negativos, además $N > K$ y $N > n$.

Notación. $X \sim H(N, K, n)$, significa que la variable aleatoria X se distribuye hipergeométrica con parámetros: N, K y n.

La forma de la función de densidad hipergeométrica puede tomar tres formas distintas, esto dependerá, como en la distribución binomial, de la proporción de éxitos que existe en la población.

Si la proporción de éxitos y la de fracasos coinciden en la población, entonces, la gráfica de la función de densidad será simétrica (en la distribución binomial es cuando p =0.5), véase la siguiente gráfica

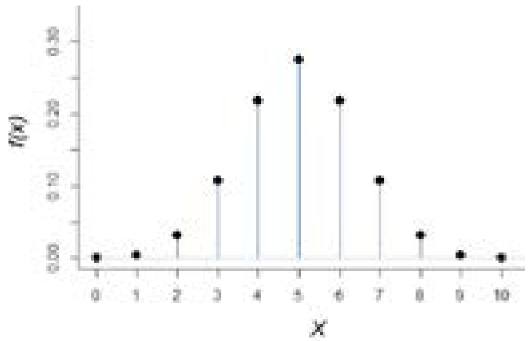


Figura 3.5. *Función de densidad. Hipergeométrica,* $K/N = 0.5$

Si la proporción de éxitos es menor que la proporción de fracasos en la población, entonces, la gráfica de la función de densidad será sesgada a la derecha.

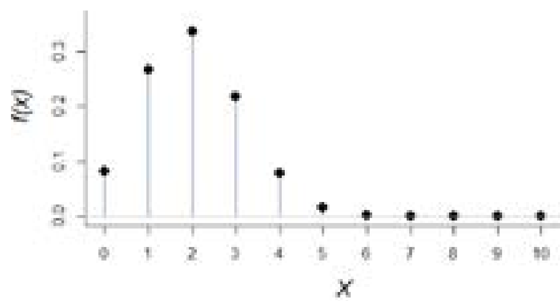


Figura 3.6. *Función de densidad. Hipergeométrica,* $K/N < 0.5$

Si la proporción de éxitos es mayor que la proporción de fracasos en la población, la gráfica de la función de densidad será segada a la izquierda, esto es,

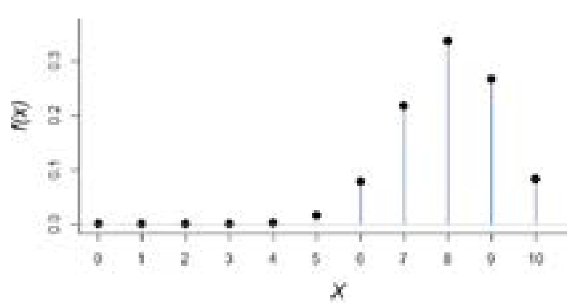


Figura 3.7. *Función de densidad. Hipergeométrica,* $K/N > 0.5$

Las diferentes formas de la función de densidad pueden ayudar al cálculo de cuantiles, mediana y moda de esta distribución.

A continuación, demostraremos que $f_X(x)$ es función de densidad.

Obsérvese

$$\binom{K}{x} \geq 0; \binom{N-K}{n-x} \geq 0; \binom{N}{n} > 0 \text{ e } I_{\{0,1,2,\ldots,n\}}(x) \geq 0.$$

Por lo tanto,

$$f_X(x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}} I_{\{0,1,2,\ldots,n\}}(x) \geq 0.$$

Por otro lado,

$$\begin{aligned}
\sum_{x=0}^{n} f_X(x) &= \sum_{x=0}^{n} \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}} \\
&= \frac{1}{\binom{N}{n}} \sum_{x=0}^{n} \binom{K}{x}\binom{N-K}{n-x} \\
&= \frac{1}{\binom{N}{n}} \binom{N}{n} = 1.
\end{aligned}$$

En la última suma, se aplicó la ecuación C.8 del apéndice C.

Por lo tanto, $f_X(x)$ es función de densidad.

La función de densidad $f_X(x)$ representa la probabilidad de que el número de éxitos en la muestra aleatoria de tamaño n, sea igual a x.

Esto es, el número de maneras diferentes de tener x éxitos y $n-x$ fracasos es $\binom{K}{x}\binom{N-K}{n-x}$, ahora, el número de maneras diferentes de obtener una muestra de tamaño n es $\binom{N}{n}$.

Por lo tanto, la probabilidad de lograr x éxitos en una muestra de tamaño n obtenida de una población de tamaño N, donde K son éxitos y $N-K$ son fracasos es

$$f_X(x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}.$$

Tanto en la distribución binomial como en la distribución hipergeométrica, la variable aleatoria se define en términos generales, como el número de éxitos en una muestra aleatoria de tamaño n.

La diferencia entre los dos modelos estriba en que, en la distribución hipergeométrica, la muestra se obtiene de una población finita (de tamaño N) y, por lo mismo, la probabilidad de éxito en el primer elemento que se selecciona es diferente a la probabilidad de éxito para el segundo elemento que se obtiene y diferente para los elementos posteriores que se seleccionan de la población. Además, la probabilidad de éxito en el r-ésimo elemento seleccionado depende de los resultados que se obtengan en los anteriores elementos que se sacaron.

En la distribución binomial, la probabilidad de éxito en cada prueba es la misma, esto es p, además, entre elemento y elemento que se selecciona, existe independencia. Esto se debe a que en el experimento binomial, la muestra se obtiene de una población infinita, y no finita como en el caso del experimento hipergeométrico.

Teorema 3.5. *Si X es una variable aleatoria con distribución hipergeométrica con parámetros N, K y n, entonces,*

$$E[X] = \frac{nK}{N},$$

$$V[X] = n\left(\frac{K}{N}\right)\left(1-\frac{K}{N}\right)\left(\frac{N-n}{N-1}\right).$$

Demostración. Se calculará el valor esperado de X

$$\begin{aligned}
E[X] &= \sum_{x=0}^{n} x f_X(x) \\
&= \sum_{x=1}^{n} x \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}} \\
&= \sum_{x=1}^{n} x \frac{\frac{K!}{x!(K-x)!}\binom{N-K}{n-x}}{\binom{N}{n}} \\
&= \sum_{x=1}^{n} \frac{\frac{K!}{(x-1)!(K-x)!}\binom{N-K}{n-x}}{\binom{N}{n}} \\
&= \sum_{x=1}^{n} \frac{K\left[\frac{(K-1)!}{(x-1)!(K-1-(x-1))!}\right]\binom{N-K}{n-x}}{\left(\frac{N}{n}\right)\left[\frac{(N-1)!}{(n-1)!((N-1)-(n-1))!}\right]}
\end{aligned}$$

$$= \frac{nK}{N} \sum_{x=1}^{n} \frac{\binom{K-1}{x-1}\binom{N-1-(K-1)}{n-1-(x-1)}}{\binom{N-1}{n-1}},$$

$$= \frac{nK}{N} \sum_{y=0}^{n-1} \frac{\binom{K-1}{y}\binom{N-1-(K-1)}{n-1-y}}{\binom{N-1}{n-1}}.$$

En la anterior suma se hizo el cambio de variable $y = x - 1$. Obsérvese que esta suma se realiza sobre los valores de una distribución hipergeométrica con parámetros $N-1, K-1$ y $n-1$.

Por lo tanto,

$$E[X] = \frac{nK}{N}. \qquad \square$$

El cálculo de la varianza se calcula a partir del segundo momento factorial, este se puede calcular de una manera similar al primer momento. A continuación, se iniciará la demostración.

$$E[X(X-1)] = \sum_{x=0}^{n} x(x-1) \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$$

$$= \sum_{x=2}^{n} x(x-1) \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$$

$$= \sum_{x=2}^{n} \frac{\frac{K!}{(x-2)!(K-x)!}\binom{N-K}{n-x}}{\binom{N}{n}}.$$

Obsérvese que se tendrá que ajustar a una suma de valores de una distribución $H(N-2, K-2, n-2)$. Se le sugiere al estudiante terminar esta comprobación.

Observación. En el teorema anterior, no se indica cuál es la función generadora de momentos, esta función sí existe, pero no es útil para obtener momentos. Por esta razón, la media y la varianza son obtenidas a partir del primer y segundo momentos factoriales, los cuales se calculan a partir de la función de densidad.

Se puede notar que las medias de las distribuciones binomial e hipergeométrica coinciden, siendo estas la multiplicación de n por la proporción de «éxitos». Por otra parte, la varianza de la binomial es la multiplicación de n por la proporción de «éxitos» y por la proporción de «fracasos». Obsérvese que la varianza de la hipergeométrica, es la varianza de la binomial multiplicada por el factor $\frac{N-n}{N-1}$.

El factor $\frac{N-n}{N-1}$ de la varianza de la distribución hipergeométrica es conocido como *factor de población finita*.

Obsérvese, cuando el tamaño de la población N es muy grande con respecto al tamaño de la muestra n, el factor de población finita es aproximadamente 1, entonces, las varianzas de ambas distribuciones serán muy parecidas. Cuando lo anterior pasa, las probabilidades de un evento calculadas con ambas distribuciones también serán muy parecidas. En resumen, las distribuciones hipergeométrica y binomial serán muy parecidas, cuando N crece con respecto a n (la población finita tiende a ser infinita) y la proporción de «éxitos» permanece constante, véase teorema 3.7.

Teorema 3.6. *La moda de una distribución hipergeométrica con parámetros N, K y n es un número entero M que cumple con la siguiente desigualdad*

$$\frac{(n+1)K+n-1-N}{N+2} \leq M \leq \frac{(n+1)K+n+1}{N+2}.$$

La demostración de este teorema se deja como ejercicio para el estudiante.

Como en el caso de la distribución binomial, también existe la posibilidad de tener más de una moda, si es el caso, cada una de ellas deberá de cumplir la anterior desigualdad.

Observación. Como se había comentado, si la muestra se obtiene de una población finita, entonces, la distribución correcta es la hipergeométrica, pero si la muestra se obtiene de una población infinita, entonces, la distribución es binomial. De este modo, si los elementos que se sacan de una población se obtienen uno por uno con reemplazo, entonces, la población se considera como infinita y el modelo correspondiente será la distribución binomial.

Ejemplo 3.6. En un grupo de doce profesionistas, se sabe que la fortaleza de la cuarta parte del grupo son las finanzas. Se saca una muestra de tamaño dos. ¿Cuál es la probabilidad de que resulte un experto en finanzas?
a) Sin reemplazo.
b) Con reemplazo.

a) X = Número de profesionistas expertos en finanzas en la muestra.
$x = 0, 1, 2$.
$X \sim H(N = 12, K = 3, n = 2)$.

$$f_X(1) = \frac{\binom{3}{1}\binom{9}{1}}{\binom{12}{2}} = \frac{27}{66} = 0.40909.$$

b) $X =$ Número de profesionistas expertos en finanzas en la muestra.
$x = 0, 1, 2.$
$X \sim B\left(2, \frac{1}{4}\right).$

$$f_X(1) = \binom{2}{1}\left(\frac{1}{4}\right)^1\left(\frac{3}{4}\right)^{2-1} = 0.375.$$

Teorema 3.7. *Si X es una variable aleatoria con distribución hipergeométrica con parámetros N, K y n, entonces,*

$$\lim_{\substack{N\to\infty \\ \frac{K}{N}=p}} \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}} I_{\{0,1,\ldots,n\}}(x) = \binom{n}{x} p^x (1-p)^{n-x} I_{\{0,1,\ldots,n\}}(x).$$

Demostración.

$$\begin{aligned}
&\lim_{\substack{N\to\infty \\ \frac{K}{N}=p}} \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}} I_{\{0,1,\ldots,n\}}(x) \\
&= \lim_{N\to\infty} \frac{\binom{Np}{x}\binom{Nq}{n-x}}{\binom{N}{n}} I_{\{0,1,\ldots,n\}}(x) \\
&= \lim_{N\to\infty} \frac{\frac{(Np)!}{x!(Np-x)!}\,\frac{(Nq)!}{(n-x)!(Nq-(n-x))!}}{\frac{N!}{n!(N-n)!}} \\
&= \lim_{N\to\infty} \binom{n}{x}\left[\frac{(Np)!}{(Np-x)!}\right]\left[\frac{(Nq)!}{(Nq-(n-x))!}\right]\left[\frac{(N-n)!}{N!}\right].
\end{aligned}$$

Ahora obsérvese

$$\begin{aligned}
\frac{(Np)!}{(Np-x)!} &= \frac{1(2)\cdots(Np-x)(Np-(x-1))\cdots(Np-1)Np}{1(2)\cdots(Np-x)} \\
&= (Np-(x-1))\cdots(Np-1)Np.
\end{aligned}$$

Es importante aclarar que el anterior resultado, tiene x factores.
Por otro lado,

$$\begin{aligned}
&\frac{(Nq)!}{(Nq-(n-x))!} \\
&= \frac{1(2)\cdots(Nq-(n-x))(Nq-(n-x-1))\cdots(Nq-1)Nq}{1(2)\cdots(Nq-(n-x))} \\
&= (Nq-(n-x-1))\cdots(Nq-1)Nq.
\end{aligned}$$

Puede verse que este último resultado tiene $n-x$ factores.

También obsérvese

$$\begin{aligned}\frac{(N-n)!}{N!} &= \frac{1(2)\cdots(N-n)}{1(2)\cdots(N-n)(N-(n-1))\cdots N} \\ &= \frac{1}{(N-(n-1))(N-(n-2))\cdots N}.\end{aligned}$$

De esta manera, el anterior límite es igual a

$$\lim_{N\to\infty}\binom{n}{x}\frac{(Np-(x-1))\cdots(Np-1)Np(Nq-(n-x-1))\cdots Nq}{(N-(n-1))(N-(n-2))\cdots N}$$

$$\times I_{\{0,1,\ldots,n\}}(x).$$

Obsérvese que el número de factores en el numerador es $x+n-x=n$, siendo el mismo número de factores que en el denominador.

Se va dividir, tanto en el numerador como en el denominador, por N^n.

$$\begin{aligned}&\lim_{N\to\infty}\binom{n}{x}\frac{(p-\frac{x-1}{N})\cdots(p-\frac{1}{N})p(q-\frac{n-x-1}{N})\cdots(q-\frac{1}{N})q}{(1-\frac{n-1}{N})(1-\frac{n-2}{N})\cdots(1-\frac{x}{N})(1-\frac{x-1}{N})\cdots(1-\frac{1}{N})1} \\ &\quad\times I_{\{0,1,\ldots,n\}}(x) \\ &= \binom{n}{x}p^x q^{n-x} I_{\{0,1,\ldots,n\}}(x).\end{aligned}$$ $\square$

El siguiente ejemplo ilustrará el teorema anterior.

Ejemplo 3.7. Considérese una población de N elementos, donde K de ellos son éxitos y el resto son fracasos. De esta población se obtendrá una muestra aleatoria de tamaño 2. Calcular la probabilidad de un éxito en la muestra para los siguientes casos

a) $N=12$ y $K=3$.

b) $N=120$ y $K=30$.

c) $N=$1,200 y $K=300$.

d) $N=$12,000 y $K=$3,000.

e) Aproximar la probabilidad de un éxito por medio de la distribución binomial.

En cada inciso, el recorrido es $x=0,1,2$.

a) Obsérvese $X\sim H(N=12, K=3, n=2)$, por lo tanto,

$$f_X(1)=\frac{\binom{3}{1}\binom{9}{1}}{\binom{12}{2}}=0.40909.$$

b) En este ejemplo $X \sim H(N = 120, K = 30, n = 2)$, de esta manera

$$f_X(1) = \frac{\binom{30}{1}\binom{90}{1}}{\binom{120}{2}} = 0.37815.$$

c) Para este caso, $X \sim H(N = 1200, K = 300, n = 2)$, de este modo

$$f_X(1) = \frac{\binom{300}{1}\binom{900}{1}}{\binom{1200}{2}} = 0.37531.$$

d) Obsérvese que $X \sim H(N = 12000, K = 3000, n = 2)$, entonces,

$$f_X(1) = \frac{\binom{3000}{1}\binom{9000}{1}}{\binom{12000}{2}} = 0.37503.$$

e) Por medio de la binomial, $X \sim B(n = 2, p = \frac{3000}{12000} = 0.25)$, así

$$f_X(1) = \binom{2}{1}(0.25)^1(0.75)^{2-1} = 0.375.$$

Se puede observar cómo la probabilidad de un éxito, por medio de la distribución hipergeométrica, se va pareciendo a la probabilidad de un éxito por medio de la distribución binomial, esto ocurre cuando N crece (y K también crece proporcionalmente) y n se queda fija. Dicho de otra manera, entre más grande sea N con respecto a n, la aproximación es más exacta.

Las anteriores probabilidades pueden calcularse por medio de R, las instrucciones son: $dhyper(1, 3, 9, 2)$; $dhyper(1, 30, 90, 2)$; $dhyper(1, 300, 900, 2)$; $dhyper(1, 3000, 9000, 2)$ y $dbinom(1, 2, 0.25)$, respectivamente. Los resultados son: 0.409091; 0.378151; 0.375313; 0.375031 y 0.375, respectivamente.

Una pregunta de interés es ¿a partir de qué valor de N, la distribución binomial es una buena aproximación para la distribución hipergeométrica? Esta pregunta no tiene una respuesta definitiva. En este sentido, algunos autores, de acuerdo a su experiencia, sugieren «reglas» o recomendaciones.

Por ejemplo, Canavos [4], página 112, opina que si $X \sim H(N, K, n)$ y se cumple $\frac{n}{N} \leq 0.1$, entonces, $X \dot{\sim} B\left(n, p = \frac{K}{N}\right)$, donde $\dot{\sim}$ significa se distribuye aproximadamente.

En la recomendación anterior, es importante mencionar que mientras más pequeño sea el cociente $\frac{n}{N}$ mejor será la aproximación. También es importante recalcar que esta regla es más que nada una sugerencia, en ocasiones se pueden tener buenas aproximaciones y no se cumple la regla, o se pueden observar aproximaciones no tan buenas, aunque la regla se cumple.

3.5. Distribución geométrica

En esta sección se tratarán las definiciones de experimento y distribución geométrica, así como sus principales propiedades, entre las más conocidas, la propiedad de pérdida de memoria. También se presentarán algunos ejemplos que ayudarán al estudiante a comprender la importancia de este modelo para aplicarlo en diferentes áreas de la vida práctica.

Definición 3.8. Un *experimento geométrico* consiste en realizar varias pruebas o ensayos idénticos e independientes. Cada prueba tiene dos opciones: «éxito» o «fracaso». La probabilidad de «éxito» en cada prueba es p y de que resulte «fracaso» es $1-p=q$. Las pruebas se realizan hasta lograr por primera vez un «éxito». La variable aleatoria se define como el número de «fracasos», antes de lograr por primera vez el primer «éxito».

Obsérvese que, en general, el recorrido de una variable aleatoria geométrica es $x = 0, 1, 2, \ldots$

Definición 3.9. Una variable aleatoria discreta X tiene *distribución geométrica* con parámetro p, si su función de densidad está dada por

$$f_X(x) = p(1-p)^x I_{\{0,1,2,\ldots\}}(x), \quad \text{donde } 0 < p < 1.$$

Notación. $X \sim$ *Geométrica*(p) significa que la variable aleatoria X se distribuye geométrica con parámetro p.

La gráfica de la función de densidad de la distribución geométrica solo tendrá una sola forma, obsérvese en la siguiente gráfica que decrece en forma geométrica.

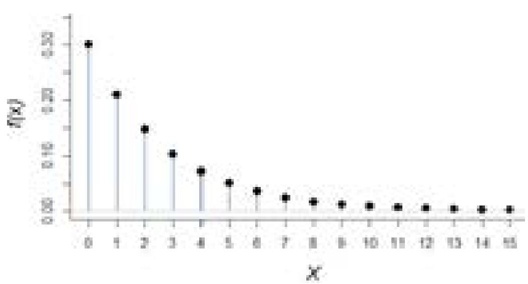


Figura 3.8. *Función de densidad. Geométrica*

Se puede verificar que $f_X(x)$ es una función de densidad.

$$p > 0; \quad 1-p > 0 \text{ e } I_{\{0,1,2,\ldots\}}(x) \geq 0.$$

Entonces,

$$f_X(x) = (1-p)^x p I_{\{0,1,2,\ldots\}}(x) \geq 0.$$

Por otro lado,

$$\begin{aligned}\sum_{x=0}^{\infty}(1-p)^x p &= p\sum_{x=0}^{\infty}(1-p)^x \\ &= p\frac{1}{p} = 1,\end{aligned}$$

donde en la última suma se aplicó el resultado C.6 del apéndice C.

Por lo tanto, $f_X(x)$ es función de densidad.

Teorema 3.8. *Si X se distribuye geométrica con parámetro p, entonces,*

$$\begin{aligned}E[X] &= \frac{q}{p}, \\ V[X] &= \frac{q}{p^2}, \\ m_X(t) &= \frac{p}{1-qe^t}, \ \ si \quad t < -ln(q).\end{aligned}$$

Demostración. Se calculará la función generadora de momentos

$$\begin{aligned}m_X(t) &= E[e^{tX}] \\ &= \sum_{x=0}^{\infty} e^{tx}(1-p)^x p \\ &= p\sum_{x=0}^{\infty}[e^t(1-p)]^x \\ &= p\left[\frac{1}{1-(1-p)e^t}\right], \ \text{ si } \ (1-p)e^t < 1,\end{aligned}$$

donde en la última suma se volvió aplicar el resultado C.6 del apéndice C.

Por lo tanto,

$$m_X(t) = \frac{p}{1-qe^t}, \ \text{ si } \ t < -ln(q).$$

Ahora,

$$\begin{aligned}\frac{d}{dt}m_X(t)\bigg|_{t=0} &= p(-1)(1-qe^t)^{-2}(-qe^t)\bigg|_{t=0} \\ &= p(1-q)^{-2}q \\ &= qp^{-1} \\ &= E[X].\end{aligned}$$

$$\begin{aligned}\frac{d^2}{dt^2}m_X(t)\bigg|_{t=0} &= pq[(1-qe^t)^{-2}e^t + e^t(-2)(1-qe^t)^{-3}(-qe^t)]\bigg|_{t=0} \\ &= pq[(1-q)^{-2} + 2q(1-q)^{-3}] \\ &= pq[p^{-2} + 2qp^{-3}] = 2q^2p^{-2} + qp^{-1} \\ &= E[X^2].\end{aligned}$$

De este modo,

$$\begin{aligned} V[X] &= 2q^2p^{-2} + qp^{-1} - q^2p^{-2} \\ &= \frac{q}{p^2}. \end{aligned}$$ □

Observación. Otra manera de obtener la media de una distribución geométrica, pero usando la función de densidad es

$$\begin{aligned} E[X] &= \sum_{x=0}^{\infty} x f_X(x) \\ &= p(1-p) \sum_{x=1}^{\infty} x(1-p)^{x-1}. \end{aligned}$$

Por otro lado, se sabe que

$$\sum_{i=0}^{\infty} a^i = (1-a)^{-1}, \quad \text{si} \quad |a| < 1.$$

Derivando con respecto a a ambos lados de la anterior igualdad se tiene

$$\sum_{i=1}^{\infty} i a^{i-1} = (1-a)^{-2}.$$

Aplicando este último resultado a la última suma del valor esperado $E[X]$, se tiene

$$\begin{aligned} E[X] &= p(1-p)(1-(1-p))^{-2} \\ &= \frac{1-p}{p}. \end{aligned}$$

Teorema 3.9. *Sea X una variable aleatoria con distribución geométrica con parámetro p, la moda de X es $M = 0$.*

Demostración. Se calculará la derivada de la función de densidad

$$\frac{d}{dx} f_X(x) = p(1-p)^x ln(1-p).$$

Se puede observar que la derivada es negativa, lo que significa que la función de densidad es decreciente. De esta manera podemos concluir que la moda es $M = 0$. □

Teorema 3.10. *Si X se distribuye geométrica con parámetro p, entonces, $P[X \geq x] = q^x$, para todo x entero no negativo.*

Demostración. Suponemos x entero no negativo,

$$\begin{aligned} P[X \geq x] &= 1 - P[X \leq x-1] \\ &= 1 - \sum_{x=o}^{x-1} pq^i \\ &= 1 - p\sum_{i=0}^{x-1} q^i \\ &= q^x, \end{aligned}$$

aplicando a la última suma el resultado C.5 del apéndice C. □

Observación. Si $X \sim$ Geométrica(p), entonces, la función de distribución es $F_X(x) = 1 - q^{x+1}$, para todo valor de x, entero no negativo.

El siguiente resultado se conoce como la *propiedad de pérdida de memoria*, y se cumple para una distribución geométrica. Cabe mencionar que esta propiedad también se cumple para una distribución exponencial, la cual se tratará en el siguiente capítulo.

Teorema 3.11. Propiedad pérdida de memoria. *Sea X una variable aleatoria con distribución geométrica con parámetro p, además sean a y b números enteros no negativos, entonces,*

$$P[X \geq a+b|X \geq a] = P[X \geq b].$$

Demostración.

$$\begin{aligned} P[X \geq a+b|X \geq a] &= \frac{P[(X \geq a+b) \cap (X \geq a)]}{P[X \geq a]} \\ &= \frac{P[X \geq a+b]}{P[X \geq a]} \\ &= \frac{q^{a+b}}{q^a} = q^b = P[X \geq b]. \end{aligned}$$ □

Ejemplo 3.8. Para un determinado seguro, se sabe que la probabilidad de que un asegurado realice una reclamo igual o menor al deducible es de $\frac{1}{6}$. Los reclamos entre los diferentes asegurados son independientes. Calcular la probabilidad de observar por primera vez un reclamo igual o menor al deducible.

a) En el quinto reclamo realizado.
b) Antes del décimo reclamo presentado.
c) Antes del décimo reclamo realizado, si se sabe que hasta el tercer reclamo no ha resultado un reclamo igual o menor al deducible.

Para cada uno de los incisos la variable aleatoria es

X = Número de reclamos mayores al deducible, antes de observar por primera vez un reclamo igual o menor al deducible.

$x = 0, 1, 2, \ldots$

$X \sim$ Geométrica($p = \frac{1}{6}$).

a) La primera probabilidad se calcula de la siguiente manera:

$$P[X = 4] = \frac{1}{6}\left(\frac{5}{6}\right)^4 = 0.08038.$$

b) La siguiente probabilidad se calcula como

$$\begin{aligned} P[X < 9] &= 1 - P[X \geq 9] \\ &= 1 - \left(\frac{5}{6}\right)^9 = 0.80619. \end{aligned}$$

c) A continuación, se calcula la última probabilidad

$$\begin{aligned} P[X < 9|X \geq 3] &= 1 - P[X \geq 9|X \geq 3] \\ &= 1 - P[X \geq 6] \\ &= 1 - \left(\frac{5}{6}\right)^6 = 0.6651. \end{aligned}$$

Las anteriores probabilidades pueden ser calculadas por medio de R, a continuación, se presentarán estos cálculos.

a) $dgeom(4, 1/6)$, dando como resultado 0.08037551.

b) $pgeom(8, 1/6)$, siendo la respuesta, 0.8061933.

c) Después de aplicar la propiedad de pérdida de memoria, se da la instrucción $pgeom(5, 1/6)$, siendo el resultado 0.665102.

Existe otra versión de la distribución geométrica, la cual difiere de la distribución geométrica que se presentó previamente, por la manera como se define la variable aleatoria.

En esta nueva versión, la variable aleatoria se define como X = número de pruebas hasta lograr el primer éxito.

Siendo el recorrido en general, $x = 1, 2, \ldots$

La función de densidad en este caso es

$$f_X(x) = (1-p)^{x-1}pI_{\{1,2,3,\ldots\}}(x).$$

La función generadora de momentos es

$$m_X(t) = \frac{pe^t}{1-qe^t}, \text{ si } t < -ln(q).$$

El valor esperado y la varianza son, respectivamente,

$$E[X] = \frac{1}{p}; V[X] = \frac{q}{p^2}.$$

La moda es $M = 1$.

La propiedad de pérdida de memoria se expresa de la siguiente manera:

$$P[X > a + b | X > a] = P[X > b],$$

donde a y b son enteros no negativos.

Las demostraciones de cada una de estas propiedades, considerando la nueva versión de densidad geométrica, se quedan como ejercicios para el estudiante, esto es, considerando la variable aleatoria como número de pruebas o ensayos hasta lograr el primer éxito.

3.6. Distribución binomial negativa

En esta parte, se presentarán las definiciones de experimento y distribución binomial negativa, se darán las propiedades más relevantes de esta distribución, además se considerarán algunos ejemplos. Cabe mencionar que esta distribución es una generalización de la distribución geométrica.

Definición 3.10. Un *experimento binomial negativo* consiste en realizar varias pruebas idénticas e independientes. Cada prueba tiene dos posibilidades: «éxito» o «fracaso». La probabilidad de «éxito» en cada prueba es p y la de «fracaso» es $1 - p = q$. Las pruebas se realizan hasta lograr por r-ésima vez un «éxito». La variable aleatoria se define como el número de «fracasos» antes de lograr por r-ésima vez un «éxito».

Obsérvese que, en general, el recorrido de la variable aleatoria binomial negativa es $x = 0, 1, 2, \ldots$

Definición 3.11. Una variable aleatoria discreta X tiene *distribución binomial negativa* con parámetros r y p, si su función de densidad está dada por

$$f_X(x) = \binom{r + x - 1}{x} p^r q^x I_{\{0,1,2,\ldots\}}(x),$$

donde $0 < p < 1$ y r es un entero positivo.

Notación. $X \sim BN(r, p)$ significa que la variable aleatoria X tiene una distribución binomial negativa con parámetros r y p.

Observación. La distribución geométrica es un caso particular de la distribución binomial negativa, dicho de otra manera, si $X \sim BN(r = 1, p)$, entonces, $X \sim \textit{Geométrica}(p)$.

En este sentido, se podía presentar primero la distribución binomial negativa y después la geométrica, con la idea de considerar varios resultados de la binomial negativa para la distribución geométrica, aunque la intención de presentar primero la distribución geométrica es para que el lector se dé cuenta de cómo las sumas geométricas están involucradas con esta distribución.

La forma de la gráfica de la función de densidad, para $r > 1$, siempre es sesgada a la derecha.

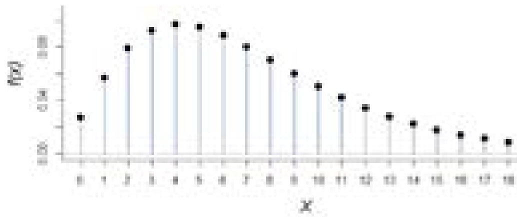


Figura 3.9. *Función de densidad. Binomial negativa*

Para demostrar que $f_X(x)$ es una función de densidad, se debe usar la igualdad C.9 del apéndice C. Se deja como ejercicio para el estudiante realizar esta demostración.

Teorema 3.12. *Si X es una variable aleatoria con distribución binomial negativa, con parámetros r y p, entonces,*

$$E[X] = \frac{r(1-p)}{p} = \frac{rq}{p},$$

$$V[X] = \frac{r(1-p)}{p^2} = \frac{rq}{p^2},$$

$$m_X(t) = \left(\frac{p}{1-qe^t}\right)^r, \quad si \;\; t < -ln(q).$$

Demostración.

$$\begin{aligned} m_X(t) &= E[e^{tX}] \\ &= \sum_{x=0}^{\infty} e^{tx}\binom{r+x-1}{x}p^r q^x \\ &= p^r \sum_{x=0}^{\infty} \binom{r+x-1}{x}(e^t q)^x \\ &= p^r(1-e^t q)^{-r}, \;\; \text{si} \;\; e^t q < 1. \end{aligned}$$

Por lo tanto,

$$m_X(t) = \frac{p^r}{(1-qe^t)^r}, \;\; \text{si} \;\; t < -ln(q). \qquad \square$$

En la anterior demostración, en la última suma se utilizó la igualdad C.9 del apéndice C.

El cálculo de la media y de la varianza se queda como ejercicio para el estudiante.

Teorema 3.13. *Si X es una variable aleatoria con distribución binomial negativa con parámetros r y p, entonces, la moda M es un número entero que se calcula de la siguiente manera:*

$$M = \begin{cases} \left[\dfrac{(r-1)(1-p)}{p}\right] & si\ r > 1, \\ \\ 0 & si\ r = 1, \end{cases}$$

donde $[a]$ es el mayor entero menor o igual a a.

Ejemplo 3.9. Considérese el ejemplo 3.8, donde la probabilidad de que un asegurado realice un reclamo menor o igual al deducible es de $\frac{1}{6}$. Supongamos que se observan los reclamos que realizan los asegurados, uno por uno hasta observar por cuarta vez un reclamo que sea menor o igual al deducible. Calcular la probabilidad de observar por cuarta vez un reclamo menor o igual al deducible.
a) En el quinto reclamo.
b) Antes del séptimo reclamo.

X = Número de reclamos mayores al deducible antes de observar por cuarta vez un reclamo igual o menor al deducible.
$x = 0, 1, 2, \ldots$
$X \sim BN(r = 4, p = \frac{1}{6})$.

$$f_X(x) = \binom{4+x-1}{x}\left(\frac{1}{6}\right)^4\left(\frac{5}{6}\right)^x I_{\{0,1,2,\ldots\}}(x).$$

a) La primera probabilidad se calcula de la siguiente manera:

$$P[X = 1] = \binom{4}{1}\left(\frac{1}{6}\right)^4\left(\frac{5}{6}\right)^1 = 0.00257.$$

b) La siguiente probabilidad se calcula como

$$P[X \leq 2] = \left(\frac{1}{6}\right)^4 + \binom{4}{1}\left(\frac{1}{6}\right)^4\left(\frac{5}{6}\right)^1 + \binom{5}{2}\left(\frac{1}{6}\right)^4\left(\frac{5}{6}\right)^2 = 0.0087.$$

Los cálculos anteriores se pueden realizar por medio de R, de la siguiente manera, respectivamente: $dnbinom(1, 4, 1/6)$ y $pnbinom(2, 4, 1/6)$, siendo las respuestas, respectivamente: 0.002572016 y 0.008701989.

A continuación, se presenta una comparación entre la distribución binomial y la distribución binomial negativa, en cuanto al número de pruebas y número de éxitos.

	Binomial	Binomial negativa
Número de pruebas	n (fijo)	X (aleatorio)
Número de éxitos	X (aleatorio)	r (fijo)

Tabla 3.2. *Diferencias entre la distribución binomial y la distribución binomial negativa*

Por otro lado, la distribución binomial negativa está relacionada con la distribución binomial, el siguiente teorema indica como es esta relación.

Teorema 3.14. *Sea X una variable aleatoria con distribución binomial negativa con parámetros r y p, entonces,*

$$P[X \leq x] = P[Y \geq r],$$

donde Y es una variable aleatoria con distribución binomial con parámetros $n = x + r$ y p.

La demostración de este teorema se deja como ejercicio para el estudiante.

Existe otra versión de la distribución binomial negativa, la cual difiere de la versión que se acaba de presentar, la diferencia es la manera como se define la variable aleatoria.

En la nueva versión, la variable aleatoria se define como $X =$ número de pruebas hasta lograr r éxitos.

El recorrido de esta nueva variable aleatoria es $x = r, r + 1, r + 2, \ldots$

La función de densidad está dada de la siguiente manera:

$$f_X(x) = \binom{x-1}{r-1} p^r q^{x-r} I_{\{r,r+1,r+2,\ldots\}}(x).$$

El valor esperado y la varianza son, respectivamente,

$$E[X] = \frac{r}{p} \quad \text{y} \quad V[X] = \frac{rq}{p^2}.$$

La función generadora de momentos es

$$m_X(t) = \left(\frac{pe^t}{1 - qe^t}\right)^r, \text{ si } t < -ln(q).$$

Las demostraciones de estos resultados son parte de los ejercicios que el estudiante tendrá que realizar.

3.7. Distribución Poisson

En esta sección se presentará la distribución Poisson la cual es muy importante por todas sus aplicaciones, en particular, es un modelo muy usado en diferentes áreas de las ciencias actuariales. Se tratará las propiedades más importantes de esta distribución, también se explicará las características que debe cumplir un fenómeno Poisson, que a diferencia de las otras distribuciones discretas, estas no son fáciles de explicar en términos prácticos.

Definición 3.12. Una variable aleatoria discreta X tiene una *distribución Poisson* con parámetro λ, si su función de densidad está dada por

$$f_X(x) = \frac{\lambda^x}{x!} e^{-\lambda} I_{\{0,1,2,\ldots,\}}(x), \text{ donde } \lambda > 0.$$

Notación. $X \sim Poisson(\lambda)$ significa que la variable aleatoria X tiene una distribución Poisson con parámetro λ.

Observación. La variable aleatoria X = número de éxitos en un intervalo de tiempo, área, volumen o espacio determinado, es candidata a distribuirse Poisson.

La forma de la gráfica de la función de densidad de una distribución Poisson es sesgada a la derecha, no importando el valor del parámetro de λ.

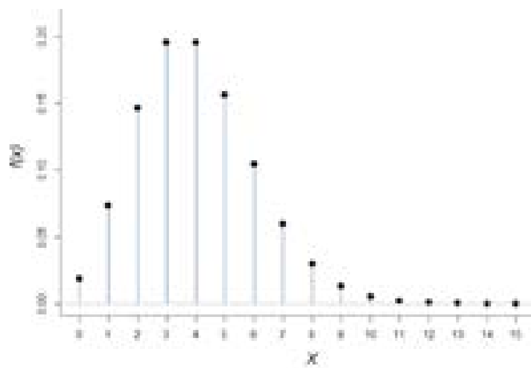


Figura 3.10. *Función de densidad. Poisson*

Se demostrará que $f_X(x)$ es una función de densidad.

$$\lambda > 0;\ x! > 0;\ e^{-\lambda} > 0 \ \text{ y } \ I_{\{0,1,2,\ldots\}}(x) \geq 0.$$

Es así como,

$$f_X(x) = \frac{\lambda^x}{x!} e^{-\lambda} I_{\{0,1,2,\ldots,\}}(x) \geq 0.$$

Por otro lado,

$$\begin{aligned} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} e^{-\lambda} &= e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} \\ &= e^{-\lambda} e^{\lambda} = 1, \end{aligned}$$

aplicando para la última suma, el desarrollo de la serie de Maclaurin para e^x (igualdad C.13 del apéndice C).

Por lo tanto, $f(x)$ es función de densidad.

Algunos ejemplos de variables aleatorias que tienen posibilidad de distribuirse Poisson son el número de accidentes en una semana; el número de llamadas que llegan a un lugar de trabajo en una hora; el número de errores tipográficos por página; el número de defectos por artículo fabricado; el número de ventas en una tienda en un día determinado, y el número de reclamos mensuales en una compañía de seguros.

La distribución Poisson sirve para modelar variables de conteo sobre el tiempo (o un espacio determinado), y no de una muestra aleatoria, como las distribuciones binomial e hipergeométrica. Para tener la seguridad de que se puede utilizar la distribución Poisson en un problema determinado, se deberá cumplir las características de un fenómeno Poisson, las cuales se presentarán más adelante.

Teorema 3.15. *Sea X una variable aleatoria con distribución Poisson con parámetro λ, entonces,*

$$E[X] = \lambda,$$
$$V[X] = \lambda,$$
$$m_X(t) = e^{\lambda(e^t-1)}.$$

Demostración.

$$\begin{aligned}
m_X(t) &= E[e^{tX}] \\
&= \sum_{x=0}^{\infty} e^{tx} \frac{\lambda^x}{x!} e^{-\lambda} \\
&= e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!} \\
&= e^{-\lambda} e^{\lambda e^t} \\
&= e^{\lambda(e^t-1)}.
\end{aligned}$$ □

Se aplicó el desarrollo de Maclaurin para e^x, igualdad C.13, apéndice C. La media y la varianza se quedan como ejercicios para el estudiante.

Ejemplo 3.10. En un proceso de producción de determinadas piezas, se sabe que el número de defectos por pieza tiene distribución Poisson con un promedio de 2.5 defectos por artículo. Calcular la probabilidad de que resulten

a) Menos de tres defectos en un artículo.
b) Dos defectos en dos piezas.
c) Más de dos defectos en dos piezas.

a) Para el cálculo de la primera probabilidad, se define la siguiente variable aleatoria con su correspondiente recorrido:

$X =$ Número de defectos por pieza.

$x = 0, 1, 2, \ldots$

En este caso,
$X \sim Poisson(2.5)$.
Entonces,

$$
\begin{aligned}
P[X < 3] &= f_X(0) + f_X(1) + f_X(2) \\
&= \frac{2.5^0}{0!}e^{-2.5} + \frac{2.5^1}{1!}e^{-2.5} + \frac{2.5^2}{2!}e^{-2.5} \\
&= e^{-2.5}\left(1 + 2.5 + \frac{2.5^2}{2!}\right) = 0.54381.
\end{aligned}
$$

b) Consideremos el siguiente evento $A = \{\text{dos defectos en dos piezas}\}$.

Este evento ocurrirá si y solo si, en una de las dos piezas hay dos defectos y en la otra no lo hay, o un defecto se tendrá en cada pieza. Es importante aclarar que se puede suponer que los defectos entre ambas piezas son independientes.

Definimos las siguientes variables aleatorias: X = número de defectos en la pieza 1 y Y = número de defectos en la pieza 2. De esta manera, la probabilidad del evento A, se calcula como

$$
\begin{aligned}
P[A] &= f_X(0)f_Y(2) + f_X(1)f_Y(1) + f_X(2)f_Y(0) \\
&= \frac{2.5^0}{0!}e^{-2.5}\frac{2.5^2}{2!}e^{-2.5} + \frac{2.5^1}{1!}e^{-2.5}\frac{2.5^1}{1!}e^{-2.5} + \frac{2.5^2}{2!}e^{-2.5}\frac{2.5^0}{0!}e^{-2.5} \\
&= e^{-5}\left\{\frac{2.5^2}{2!} + 2.5^2 + \frac{2.5^2}{2!}\right\} \\
&= e^{-5}\frac{5^2}{2!} = 0.08422.
\end{aligned}
$$

c) Obsérvese que para el cálculo de esta probabilidad se volverán a considerar las dos variables aleatorias X y Y que fueron consideradas en el inciso anterior.

De esta modo,

$P[\text{Más de dos defectos en dos piezas seleccionadas}] =$

$$
\begin{aligned}
&= 1 - \{f_X(0)f_y(0) + f_X(1)f_y(0) + f_X(0)f_y(1) + P[\text{2 defectos en 2 piezas}]\} \\
&= 1 - \left\{e^{-2.5}e^{-2.5} + 2.5e^{-2.5}e^{-2.5} + e^{-2.5}2.5e^{-2.5} + e^{-5}\frac{5^2}{2!}\right\} \\
&= 1 - e^{-5}\left\{1 + 5 + \frac{5^2}{2!}\right\} = 0.87535.
\end{aligned}
$$

La primera probabilidad por medio de R, se calcula como $ppois(2, 2.5)$. siendo el resultado 0.5438131.

A continuación, se presentarán las propiedades generales que debe cumplir un fenómeno Poisson.

Fenómeno Poisson

Sea X una variable aleatoria la cual se define como $X =$ número de sucesos en un intervalo de tiempo (área, volumen o espacio), sea λ el promedio de sucesos en la unidad de tiempo (área, volumen o espacio). Un *fenómeno Poisson* es aquel que cumple las siguientes características:

(I) $P[X = 1$ en un intervalo de longitud $h] = \lambda h + o(h)$,

(II) $P[X > 1]$ en un intervalo de longitud $h] = o(h)$,

(III) El número de sucesos en intervalos de tiempos disjuntos son independientes.

La función $o(h)$ es conocida como *función de orden más pequeño que* h, y cumple la siguiente propiedad:

$$\lim_{h\to 0} \frac{o(h)}{h} = 0.$$

Observación. La función de orden más pequeño que h, $o(h)$, es una función que decrece más rápido que h, cuando h tiende a 0.

Un experimento que conlleva a la realización de un fenómeno Poisson es llamado *Experimento Poisson*.

La interpretación de la primera propiedad de un fenómeno Poisson es la probabilidad de que ocurra un suceso en un intervalo de longitud h, es aproximadamente igual a λh, si h es «pequeño». En el caso de la segunda propiedad, la probabilidad de que ocurran dos o más sucesos en un intervalo de longitud h es aproximadamente cero, si h es «pequeño». Y finalmente, la tercera propiedad dice que el número de sucesos que ocurran en dos intervalos ajenos son independientes entre sí.

El fenómeno Poisson está relacionado con la distribución Poisson.

Teorema 3.16. *Sea X una variable aleatoria la cual se define como $X =$ número de sucesos en un intervalo de tiempo de longitud t (área, volumen o espacio), si la variable X cumple las condiciones de fenómeno Poisson, entonces, X tiene distribución Poisson con media igual a λt.*

La demostración de este teorema se puede encontrar en Mood, Graybill y Boes [21] y Ross [26]. Se sugiere al estudiante, revisar con detalle esta demostración, es una forma de entender de dónde proviene la función de densidad de esta distribución.

Una característica más de un fenómeno Poisson se indica en el anterior teorema, el promedio de sucesos en un intervalo de tiempo de longitud t es λt, esto es, el promedio de sucesos es proporcional al tamaño de longitud del intervalo (área, volumen o espacio).

Los incisos b y c del ejemplo 3.10, se pueden resolver usando la distribución Poisson, pero considerando un nuevo promedio, este promedio es igual a $2\lambda = 2(2.5) = 5$. Dicho de otra manera, si $W =$ número de defectos en dos piezas, entonces, $W \sim Poisson(5)$ y de esta manera el inciso b se resolverá como $P[W = 2]$ y el inciso c como $P[W > 2]$.

El siguiente teorema trata sobre la moda de una distribución Poisson.

Teorema 3.17. *Sea X una variable aleatoria con distribución Poisson con parámetro λ, entonces, la moda M es un valor del recorrido que cumple con la siguiente desigualdad*

$$\lambda - 1 \leq M \leq \lambda.$$

Demostración. Obsérvese

$$\begin{aligned} \frac{f_X(x)}{f_X(x-1)} &= \frac{\dfrac{\lambda^x}{x!} e^{-\lambda}}{\dfrac{\lambda^{x-1}}{(x-1)!} e^{-\lambda}} \\ &= \frac{\lambda}{x}. \end{aligned}$$

Entonces,

$$\begin{aligned} f_X(x-1) &< f_X(x) \quad \text{si } x < \lambda, \\ f_X(x-1) &> f_X(x) \quad \text{si } x > \lambda, \\ f_X(x-1) = f_X(x) &\quad \text{si } x = \lambda, \text{ y } \lambda \quad \text{es un número entero positivo.} \end{aligned}$$

Supongamos que M es la moda.

Si $M > \lambda$, entonces, $f_X(M-1) > f_X(M)$, lo cual no puede ser.

Ahora, si $M < \lambda - 1$, entonces, $M + 1 < \lambda$, lo que implicaría que $f_X(M) < f_X(M+1)$, lo cual tampoco puede ser.

De esta manera,

$$\lambda - 1 \leq M \leq \lambda. \qquad \square$$

Observación. En la anterior demostración se obtuvo

$$\frac{f_X(x)}{f_X(x-1)} = \frac{\lambda}{x}.$$

De esta manera, se logra una fórmula recursiva para calcular probabilidades con la distribución Poisson, esto es,

$$f_X(x) = \frac{\lambda}{x} f_X(x-1).$$

Esta igualdad es una fórmula recursiva para calcular probabilidades de una distribución Poisson y puede utilizarse, como en la distribución binomial, para programar en la computadora el cálculo de probabilidades de una distribución Poisson.

El siguiente resultado trata sobre la convergencia de la distribución binomial a una distribución Poisson, y en principio será útil para aproximar probabilidades de una distribución binomial por medio de una distribución Poisson.

Teorema 3.18. *Si X es una variable aleatoria con distribución binomial, con parámetros n y p, entonces,*

$$\lim_{\substack{n\to\infty \\ np=\lambda}} \binom{n}{x} p^x (1-p)^{n-x} I_{\{0,1,2\ldots,n\}}(x) = \frac{\lambda^x}{x!} e^{-\lambda} I_{\{0,1,2,\ldots\}}(x).$$

Demostración. Obsérvese

$$\lim_{\substack{n\to\infty \\ np=\lambda}} \binom{n}{x} p^x (1-p)^{n-x} I_{\{0,1,2\ldots,n\}}(x)$$

$$\begin{aligned}
&= \lim_{n\to\infty} \frac{n!}{x!(n-x)!} \left(\frac{\lambda}{n}\right)^x \left(1-\frac{\lambda}{n}\right)^{n-x} I_{\{0,1,2,\ldots,n\}}(x) \\
&= \frac{\lambda^x}{x!} \lim_{n\to\infty} \frac{n!}{(n-x)!} \frac{1}{n^x} \left(1-\frac{\lambda}{n}\right)^n \left(1-\frac{\lambda}{n}\right)^{-x} I_{\{0,1,2\ldots,n\}}(x) \\
&= \frac{\lambda^x}{x!} \lim_{n\to\infty} \frac{(n-x+1)\cdots(n-1)n}{n^x} \left(1-\frac{\lambda}{n}\right)^n \left(1-\frac{\lambda}{n}\right)^{-x} I_{\{0,1,2\ldots,n\}}(x) \\
&= \frac{\lambda^x}{x!} \lim_{n\to\infty} \left(1-\frac{x-1}{n}\right)\cdots\left(1-\frac{1}{n}\right)(1)\left(1-\frac{\lambda}{n}\right)^n \left(1-\frac{\lambda}{n}\right)^{-x} I_{\{0,1,\ldots,n\}}(x) \\
&= \frac{\lambda^x}{x!} e^{-\lambda} I_{\{0,1,2\ldots\}}(x).
\end{aligned}$$

□

Obsérvese que se aplicó el resultado límite C.2 del apéndice C.

Este último resultado es útil para aproximar probabilidades de la distribución binomial por medio de la distribución Poisson. De acuerdo al teorema anterior, una aproximación de la binomial por la distribución Poisson será recomendable, en general, cuando el tamaño de la muestra sea «grande» y la probabilidad de éxito sea «pequeña».

La pregunta es ¿cuáles serán los valores apropiados de n y p para lograr una buena aproximación?

Algunos autores dan recomendaciones para lograr una buena aproximación. La siguiente sugerencia se encuentra en Wackerly, *et al* [28], página 132: Si $X \sim B(n,p)$, donde n es «grande» y se cumple $\lambda = np \leq 7$, entonces, $X \dot{\sim} Poisson(\lambda = np)$, donde $\dot{\sim}$ significa que se distribuye aproximadamente.

En la anterior recomendación, entre más grande sea n y más pequeña sea p, la aproximación se va a dar de mejor manera. Queda la duda, a partir de qué valor n es grande, en este sentido, algunos autores consideran que n es grande a partir de 25 o 30, pero de nuevo solo son opiniones.

Ejemplo 3.11. En una línea de producción se sabe que el 1.23 % de artículos son defectuosos. Se escogen 60 artículos al azar de la producción. Calcular la probabilidad de tener 5 artículos defectuosos, por medio de la distribución

a) Poisson.

b) Binomial.

X = Número de artículos defectuosos en la muestra de tamaño 60.
$x = 0, 1, \ldots, 60$.
$X \sim B(60, 0.0123)$.
Obsérvese que $60 > 30$, y $\lambda = np = 60(0.0123) = 0.738 \leq 7$.
a) Por medio de la distribución Poisson

$$P[X = 5] \approx \frac{0.738^5}{5!} e^{-0.738} = 0.00087.$$

b) Por medio de la distribución binomial

$$P[X = 5] = \binom{60}{5} 0.0123^5 (1 - 0.0123)^{55} = 0.00078.$$

Los anteriores cálculos se realizarán por medio de R, con los códigos, $dpois(5, 0.738)$ y $dbinom(5, 60, 0.0123)$, dando como resultados, respectivamente, 0.0008721501 y 0.0007784243.

3.8. Ejemplos diversos

A continuación, se presentarán ejemplos donde se involucran algunas de las distribuciones discretas tratadas en este capítulo.

Ejemplo 3.12. Se lanzan varias veces tres monedas simultáneamente. Calcular la probabilidad de obtener tres águilas o tres soles por segunda vez en el quinto tiro.

X = Número de tiros con resultados disparejos antes de obtener por segunda vez resultados iguales.
$x = 0, 1, 2, \ldots$
$X \sim BN\left(r = 2, p = \frac{1}{4}\right)$.

$$P[X = 3] = \binom{4}{1} \left(\frac{1}{4}\right)^2 \left(\frac{3}{4}\right)^3 = 0.10547.$$

Por medio de R, $dnbinom(3, 2, 1/4)$, siendo la respuesta 0.1054687.

Ejemplo 3.13. Una empresa de manufactura emplea un procedimiento de aceptación o rechazo, en los artículos fabricados antes de sacarlos al mercado. De cada lote de 30 artículos, se toman al azar tres artículos para su revisión, si se encuentra algún elemento defectuoso, el lote se regresa para inspeccionar el 100 % de su contenido. Calcular la probabilidad
a) De aceptar una caja que contiene cinco artículos defectuosos.
b) De regresar una caja que contiene solo un artículo defectuoso.

a) Obsérvese que para este problema
X = Número de artículos defectuosos encontrados en la muestra.
$x = 0, 1, 2, 3.$
$X \sim H(N = 30, K = 5, n = 3).$

$$P[X = 0] = \frac{\binom{5}{0}\binom{25}{3}}{\binom{30}{3}} = 0.5665.$$

b) En este caso
X = Número de artículos defectuosos encontrados en la muestra.
$x = 0, 1.$
$X \sim H(N = 30, K = 1, n = 3).$

$$P[X = 1] = \frac{\binom{1}{1}\binom{29}{2}}{\binom{30}{3}} = 0.1.$$

Estos cálculos por medio de R, se calculan, respectivamente, de la siguiente manera: $dhyper(0, 5, 25, 3)$ y $dhyper(1, 1, 29, 3)$; las respuestas son, respectivamente, 0.5665025 y 0.1.

Obsérvese que las anteriores probabilidades calculadas por medio de la distribución binomial son respectivamente 0.5787 y 0.09344.

Ejemplo 3.14. En un banco, se sabe que el número de créditos personales asignados son, en promedio, siete por día. Se sabe que el número de créditos personales es una variable aleatoria con distribución Poisson. Calcular la probabilidad de que
a) Se otorguen más de cuatro créditos en un día dado.
b) No se otorgue algún crédito en un día dado.
c) En cinco días se otorguen 20 créditos.
d) De observar por lo menos un día en donde no se otorguen créditos, de 30 días seleccionados al azar.

a) Para este problema,
X = Número de créditos que se otorgan en un día.
$x = 0, 1, 2, \ldots$
$X \sim Poisson(\lambda = 7).$

$$\begin{aligned} P[X > 4] &= 1 - P[X \leq 4] \\ &= 1 - e^{-7}\left(\frac{7^0}{0!} + \frac{7^1}{1!} + \frac{7^2}{2!} + \frac{7^3}{3!} + \frac{7^4}{4!}\right) = 0.82701. \end{aligned}$$

b) Es la misma variable definida en el inciso anterior.

$$P[X=0]=\frac{7^0}{0!}e^{-7}=0.00091.$$

c) $X=$ Número de créditos que se otorgan en 5 días.
$x=0,1,2,\ldots$
$X\sim Poisson(\lambda=35)$.

$$P[X=20]=\frac{35^{20}}{20!}e^{-35}=0.00197.$$

d) $X=$ Número de días de los 30 que no se dan créditos.
$x=0,1,\ldots,30$.
$X\sim B(30,p)$.
Obsérvese

$$p=e^{-7}=0.00091.$$

Por lo tanto,

$$X\sim B(30,p=0.00091).$$

Obsérvese que $n=30$ y además, $\lambda=np=30(0.00091)=0.0273\leq 7$. La probabilidad se aproximará por medio de la Poisson con $\lambda=0.0273$.

De esta manera,

$$\begin{aligned}P[X\geq 1]&=1-P[Y=0]\\&=1-e^{-0.0273}=0.02693.\end{aligned}$$

Esta última probabilidad, usando la distribución binomial, es 0.02694.
A continuación, se realizan los cálculos anteriores, usando R:
a) $1-ppois(4,7)$, siendo la respuesta 0.8270084.
b) $dpois(0,7)$, donde la probabilidad es 0.000911882.
c) $dpois(20,35)$, donde el resultado es 0.001972102.
d) $1-dpois(0,0.0273)$, así la respuesta es 0.02693072.

Ejemplo 3.15. En una empresa de manufactura, los artículos producidos se guardan en cajas de 100, de los cuales se sabe que 10 son defectuosos. Antes de sacar las cajas a la venta, de cada caja se obtienen 3 artículos aleatoriamente para revisarlos. Esta revisión se realiza caja por caja y en forma independiente. Calcular la probabilidad de encontrar

a) Por tercera vez una caja que no tenga artículos defectuosos en la muestra, en la quinta caja revisada.
b) Tres cajas sin artículos defectuosos en la muestra, en las cinco primeras cajas revisadas.
c) Una caja con tres artículos defectuosos en la muestra por primera vez, antes de la octava caja revisada, dado que se sabe que hasta la quinta caja revisada, no se encontraron en la muestra, tres artículos defectuosos.

a) $X =$ Número de cajas revisadas con uno o más artículos defectuosos en la muestra, antes de encontrar por tercera vez una caja sin artículos defectuosos.

$x = 0, 1, 2, \ldots$

$X \sim BN(r = 3, p)$, donde p se calcula de la siguiente manera:

$Y =$ Número de artículos defectuosos encontrados en la muestra obtenida de una caja.

$y = 0, 1, 2, 3.$

$Y \sim H(N = 100, K = 10, n = 3)$.

Obsérvese que se cumple $\frac{n}{N} = 0.03 < 0.1$. Por lo que aproximaremos la probabilidad por medio de la binomial, esto es, $Y \dot{\sim} B(n = 3, p = 0.1)$.

Entonces,

$$\begin{aligned} p &= P[Y = 0] \\ &= \binom{3}{0}(0.1)^0(0.9)^3 = 0.729. \end{aligned}$$

Regresando al problema, $X \sim BN(r = 3, p = 0.729)$.

De esta manera,

$$P[X = 2] = \binom{4}{2}(0.729)^3(0.271)^2 = 0.17072.$$

b) $X =$ Número de cajas sin artículos defectuosos en la muestra, entre las primeras 5 cajas revisadas.

$x = 0, 1, \ldots, 5.$

$X \sim B(n = 5, p = 0.729)$.

$$P[X = 3] = \binom{5}{3}(0.729)^3(0.271)^2 = 0.28452.$$

c) $X =$ Número de cajas revisadas con un número de artículos defectuosos en la muestra menor a tres, antes de lograr por primera vez una caja con tres artículos defectuosos en la muestra.

$x = 0, 1, 2, \ldots$

$X \sim$ Geométrica(p).

Donde p se calcula por medio de la variable Y, de la siguiente manera:

$Y =$ Número de artículos defectuosos encontrados en la muestra obtenida de una caja.

$y = 0, 1, 2, 3.$

$Y \sim H(N = 100, K = 10, n = 3)$.

Como en el inciso a, la probabilidad se aproximará por medio de la distribución binomial, esto es, $Y \dot{\sim} B(n = 3, p = 0.1)$.

Entonces,

$$\begin{aligned} p &= P[Y = 3] \\ &= \binom{3}{3}(0.1)^3(0.9)^0 = 0.001. \end{aligned}$$

De esta manera,

$$\begin{aligned} P[X < 7|X \geq 5] &= 1 - P[X \geq 7|X \geq 5] \\ &= 1 - P[X \geq 2] \\ &= 1 - (1 - 0.001)^2 = 0.002. \end{aligned}$$

Se presentarán los anteriores cálculos utilizando R:

a) Para el cálculo de p, el código en R es $dbinom(0, 3, 0.1)$, siendo el resultado 0.729.

El cálculo de la probabilidad de este inciso, es por medio de la instrucción en R $dnbinom(2, 3, 0.729)$, donde la respuesta es 0.1707153.

b) En este caso, el código es $dbinom(3, 5, 0.729)$, y la probabilidad es 0.2845255.

c) El cálculo de p se realiza de la siguiente manera: $dbinom(3, 3, 0.1)$, donde 0.001 es la respuesta.

Después de aplicar la propiedad de pérdida de memoria, se da la instrucción $pgeom(1, 0.001)$, siendo el resultado 0.001999.

Ejemplo 3.16. Supongamos que en promedio dos de cada mil personas cometen una infracción de tránsito en cierto crucero muy transitado. Si se seleccionan al azar cinco mil conductores que pasan por el crucero en un día dado, calcular la probabilidad de que más de 9 conductores, pero menos de 13 cometan una infracción.

a) Por medio de la distribución Poisson.

b) Por medio de la distribución binomial.

La variable aleatoria se define como X =número de conductores que cometen una infracción de los cinco mil seleccionados. Para esta variable,

$x = 0, 1, 2, \ldots, 5000$.

$X \sim B(n = 5000, p = 0.002)$.

Obsérvese que el promedio $np = 5000(0.002) = 10 > 7$.

a) Se realizará el cálculo de probabilidad por medio de la distribución Poisson, esto es, suponemos $X \dot{\sim} Poisson(\lambda = 10)$.

$$P[10 \leq X \leq 12] = e^{-10}\left(\frac{10^{10}}{10!} + \frac{10^{11}}{11!} + \frac{10^{12}}{12!}\right) = 0.33363.$$

b) El cálculo por medio de la distribución binomial.

$$\begin{aligned} P[10 \leq X \leq 12] = & \binom{5000}{10}(0.002)^{10}(0.998)^{4990} \\ & + \binom{5000}{11}(0.002)^{11}(0.998)^{4989} \\ & + \binom{5000}{12}(0.002)^{12}(0.998)^{4988} = 0.33394. \end{aligned}$$

Utilizando R, los cálculos anteriores respectivamente se realizan como $ppois(12,10) - ppois(9,10)$ y $pbinom(12,5000,0.002) - pbinom(9,5000,0.002)$, donde los resultados son respectivamente 0.3336268 y 0.3339419.

Obsérvese que ambas probabilidades resultaron ser muy parecidas, a pesar de que el promedio $np = 10 > 7$.

Ejercicios del capítulo 3

1.- Un examen de opción múltiple está compuesto de 20 preguntas, con cuatro respuestas posibles cada una, de las cuales solamente una es la correcta. Supóngase que uno de los estudiantes que realiza el examen contesta las preguntas al azar. Calcular

  a) La probabilidad de que conteste correctamente 18 preguntas.
  b) La probabilidad de que obtenga una nota superior a 7, si la máxima nota es 10.
  c) El 0.2-ésimo cuantil.
  d) La moda.

2.- Un sistema de seguridad está compuesto por 10 dispositivos, los cuales funcionan en forma independiente. Cada dispositivo detecta un intento de robo con una probabilidad de 0.6. Encontrar la probabilidad de que el sistema de seguridad detecte un intento de robo.

3.- El número de baches en cierta carretera tiene una distribución Poisson con un promedio de 3 baches por km. ¿Cuál es la probabilidad de que un tramo de carretera de 2 km no tenga más de 3 baches?

4.- En un grupo de 100 asegurados, la mitad de ellos han tenido al menos un accidente en un periodo menor a un año, este grupo es llamado grupo A. Se seleccionan al azar 20 personas de este grupo. Calcular

  a) La probabilidad de obtener en la muestra, menos de 17 asegurados del grupo A.
  b) La probabilidad de obtener en la muestra, menos de 5 asegurados del grupo A, dado que se sabe que al menos dos resultaron de este grupo.

c) La moda del número de personas del grupo A.

5.- El número de imperfecciones en determinada tela tiene una distribución de Poisson con una media de 2.5 imperfecciones por metro cuadrado. Encontrar la probabilidad de tener a lo más una imperfección en un pedazo de tela de medio metro cuadrado.

6.- Un jugador de futbol profesional realizará una serie de tiros penal. La probabilidad de que un tiro penal sea exitoso es de 0.7. Suponiendo que los tiros los realice en forma independiente. Calcular la probabilidad de que el jugador

a) Logre su primer tiro penal exitoso, en el cuarto tiro realizado.
b) No tenga tiros exitosos de 10 tiros que realiza.
c) Logre el segundo tiro exitoso en el quinto tiro realizado.
d) Logre su tercer tiro exitoso, después del quinto intento realizado.
e) Logre su primer tiro exitoso antes del décimo intento, dado que se sabe que hasta el quinto intento realizado no había logrado un tiro exitoso.

7.- Un aparato está conformado con 6 componentes idénticos, donde cada uno tiene una probabilidad de 0.09 de fallar en menos de 1,000 horas. El aparato funciona si más de 4 componentes trabajan en forma adecuada. Si los componentes operan en forma independientemente, encontrar

a) La probabilidad de que exactamente tres componentes funcionen más de 1,000 horas.
b) La probabilidad de que el sistema funcione por más de 1,000 horas.
c) La media y varianza del número de componentes que funcionan más de 1,000 horas.

8.- Sea X una variable aleatoria con función de densidad dada por $f_X(x) = (1-p)^{x-1}pI_{\{1,2,3,\ldots\}}(x), 0 < p < 1$ (segunda versión de la distribución geométrica).

a) Obtener la función de distribución para todo valor del recorrido.
b) Demostrar que $P[X > a] = q^a$, donde a es un entero positivo.
c) Demostrar que $P[X > a + b|X > a] = P[X > b]$, donde a y b son números enteros positivos.
d) Obtener la función generadora de momentos de X.
e) Obtener la media y la varianza de X.

9.- Sea $X \sim Poisson(\lambda)$, demostrar que $E[X] = V[X] = \lambda$.

10.- Terminar la demostración del teorema 3.5, esto es, calcular la varianza de una variable aleatoria X la cual tiene distribución hipergeométrica.

11.- Una persona tiene 14 opciones de invertir su dinero, de las cuales 5 son de rendimiento fijo y las demás de rendimiento variable. La persona va a seleccionar 5 opciones de inversión para formar un portafolio. Supongamos que la selección la realiza aleatoriamente. Calcular

a) La probabilidad de observar todas sus inversiones con rendimiento variable.
b) La probabilidad de observar menos de 4 inversiones con rendimiento fijo.
c) La media y la desviación estándar del número de inversiones con rendimiento fijo.
d) La moda del número de inversiones con rendimiento fijo.

12.- Para una determinada acción financiera, se sabe que en un periodo, la acción sube con una probabilidad del 0.2 y baja con una probabilidad del 0.8. Se supone independencia entre los periodos. Encontrar la probabilidad de que

a) El primer periodo que sube la acción sea en el quinto periodo.
b) El primer periodo que sube la acción sea después del quinto periodo.
c) El tercer periodo que sube, sea en el quinto periodo.
d) El primer periodo que sube sea antes del décimo periodo, dado que se sabe que hasta el quinto periodo, la acción no había subido.

13.- Encontrar la función generadora de probabilidades de la distribución geométrica.

14.- El número de errores tipográficos encontrados en una página tiene una distribución de Poisson con una media de tres errores por página. Calcular la probabilidad de encontrar

a) En una página menos de 3 errores tipográficos.
b) Diez errores en dos páginas.
c) Por lo menos un error en una página.
d) De tener 15 páginas con algún error tipográfico, de 20 páginas seleccionadas al azar.

15.- Una fábrica de chocolates prepara para su venta cajas de 20 chocolates, donde 5 de ellos son amargos. Cada caja es revisada, se seleccionan al azar 3 chocolates para contar el número de chocolates amargos. Calcular la probabilidad de observar

a) Una caja con tres chocolates amargos en la muestra, dado que se sabe que al menos un chocolate amargo resultó en la muestra.
b) La primera caja con todos los chocolates amargos en la muestra, después de la octava caja revisada.
c) Al menos dos cajas con todos los chocolates amargos en la muestra, en las primeras 10 cajas revisadas.

d) La segunda caja con ningún chocolate amargo en la muestra, sea después de la quinta caja revisada.

16.- Encontrar la función generadora de probabilidades de la distribución binomial negativa.

17.- En un seguro de gastos médicos mayores, un asegurado tiene 12 opciones para seleccionar un hospital, de los cuales seis son de nivel 1, cuatro de nivel 2 y dos de nivel 3. Un asegurado en su póliza debe de especificar tres hospitales. Si el asegurado realiza su selección aleatoriamente, calcular la probabilidad de que los tres hospitales sean de nivel 1, si

a) La muestra es sin reemplazo.
b) La muestra es con reemplazo.

18.- Cierta maceta de cerámica es fabricada en hornos de altas temperaturas. Se sabe que el 20 % de las piezas fabricadas tienen un desperfecto importante. Un inspector revisa varias macetas recién fabricadas en forma independiente. Calcular la probabilidad de que

a) La primera maceta con problemas importantes de fabricación sea la cuarta maceta revisada.
b) La cuarta maceta con desperfectos importantes sea encontrada después de revisar la séptima maceta.
c) La primera maceta con problemas importantes se encuentre después de revisar la octava maceta, dado que hasta la tercera maceta revisada no se había encontrado una maceta con problemas.
d) Se encuentren 4 macetas con desperfectos importantes de 10 que se revisaron en un día.

19.- Encontrar la función generadora de probabilidades de la distribución binomial.

20.- En un banco hay dos módulos de atención para sus clientes, el módulo A y el módulo B. El número de clientes que llegan al módulo A y módulo B tiene una distribución de Poisson con medias de dos y tres clientes por hora, respectivamente. El número de clientes que llegan a ambos módulos son independientes. Calcular la probabilidad de que

a) Cuatro clientes sean atendidos durante un periodo de media hora.
b) Menos de 3 clientes sean atendidos durante una hora.

21.- El número de personas que utilizan un puente colgante tiene una distribución Poisson con un promedio de cuatro personas por un periodo de cinco minutos. Un número excesivo de personas que utilizan al mismo tiempo el puente puede ocasionar una sensación de pánico para las personas que en ese momento utilizan el puente. Encontrar la probabilidad de que el número de personas que usan el puente colgante durante un periodo de 1 minuto exceda a cuatro personas.

22.- Considerando el ejercicio 21, supóngase que se observa el número de personas que utilizan el puente durante 10 intervalos de un minuto (no traslapados), obteniendo así, 10 observaciones independientes $X_1, X_2, \dots, X_{10}$ del número de personas que usan el puente por minuto. Encontrar la probabilidad de $X_i > 4$ en al menos uno de los 10 intervalos de un minuto.

23.- En cierta población, la probabilidad de que un crédito proporcionado por un banco no sea pagado por el prestatario es del 10 %. Se seleccionan en forma aleatoria créditos otorgados por el banco. Calcular la probabilidad de encontrar

a) Más de tres créditos no pagados, de 40 créditos seleccionados.
b) El primer crédito no pagado, después del quinto crédito seleccionado.
c) El cuarto crédito no pagado, después del octavo crédito seleccionado.

24.- Contestar las siguientes preguntas:

a) Si X es una variable aleatoria con una distribución Poisson que satisface $P[X = 0] = 2P[X = 2]$, ¿cuál es el $E[X]$?
b) Si X se distribuye binomial con parámetros n y p y, además, $E[X] = 10$ y $V[X] = 5$. Encontrar n y p.
c) Si X tiene una distribución Poisson y $P[X \geq 1] = \frac{1}{3}$, ¿cuál es el valor de λ?
d) Suponga que X tiene una distribución Bernoulli con parámetro p. ¿Para qué valores de p la varianza se maximiza?
e) Si X es una variable aleatoria con una función generadora de momentos e^{e^t-1}, sin derivar esta función, ¿cuál es el $E[X]$? Justificar la respuesta.
f) Si X tiene una distribución hipergeométrica con parámetros N, K y $n = 3$, y se sabe que el valor esperado y la varianza son igual a 1.5 y $\frac{7}{12}$, respectivamente, encontrar N y K.

25.- Encontrar la mediana para:

a) $X \sim B(n = 4, p = 0.5)$.
b) $X \sim B(n = 5, p = 0.5)$.
c) $X \sim Poisson(\lambda = 2)$.
d) $X \sim$ Geométrica$(p = 0.3)$.
e) $X \sim H(N = 10, K = 5, n = 3)$.

26.- Encontrar la moda para cada uno de los casos del ejercicio 25.

27.- Sea X una variable aleatoria con función de densidad dada por $f_X(x) = \binom{x-1}{r-1} p^r q^{x-r} I_{\{r,r+1,r+2,\dots\}}(x)$.

a) Demostrar que $f_X(x)$ es función de densidad.
b) Encontrar la función generadora de momentos de X.
c) Encontrar la media y la varianza de X.

28.- El número de artículos que vende un comerciante en un día es una variable aleatoria con distribución Poisson de media igual a 4.

a) ¿Cuántos artículos el comerciante debería tener en existencia para estar un 95 % seguro de que tendrá los suficientes artículos para los siguientes 25 días?
b) De 25 días, ¿cuál es la probabilidad de tener por lo menos un día sin artículos vendidos?
c) De 25 días, ¿cuál es el valor esperado de números de días, que no vende artículos?

29.- En una población de 100 ciudadanos, se sabe que el 50 % está a favor con las políticas de su alcalde. Una muestra aleatoria de tamaño n es obtenida de esta población. Calcular la probabilidad de que la mayoría de los ciudadanos seleccionados estén a favor con las políticas de su alcalde.

a) Si $n = 10$.
b) Si $n = 4$.

30.- Una compañía de seguros encuentra que el 3 % de la población tiene un accidente automovilístico más alto que el deducible en un año. ¿Cuál es la probabilidad de que la compañía deba de pagar a más de 5 de los 100 asegurados contra accidentes de automóvil en un año determinado?

31.- Considerando el ejercicio 30, calcular

a) La moda.
b) El 0.1-ésimo cuantil.
c) La mediana.

32.- Si X se distribuye Poisson y $P[X = 2] = P[X = 3]$, calcular $P[X \geq 2]$.

33.- Si X se distribuye $Poisson(\lambda = 2)$. Demostrar que $E[|X - 2|] = \frac{4\sigma_X^2}{e^2}$.

34.- El número de asegurados hospitalizados en cierta ciudad tiene una distribución Poisson con un promedio de una hospitalización por hora. Calcular la probabilidad de que

a) Ocurran más de 3 hospitalizaciones en una hora, dado que se sabe que al menos una hospitalización ocurrió.
b) Ocurran más de 5 hospitalizaciones en un lapso de 7 horas.
c) Transcurran más de tres horas entre dos hospitalizaciones.

35.- La fabricación de cierto mueble difícilmente se puede lograr sin desperfecto alguno. Se sabe que el número de desperfectos en un mueble tiene distribución Poisson con un promedio de 5 defectos por mueble.

a) Calcular la probabilidad de encontrar más de 5 desperfectos en un mueble.
b) Dado que se sabe que un mueble tiene al menos dos desperfectos, calcular la probabilidad de que tenga menos de 5 desperfectos.
c) Se revisa mueble por mueble. Calcular la probabilidad de encontrar el sexto mueble con menos de tres desperfectos, después del décimo mueble revisado.
d) Se revisa mueble por mueble, hasta el octavo mueble revisado, no se ha observado uno con menos de tres desperfectos, calcular la probabilidad de encontrar antes del décimo segundo mueble revisado el primer mueble con menos de tres desperfectos.

36.- El número de siniestros reportados en una oficina de seguros tiene una distribución Poisson con un promedio de 5 reportes por hora. Se sabe que un exceso de reportes en una hora puede causar problemas de atención a sus clientes. Calcular la probabilidad de

a) Observar más de 4 reportes en una hora de trabajo.
b) Observar al menos una hora con más de 4 reportes, en las siguientes 7 horas de trabajo.
c) Observar la segunda hora de trabajo con más de 4 reportes, antes de la quinta hora de trabajo.

37.- Si la función generadora de momentos de una variable aleatoria X es $m_X(t) = 0.0081\left(1 - 0.7e^t\right)^{-4}$. Encontrar $P[X = 2 \text{ ó } 3]$.

38.- Sea X el número de sucesos exitosos de n repeticiones independientes de un experimento aleatorio, siendo la probabilidad de éxito $p = \frac{1}{3}$. Determinar el valor más pequeño de n, que cumple con $P[X \geq 1] \geq 0.80$.

39.- Sea X una variable aleatoria con distribución $B(5, p)$ y Y otra variable aleatoria con distribución $B(6, p)$. Si $P[X \geq 1] = \frac{5}{16}$, calcular $P[Y \geq 1]$.

40.- Sea $f_X(x)$ una función de densidad para enteros no negativos, tal que $f_X(x) = \frac{7}{x} f_X(x-1)$, $x = 1, 2, 3, \ldots$ ¿Cómo se distribuye X? Justificar la respuesta.

41.- Sea X una variable aleatoria con distribución Bernoulli con probabilidad de éxito igual a p, si la media es dos veces la varianza. Calcular la probabilidad de fracaso.

42.- Dos dados son lanzados 8 veces. ¿Cuál es la probabilidad de que menos de 4 veces tengan resultados disparejos?

43.- Considerando la variable aleatoria del ejemplo 42, calcular

a) La media y la varianza.
b) La moda.

c) La mediana.
d) El 0.3-ésimo cuantil.

44.- Una moneda es lanzada 6 veces. ¿Cuál es la probabilidad de que el número de águilas exceda al número de soles?

45.- Sea $X \sim UD\{1, 2, \ldots, N\}$, si la varianza es $\frac{5}{6}$ la media. Encontrar N.

46.- El número de reclamaciones por gastos médicos por día en una compañía de seguros, tiene una distribución Poisson con un promedio de 10 reclamaciones por día. Calcular la probabilidad de

a) Tener menos de 8 reclamaciones en un día, dado que se sabe que al menos se han realizado 4 reclamaciones.
b) Observar el quinto día sin reclamaciones después del octavo día observado.

47.- Una familia tiene siete hijos. Suponiendo que la probabilidad de tener una niña en cada nacimiento es de 0.52 y que los siete nacimientos son independientes, ¿cuál es la probabilidad de que tengan al menos una niña, dado que se sabe que ellos tuvieron al menos dos niños?

48.- Suponga que X es una variable aleatoria con distribución binomial basada en n pruebas independientes, con una probabilidad de éxito p en cada prueba. Si $P[X = n] = 0.004096$ y $P[X = n - 1] = 0.006144n$, ¿cuál es el valor de p?, ¿cuál es el valor de n?

49.- Un dado es lanzado hasta que un número impar aparece. Dado que se sabe que el primer número impar apareció en un lanzamiento par, ¿cuál es la probabilidad de que el número impar sea en el cuarto lanzamiento?

50.- Un sistema de inspección para los artículos fabricados en una empresa, consiste en la revisión de estos por medio de varios inspectores. Se sabe que la probabilidad de que un inspector detecte un artículo defectuoso como artículo no exitoso es de 0.8. Supongamos que n inspectores revisan en forma independiente un artículo defectuoso.

a) Si $n = 5$, ¿cuál es la probabilidad de que todos los inspectores detecten el artículo como defectuoso?, ¿cuál es la probabilidad de que al menos un inspector detecte el artículo como defectuoso?
b) ¿Cuál debe ser el valor de n para que la probabilidad de detectar el artículo como defectuoso sea de 0.9999?

51.- En determinada ciudad existen tres carreteras. El número de accidentes por semana tiene distribución Poisson, con medias respectivamente de 4, 5 y 2 para las carreteras I, II y III. ¿Cuál es la probabilidad de que ocurran exactamente dos accidentes en una semana?

52.- Si un dado es lanzado repetidamente, dar una expresión de la probabilidad de que el tercer número 2 ocurra después del n-ésimo lanzamiento.

53.- Considerando la variable aleatoria del ejercicio, calcular 52,

a) La media y varianza.
b) La moda.
c) El 0.12-ésimo cuantil.

54.- Sea X una variable aleatoria con distribución geométrica con $p = 0.2$. ¿Cuál es la probabilidad de que $X \geq 2$, dado que se sabe que $X \leq 5$?

55.- Sea X el número de éxitos en 10 pruebas repetidas en forma idéntica e independiente, y p la probabilidad de éxito en cada una de las pruebas. Si el valor esperado y la desviación estándar de la variable X son iguales, ¿cuál es el valor esperado de X?, ¿cuál es el valor de la varianza de X?

56.- Sea X una variable aleatoria con distribución Poisson con media λ. Si $P[X = 1|X \leq 2] = 0.1$, ¿cuál es el valor de λ?

57.- Encontrar la función generadora de probabilidades de una distribución Poisson con parámetro λ.

58.- Sea X una variable aleatoria con distribución binomial con parámetros n y p, demostrar $P[X \text{ es par}] = \frac{1}{2}\{1 + (q - p)^n\}$.

59.- Sea X una variable aleatoria con distribución $Poisson(\lambda)$, demostrar $P[X\text{es impar}] = \frac{1}{2}\{1 - e^{-2\lambda}\}$.

60.- Sea X una variable aleatoria con distribución binomial negativa con parámetros r y p, y que cumple con $P[X = 0] = 0.004096$ y $P[X = 1] = 0.024576(1 - p)$. Encontrar r y p.

61.- Un dado es lanzado varias veces hasta lograr por primera vez un número menor a 3. Se definen los siguientes eventos: $A =$ {se logre por primera vez, un número menor a 3 en un lanzamiento múltiplo a 4} y $B =$ {se logre por primera vez, un número menor a 3 en un lanzamiento múltiplo a 5}. Encontrar la probabilidad de que resulte

a) A y B a la vez.
b) A o B.

62.- Sea X una variable aleatoria con distribución *Geométrica*(p), demostrar que se cumple la siguiente igualdad: $P[X = n + k|X \geq n] = P[X = k]$.

63.- Sea X una variable aleatoria con distribución binomial negativa con parámetros r y p, demostrar la siguiente fórmula recursiva para calcular probabilidades $f_X(x) = (1 - p)\left(1 + \frac{r-1}{x}\right) f_X(x - 1)$.

64.- Sea X una variable aleatoria con distribución $H(N = 10, K = 4, n = 4)$. Encontrar la función de densidad truncada en 0.

65.- Considerando la función de densidad encontrada en el ejercicio 64, encontrar la media y la varianza.

66.- Sea X una variable aleatoria con distribución $BN(r, p)$. Encontrar la función de densidad truncada en el conjunto $\{5, 6, \ldots\}$.

67.- Encontrar el valor esperado y la varianza de X con la función de densidad encontrada en 66, considerando $r = 2$ y $p = 0.5$.

68.- El número de créditos que una sucursal bancaria otorga al día, se distribuye Poisson con medias, respectivamente, de 3, 2 y 4, para los créditos de autos, hipotecarios y personales. Calcular

a) La probabilidad de tener menos de 2 créditos al día.
b) La probabilidad de tener menos de 5 créditos hipotecarios en un día, dado que al menos 2 créditos de este tipo han sido otorgados.
c) El 0.005-ésimo cuantil de la variable aleatoria número de créditos otorgados por el banco en un día.

69.- Una universidad está formada por 220 académicos de diferentes áreas, donde se sabe que 10 de ellos tienen conocimientos en probabilidad y estadística. Cada mes se celebra una reunión del comité de análisis de información académica, para esta reunión se seleccionan aleatoriamente 10 académicos de esta universidad. Calcular

a) La probabilidad de que asista a la siguiente reunión por lo menos un académico con conocimientos en probabilidad y estadística.
b) La media y la varianza del número de académicos con conocimiento en probabilidad y estadística que asisten a la reunión.
c) La probabilidad de observar por primera vez una reunión sin académicos en conocimientos en probabilidad y estadística, después de la quinta reunión realizada.

70.- El total reclamado a una compañía de seguros en un mes por un siniestro es una variable aleatoria X con función de densidad $f_X(x) = 0.2e^{-0.2x}I_{(0,\infty)}(x)$. Para la compañía es preocupante si el total reclamado en un mes excede de 6. Para 10 meses seleccionados al azar, calcular la probabilidad de tener más de 3 meses en una situación preocupante.

71.- En una acción financiera, el retorno de inversión mensual, en porcentaje, es una variable aleatoria X con función de densidad dada por $f_X(x) = \frac{2}{9}(x + 1)I_{(-1,2)}(x)$. Se observa mes con mes el retorno de inversión, suponiendo independencia entre los meses, calcular la probabilidad de observar por primera vez un mes con retorno inferior a -0.5, después del décimo mes.

Distribuciones continuas

Capítulo

Introducción

En este capítulo se estudiarán algunas de las distribuciones continuas más importantes que existen en la literatura de probabilidad y estadística. La importancia de estas distribuciones estriba, principalmente, en sus aplicaciones en problemas de diversas áreas como la industria, las ciencias sociales, las ciencias actuariales, problemas de economía, problemas demográficos, por mencionar algunas áreas de aplicación.

Para cada modelo continuo, se definirá la función de densidad, la media, la varianza y, si es el caso, la función generadora de momentos, además, en general, se presentarán las principales propiedades de cada distribución.

Es importante señalar que la búsqueda de un modelo particular para una variable aleatoria continua no es un problema de probabilidad, como es el caso de la mayoría de las variables aleatorias discretas, sino es un problema de estadística, donde se necesitan datos específicos de la variable aleatoria, para inferir el modelo continuo; siendo más específicos se tiene que aplicar métodos de bondad de ajuste. Por lo que no es propósito de este capítulo, investigar el modelo continuo que le corresponde a una variable aleatoria particular.

Antes de empezar a definir las distribuciones continuas más importantes, se darán algunos conceptos relevantes sobre el tipo de parámetros que una distribución continua puede tener.

Definición 4.1. Se dice que un parámetro α de una distribución es de *escala*, si los valores del parámetro influyen en la dispersión de la distribución, en lo particular, en la dispersión de la función de densidad.

Definición 4.2. Se dice que un parámetro β de una distribución es de *forma*, si para valores diferentes del parámetro, es posible que la forma de la gráfica de la función de densidad cambie.

Definición 4.3. Se dice que un parámetro θ de una distribución es de *localización*, si para valores diferentes del parámetro, la gráfica de la función de densidad se desplaza hacia la izquierda o a la derecha.

4.1. Distribución uniforme continua

En esta sección se tratará con la definición de una distribución uniforme continua o llamada simplemente distribución uniforme, además, se presentarán sus principales propiedades, aplicaciones y algunos ejemplos.

Definición 4.4. Una variable aleatoria continua X tiene *distribución uniforme* (*distribución uniforme continua*) sobre el intervalo (a, b) (con parámetros a y b) si su función de densidad está dada por:

$$f_X(x) = \frac{1}{b-a} I_{(a,b)}(x), \quad \text{donde} \quad a < b.$$

Los parámetros de la distribución uniforme no son ni de escala, ni de forma y tampoco de localización.

Notación. $X \sim U(a, b)$ significa que la variable aleatoria X tiene distribución uniforme sobre el intervalo (a, b).

La gráfica de la función de densidad de una distribución uniforme es

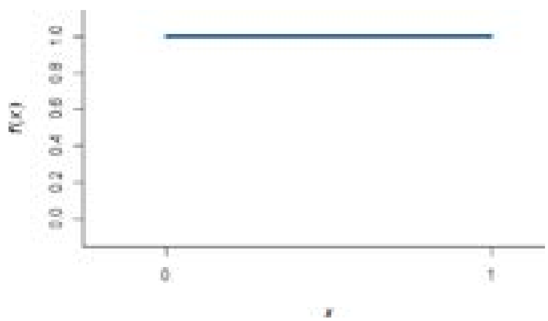


Figura 4.1. *Función de densidad. Uniforme sobre el intervalo (0,1)*

Obsérvese que $f_X(x)$ es función de densidad.

$$\frac{1}{b-a} > 0 \text{ e } I_{(a,b)}(x) \geq 0.$$

Por lo tanto,

$$f_X(x) = \frac{1}{b-a} I_{(a,b)}(x) \geq 0.$$

Por otro lado,

$$\begin{aligned} \int_a^b f_X(x)dx &= \int_a^b \frac{1}{b-a} dx \\ &= \frac{1}{b-a} \int_a^b dx = 1. \end{aligned}$$

Por lo tanto, $f_X(x)$ es función de densidad.

Una aplicación de suma importancia de la distribución uniforme se presenta en la simulación de valores de variables aleatorias. De hecho, todo algoritmo de simulación estocástica (simulación de Monte Carlo), se basa en la simulación de valores de una variable aleatoria con distribución $U(0,1)$.

Teorema 4.1. *Si X es una variable aleatoria con distribución uniforme sobre el intervalo (a,b), entonces,*

$$E[X] = \frac{a+b}{2},$$

$$V[X] = \frac{(b-a)^2}{12},$$

$$m_X(t) = \frac{e^{bt}-e^{at}}{t(b-a)} I_{\{t|t\neq 0\}}(t) + I_{\{0\}}(t).$$

Demostración. Primero la función generadora de momentos

$$\begin{aligned} E[e^{tX}] &= \int_a^b \left(\frac{1}{b-a}\right) e^{tx} dx \\ &= \left.\frac{e^{tx}}{(b-a)t}\right|_a^b, \text{ si } t \neq 0. \end{aligned}$$

Entonces, para $t \neq 0$,

$$E[e^{tX}] = \frac{e^{bt}-e^{at}}{(b-a)t}.$$

Se calculará el límite de esta función cuando t tiende a 0.

$$\begin{aligned} \lim_{t\to 0} E[e^{tX}] &= \lim_{t\to 0} \frac{e^{bt}-e^{at}}{t(b-a)} \\ &= \lim_{t\to 0} \frac{be^{bt}-ae^{at}}{b-a} = 1. \end{aligned}$$

Por lo tanto,

$$m_X(t) = \frac{e^{bt}-e^{at}}{t(b-a)} I_{\{t|t\neq 0\}}(t) + I_{\{0\}}(t).$$

Por otro lado,

$$\begin{aligned} E[X] &= \int_a^b x \frac{1}{b-a} dx \\ &= \left.\frac{1}{b-a}\frac{x^2}{2}\right|_a^b \\ &= \frac{1}{b-a}\left(\frac{b^2-a^2}{2}\right) \\ &= \frac{b+a}{2}. \end{aligned}$$

$$
\begin{aligned}
E[X^2] &= \int_a^b x^2 \frac{1}{b-a} dx \\
&= \frac{1}{b-a} \frac{x^3}{3}\bigg|_a^b \\
&= \frac{1}{b-a}\left(\frac{b^3-a^3}{3}\right) \\
&= \frac{a^2+ab+b^2}{3}.
\end{aligned}
$$

De esta manera,

$$
\begin{aligned}
V[X] &= \frac{a^2+ab+b^2}{3} - \frac{(b+a)^2}{4} \\
&= \frac{4b^2+4ab+4a^2-3a^2-6ab-3b^2}{12} \\
&= \frac{(b-a)^2}{12}.
\end{aligned}
$$

□

Para distribuciones uniformes, no es práctico calcular los momentos por medio de la función generadora de momentos, se sugiere para esta distribución, hacerlo por medio de la función de densidad.

Observación. Para un valor de x del recorrido, la función de distribución $F_X(x)$ se puede calcular de la siguiente forma:

$$
\begin{aligned}
F_X(x) &= \int_a^x \frac{1}{b-a} dt \\
&= \frac{x-a}{b-a}.
\end{aligned}
$$

De esta manera, la función de distribución para todo número real se puede expresar de la siguiente manera:

$$F_X(x) = \frac{x-a}{b-a} I_{[a,b]}(x) + I_{(b,\infty)}(x).$$

Observación. Consideremos a c y d, tal que $c < d$, además, ambos se encuentran dentro del intervalo (a, b), entonces, $P[c < X < d]$ se puede calcular de diferentes formas,

1) $\displaystyle\int_c^d \frac{1}{b-a} dx = \frac{d-c}{b-a}.$

2) Por medio de la función de distribución, esto es, $F_X(d) - F_X(c)$.

3) Como el área por debajo de la gráfica de la función de densidad entre c y d, en este caso es el área del rectángulo con dimensiones $\dfrac{1}{b-a}$ y $d-c$, siendo el área igual a $\dfrac{1}{b-a}(d-c)$.

De las tres formas, la más práctica es la tercera.

Ejemplo 4.1. Se sabe que un autobús llegará de forma aleatoria en un lapso de entre 0 y 15 minutos. Calcular
a) La probabilidad de que el autobús llegue después de 10 minutos.
b) La probabilidad de que el autobús llegue después de 10 minutos, dado que se sabe que a los 5 minutos no llegó.
c) La media y desviación estándar del momento en que llegará el autobús.

$X =$ Momento en el que llega el autobús.
$X \sim U(0, 15)$.
La función de densidad es $f_X(x) = \frac{1}{15} I_{(0,15)}(x)$.
a) La primera probabilidad

$$P[X > 10] = 5\left(\frac{1}{15}\right) = \frac{1}{3}.$$

b) La segunda probabilidad

$$\begin{aligned} P[X > 10 | X > 5] &= \frac{P[(X > 10) \cap (X > 5)]}{P[X > 5]} \\ &= \frac{PX > 10]}{P[X > 5]} = \frac{1}{2}. \end{aligned}$$

c) La media y desviación estándar respectivamente $E[X] = \frac{15+0}{2} = 7.5$ y $\sigma = \sqrt{\frac{(15-0)^2}{12}} = 4.33333$.

La primera probabilidad por medio de R es usando la siguiente instrucción $1 - punif(10, 0, 15)$, siendo el resultado 0.3333333.

4.2. Distribución gamma

En esta sección se tratará la familia de distribuciones gamma, la cual tiene muchas aplicaciones en diversas áreas. Se dará la definición de esta distribución y se estudiarán las propiedades más importantes de esta familia.

Definición 4.5. Una variable aleatoria continua X tiene *distribución gamma* con parámetros r y λ si su función de densidad está dada por

$$f_X(x) = \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x} I_{(0,\infty)}(x), \quad \text{donde} \quad r > 0, \lambda > 0 \quad \text{y}$$

$$\Gamma(r) = \int_0^\infty x^{r-1} e^{-x} dx.$$

La función $\Gamma(r)$ es conocida como la *función gamma de Euler* o simplemente *función gamma* y tiene las siguientes propiedades:

1) $\Gamma(r) = (r-1)\Gamma(r-1)$.

2) $\Gamma(n) = (n-1)!$, si $n \in \mathbb{N}$.

3) $\Gamma(1/2) = \sqrt{\pi}$.

En el caso de la distribución gamma, el parámetro r es de forma y el parámetro λ es de escala, véase las siguientes dos gráficas.

Se puede generalizar la función de densidad gamma, introduciendo un parámetro de localización, esto es,

$$f_X(x) = \frac{\lambda^r}{\Gamma(r)}(x-\theta)^{r-1}e^{-\lambda(x-\theta)}I_{(\theta,\infty)}(x).$$

Esta versión de función de densidad gamma no será utilizada, a menos que se indique.

Notación. $X \sim Gamma(r, \lambda)$ denotará que la variable aleatoria X tiene una distribución gamma con parámetros r y λ.

La función de densidad gamma tiene diferentes formas, estas dependen de los valores de los parámetros. A continuación, se muestran algunos casos.

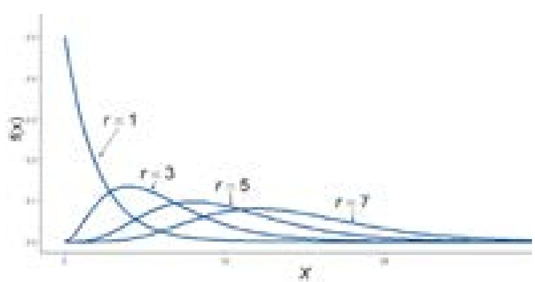


Figura 4.2. *Funciones de densidad gamma. Diferentes valores de* r *y* λ *fijo* ($\lambda = 0.5$)

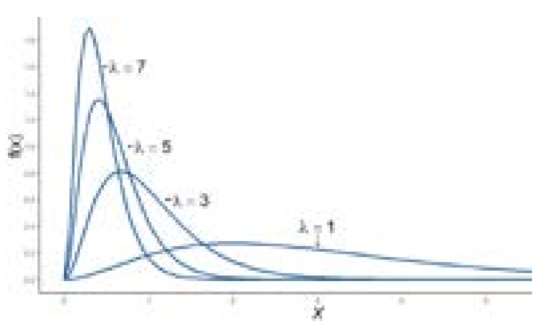


Figura 4.3. *Funciones de densidad gamma. Diferentes valores de* λ *y* r *fijo* ($r = 3$)

Es importante mencionar, en el caso donde r no es entero, la integral $\Gamma(r) = \int_0^\infty x^{r-1} e^{-x} dx$ no se puede realizar en forma analítica.

Para conocer un valor de la función $\Gamma(r)$, donde r no es entero, se pueden aplicar métodos numéricos o utilizar algunos paquetes matemáticos o estadísticos. Como consecuencia de este problema, para r no entero, no se puede resolver en forma analítica la siguiente probabilidad

$$P[a < X < b] = \int_a^b \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x} dx.$$

No obstante, a continuación, se demostrará que $f_X(x)$ es una función de densidad, obsérvese

$$\lambda^r > 0; \quad x^{r-1} > 0; \quad e^{-\lambda x} > 0; \quad \Gamma(r) > 0 \text{ e } I_{(0,\infty)}(x) \geq 0.$$

De esta manera,

$$f_X(x) = \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x} I_{(0,\infty)}(x) \geq 0.$$

Por otro lado,

$$\begin{aligned} \int_0^\infty \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x} dx &= \frac{1}{\Gamma(r)} \int_0^\infty \lambda^r x^{r-1} e^{-\lambda x} dx \\ &= \frac{1}{\Gamma(r)} \int_0^\infty (\lambda x)^{r-1} e^{-\lambda x} \lambda dx. \end{aligned}$$

Se realizará el cambio de variable: $u = \lambda x$. De esta manera, $du = \lambda dx$. Por lo tanto,

$$\begin{aligned} \frac{1}{\Gamma(r)} \int_0^\infty (\lambda x)^{r-1} e^{-\lambda x} \lambda dx &= \frac{1}{\Gamma(r)} \int_0^\infty u^{r-1} e^{-u} du \\ &= \frac{1}{\Gamma(r)} \Gamma(r) = 1. \end{aligned}$$

Se concluye que $f_X(x)$ es función de densidad.

Las aplicaciones de la distribución gamma son muy diversas, en principio si la variable aleatoria tiene recorrido los números reales no negativos, existe una posibilidad de ser modelada por medio de la distribución gamma. Por mencionar algunos ejemplos, el tiempo en el que una máquina requiere mantenimiento, el tiempo de vida de un componente electrónico, las ventas mensuales de una empresa, la vida de un ser viviente, la cantidad reclamada mensual en una compañía de seguros, tienen posibilidades de ser modelados por medio de la distribución gamma.

Por otro lado, consideremos la variable aleatoria *número de eventos en un determinado lapso de tiempo* y supongamos que tiene distribución Poisson con parámetro λ, entonces, la variable aleatoria definida como el tiempo para observar el r-ésimo evento tiene distribución $Gamma(r, \lambda)$, resultado que será demostrado más adelante (véase teorema 4.29).

El siguiente resultado trata con el valor esperado, varianza y función generadora de momentos de una distribución gamma.

Teorema 4.2. *Sea X una variable aleatoria con distribución gamma con parámetros r y λ, entonces,*

$$E[X] = \frac{r}{\lambda},$$
$$V[X] = \frac{r}{\lambda^2},$$
$$m_X(t) = \left(\frac{\lambda}{\lambda - t}\right)^r = \left(1 - \frac{t}{\lambda}\right)^{-r}, \quad \textit{si } t < \lambda.$$

Demostración. Se calculará la función generadora de momentos

$$\begin{aligned} m_X(t) &= E[e^{tX}] \\ &= \int_0^\infty e^{tx} \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x} dx \\ &= \lambda^r \int_0^\infty \frac{1}{\Gamma(r)} x^{r-1} e^{-(\lambda - t)x} dx \\ &= \frac{\lambda^r}{(\lambda - t)^r} \int_0^\infty \frac{(\lambda - t)^r}{\Gamma(r)} x^{r-1} e^{-(\lambda - t)x} dx \\ &= \frac{\lambda^r}{(\lambda - t)^r}, \quad \text{si } t < \lambda. \end{aligned}$$

□

La última integral corresponde a una función de densidad de una distribución gamma con parámetros r y $\lambda - t$, siempre y cuando se cumpla $\lambda - t > 0$. Por lo que el resultado es válido siempre y cuando $t < \lambda$.

La media y la varianza se pueden obtener a partir de función generadora de momentos y se dejará como ejercicio para el estudiante.

Observación. De acuerdo a la definición 2.23, la distribución gamma es de cola ligera.

Teorema 4.3. *Sea X una variable aleatoria con distribución gamma con parámetros r y λ, entonces, la moda M es igual a*

$$M = \frac{r-1}{\lambda}, \quad \textit{si } r > 1.$$

Demostración. Se obtendrá la primera derivada de la función de densidad y se igualará a cero,

$$\begin{aligned} \frac{d}{dx} f_X(x) &= \frac{\lambda^r}{\Gamma(r)} \left[x^{r-1} e^{-\lambda x}(-\lambda) + e^{-\lambda x}(r-1)x^{r-2}\right] \\ &= \frac{\lambda^r}{\Gamma(r)} e^{-\lambda x} x^{r-2} [-\lambda x + r - 1] = 0. \end{aligned}$$

Obsérvese

$$\frac{\lambda^r}{\Gamma(r)} e^{-\lambda x} x^{r-2} > 0.$$

Entonces,

$$-\lambda x + r - 1 = 0.$$

Por lo tanto,

$$x = \frac{r-1}{\lambda}.$$

Obsérvese, si $x > 0$, entonces, se debe de cumplir $r > 1$.

Ahora, la primera derivada es negativa, si y solo si, $-\lambda x + r - 1 < 0$, esto es, si y solo si, $x > \dfrac{r-1}{\lambda}$.

Por otro lado, la primera derivada es positiva, si y solo si, $-\lambda x + r - 1 > 0$, esto es, si $x < \dfrac{r-1}{\lambda}$.

Por lo tanto, en $x = \dfrac{r-1}{\lambda}$ hay un máximo, dicho de otra manera,

$$M = \frac{r-1}{\lambda}, \quad \text{si } r > 1. \qquad \square$$

Definición 4.6. Sea X una variable aleatoria con distribución gamma con parámetros n y λ donde n es un número entero, entonces, se dice que X tiene distribución Erlang con parámetros n y λ .

Notación. $X \sim \text{Erlang}(n, \lambda)$, significa que la variable aleatoria X tiene distribución Erlang con parámetros n y λ.

La distribución Erlang tiene una importante relación con la distribución Poisson, resultado que se va a explicar más adelante, véase el terorema 4.29 y su demostración correspondiente.

El siguiente teorema trata sobre la función de distribución de una variable aleatoria la cual se distribuye Erlang.

Teorema 4.4. *Si X es una variable aleatoria con distribución Erlang con parámetros n y λ, entonces, la función de distribución de X es igual a*

$$F_X(x) = 1 - \sum_{i=0}^{n-1} \frac{(\lambda x)^i}{i!} e^{-\lambda x}.$$

La demostración se deja para el estudiante, y se le sugiere que realice esta por medio de inducción matemática, utilizando, en su momento, integración por partes considerando $u = t^{n-1}$ y $dv = e^{-\lambda t} dt$.

Del teorema anterior, se puede observar que, para el caso de la distribución Erlang, la función de supervivencia es

$$S_X(x) = \sum_{i=0}^{n-1} \frac{(\lambda x)^i}{i!} e^{-\lambda x}.$$

Observación. De los dos anteriores resultados, se puede ver

$$F_X(x) = 1 - P[Y \leq n-1] = P[Y \geq n],$$
$$S_X(x) = P[Y \leq n-1],$$

donde $Y \sim Poisson(\lambda x)$.

Ejemplo 4.2. La duración (vida) en cientos de horas de determinado componente es una variable aleatoria con distribución gamma (Erlang) con parámetros $r = 5$ y $\lambda = 2$. Calcular

a) La probabilidad de que un componente dure más de 300 horas.
b) La probabilidad de que un componente dure entre 250 y 300 horas.
c) El número de horas promedio de la vida de un componente, así como su desviación estándar.
d) La moda del tiempo de duración.

$X =$ Vida en cientos de horas de un componente.
$X \sim Gamma(r = 5, \lambda = 2)$.
a) La primera probabilidad se puede calcular

$$P[X > 3] = \int_3^\infty \frac{2^5}{\Gamma(5)} x^4 e^{-2x} dx.$$

Para resolver esta integral, se tendrá que aplicar 4 veces integración por partes. Otra manera es aplicando el teorema 4.4.

$$\begin{aligned} P[X > 3] &= \sum_{i=0}^{4} \frac{[(2)(3)]^i}{i!} e^{-2(3)} \\ &= \sum_{i=0}^{4} \frac{6^i}{i!} e^{-6} \\ &= e^{-6}[1 + 6 + 18 + 36 + 54] = 0.28506. \end{aligned}$$

b) La siguiente probabilidad

$$\begin{aligned} P[2.5 < X < 3] &= F_X(3) - F_X(2.5) \\ &= 1 - \sum_{i=0}^{4} \frac{[(2)(3)]^i}{i!} e^{-2(3)} - 1 + \sum_{i=0}^{4} \frac{[(2)(2.5)]^i}{i!} e^{-2(2.5)} \\ &= 0.15544. \end{aligned}$$

c) La media y desviación estándar son, respectivamente, $E[X] = 5/2 = 2.5$ y $\sigma = \sqrt{1.25} = 1.11803$.

En promedio dura 250 horas con una desviación estándar de 111.8 horas.

d) La moda es $M = \frac{r-1}{\lambda} = 2$. Por lo tanto, la moda en horas es 200.

Por medio de R, las probabilidades anteriores se calculan respectivamente: $1 - pgamma(3, 5, 2)$ y $pgamma(3, 5, 2) - pgamma(2.5, 5, 2)$, siendo los resultados respectivamente: 0.2850565 y 0.1554368.

Definición 4.7. Una variable aleatoria continua X tiene *distribución ji cuadrada* con parámetro ν, si X se distribuye gamma con parámetros $r = \frac{\nu}{2}$ y $\lambda = \frac{1}{2}$. El parámetro ν se le conoce como los *grados de libertad* de la distribución.

Notación. $X \sim \chi^2(\nu)$ significa que la variable aleatoria X tiene distribución ji cuadrada con ν grados de libertad.

Obsérvese que si $X \sim \chi^2(\nu)$, entonces, la función de densidad, la media, la varianza y la función generadora de momentos respectivamente son

$$f_X(x) = \frac{\left(\frac{1}{2}\right)^{\frac{\nu}{2}}}{\Gamma(\frac{\nu}{2})} x^{\frac{\nu}{2}-1} e^{-\frac{x}{2}} I_{(0,\infty)}(x),$$
$$E[X] = \nu,$$
$$V[X] = 2\nu,$$
$$m_X(t) = (1-2t)^{-\nu/2}, \text{ si } t < 1/2.$$

La gráfica de la función de densidad de una distribución ji cuadrada, siempre es sesgada a la derecha, véase los siguientes casos, para diferentes grados de libertad.

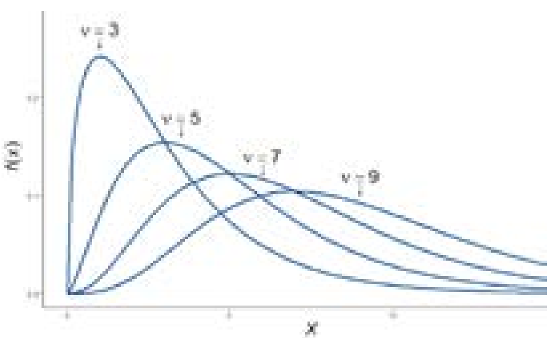


Figura 4.4. *Funciones de densidad de una distribución ji cuadrada con diferentes valores de grados de libertad*

Esta distribución adquiere mucha importancia en varios tópicos de inferencia estadística, por mencionar algunos, cuando se realiza inferencia estadística para una varianza, en pruebas estadísticas para datos categóricos, en problemas de bondad de ajuste, entre otras aplicaciones.

Para calcular las probabilidades o cuantiles de una distribución ji cuadrada, se puede utilizar una tabla de esta distribución (véanse tablas 3 y 4 del apéndice A) o se puede realizar el cálculo por medio de algún *software* como Excel o R.

Existe otra versión de la distribución gamma, la cual se denota como $Gamma(\alpha, \beta)$ y su función de densidad está dada por

$$f_X(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} I_{(0,\infty)}(x).$$

La relación que existe entre los parámetros r y λ de la primera versión, con los parámetros α y β de la segunda versión es

$$\alpha = r \quad \text{y} \quad \beta = \frac{1}{\lambda}.$$

La media, la varianza y la función generadora de momentos, para la nueva versión de la distribución gamma, son, respectivamente,

$$m_X(t) = (1 - \beta t)^{-\alpha}, \ \text{si } t < 1/\beta,$$

$$E[X] = \alpha\beta,$$

$$V[X] = \alpha\beta^2.$$

Es importante que el estudiante, cuando consulte algún libro, artículo, incluso algún *software*, observe que versión de la distribución gamma se está considerando.

4.3. Distribución exponencial

La distribución exponencial, como se verá, es un caso particular de la distribución gamma. En esta sección se presentarán la definición y sus principales propiedades, entre las más importantes, la propiedad de pérdida de memoria. Por otra parte, se tratarán algunos ejemplos que ayudarán a entender la importancia que tiene este modelo para ser aplicado en problemas reales.

Definición 4.8. Una variable aleatoria continua X tiene una *distribución exponencial* con parámetro λ si esta se distribuye gamma con parámetros $r = 1$ y λ.

Para el caso particular de la distribución exponencial, se incluirá el cero en el recorrido de la variable. Entonces, la función de densidad de una distribución exponencial con parámetro λ es

$$f_X(x) = \lambda e^{-\lambda x} I_{[0,\infty)}(x), \ \text{donde } \lambda > 0.$$

Como se había comentado en la anterior sección, λ es un parámetro de escala.

Se puede generalizar la definición de función de densidad exponencial, considerando un parámetro de localización de la siguiente manera:

$$f_X(x) = \lambda e^{-\lambda(x-\theta)} I_{[\theta,\infty)}(x), \ \text{donde } \lambda > 0 \ \text{ y } \ \theta > 0.$$

Las propiedades y ejercicios que se tratarán de la distribución exponencial serán considerando un parámetro de localización igual a cero, a menos que se indique lo contrario.

Notación. $X \sim Exp(\lambda)$ significa que la variable aleatoria X tiene una distribución exponencial con parámetro λ.

Existe una única forma de la gráfica de la función de densidad exponencial, esta se presenta a continuación:

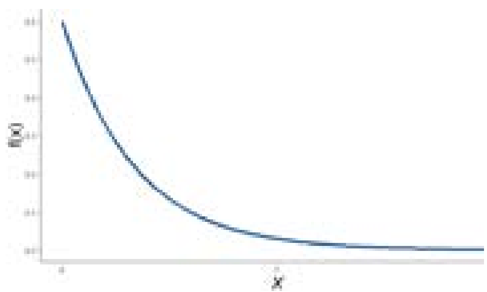


Figura 4.5. *Función de densidad. Exponencial*

En la sección anterior, para una distribución gamma, se demostró que la función $f_X(x)$ es de densidad.

De una manera similar que la distribución gamma, la distribución exponencial es un posible modelo para las variables aleatorias continuas con recorrido no negativo, por mencionar algunos ejemplos, tiempo para observar un suceso, tiempo entre suceso y suceso, tiempo de espera, tiempo de atención de un servicio.

Teorema 4.5. *Sea X una variable aleatoria con distribución exponencial con parámetro λ, entonces,*

$$E[X] = \frac{1}{\lambda},$$

$$V[X] = \frac{1}{\lambda^2},$$

$$m_X(t) = \frac{\lambda}{\lambda - t} = \left(1 - \frac{t}{\lambda}\right)^{-1}, \quad si \ \ t < \lambda.$$

Este teorema es un caso particular del teorema 4.2, debido a que la distribución exponencial pertenece a la familia de distribuciones gamma.

También, se puede afirmar que la distribución exponencial es de cola ligera, véase definición 2.23.

Obsérvese que la distribución del ejemplo 2.16 es un caso particular de la distribución exponencial con $\lambda = 1$.

Teorema 4.6. *Si X es una variable aleatoria con distribución exponencial con parámetro λ, entonces, la moda es $M = 0$.*

Demostración. Obsérvese $\frac{d}{dx} f_X(x) = \lambda e^{-\lambda x}(-\lambda) < 0$, lo que significa que la función de densidad es decreciente, por lo tanto, en $x = 0$ está el máximo de la función, por lo tanto, la moda es $M = 0$. □

La función de distribución puede encontrarse en forma explícita, véase el siguiente teorema.

Teorema 4.7. *Si X es una variable aleatoria con distribución exponencial y parámetro λ, entonces,*

$$F_X(x) = (1 - e^{-\lambda x}) I_{[0,\infty)}(x).$$

Demostración. Para $x \geq 0$,

$$F_X(x) = \int_0^x \lambda e^{-\lambda t} dt = -e^{-\lambda t}\Big|_0^x = -e^{-\lambda x} + 1.$$

Entonces,

$$F_X(x) = \left(1 - e^{-\lambda x}\right) I_{[0,\infty)}(x). \qquad \square$$

Como consecuencia del teorema anterior, la función de supervivencia es

$$S_X(x) = e^{-\lambda x} I_{[0,\infty)}(x).$$

También, se puede observar

$$P[a < X < b] = e^{-a\lambda} - e^{-b\lambda}.$$

Igual que la distribución geométrica, la distribución exponencial cumple con la propiedad de pérdida de memoria, véase siguiente teorema.

Teorema 4.8. *Sea X una variable aleatoria con distribución exponencial con parámetro λ, además sea $a > 0$ y $b > 0$, entonces,*

$$P[X > a + b | X > a] = P[X > b].$$

Demostración. Obsérvese

$$\begin{aligned}
P[X > a + b | X > a] &= \frac{P[(X > a + b) \cap (X > a)]}{P[X > a]} \\
&= \frac{P[X > a + b]}{P[X > a]} \\
&= \frac{e^{-\lambda(a+b)}}{e^{-\lambda a}} \\
&= e^{-\lambda b} \\
&= P[X > b]. \qquad \square
\end{aligned}$$

Supongamos que la variable X se define como el tiempo hasta que ocurra un determinado suceso. Entonces, el evento $\{X > c\}$ se puede interpretar: hasta el momento c, el suceso no ha ocurrido.

De esta manera, la propiedad de pérdida de memoria se puede interpretar de la siguiente manera: dado que se sabe que el suceso no ha ocurrido hasta el momento a, esto es, que $\{X > a\}$, la probabilidad de que el suceso todavía no ocurra en b tiempo más, esto es, que $\{X > a+b\}$ es igual a la probabilidad incondicional de que el suceso no ocurra hasta el momento b, esto es, que $\{X > b\}$.

Observemos que la información que se da, de que el suceso no ha ocurrido hasta el momento a, no es una información relevante en ese momento, es como volver a empezar a observar la ocurrencia del evento, esto es, desde el momento cero.

Ejemplo 4.3. El tiempo en minutos de la llegada entre un cliente y un cliente a un negocio es una variable aleatoria con distribución exponencial y parámetro $\lambda = \frac{1}{15}$. Calcular
a) La probabilidad de que el siguiente cliente llegue después de 10 minutos.
b) De que el siguiente cliente llegue antes de los 20 minutos, dado que se sabe que hasta los 10 minutos no había llegado.
c) El promedio y desviación estándar de la llegada del siguiente cliente.

$X =$ Tiempo entre cliente y cliente.
$X \sim Exp(\lambda = \frac{1}{15})$.
a) La primera probabilidad

$$P[X > 10] = e^{-\frac{10}{15}} = 0.513417.$$

b) La segunda probabilidad

$$\begin{aligned} P[X < 20|X > 10] &= 1 - P[X > 20|X > 10] \\ &= 1 - P[X > 10] \\ &= 1 - e^{-\frac{1}{15}10} = 0.48658. \end{aligned}$$

c) El valor esperado y la desviación estándar son, respectivamente, $E[X] = 15$ minutos y $\sigma = 15$.

La primera probabilidad por medio de R se calcula como $1 - pexp(10, 1/15)$, siendo la respuesta 0.5134171. La segunda probabilidad después de aplicar la propiedad de pérdida es el complemento de la anterior probabilidad.

En el siguiente ejemplo, se considera la mezcla de dos distribuciones, donde una de ellas es la distribución exponencial y la otra es una distribución discreta.

Ejemplo 4.4. Los gastos médicos de un asegurado por contraer una enfermedad se distribuyen exponencialmente con media igual a 1,000. La póliza de seguro indica que solo se pagarán los gastos médicos al asegurado hasta una cantidad de 1,200. Encontrar
a) La función de densidad de los gastos médicos que pagará la compañía al asegurado.
b) La mediana de la cantidad que la compañía pagará al asegurado.

c) Dado que se sabe que el reembolso para un asegurado es menor a 800, calcular la probabilidad de que este exceda a 500.
d) La media de la cantidad pagada por la compañía al asegurado.

a) Sea Y los gastos médicos de un paciente y X la cantidad pagada por la compañía al asegurado.

Se sabe que $Y \sim Exp(\lambda = 0.001)$. Por otro lado, obsérvese que la distribución de X es una mezcla de dos distribuciones, una distribución continua con recorrido el intervalo $(0, 1200)$ y otra discreta que toma solo un valor en $x = 1200$. Se recomienda revisar definiciones y resultados de la sección 2.14.

Sea $f_1(x)$ la función de densidad de la distribución continua y $f_2(x)$ la función de densidad discreta.

La probabilidad de que la compañía pague al asegurado 1,200 es igual a

$$P[Y \geq 1200] = e^{-\frac{1200}{1000}} = 0.30119.$$

Por lo tanto, la probabilidad de que la compañía pague una cantidad menor a 1,200 es $P[Y < 1200] = 1 - 0.30119 = 0.69881$.

De esta manera, la función de densidad de X es

$$\begin{aligned} f_X(x) &= p_1 f_1(x) + p_2 f_2(x) \\ &= 0.69881 f_1(x) + 0.30119 f_2(x), \text{ donde} \end{aligned}$$

$$f_1(x) = \left(\frac{1}{0.69881}\right) \frac{1}{1000} e^{-\frac{1}{1000}} I_{(0,1200)}(x) = \frac{1}{698.81} e^{-\frac{x}{1000}} I_{(0,1200)}(x),$$

$$f_2(x) = I_{\{1200\}}(x).$$

Por lo tanto,

$$\begin{aligned} f_X(x) &= 0.69881 \left(\frac{1}{698.81} e^{-\frac{x}{1000}} I_{(0,1200)}(x)\right) + 0.30119 I_{\{1200\}}(x) \\ &= \frac{1}{1000} e^{-\frac{x}{1000}} I_{(0,1200)}(x) + 0.30119 I_{\{1200\}}(x). \end{aligned}$$

b) Si $P[X < 1200] = 0.69881$, entonces, la mediana M se encuentra en el intervalo $(0, 1200)$.

De esta manera,

$$P[X < M] = 0.5.$$

Esto es,

$$\int_0^M \frac{1}{1000} e^{-\frac{x}{1000}} dx = 0.5.$$

Por lo tanto,

$$1 - e^{-\frac{M}{1000}} = 0.5.$$

Se concluye que la mediana $M = -1000ln(0.5) = 693.14718$.

c) La probabilidad se calculará considerando la parte continua de la distribución, esto es,

$$P[X > 500|X < 800] = \frac{P[500 < X < 800]}{P[X < 800]}$$
$$= \frac{e^{-0.5} - e^{-0.8}}{1 - e^{-0.8}} = 0.28547.$$

d) El cálculo de la media se da a continuación:

$$E[X] = \int_0^{1200} x \frac{1}{1000} e^{-\frac{x}{1000}} dx + (0.30119)(1200) = 698.80073.$$

Es importante mencionar que existe otra versión de la distribución exponencial, la cual tiene la siguiente función de densidad

$$f_X(x) = \frac{1}{\beta} e^{-\frac{x}{\beta}} I_{[0,\infty)}(x), \quad \text{donde} \quad \beta > 0.$$

Obsérvese que el parámetro β es igual al recíproco del parámetro λ, esto es, $\beta = \frac{1}{\lambda}$.

La media, varianza y función generadora de momentos de esta nueva versión, respectivamente, son

$$E[X] = \beta,$$

$$V[X] = \beta^2,$$

$$m_X(t) = (1 - \beta t)^{-1}, \text{ si } t < \frac{1}{\beta}.$$

Esta versión de la distribución exponencial no será utilizada a menos que se indique.

Cuando se consulte literatura o algún *software* donde este involucrada la distribución exponencial, es importante revisar cuál de las dos versiones se está considerando.

4.4. Distribución beta

En esta sección se va a definir la distribución beta, se tratarán sus propiedades más importantes, además, se presentarán las diversas formas que tiene su función de densidad, las cuales son importantes en las posibles aplicaciones que tiene este modelo. Con la idea de dar a conocer algunas aplicaciones, se presentarán ejemplos.

Definición 4.9. Una variable aleatoria continua X tiene *distribución beta* con parámetros a y b, si su función de densidad está dada por

$$f_X(x) = \frac{1}{\beta(a,b)} x^{a-1}(1-x)^{b-1} I_{(0,1)}(x), \quad \text{donde} \quad a > 0, b > 0 \quad \text{y}$$

$$\beta(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1} dx.$$

La función $\beta(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1} dx$ es conocida como *función beta de Euler* o simplemente *función beta*.

Se puede demostrar que

$$\beta(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}.$$

Ambos parámetros a y b son considerados parámetros de forma, véase las siguientes gráficas.

Notación. $X \sim Beta(a,b)$ significa que X es una variable aleatoria con distribución beta con parámetros a y b.

Observación. Si $X \sim Beta(1,1)$, entonces, $X \sim U(0,1)$.

A continuación, se presentan algunas formas de su función de densidad

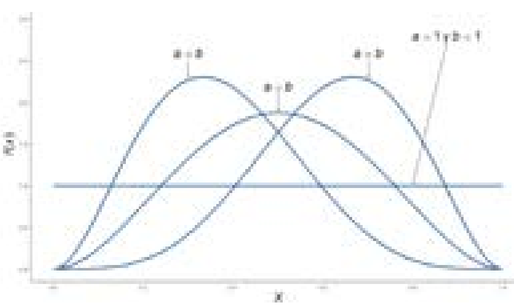


Figura 4.6. *Funciones de densidad, beta, $a \geq 1$ y $b \geq 1$*

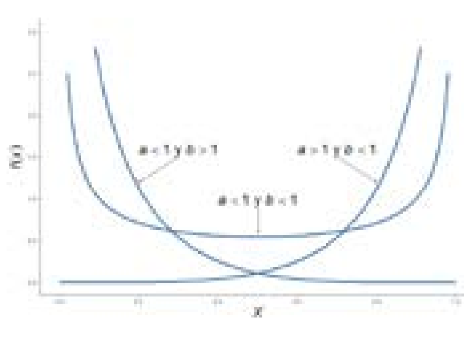


Figura 4.7. *Funciones de densidad, beta, $a < 1$ o $b < 1$*

Es posible demostrar que la función $f_X(x)$ es de densidad, obsérvese

$$x^{a-1} > 0 \text{ y } (1-x)^{b-1} > 0.$$

Por lo tanto,

$$\beta(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx > 0.$$

Además,

$$I_{(0,1)}(x) \geq 0.$$

De esta manera,

$$f_X(x) = \frac{1}{\beta(a,b)} x^{a-1}(1-x)^{b-1} I_{(0,1)}(x) \geq 0.$$

Por otro lado,

$$\begin{aligned}\int_0^1 \frac{1}{\beta(a,b)} x^{a-1}(1-x)^{b-1}dx &= \frac{1}{\beta(a,b)} \int_0^1 x^{a-1}(1-x)^{b-1}dx \\ &= \frac{1}{\beta(a,b)} \beta(a,b) = 1.\end{aligned}$$

Por lo tanto, $f_X(x)$ es función de densidad.

Las aplicaciones de la distribución beta son muy diversas, pueden modelar variables aleatorias como la proporción de impurezas que tiene un determinado líquido o la proporción de tiempo que una máquina permanece parada por mantenimiento, por mencionar algunas. En general, si la variable es una proporción (continua), la distribución beta es un modelo posible para esta variable.

Cabe mencionar que es posible modificar el modelo beta a una distribución con recorrido diferente al intervalo (0,1). De esta manera, al tener la función de densidad de la distribución beta una gran diversidad de formas (véase anteriores gráficas), la función de densidad modificada se podrá también considerar como un posible modelo para variables con recorrido diferente al intervalo $(0, 1)$, véase ejercicio 73 del capítulo 6.

Por otro lado, la distribución beta ha sido un modelo muy importante en la aplicación de la estadística bayesiana, donde los parámetros de una distribución son considerados como variables aleatorias, tal es el caso de la distribución binomial, donde uno de sus parámetros es $p =$ proporción de éxitos, y es común que la distribución beta se considere para esta variable.

Para el caso de la distribución beta, igual que la distribución hipergeométrica, no es útil la función generadora de momentos. El siguiente resultado tiene que ver con el cálculo del k-ésimo momento y, como se verá, este se calcula usando la función de densidad.

Teorema 4.9. *Si X es una variable aleatoria con distribución beta con parámetros a y b, entonces,*

$$E[X^k] = \frac{\beta(a+k,b)}{\beta(a,b)}.$$

Demostración. Obsérvese

$$\begin{aligned} E[X^k] &= \int_o^1 x^k \frac{1}{\beta(a,b)} x^{a-1}(1-x)^{b-1}dx \\ &= \frac{1}{\beta(a,b)} \int_o^1 x^{a+k-1}(1-x)^{b-1}dx \\ &= \frac{1}{\beta(a,b)}\beta(a+k,b). \end{aligned}$$

□

A partir de este resultado, es posible calcular la media y la varianza.

Teorema 4.10. *Si X es una variable aleatoria con distribución beta con parámetros a y b, entonces,*

$$E[X] = \frac{a}{a+b},$$

$$V[X] = \frac{ab}{(a+b)^2(a+b+1)}.$$

Demostración. Usando el resultado del anterior teorema, se calculará $E[X]$

$$\begin{aligned} E[X] &= \frac{\beta(a+1,b)}{\beta(a,b)} \\ &= \frac{\frac{\Gamma(a+1)\Gamma(b)}{\Gamma(a+b+1)}}{\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}} \\ &= \frac{a\Gamma(a)\Gamma(a+b)}{\Gamma(a)(a+b)\Gamma(a+b)} \\ &= \frac{a}{a+b}. \end{aligned}$$

El segundo momento

$$\begin{aligned} E[X^2] &= \frac{\beta(a+2,b)}{\beta(a,b)} \\ &= \frac{\frac{\Gamma(a+2)\Gamma(b)}{\Gamma(a+b+2)}}{\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}} \\ &= \frac{(a+1)a\Gamma(a)\Gamma(a+b)}{\Gamma(a)(a+b+1)(a+b)\Gamma(a+b)} \end{aligned}$$

$$= \frac{a(a+1)}{(a+b)(a+b+1)}.$$

El cálculo de la varianza

$$\begin{aligned} V[X] &= \frac{a(a+1)}{(a+b)(a+b+1)} - \frac{a^2}{(a+b)^2} \\ &= \frac{a(a+1)(a+b) - a^2(a+b+1)}{(a+b)^2(a+b+1)} \\ &= \frac{ab}{(a+b)^2(a+b+1)}. \end{aligned}$$ □

Teorema 4.11. *Si X es una variable aleatoria con distribución beta con parámetros a y b, entonces, la moda M es*

$$M = \frac{a-1}{a+b-2}, \quad si \quad a > 1 \text{ y } b > 1.$$

Demostración. Se calculará la primera derivada de la función de densidad y se igualará a cero

$$\begin{aligned} \frac{d}{dx} f_X(x) &= \frac{1}{\beta(a,b)}[-x^{a-1}(b-1)(1-x)^{b-2} + (1-x)^{b-1}(a-1)x^{a-2}] \\ &= \frac{1}{\beta(a,b)} x^{a-2}(1-x)^{b-2}[-x(b-1) + (1-x)(a-1)] = 0. \end{aligned}$$

Obsérvese

$$\frac{1}{\beta(a,b)} x^{a-2}(1-x)^{b-2} > 0.$$

De esta manera,

$$-x(b-1) + (1-x)(a-1) = 0.$$

Por lo tanto,

$$x = \frac{a-1}{a+b-2}.$$

Se verificará que existe un máximo o un mínimo en el valor encontrado. El valor encontrado debe estar entre 0 y 1, esto es,

$$0 < \frac{a-1}{a+b-2} < 1.$$

Una opción para que se cumpla la anterior desigualdad es que $0 < a-1$ y también $0 < a+b-2$, esto es, $1 < a$ y $2 < a+b$. Además, se tendrá que cumplir $a - 1 < a + b - 2$, esto es, $b > 1$. En conclusión, se deberá cumplir que $a > 1$ y $b > 1$.

Otra opción es que $a - 1 < 0$ y $a + b - 2 < 0$, esto es, $a < 1$ y $a + b < 2$. Además, se deberá cumplir $a - 1 > a + b - 2$, esto es que $b < 1$. En conclusión, se deberá cumplir $a < 1$ y $b < 1$.

En resumen, se deberá cumplir $a > 1$ y $b > 1$ o $a < 1$ y $b < 1$.

Obsérvese que la primera derivada de la función de densidad es negativa si y solo si

$$-x(b-1) + (1-x)(a-1) < 0.$$

Lo anterior se cumple, si y solo si

$$a - 1 < x(a + b - 2).$$

Si $a > 1$ y $b > 1$, entonces,

$$x > \frac{a-1}{a+b-2}.$$

Si $a < 1$ y $b < 1$, entonces,

$$x < \frac{a-1}{a+b-2}.$$

Por otro lado, la primera derivada de la función de densidad es positiva, si y solo si

$$-x(b-1) + (1-x)(a-1) > 0.$$

La anterior desigualdad se cumple, si y solo si

$$a - 1 > x(a + b - 2).$$

Si $a > 1$ y $b > 1$, entonces,

$$x < \frac{a-1}{a+b-2}.$$

Si $a < 1$ y $b < 1$, entonces,

$$x > \frac{a-1}{a+b-2}.$$

Esto es, si $a > 1$ y $b > 1$, entonces, en $x = \dfrac{a-1}{a+b-2}$ existe un máximo.

Por otra lado, si $a < 1$ y $b < 1$, entonces, en $x = \dfrac{a-1}{a+b-2}$, hay un mínimo.

En conclusión, la moda es

$$M = \frac{a-1}{a+b-2}, \text{ si } a > 1 \text{ y } b > 1. \qquad \square$$

Observación. Se puede apreciar que varias de las gráficas que se presentan en la figura 4.6 tienen una moda, estos son casos donde $a > 1$ y $b > 1$. Por otra parte, en la figura 4.7, obsérvese la gráfica en forma de «**U**», esta tiene un mínimo y es un caso donde $a < 1$ y $b < 1$. En ambos casos, máximo y mínimo, se encuentran en $\dfrac{a-1}{a+b-2}$.

El siguiente resultado da una solución de la función de distribución para el modelo beta, cuando a y b son números enteros.

Teorema 4.12. *Si X es una variable aleatoria con distribución beta con parámetros $a = k$ y $b = n - k + 1$, donde n y k son números enteros no negativos, y $n \geq k$. Entonces, la función de distribución de X es igual a*

$$F_X(x) = \sum_{j=k}^{n} \binom{n}{j} x^j (1-x)^{n-j}.$$

Se le sugiere al estudiante que realice la demostración del anterior teorema.

Observación. En el anterior resultado, se puede ver que

$$\sum_{j=k}^{n} \binom{n}{j} x^j (1-x)^{n-j} = P[Y \geq k] = 1 - F_Y(k-1),$$

donde $Y \sim B(n, p = x)$.

Para el cálculo de probabilidades con la distribución beta, en el caso donde a y b son enteros, se podrá aplicar el teorema anterior o realizar la integral correspondiente de la función de densidad, la cual será la integral de un polinomio. En el caso donde los parámetros no son números enteros, se sugiere usar un *software*, en particular puede ser R.

Ejemplo 4.5. Se sabe que la proporción de impureza que tiene cierta sustancia líquida es una variable aleatoria con distribución beta con parámetros $a = 2$ y $b = 4$. Para un envase que contiene esta sustancia, contestar las siguientes preguntas

a) Calcular la probabilidad de que contenga entre 15 % y 30 % de impureza.
b) Calcular el promedio y desviación estándar del porcentaje de líquido impuro que contiene un envase.
c) Los envases que contengan cierto porcentaje de líquido impuro o más, se regresarán como mercancía de mala calidad. Si se pretende solo regresar el 20 % de los envases, calcular a partir de qué porcentaje de líquido impuro se regresará un envase.

a) Obsérvese para la primera probabilidad
$X =$ Porcentaje de líquido impuro en un envase.
$X \sim Beta(2, 4)$.

De esta manera, $f_X(x) = 20(x - 3x^2 + 3x^3 - x^4)I_{(0,1)}(x)$.

$$\begin{aligned} P[0.15 < X < 0.30] &= \int_{0.15}^{0.30} 20(x - 3x^2 + 3x^3 - x^4)dx \\ &= 10x^2 - 20x^3 + 15x^4 - 4x^5 \Big|_{0.15}^{0.30} = 0.30699. \end{aligned}$$

b) La media y desviación estándar son, respectivamente,

$$E[X] = \tfrac{2}{2+4} = \tfrac{1}{3},$$

$$\sigma = \sqrt{\frac{ab}{(a+b)^2(a+b+1)}} = \sqrt{\tfrac{2}{63}} = 0.17819.$$

c) La función de distribución

$$\begin{aligned} F_X(x) &= \int_0^x 20(t - 3t^2 + 3t^3 - t^4)dt \\ &= 10x^2 - 20x^3 + 15x^4 - 4x^5. \end{aligned}$$

Es necesario encontrar las raíces del polinomio

$$10x^2 - 20x^3 + 15x^4 - 4x^5 - 0.8 = 0.$$

Para encontrar las soluciones de esta ecuación, en forma aproximada, se puede aplicar un método numérico, en este caso sería el método de Newton. También se puede usar algún *software* de matemáticas o estadística. En este caso, usaremos R para encontrar el 0.80-ésimo cuantil de la distribución $Beta(2, 4)$, el cual es una de las raíces del polinomio anterior, en este caso, es aquella raíz que se encuentra en el intervalo $(0, 1)$. La instrucción en R es $qbeta(0.80, 2, 4)$, siendo la solución 0.49019.

La probabilidad del inciso a se calcula por medio de R usando el código: $pbeta(0.30, 2, 4) - pbeta(0.15, 2, 4)$, siendo la respuesta 0.30699.

4.5. Distribución de Weibull

En esta sección se tratará la familia de distribuciones de Weibull, se definirá su función de densidad, se presentarán sus propiedades más importantes y las posibles aplicaciones que esta distribución tiene en la vida real.

Definición 4.10. Una variable aleatoria continua X tiene una *distribución de Weibull* con parámetros a y b, si su función de densidad está dada por

$$f_X(x) = \frac{b}{a^b} x^{b-1} e^{-\left(\frac{x}{a}\right)^b} I_{(0,\infty)}(x), \quad \text{donde } a > 0 \text{ y } b > 0.$$

El parámetro a es de escala y el parámetro b es de forma, se puede apreciar esto en la siguiente gráfica.

Notación. $X \sim Weibull(a, b)$ significa que la variable aleatoria X tiene una distribución de Weibull con parámetros de escala y forma a y b, respectivamente.

Observación. La distribución exponencial es un caso particular de la distribución de Weibull, esto es, $Weibull(a = \lambda^{-1}, b = 1) = Exp(\lambda)$.

La definición de una función de densidad de Weibull se puede generalizar considerando un parámetro de localización de la siguiente manera:

$$f_X(x) = \frac{b}{a^b}(x-\theta)^{b-1} e^{-\left(\frac{x-\theta}{a}\right)^b} I_{(\theta,\infty)}(x).$$

Mientras no se aclare, el parámetro de localización no será considerado.

Se presentarán diferentes formas de la función de densidad de Weibull, donde se puede observar que uno de los casos es la distribución exponencial.

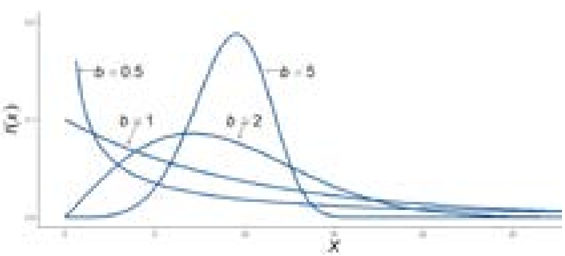


Figura 4.8. *Funciones de densidad de la distribución de Weibull. Diferentes valores de b y a fijo (a = 10)*

Se demostrará que la función $f_X(x)$ es de densidad, obsérvese

$$b > 0;\ \frac{1}{a^b} > 0;\ x^{b-1} > 0;\ e^{-\left(\frac{x}{a}\right)^b} > 0 \ \text{ e } \ I_{(0,\infty)}(x) \geq 0.$$

Por lo tanto,

$$f_X(x) = \frac{b}{a^b} x^{b-1} e^{-\left(\frac{x}{a}\right)^b} I_{(0,\infty)}(x) \geq 0.$$

Por otro lado,

$$\int_0^\infty \frac{b}{a^b} x^{b-1} e^{-\left(\frac{x}{a}\right)^b} dx = \int_0^\infty \frac{b}{a}\left(\frac{x}{a}\right)^{b-1} e^{-\left(\frac{x}{a}\right)^b} dx.$$

Se considera el siguiente cambio de variable

$$u = \left(\frac{x}{a}\right)^b; \quad du = \frac{b}{a}\left(\frac{x}{a}\right)^{b-1} dx.$$

De esta manera, la última integral es igual a

$$\int_0^\infty e^{-u} du = 1.$$

Por lo tanto, $f_X(x)$ es función de densidad.

Las aplicaciones de la distribución de Weibull son importantes en temas de confiabilidad y supervivencia. Es un modelo candidato para variables aleatorias como el tiempo de duración de algún componente, el tiempo necesario para que una máquina necesite mantenimiento o el tiempo de vida de un ser vivo.

Como se verá más adelante, la distribución de Weibull para algunos casos es una distribución de cola pesada, y en este sentido, es un modelo importante para variables que toman valores extremos con probabilidades no tan pequeñas o despreciables. Por ejemplo, el reclamo excesivo de un siniestro a una compañía de seguros debido a la pérdida considerable por un evento natural catastrófico, como un huracán o un sismo. También, el alza considerable o caída de una acción bursátil.

En general, la distribución de Weibull es un modelo candidato para variables aleatorias con recorrido, los números reales no negativos.

A continuación, se tratarán los resultados más importantes de esta distribución.

Teorema 4.13. *Si X es una variable aleatoria con distribución de Weibull con parámetros a y b, entonces,*

$$E[X^k] = a^k \Gamma\left(\frac{k}{b} + 1\right).$$

Demostración. Se calculará $E[X^k]$

$$\begin{aligned} E[X^k] &= \int_0^\infty x^k \frac{b}{a^b} x^{b-1} e^{-\left(\frac{x}{a}\right)^b} dx \\ &= \int_0^\infty x^k \frac{b}{a}\left(\frac{x}{a}\right)^{b-1} e^{-\left(\frac{x}{a}\right)^b} dx. \end{aligned}$$

Donde se considera el siguiente cambio de variable

$$u = \left(\frac{x}{a}\right)^b; \quad du = \frac{b}{a}\left(\frac{x}{a}\right)^{b-1} dx.$$

De donde,

$$x^k = u^{\frac{k}{b}} a^k.$$

Entonces,

$$\begin{aligned} E[X^k] &= a^k \int_0^\infty u^{\frac{k}{b}} e^{-u} du \\ &= a^k \Gamma\left(\frac{k}{b} + 1\right). \end{aligned}$$

□

Teorema 4.14. *Si X es una variable aleatoria con distribución de Weibull con parámetros a y b, entonces,*

$$E[X] = a\Gamma\left(\frac{1}{b}+1\right),$$

$$V[X] = a^2\left(\Gamma\left(\frac{2}{b}+1\right) - \Gamma^2\left(\frac{1}{b}+1\right)\right).$$

Las demostraciones de estos últimos resultados se logran aplicando el teorema 4.13.

Teorema 4.15. *Si X es una variable aleatoria con distribución de Weibull con parámetros a y b, entonces,*

$$F_X(x) = \left[1 - e^{-\left(\frac{x}{a}\right)^b}\right] I_{(0,\infty)}(x).$$

Demostración. Para $x > 0$,

$$\begin{aligned} F_X(x) &= \int_0^x \frac{b}{a^b} t^{b-1} e^{-\left(\frac{t}{a}\right)^b} dt \\ &= \int_0^x \frac{b}{a}\left(\frac{t}{a}\right)^{b-1} e^{-\left(\frac{t}{a}\right)^b} dt \\ &= -e^{-\left(\frac{t}{a}\right)^b}\Big|_0^x \\ &= 1 - e^{-\left(\frac{x}{a}\right)^b}. \end{aligned}$$

Por lo tanto,

$$F_X(x) = \left[1 - e^{-\left(\frac{x}{a}\right)^b}\right] I_{(0,\infty)}(x). \qquad \square$$

Del resultado anterior, se puede obtener la función de supervivencia,

$$S_X(x) = e^{-\left(\frac{x}{a}\right)^b} I_{(0,\infty)}(x).$$

Teorema 4.16. *Si X es una variable aleatoria con distribución de Weibull con parámetros a y b y además $c > 0$ y $d > 0$, entonces,*

$$P[X > c+d | X > c] = e^{-\left[\left(\frac{c+d}{a}\right)^b - \left(\frac{c}{a}\right)^b\right]}.$$

Demostración. Obsérvese

$$
\begin{aligned}
P[X > c + d | X > c] &= \frac{P[(X > c + d) \cap (X > c)]}{P[X > c]} \\
&= \frac{P[X > c + d]}{P[X > c]} \\
&= \frac{e^{-\left(\frac{c+d}{a}\right)^b}}{e^{-\left(\frac{c}{a}\right)^b}} \\
&= e^{-\left[\left(\frac{c+d}{a}\right)^b - \left(\frac{c}{a}\right)^b\right]}. \qquad \square
\end{aligned}
$$

Teorema 4.17. *Si X es una variable aleatoria con distribución de Weibull con parámetros a y b, entonces, la moda M es igual a*

$$M = a\left(1 - \frac{1}{b}\right)^{\frac{1}{b}}, \quad si \quad b > 1.$$

Demostración. Se obtendrá la primera derivada de la función de densidad y se igualará a cero.

$$\frac{d}{dx} f_X(x) = \frac{b}{a^b}\left[x^{b-1} e^{-\left(\frac{x}{a}\right)^b}\left(-\frac{b}{a}\left(\frac{x}{a}\right)^{b-1}\right) + e^{-\left(\frac{x}{a}\right)^b}(b-1)x^{b-2}\right] = 0.$$

Esto es,

$$\frac{b}{a^b}\left[x^{b-2} e^{-\left(\frac{x}{a}\right)^b}\right]\left[x\left(-\frac{b}{a}\right)\left(\frac{x}{a}\right)^{b-1} + b - 1\right] = 0.$$

Se puede observar

$$\frac{b}{a^b}\left[x^{b-2} e^{-\left(\frac{x}{a}\right)^b}\right] > 0.$$

De esta manera,

$$\left[x\left(-\frac{b}{a}\right)\left(\frac{x}{a}\right)^{b-1} + b - 1\right] = 0.$$

Por lo tanto,

$$x = a\left[1 - \frac{1}{b}\right]^{\frac{1}{b}}.$$

El valor crítico encontrado tiene que ser positivo. Esto se cumple, si y solo si $b > 1$. Véase las diferentes formas de la función de densidad de la distribución de Weibull, figura 4.8, y en los casos donde hay una moda, es cuando el parámetro b es mayor que 1.

La primera derivada es negativa si

$$x\left(-\frac{b}{a}\right)\left(\frac{x}{a}\right)^{b-1}+b-1<0.$$

Esto es, cuando

$$x>a\left(1-\frac{1}{b}\right)^{\frac{1}{b}}.$$

La primera derivada es positiva cuando

$$x\left(-\frac{b}{a}\right)\left(\frac{x}{a}\right)^{b-1}+b-1>0.$$

Esto es, si

$$x<a\left(1-\frac{1}{b}\right)^{\frac{1}{b}}.$$

En conclusión, hay un máximo en

$$x=a\left[1-\frac{1}{b}\right]^{\frac{1}{b}}.$$

Por lo tanto, la moda M es

$$M=a\left(1-\frac{1}{b}\right)^{\frac{1}{b}}, \text{ si } \quad b>1. \qquad \square$$

Observación. Se puede demostrar para la distribución de Weibull, en los casos que $0<b<1$, su función generadora de momentos no existe para $t>0$. De esta manera, se puede afirmar que la distribución de Weibull, cuando $0<b<1$, es una distribución de cola pesada.

Obsérvese las diferentes formas de función de densidad de la distribución de Weibull en la figura 4.8, es notorio que la gráfica con parámetro de forma $b=0.5$ tiene una cola derecha con más volumen que las demás funciones de densidad.

Ejemplo 4.6. La duración en horas de cierto componente es una variable aleatoria con distribución de Weibull con parámetros $a=100$ y $b=0.5$. Calcular

a) La media y desviación estándar de la vida en horas de un componente.
b) La probabilidad de que un componente dure más de 300 horas.
c) La probabilidad de que un componente deje de funcionar antes de las 300 horas, dado que se sabe que hasta las 200 horas el componente seguía funcionando.
d) El tiempo en horas, a partir del cual, la vida de un componente es mayor a este tiempo con una probabilidad del 0.10.

a) La media y desviación estándar son, respectivamente,

$$E[X] = a\Gamma\left(\frac{1}{b}+1\right) = 100\Gamma(3) = 200,$$

$$\sigma = \sqrt{a^2\left(\Gamma\left(\frac{2}{b}+1\right) - \Gamma^2\left(\frac{1}{b}+1\right)\right)} = \sqrt{200000} = 447.2136.$$

b) La primera probabilidad

$$P[X > 300] = e^{-\left(\frac{300}{100}\right)^{0.5}} = 0.17692.$$

c) La probabilidad condicional

$$\begin{aligned} P[X < 300|X > 200] &= 1 - P[X \geq 300|X > 200] \\ &= 1 - e^{-\left[\left(\frac{300}{100}\right)^{0.5} - \left(\frac{200}{100}\right)^{0.5}\right]} = 0.27228. \end{aligned}$$

d) Se debe de cumplir

$$0.90 = P[X \leq t] = 1 - e^{-\left(\frac{t}{100}\right)^{0.5}}.$$

Despejando t,

$$t = 100[-ln(0.1)]^2 = 530.18981.$$

La probabilidad del inciso b por medio de R, se puede calcular como: $1 - pweibull(300, 0.5, 100)$, dando como resultado 0.1769212.

Para resolver el inciso d por R, el código es $qweibull(0.90, 0.5, 100)$, siendo la respuesta 530.1898.

4.6. Distribución normal

En esta sección se definirá la distribución normal, una de las familias de distribuciones más importantes debido a sus aplicaciones, ya que es un modelo muy usado en la práctica. Además, es muy utilizado en forma aproximada para variables, tanto discretas como continuas. También, es una distribución fundamental en varios tópicos de métodos de inferencia estadística.

Definición 4.11. Una variable aleatoria continua X tiene *distribución normal* con parámetros μ y σ^2, si su función de densidad está dada por

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad \text{donde } -\infty < \mu < \infty \text{ y } \sigma > 0.$$

El parámetro μ es de localización y el parámetro σ^2 es de escala. No tiene parámetro de forma, ya que la función de densidad siempre tendrá la forma de una «campana».

Notación. $X \sim N(\mu, \sigma^2)$ significa que la variable aleatoria X tiene distribución normal con parámetros μ y σ^2.

Esta distribución también se le conoce como «campana de Gauss» o «distribución Gaussiana». En las siguientes gráficas, se muestran diferentes casos de la función de densidad.

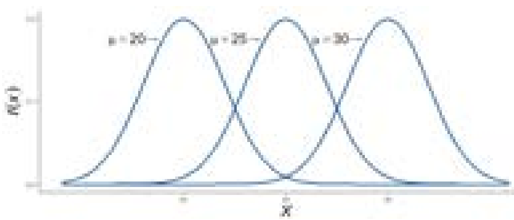


Figura 4.9. *Funciones de densidad de una distribución normal. Diferentes valores de μ y σ^2 fijo ($\sigma^2 = 4$)*

Se puede notar en las gráficas anteriores, cómo se desplaza la gráfica de la función de densidad por cambiar el valor del parámetro μ.

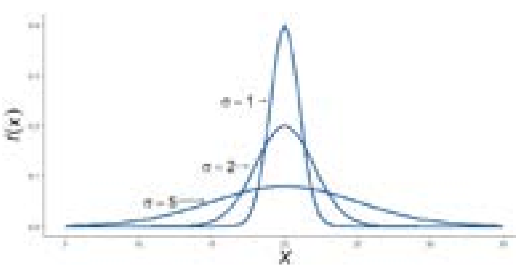


Figura 4.10. *Funciones de densidad de una distribución normal. Diferentes valores de σ y μ fijo ($\mu = 20$)*

Entre más grande sea σ, más ancha y más baja es la gráfica de la función de densidad. Entre más pequeña sea σ, la gráfica es más delgada y más alta.

El hecho de que no aparece una función indicadora en la función de densidad significa que el recorrido de la variable aleatoria es el conjunto de todos los números reales. Mas es importante aclarar que difícilmente podemos encontrar ejemplos prácticos, donde la variable aleatoria tenga como recorrido los números reales. No obstante, a pesar de esta situación, las aplicaciones de la distribución normal son muy diversas, situación que se va aclarar más adelante.

Un primer problema con la distribución normal es que la función de densidad no tiene una función primitiva o antiderivada elemental, por lo que el cálculo de probabilidades no se podrá realizar por medio de integrales analíticas de la función de densidad.

A pesar de la situación anterior, es posible demostrar que la función $f_X(x)$ dada en la definición 4.11 es de densidad.

Obsérvese

$$\sqrt{2\pi} > 0; \quad \sigma > 0 \quad \text{y} \quad e^{-\frac{(x-\mu)^2}{2\sigma^2}} > 0.$$

Por lo tanto,

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} > 0.$$

Por otro lado, se demostrará

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = 1.$$

En la anterior integral, se considera el siguiente cambio de variable

$$z = \frac{x-\mu}{\sigma}; \quad dz = \frac{1}{\sigma} dx.$$

Por lo tanto, el problema se reduce a demostrar que

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = 1.$$

Ahora, obsérvese

$$\begin{aligned}
\left[\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz\right]^2 &= \left[\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz\right]\left[\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}} dv\right] \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-\frac{z^2+v^2}{2}} dzdv.
\end{aligned}$$

Se resolverá la anterior integral doble, y para su solución se aplicará coordenadas polares, esto es,

$$z = rsen(\theta); \quad v = rcos(\theta).$$

De esta transformación, se calculará el jacobiano correspondiente,

$$\begin{aligned}
J &= \begin{vmatrix} \dfrac{\partial z}{\partial r} & \dfrac{\partial z}{\partial \theta} \\[2ex] \dfrac{\partial v}{\partial r} & \dfrac{\partial v}{\partial \theta} \end{vmatrix} \\
&= \begin{vmatrix} sen(\theta) & rcos(\theta) \\ cos(\theta) & -rsen(\theta) \end{vmatrix} \\
&= -rsen^2(\theta) - rcos^2(\theta) = -r.
\end{aligned}$$

Por lo tanto, $|J| = r$.
Entonces,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{2\pi}e^{-\frac{z^2+v^2}{2}}dzdv = \int_{0}^{\infty}\int_{0}^{2\pi} r\frac{1}{2\pi}e^{-\frac{r^2}{2}}d\theta dr$$
$$= \int_{0}^{\infty} re^{-\frac{r^2}{2}}dr.$$

Para esta última integral, se propone el siguiente cambio de variable

$$y = r^2;\ dy = 2rdr.$$

Entonces,

$$\int_{0}^{\infty} re^{-\frac{r^2}{2}}dr = \int_{0}^{\infty}\frac{1}{2}e^{-\frac{y}{2}}dy = 1.$$

Obsérvese que la última integral se realiza sobre todos los valores de una función de densidad de una distribución exponencial con media igual a 2.
Se concluye

$$\left[\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz\right]^2 = 1.$$

Por lo tanto,

$$\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz = \pm 1.$$

Como

$$\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}} > 0.$$

Entonces,

$$\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz = 1.$$

Por lo tanto, la función $f_X(x)$ es de densidad.

Algunos ejemplos donde es posible aplicar este modelo son las estaturas o pesos de seres vivos de una misma especie o raza; las longitudes, diámetros o pesos de artículos fabricados de una misma línea de producción bajo un proceso estable; también, los ingresos familiares de una población, por mencionar algunos ejemplos.

Además, como ya se mencionó, muchos modelos tanto discretos como continuos pueden ser aproximados por la distribución normal.

También por el teorema de límite central (que se tratará en el capítulo 7), las variables aleatorias como la media muestral, suma muestral, proporción muestral, entre otras, tienen aproximadamente una distribución normal, cuando el tamaño de la muestra es «suficientemente grande».

En los ejemplos anteriormente mencionados, es posible que el recorrido de la variable aleatoria involucrada no sea el conjunto de los números reales, mas esta situación tiene una explicación, el área debajo de la curva de la cola derecha (izquierda) de la función de densidad, a partir de cierto valor (hasta cierto valor), es despreciable, prácticamente nulo. Por lo anterior, será posible considerar a la distribución normal como modelo (en forma aproximada) para varias variables aleatorias, a pesar de que el recorrido no sea los números reales.

El siguiente resultado trata de la media, varianza y función generadora de momentos de una distribución normal.

Teorema 4.18. *Si X es una variable aleatoria con distribución normal con parámetros μ y σ^2, entonces,*

$$E[X] = \mu,$$

$$V[X] = \sigma^2,$$

$$m_X(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}.$$

Demostración. Se obtendrá la función generadora de momentos

$$\begin{aligned}
m_X(t) &= \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{tx} e^{-\frac{x^2-2\mu x+\mu^2}{2\sigma^2}} dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2-2x(\mu+\sigma^2 t)+\mu^2}{2\sigma^2}} dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2-2x(\mu+\sigma^2 t)+(\mu+\sigma^2 t)^2}{2\sigma^2}} e^{\frac{(\mu+\sigma^2 t)^2-\mu^2}{2\sigma^2}} dx \\
&= e^{\frac{(\mu+\sigma^2 t)^2-\mu^2}{2\sigma^2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-(\mu+\sigma^2 t))^2}{2\sigma^2}} dx \\
&= e^{\frac{\mu^2+2\mu\sigma^2 t+\sigma^4 t^2-\mu^2}{2\sigma^2}} (1) \\
&= e^{\frac{2\mu\sigma^2 t+\sigma^4 t^2}{2\sigma^2}} = e^{\mu t+\frac{\sigma^2 t^2}{2}}.
\end{aligned}$$

□

La última integral valió 1, ya que se integró sobre todos los valores de una función de densidad de una distribución $N(\mu + \sigma^2 t, \sigma^2)$.

El cálculo de la media y varianza queda como ejercicio para el estudiante.

Teorema 4.19. *Si X es una variable aleatoria con distribución normal con media μ y varianza σ^2, entonces, $Z = \frac{X-\mu}{\sigma}$ se distribuye normal con media 0 y varianza 1.*

Demostración. Se calculará la función generadora de momentos de Z.

$$\begin{aligned}
m_Z(t) &= E[e^{Zt}] \\
&= E\left[e^{\left(\frac{X-\mu}{\sigma}\right)t}\right] \\
&= E\left[e^{\frac{Xt}{\sigma}-\frac{\mu t}{\sigma}}\right] \\
&= e^{-\frac{\mu t}{\sigma}} E\left[e^{\frac{Xt}{\sigma}}\right] \\
&= e^{-\frac{\mu t}{\sigma}} m_X\left(\tfrac{t}{\sigma}\right) \\
&= e^{-\frac{\mu t}{\sigma}} e^{\mu\frac{t}{\sigma}+\frac{\sigma^2}{2}\frac{t^2}{\sigma^2}} = e^{\frac{t^2}{2}}.
\end{aligned}$$

Esta última es la función generadora de momentos de una variable aleatoria con distribución $N(0,1)$, por el teorema de unicidad (teorema 2.6), entonces, $Z \sim N(0,1)$. □

Una variable aleatoria que tiene distribución $N(0,1)$ es denotada como Z, y su distribución es llamada *normal estándar*. Al proceso de transformar una variable aleatoria con distribución $N(\mu,\sigma^2)$ a una variable con distribución normal estándar, se le llama *proceso de estandarización.*

En principio, la probabilidad de una distribución normal debería ser calculada por medio de una integral definida de la función de densidad correspondiente, mas ya se había comentado que la integral de una función de densidad de una distribución normal no tiene una solución explícita en término de funciones elementales. Entonces, el cálculo de probabilidades de una distribución normal se calculará por medio de una tabla de probabilidades de una distribución normal estándar o por medio de algún *software*. Es importante aclarar que, en cualquiera de los dos casos, lo cálculos se realizan por medio de métodos numéricos.

La tabla que se utilizará para encontrar probabilidades de la distribución normal proporciona valores de la función de distribución de una distribución normal estándar, para diferentes valores de la variable aleatoria. Esto es, para un z específico, la tabla dará el valor de

$$F_Z(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\, dt.$$

En los siguientes ejemplos, se calcularán primero probabilidades y cuantiles, para una variable aleatoria Z con distribución normal estándar. Para estos cálculos se usará la tabla de la función de distribución normal estándar que aparece en el apéndice A. En esta tabla aparecen valores de la función de distribución para valores de z entre 0 a 3.49. Para valores de z negativos, se utilizará la misma tabla, aprovechando que la función de densidad es simétrica con respecto al 0.

Ejemplo 4.7. Si $Z \sim N(0, 1)$, entonces, calcular
a) $P[Z < 1.5]$.
b) $P[Z > 0.68]$.
c) $P[Z < -1.2]$.
d) $P[1.5 < Z < 2.0]$.
e) $P[-2.2 < Z < 0.97]$.
f) $P[-2.2 < Z < -1.5]$.
g) z tal que $P[Z < z] = 0.95$.

a) Para calcular $P[Z < 1.5]$ se considera en la tabla, la intersección entre el renglón de $z = 1.5$ y la columna de 0.00. Por lo tanto, $P[Z < 1.5] = 0.9332$. Véase la siguiente gráfica

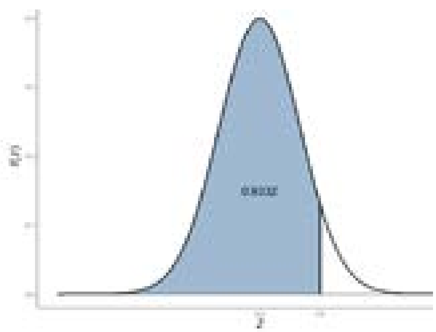


Figura 4.11. *Probabilidad, ejemplo 4.7, a)*

b) La probabilidad pedida

$$\begin{aligned} P[Z > 0.68] &= 1 - P[Z < 0.68] \\ &= 1 - 0.7517 = 0.2483. \end{aligned}$$

Obsérvese en las siguientes gráficas la probabilidad calculada y su complemento, respectivamente,

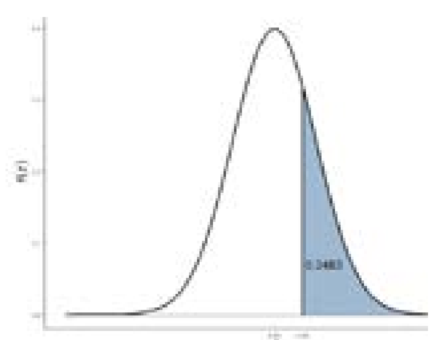


Figura 4.12. *Probabilidad, ejemplo 4.7, b)*

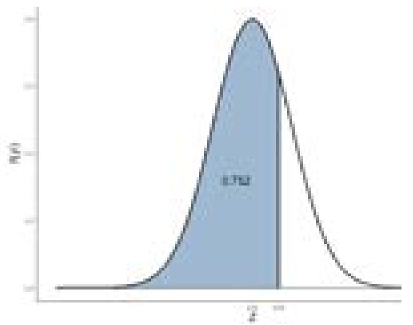


Figura 4.13. *Probabilidad complemento, ejemplo 4.7, b)*

c) La probabilidad se calcula de la siguiente manera:

$$\begin{aligned} P[Z < -1.2] &= P[Z > 1.2] \\ &= 1 - P[Z < 1.2] \\ &= 1 - 0.8849 = 0.1151. \end{aligned}$$

En la siguiente gráfica, se observa la probabilidad calculada

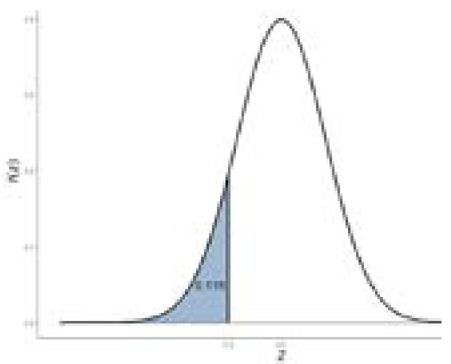


Figura 4.14. *Probabilidad, ejemplo 4.7, c)*

Obsérvese que el área anterior equivale a la siguiente área

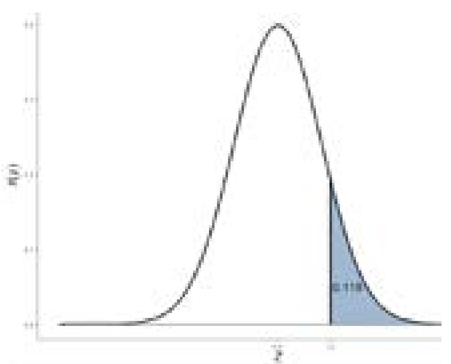


Figura 4.15. *Probabilidad, ejemplo 4.7, c)*

d) En este caso, la probabilidad se calcula

$$\begin{aligned} P[1.5 < Z < 2.0] &= P[Z < 2.0] - P[Z < 1.5] \\ &= 0.9772 - 0.9332 = 0.044. \end{aligned}$$

Esta última probabilidad se puede apreciar en la siguiente gráfica:

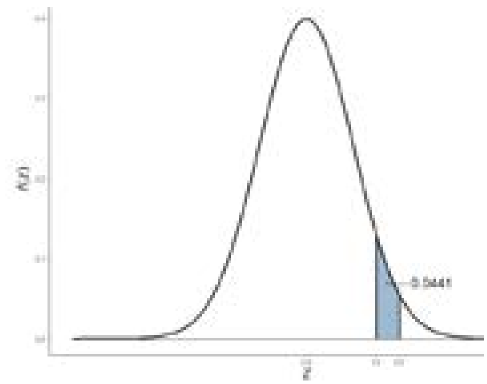


Figura 4.16. *Probabilidad, ejemplo 4.7, d)*

e) La probabilidad se calcula como

$$\begin{aligned} P[-2.2 < Z < 0.97] &= P[Z < 0.97] - P[Z < -2.2] \\ &= P[Z < 0.97] - (1 - P[Z < 2.2]) \\ &= 0.8340 - 1 + 0.9861 = 0.8201. \end{aligned}$$

Obsérvese la siguiente gráfica:

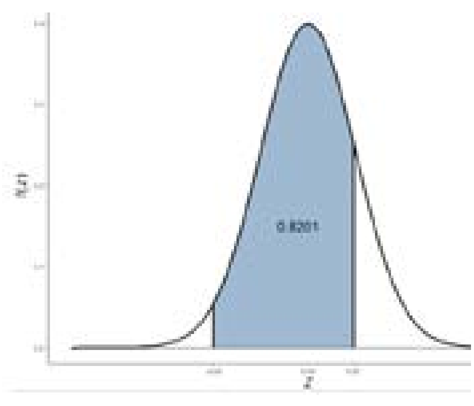


Figura 4.17. *Probabilidad, ejemplo 4.7, e)*

f) Obsérvese

$$\begin{aligned} P[-2.2 < Z < -1.5] &= P[1.5 < Z < 2.2] \\ &= P[Z < 2.2] - P[Z < 1.5] \\ &= 0.9861 - 0.9332 = 0.0529. \end{aligned}$$

En la siguiente gráfica, se puede ver esta probabilidad

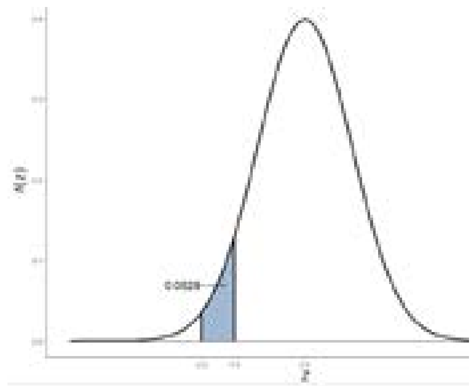


Figura 4.18. *Probabilidad, ejemplo 4.7, f)*

g) La búsqueda del 0.95-ésimo cuantil se hace de manera inversa en la tabla. Esto es, se busca la probabilidad acumulada más próxima a 0.95. Obsérvese que existen dos valores, 0.9495 y 0.9505, siendo 1.64 y 1.65, los cuantiles, respectivamente. Como ambas probabilidades quedan igualmente distantes a 0.95, el 0.95-ésimo cuantil se aproxima al promedio de ambos cuantiles, esto es, $z = \frac{1.64+1.65}{2} = 1.645$. Véase la siguiente gráfica.

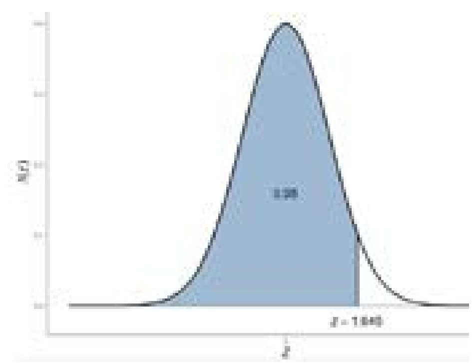


Figura 4.19. *Cuantil, ejemplo 4.7, g)*

Los cálculos anteriores se pueden realizar por medio de R, los códigos respectivamente se dan a continuación: $pnorm(1.5, 0, 1)$; $1 - pnorm(0.68, 0, 1)$; $pnorm(-1.2, 0, 1)$; $pnorm(2, 0, 1) - pnorm(1.5, 0, 1)$; $pnorm(0.97, 0, 1) - pnorm(-2.2, 0, 1)$; $pnorm(-1.5, 0, 1) - pnorm(-2.2, 0, 1)$ y $qnorm(0.95, 0, 1)$. Y las respuestas son: 0.9331928; 0.2482522; 0.1150697; 0.04405707; 0.8200733; 0.05290375 y 1.644854, respectivamente.

El siguiente ejercicio tiene que ver con el cálculo de probabilidades y cuantiles de una normal no estándar. Para calcular este tipo de probabilidades nos apoyaremos en el proceso de estandarización, véase teorema 4.19.

Ejemplo 4.8. Después de un estudio que se realizó sobre un bono financiero, se puede afirmar que la distribución del rendimiento mensual del bono tiene distribución normal con una media de 5.6 y una varianza de 1.44. Calcular para un mes dado

a) La probabilidad de que el rendimiento esté entre 2 y 3.2.

b) La probabilidad de que el rendimiento sea superior 4.64.

c) La constante c tal que la probabilidad de que el rendimiento sea menor a c, y sea igual a 0.45.

La variable aleatoria es

$X =$ Rendimiento mensual del bono.

$X \sim N(5.6, 1.44)$.

a) La primera probabilidad

$$\begin{aligned} P[2 < X < 3.2] &= P\left[\frac{2-5.6}{1.2} < \frac{X-5.6}{1.2} < \frac{3.2-5.6}{1.2}\right] \\ &= P[-3 < Z < -2] \\ &= P[2 < Z < 3] \\ &= 0.9987 - 0.9772 = 0.0215. \end{aligned}$$

b) La siguiente probabilidad

$$\begin{aligned} P[X > 4.64] &= \left[\frac{X-5.6}{1.2} > \frac{4.64-5.6}{1.2}\right] \\ &= P[Z > -0.8] \\ &= P[Z < 0.8] = 0.7881. \end{aligned}$$

c) Obsérvese

$$P[X < c] = 0.45.$$

Esto es,

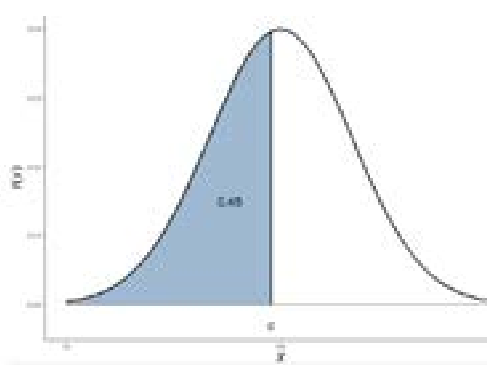


Figura 4.20. *Cuantil, ejemplo 4.8, c)*

Estandarizando

$$P\left[\frac{X-100}{10}<\frac{c-5.6}{1.2}\right]=0.45.$$

De esta manera,

$$P[Z<c']=0.45, \text{ donde } c'=\frac{c-5.6}{1.2}.$$

Obsérvese

$$P[Z>c']=0.55.$$

Esto es,

$$P[Z<-c']=0.55.$$

Buscando en la tabla de la normal estándar, el valor que más se aproxima es $-c'=0.13$, por lo tanto, $c=-0.13(1.2)+5.6=5.444$.

Para los resultados anteriores, por medio de R se utilizan, respectivamente, los siguientes códigos: $pnorm(3.2, 5.6, 1.2) - pnorm(2, 5.6, 1.2)$; $1 - pnorm(4.64, 5.6, 1.2)$ y $qnorm(0.45, 5.6, 1.2)$. Los resultados son 0.02140023; 0.7881446 y 5.449206, respectivamente.

Teorema 4.20. *Si X es una variable aleatoria con distribución normal con media μ y varianza σ^2, entonces,*

$$Moda = Mediana = Media.$$

Demostración. Por el teorema 4.18, se sabe que la media es μ.

Ahora, sea ν la mediana, entonces,

$$P[X\leq \nu]=0.5.$$

Esto es,

$$P\left[\frac{X-\mu}{\sigma}\leq\frac{\nu-\mu}{\sigma}\right]=P\left[Z\leq\frac{\nu-\mu}{\sigma}\right]=0.5.$$

Entonces,

$$\frac{\nu-\mu}{\sigma}=0.$$

Por lo tanto, $\nu=\mu$.

Por otro lado, se calculará la primera derivada de la función de densidad y se igualará a cero

$$\frac{d}{dx}f_X(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\left[-\frac{x-\mu}{\sigma^2}\right]=0.$$

De esta igualdad, se obtiene como punto crítico, $x=\mu$.

Obsérvese, de la primera derivada que

$$\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}} > 0.$$

Entonces, la primera derivada es positiva si y solo si $-\dfrac{x-\mu}{\sigma^2} > 0$, esto es, si $x < \mu$. Y es negativa si y solo si $-\dfrac{x-\mu}{\sigma^2} < 0$, esto es, si $x > \mu$.

Lo que significa que en $x = \mu$ hay un máximo, por lo tanto, la moda es $Moda = \mu$.

De esta manera, $Moda = Mediana = Media$. □

Teorema 4.21. *Si X es una variable aleatoria con distribución normal con media μ y varianza σ^2, entonces, los puntos de inflexión de la función de densidad se encuentran en $\mu - \sigma$ y $\mu + \sigma$.*

Demostración. Se obtendrá la segunda derivada de la función de densidad y esta se igualará a 0

$$\begin{aligned}\frac{d^2}{dx^2}f_X(x) &= \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\left[-\frac{1}{\sigma^2}\right] \\ &\quad + \left[-\frac{x-\mu}{\sigma^2}\right]\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\left[-\frac{x-\mu}{\sigma^2}\right] = 0.\end{aligned}$$

Esto es,

$$\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\left[-\frac{1}{\sigma^2}\right]\left[1 - \frac{(x-\mu)^2}{\sigma^2}\right] = 0.$$

Se puede observar

$$\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\left[-\frac{1}{\sigma^2}\right] \neq 0.$$

Por lo tanto,

$$1 - \frac{(x-\mu)^2}{\sigma^2} = 0.$$

Resolviendo la ecuación, $x = \mu \pm \sigma$.

Esto es, los puntos de inflexión se encuentran en $\mu - \sigma$ y $\mu + \sigma$. □

Observación. Si $X \sim N(\mu, \sigma^2)$, entonces, se cumplen las siguientes igualdades

$$P[\mu - \sigma < X < \mu + \sigma] = 0.6826,$$

$$P[\mu - 2\sigma < X < \mu + 2\sigma] = 0.9544,$$

$$P[\mu - 3\sigma < X < \mu + 3\sigma] = 0.9974.$$

Se calcularán las anteriores probabilidades

$$
\begin{aligned}
P[\mu-\sigma<X<\mu+\sigma] &= P\left[\frac{\mu-\sigma-\mu}{\sigma}<\frac{X-\mu}{\sigma}<\frac{\mu+\sigma-\mu}{\sigma}\right]\\
&= P[-1<Z<1]\\
&= 2P[Z<1]-1=0.6826.
\end{aligned}
$$

$$
\begin{aligned}
P[\mu-2\sigma<X<\mu+2\sigma] &= P\left[\frac{\mu-2\sigma-\mu}{\sigma}<\frac{X-\mu}{\sigma}<\frac{\mu+2\sigma-\mu}{\sigma}\right]\\
&= P[-2<Z<2]\\
&= 2P[Z<2]-1=0.9544.\\
P[\mu-3\sigma<X<\mu+3\sigma] &= P\left[\frac{\mu-3\sigma-\mu}{\sigma}<\frac{X-\mu}{\sigma}<\frac{\mu+3\sigma-\mu}{\sigma}\right]\\
&= P[-3<Z<3]\\
&= 2P[Z<3]-1=0.9974.
\end{aligned}
$$

Una interpretación de las anteriores probabilidades es que si una variable aleatoria tiene distribución normal con media μ y varianza σ^2, y es posible obtener en forma independiente varios de sus valores (muestra aleatoria), el 68 % aproximadamente de estos valores se encontrarán a lo más una desviación estándar de la media, por otro lado, el 95 % aproximadamente se encontrarán a lo más dos desviaciones estándar de la media, y casi todos los valores se encontrarán dentro de 3 desviaciones estándar de la media.

De acuerdo a lo anterior, para una variable aleatoria con distribución $N(\mu, \sigma^2)$, la probabilidad de que tome valores a una distancia más grande que 3 desviaciones estándar de la media es muy pequeña.

4.7. Distribución lognormal

A continuación, se va a definir la distribución lognormal, se presentarán sus principales propiedades y algunos ejemplos donde se ilustrarán las aplicaciones de este modelo.

Definición 4.12. Una variable aleatoria continua X tiene *distribución lognormal* con parámetros μ y σ, si su función de densidad está dada por

$$
f_X(x)=\frac{1}{x\sqrt{2\pi}\sigma}e^{-\frac{(ln(x)-\mu)^2}{2\sigma^2}}I_{(0,\infty)}(x),
$$

donde $-\infty<\mu<\infty$ y $\sigma>0$.

Notación. $X\sim Logn(\mu,\sigma^2)$ significa que la variable aleatoria X tiene una distribución lognormal con parámetros μ y σ.

La forma de la función de densidad de una distribución lognormal es

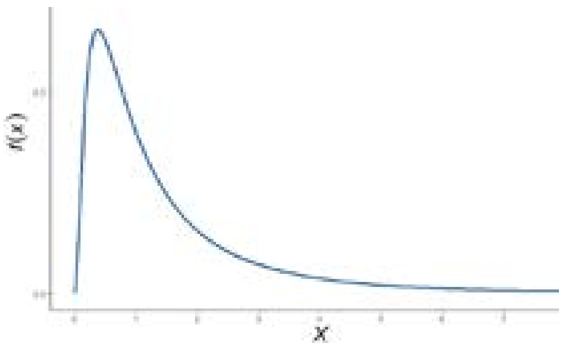


Figura 4.21. *Funciones de densidad de una distribución lognormal*

Observación. Una variable aleatoria con distribución lognormal se puede expresar en función de una variable aleatoria con distribución normal, esto es, se puede demostrar que si $Y \sim N(\mu, \sigma^2)$, entonces, $X = e^Y \sim Logn(\mu, \sigma^2)$.

Este resultado será útil para demostrar resultados propios de la distribución lognormal y también para el cálculo de probabilidades de esta distribución. La demostración de este resultado quedará pendiente, ya que se necesita de metodologías para encontrar distribuciones de funciones de variables aleatorias, material que se presentará en el capítulo 6.

Se puede demostrar que la función $f_X(x)$ de la anterior definición es de densidad, esto es,

$$\frac{1}{x} > 0; \ \frac{1}{\sqrt{2\pi}\sigma} > 0; \ e^{-\frac{(ln(x)-\mu)^2}{2\sigma^2}} > 0 \text{ y } I_{(0,\infty)}(x) \geq 0.$$

De esta manera,

$$f_X(x) = \frac{1}{x\sqrt{2\pi}\sigma} e^{-\frac{(ln(x)-\mu)^2}{2\sigma^2}} I_{(0,\infty)}(x) \geq 0.$$

Se resolverá la siguiente integral

$$\int_0^{\infty} \frac{1}{x\sqrt{2\pi}\sigma} e^{-\frac{(ln(x)-\mu)^2}{2\sigma^2}} dx.$$

Consideremos el siguiente cambio de variable $y = ln(x)$, de esta forma, $dy = \frac{1}{x}dx$. Además, se puede notar que el recorrido de Y son todos los números reales. Por lo tanto,

$$\int_0^{\infty} \frac{1}{x\sqrt{2\pi}\sigma} e^{-\frac{(ln(x)-\mu)^2}{2\sigma^2}} dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy = 1.$$

Por lo tanto, $f_X(x)$ es función de densidad.

Obsérvese que la última integral es la integral de una función de densidad de una distribución $N(\mu, \sigma^2)$, sobre todos los reales.

El siguiente teorema trata sobre la media y varianza de esta distribución.

Teorema 4.22. *Sea X una variable aleatoria con distribución lognormal con parámetros μ y σ^2, entonces,*

$$E[X] = e^{\mu+\frac{\sigma^2}{2}},$$
$$V[X] = e^{2\mu+2\sigma^2} - e^{2\mu+\sigma^2}.$$

Demostración. Se sabe

$$E[X] = E[e^Y], \text{ donde } Y \sim N(\mu, \sigma^2).$$

Entonces,

$$\begin{aligned} E[X] &= E[e^Y] \\ &= \int_{-\infty}^{\infty} e^y \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy. \end{aligned}$$

Esta última integral es un caso particular de la integral que se calculó para encontrar la función generadora de momentos de una distribución $N(\mu, \sigma^2)$ con $t = 1$, por lo tanto,

$$E[X] = e^{\mu+\frac{\sigma^2}{2}}.$$

Por otro lado,

$$\begin{aligned} E[X^2] &= E[e^{2Y}] \\ &= \int_{-\infty}^{\infty} e^{2y} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy. \end{aligned}$$

Obsérvese que esta última integral es un caso particular de la función generadora de momentos de la distribución $N(\mu, \sigma^2)$, cuando $t = 2$, de esta manera,

$$E[X^2] = e^{\mu(2)+\frac{\sigma^2(4)}{2}} = e^{2\mu+2\sigma^2}.$$

Por lo tanto,

$$\begin{aligned} V[X] &= e^{2\mu+2\sigma^2} - \left[e^{\mu+\frac{\sigma^2}{2}}\right]^2 \\ &= e^{2\mu+2\sigma^2} - e^{2\mu+\sigma^2}. \end{aligned}$$ □

Observación. Para una distribución lognormal, no existe la función generadora de momentos para valores de $t > 0$. Lo que significa que la distribución lognormal es de cola pesada .

Los momentos de la distribución lognormal no serán obtenidos a través de la función generadora de momentos, mas el siguiente teorema trata con el $k-$ésimo momento de esta distribución.

Teorema 4.23. *Sea X una variable aleatoria con distribución lognormal con parámetros μ y σ^2, entonces,*

$$E[X^k] = e^{k\mu + \frac{k^2\sigma^2}{2}}.$$

Se queda como ejercicio, la comprobación de este resultado.

La distribución lognormal tiene aplicaciones en diversas áreas, por mencionar algunas, en confiabilidad, en seguros, en finanzas. Siendo más específicos, puede modelar variables aleatorias, como los daños ocasionados por un siniestro, también el precio de una acción financiera, entre varios ejemplos. En general, puede modelar variables aleatorias continuas no negativas que pueden tomar valores muy grandes o valores extremos con una probabilidad no tan «pequeña».

Observación. Si X es una variable aleatoria con distribución lognormal con parámetros μ y σ, entonces, $E[ln(X)] = \mu$ y $V[ln(X)] = \sigma^2$.

Observación. El cálculo de valores de la función de distribución $F_X(x)$, de la lognormal, se puede realizar por medio de la función de distribución normal, esto es,

$$\begin{aligned} F_X(x) &= P[X \leq x] \\ &= P[ln(X) \leq ln(x)] \\ &= P[Y \leq ln(x)] \\ &= P\Big[Z \leq \frac{ln(x) - \mu}{\sigma}\Big], \end{aligned}$$

donde $Y \sim N(\mu, \sigma^2)$ y $Z \sim N(0, 1)$.

Usando el anterior resultado, también es posible calcular valores de la función de supervivencia $S_X(x)$.

Teorema 4.24. *Si X es una variable aleatoria con distribución lognormal con parámetros μ y σ, entonces, la moda M es igual a $M = e^{\mu - \sigma^2}$.*

La demostración se deja como ejercicio para el estudiante.

Ejemplo 4.9. La cantidad reclamada por un asegurado por un tipo de seguro es una variable aleatoria X la cual se distribuye lognormal con parámetros $\mu = 3$ y $\sigma^2 = 4$. Para un asegurado seleccionado al azar, calcular

a) La probabilidad de que reclame más de 200.

b) La probabilidad de que reclame entre 200 y 300.

c) La media y la desviación estándar de la cantidad reclamada.

$X =$ Cantidad reclamada por un asegurado.

$X \sim Logn(3, 4)$.

a) La primera probabilidad

$$\begin{aligned}
P[X > 200] &= 1 - P[X \leq 200] \\
&= 1 - P[ln(X) \leq ln(200)] \\
&= 1 - P[Y \leq ln(200)] \\
&= 1 - P\left[Z \leq \frac{ln(200) - 3}{2}\right] = 0.1251.
\end{aligned}$$

b) La segunda probabilidad

$$\begin{aligned}
P[200 \leq X \leq 300] =&P[ln(200) \leq Y \leq ln(300)] \\
&= P[\frac{ln(200) - 3}{2} \leq Z \leq \frac{ln(300) - 3}{2}] \\
&= P[1.15 \leq Z \leq 1.35] = 0.0366.
\end{aligned}$$

c) La media y la desviación estándar de la cantidad reclamada son, respectivamente, $E[X] = e^{\mu + \frac{\sigma^2}{2}} = 148.41316$ y $\sigma = \sqrt{e^{2\mu + 2\sigma^2} - e^{2\mu + \sigma^2}} = 1086.54398$.

Utilizando R, las instrucciones para calcular las probabilidades de los incisos a y b son, respectivamente, $1 - plnorm(200, 3, 2)$ y $plnorm(300, 3, 2) - plnorm(200, 3, 2)$, siendo los resultados 0.1252453 y 0.03704022, respectivamente.

Ejemplo 4.10. Se ha estudiado el comportamiento del precio de una acción financiera, y se sabe que en un periodo no muy largo, la distribución del precio de la acción se distribuye lognormal con media igual a 25 y varianza igual a 9. Calcular

a) El valor de sus parámetros μ y σ.

b) La probabilidad de que el precio sea superior a 20.

c) El 0.9-ésimo cuantil.

a) Obsérvese,

$$e^{\mu + \frac{\sigma^2}{2}} = 25,$$

de esta manera,

$$2\mu + \sigma^2 = 2ln(25) = ln(625).$$

Por otro lado,

$$e^{2\mu + 2\sigma^2} = 9 + 625 = 634,$$

entonces,

$$2\mu + 2\sigma^2 = ln(634).$$

Por lo tanto,

$$\sigma^2 = ln(634) - ln(625) = 0.014297,$$

$$\mu = ln(25) - 0.5(ln(634) - ln(625)) = 3.211727.$$

b) La probabilidad se calcula

$$\begin{aligned} P[X > 20] &= 1 - P[X \leq 20] \\ &= 1 - P[Y \leq ln(20)] \\ &= 1 - P\left[Z \leq \frac{ln(20) - 3.211727}{0.11957}\right] \\ &= 1 - P[Z \leq -1.8064] \\ &= P[Z \leq 1.8064] = 0.9649. \end{aligned}$$

c) El 0.9-ésimo cuantil= $\zeta_{0.9}$ se calcula de la siguiente manera:

$$F_X(\zeta_{0.9}) = 0.9.$$

Esto es,

$$\begin{aligned} F_X(\zeta_{0.9}) &= P[X \leq \zeta_{0.9}] \\ &= P[Y \leq ln(\zeta_{0.9})] \\ &= P\left[Z \leq \frac{ln(\zeta_{0.9}) - 3.211727}{0.11957}\right] = 0.9. \end{aligned}$$

Por lo tanto,

$$\frac{ln(\zeta_{0.9}) - 3.211727}{0.11957} = 1.28.$$

De esta manera,

$$ln(\zeta_{0.9}) = 1.28(0.11957) + 3.211727 = 3.364777.$$

En conclusión

$$\zeta_{0.9} = e^{3.364777} = 28.927046.$$

Por medio de R, la probabilidad se calcula $1-plnorm(20, 3.211727, 0.11957)$, donde la respuesta es 0.9645743. Para el cálculo del 0.90-ésimo cuantil, el código es $qlnorm(0.90, 3.211727, 0.11957)$, siendo el resultado 28.9324.

4.8. Distribución de Pareto

En esta parte se presentará la definición de la distribución de Pareto y sus principales propiedades, y se explicará la importancia de este modelo para variables aleatorias que toman «valores extremos».

Definición 4.13. Una variable aleatoria continua X tiene *distribución de Pareto* con parámetros a y b, si su función de densidad está dada por

$$f_X(x) = \frac{ab^a}{x^{a+1}} I_{(b,\infty)}(x), \text{ donde } a > 0 \text{ y } b > 0.$$

Notación. $X \sim Pareto(a, b)$ significa que la variable aleatoria X tiene una distribución de Pareto con parámetros a y b.

A continuación, se presenta algunas gráficas de la función de densidad de una distribución de Pareto para diferentes valores de sus parámetros.

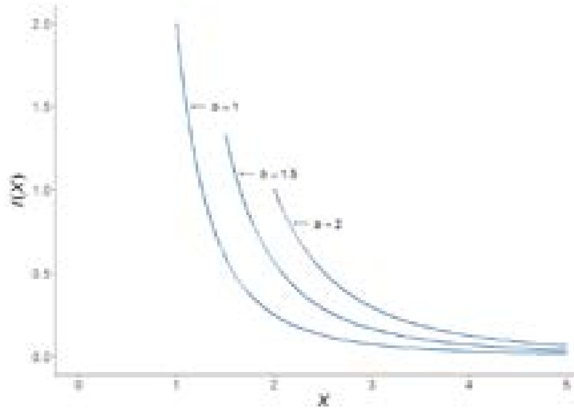


Figura 4.22. *Funciones de densidad de una distribución de Pareto. Diferentes valores de b y a fijo (a = 2)*

Se demostrará que la función $f_X(x)$ es de densidad.
Obsérvese que

$$a > 0;\ b^a > 0;\ \frac{1}{x^{a+1}} > 0 \text{ y } I_{(b,\infty)} \geq 0.$$

De este modo,

$$f_X(x) = \frac{ab^a}{x^{a+1}} I_{(b,\infty)}(x) \geq 0.$$

Además,

$$\begin{aligned} \int_b^\infty ab^a x^{-a-1} dx &= ab^a \int_b^\infty x^{-a-1} dx \\ &= -ab^a \frac{1}{a} x^{-a} \Big|_b^\infty = 1. \end{aligned}$$

Por lo tanto, $f_X(x)$ es función de densidad.

Observación. Es posible demostrar, para una distribución de Pareto, que la función generadora de momentos no existe para valores de $t > 0$, lo que significa que es una distribución de cola pesada .

No obstante, el $k-$ésimo momento con respecto al origen sí existe, véase siguiente teorema.

Teorema 4.25. *Si X tiene distribución de Pareto con parámetros a y b, entonces,*

$$E[X^k] = \frac{ab^k}{a-k}, \quad si \quad a > k.$$

La comprobación de este resultado se deja como ejercicio para el estudiante.

Teorema 4.26. *Sea X una variable aleatoria con distribución de Pareto con parámetros a y b, entonces,*

$$E[X] = \frac{ab}{a-1}, \ si \ a > 1,$$

$$V[X] = \frac{ab^2}{(a-1)^2(a-2)}, \ si \ a > 2.$$

La demostración del anterior teorema se deja como ejercicio para el estudiante. Se sugiere aplicar el teorema 4.25.

El siguiente teorema trata con la función de distribución de este modelo.

Teorema 4.27. *Sea X una variable aleatoria con distribución de Pareto con parámetros a y b, entonces, la función de distribución está dada por*

$$F_X(x) = \left\{1 - \left(\frac{b}{x}\right)^a\right\} I_{(b,\infty)}(x).$$

Demostración.

$$\begin{aligned} F_X(x) &= \int_b^x \frac{ab^a}{t^{a+1}} dt \\ &= ab^a \int_b^x t^{-a-1} dt \\ &= -b^a t^{-a} \Big|_b^x \\ &= -b^a x^{-a} + 1 \quad \text{si} \quad x > b. \end{aligned}$$

Por lo tanto,

$$F_X(x) = \left\{1 - \left(\tfrac{b}{x}\right)^a\right\} I_{(b,\infty)}(x). \qquad \square$$

De esta manera, la función de supervivencia está dada por

$$S_X(x) = \left(\frac{b}{x}\right)^a I_{(b,\infty)}(x).$$

Ejemplo 4.11. La cantidad reclamada mensualmente por una empresa a una aseguradora por las pólizas de seguros de gastos médicos mayores de sus empleados es una variable aleatoria X, la cual tiene una distribución de Pareto con parámetros $a = 20$ y $b = 20$. Calcular

a) La media y la desviación estándar de la cantidad reclamada por mes.

b) La probabilidad de que un reclamo sea más de 35, dado que se sabe que es mayor a 30.

c) A partir de qué cantidad, la probabilidad de tener un total de reclamos mayor a esta cantidad es con una probabilidad del 0.02.

a) La media, la varianza y la desviación estándar, respectivamente

$$E[X] = \frac{ab}{a-1} = \frac{400}{19} = 21.05263,$$

$$V[X] = \frac{ab^2}{(a-1)^2(a-2)} = \frac{(20)(400)}{(19^2)(18)} = 1.23115,$$

$$\sigma = 1.10957.$$

b) La probabilidad pedida se calcula de la siguiente manera:

$$\begin{aligned} P[X > 35|X > 30] &= \frac{P[X > 35]}{P[X > 30]} \\ &= \left(\frac{30}{35}\right)^{20} = 0.04582. \end{aligned}$$

c) La cantidad reclamada c debe de cumplir con la siguiente igualdad $P[X > c] = 0.02$, esto es,

$$\left(\frac{20}{c}\right)^{20} = 0.02.$$

Por lo tanto, $c = 24.320836$.

Entonces, la probabilidad de tener un reclamo mayor a 24.320836 es de 0.02.

4.9. Relaciones entre distribuciones

En esta sección se presentarán algunos resultados donde se relacionan dos distribuciones diferentes. Primero se presentarán y se comprobarán las relaciones que existen entre la distribución Poisson con las distribuciones exponencial y gamma. Posteriormente, se presentarán resultados que serán demostrados en el capítulo de convergencia, donde se relacionan las distribuciones Poisson, ji cuadrada y binomial con la distribución normal.

A continuación, se presentará la relación que hay entre la distribución Poisson y la distribución exponencial.

Teorema 4.28. *Se definen las variables aleatorias X = número de sucesos en un determinado intervalo de tiempo y Y = tiempo entre suceso y suceso (o tiempo hasta que ocurra el primer suceso). Si X tiene distribución Poisson con parámetro λ entonces, Y se distribuye exponencial con parámetro λ.*

Demostración. Supongamos que X tiene distribución $Poisson(\lambda)$.

Ahora, obsérvese que el recorrido de la variable Y es el intervalo $[0, \infty)$, y su función de distribución se calcula de la siguiente manera:

$$\begin{aligned} F_Y(y) &= P[Y \leq y] \\ &= 1 - P[Y > y] \\ &= 1 - P[\text{Número de sucesos en el intervalo } (0, y] \text{ es igual a } 0] \\ &= 1 - P[X = 0 \text{ (en un intervalo de longitud } y)] \\ &= 1 - \frac{(\lambda y)^0}{0!} e^{-\lambda y} \\ &= 1 - e^{-\lambda y}. \end{aligned}$$

Por lo tanto,

$$F_Y(y) = [1 - e^{-\lambda y}] I_{[0,\infty)}(y).$$

Esto es,

$$f_Y(y) = \lambda e^{-\lambda y} I_{[0,\infty)}(y).$$

Se concluye que Y se distribuye exponencial con parámetro λ. □

El siguiente resultado relaciona las distribuciones Poisson y gamma.

Teorema 4.29. *Se definen l as siguientes variables aleatorias X = número de sucesos en un determinado intervalo de tiempo, y Y = tiempo hasta observar el n-ésimo suceso. Si X tiene distribución Poisson con parámetro λ, entonces, Y se distribuye gamma (Erlang) con parámetros n y λ.*

Demostración. Supongamos que X tiene distribución $Poisson(\lambda)$.

Ahora obsérvese que el recorrido de Y es el intervalo $(0, \infty)$, además, su función de distribución se obtiene de la siguiente manera:

$$\begin{aligned} F_Y(y) &= P[Y \leq y] \\ &= 1 - P[Y > y] \\ &= 1 - P[\text{ Número de sucesos en el intervalo } (0, y] \leq n - 1] \\ &= 1 - P[X \leq n - 1 (\text{ en un intervalo de longitud } y)] \\ &= 1 - \sum_{i=0}^{n-1} \frac{(\lambda y)^i}{i!} e^{-\lambda y}. \end{aligned}$$

Por lo tanto,

$$F_Y(y) = \left(1 - \sum_{i=0}^{n-1} \frac{(\lambda y)^i}{i!} e^{-\lambda y}\right) I_{(0,\infty)}(y).$$

Obsérvese que esta es la función de distribución de una $Gamma(n, \lambda)$, véase teorema 4.4.

Se concluye que Y se distribuye gamma con parámetros n y λ. $\square$

Ejemplo 4.12. El número de artículos vendidos en un almacén es una variable aleatoria con distribución Poisson con un promedio de 4 artículos vendidos en una hora. Calcular la probabilidad de que
a) El primer artículo vendido sea antes de la media hora.
b) El segundo artículo vendido sea después de una hora.

a) Obsérvese
Y = Tiempo en el cual es vendido el primer artículo.
$Y \sim Exp(\lambda = 4)$.

Cabe mencionar que el tiempo promedio en el cual se vende el primer artículo es de 0.25 horas, esto es, el tiempo promedio es de 15 minutos.

Ahora,

$$\begin{aligned} P[Y < 0.5] &= F_Y(0.5) \\ &= 1 - e^{-4(0.5)} = 0.86466. \end{aligned}$$

b) En este caso,
Y = Tiempo en el cual se venden dos artículos.
$Y \sim Gamma(r = 2, \lambda = 4)$.

El tiempo promedio para vender dos artículos es 0.5 horas, esto es, de media hora.

De esta manera,

$$P[Y > 1] = \sum_{i=0}^{1} \frac{(4(1))^i}{i!} e^{-4(1)} = 0.091578.$$

Los dos resultados anteriores se calculan por medio de R de la siguiente manera, respectivamente: $pexp(0.5, 4)$ y $1 - pgamma(1, 2, 4)$, donde los resultados son 0.8646647 y 0.09157819 respectivamente.

El siguiente teorema se puede ver como el inverso de los anteriores resultados.

Teorema 4.30. *Se definen las siguientes variables aleatorias, X = número de sucesos en un determinado intervalo de tiempo (en una unidad de tiempo), Y = tiempo hasta que ocurra el n-ésimo suceso, y Y_i = tiempo entre el suceso i y el suceso $i - 1$, donde el suceso 0 es el inicio (momento cero). Si $Y_1, Y_2, \ldots, Y_n$ son variables aleatorias independientes con distribución común exponencial con parámetro λ, entonces, X se distribuye Poisson con parámetro λ.*

No se realizará la demostración de este teorema, ya que se necesita de teoría que por el momento no se ha visto. En particular, es necesario usar la definición de variables aleatorias independientes, la cual se verá en el siguiente capítulo, definición 5.8, además, se debe usar el siguiente resultado: la suma de n variables aleatorias independientes con distribución común exponencial con parámetro λ tiene distribución gamma con parámetros n y λ (resultado que se presentará en el capítulo 6).

Cabe mencionar que el anterior resultado tiene una aplicación muy concreta en simulación estocástica, es usado para generar valores de una variable aleatoria Poisson, por medio de la distribución exponencial.

Las siguientes relaciones entre distribuciones se comprobarán en el capítulo de convergencia de sucesiones de variables aleatorias, cada uno de estos resultados es una aproximación de la distribución normal a otras distribuciones, como la distribución Poisson, ji cuadrada y binomial.

Para propósitos de explicación, la función de distribución de una normal estándar se denotará como $\Phi(z)$.

Primero se presentará la relación entre la distribución Poisson y la distribución normal. Este resultado ayudará a aproximar probabilidades de la distribución Poisson por medio de la distribución normal.

Si X es una variable aleatoria con distribución Poisson con parámetro λ y $Z = \frac{X-\lambda}{\sqrt{\lambda}}$, entonces, $F_Z(z) \xrightarrow[\lambda\to\infty]{} \Phi(z)$, donde $a \xrightarrow[c\to\infty]{} b$ significa: a converge a b cuando $c \to \infty$.

La siguiente relación es entre la distribución ji-cuadrada y la distribución normal. Este resultado servirá para aproximar probabilidades de la ji cuadrada por medio de la distribución normal.

Si X es una variable aleatoria con distribución ji cuadrada con ν grados de libertad y $Z = \frac{X-\nu}{\sqrt{2\nu}}$, entonces, $F_Z(z) \xrightarrow[\nu\to\infty]{} \Phi(z)$.

Para ilustrar el resultado anterior, se puede apreciar en la gráfica 4.4, cómo la forma de la función de densidad de la distribución ji cuadrada va tendiendo a una gráfica en forma de «campana», cuando los grados de libertad van creciendo.

El siguiente resultado trata sobre la convergencia de una distribución binomial a una distribución normal. La comprobación de este resultado es una consecuencia del teorema de límite central, el cual se presentará y se demostrará también en el capítulo de convergencia.

Si X es una variable aleatoria con distribución binomial con parámetros n y p y $Z = \frac{X-np}{\sqrt{np(1-p)}}$, entonces, $F_Z(z) \xrightarrow[n\to\infty]{} \Phi(z)$.

Este último resultado, además de aproximar probabilidades de una distribución binomial por una normal, también tiene aplicaciones en inferencia estadística, en particular, en la construcción de intervalos de confianza y pruebas de hipótesis para una proporción p.

Por último, recordemos que en el capítulo anterior, se presentaron dos relaciones más entre distribuciones, la convergencia de la distribución hipergeométrica a la distribución binomial (véase teorema 3.7) y la convergencia de la distribución binomial a la distribución Poisson (véase teorema 3.18).

4.10. Distribuciones usadas en estadística

Existen distribuciones que son usadas con mucha frecuencia en estadística, como la distribución normal, la cual ya se presentó con sus principales propiedades, la distribución ji cuadrada, la cual es un caso particular de la distribución gamma, que también ya se presentó. En esta sección se presentarán otras dos distribuciones, la distribución t de student y la distribución F, las cuales son también muy usadas en diferentes temas de estadística.

Definición 4.14. Una variable aleatoria continua X tiene *distribución t de student* con ν grados de libertad, si su función de densidad está dada por

$$f_X(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{1}{\sqrt{\pi\nu}\left(1+\frac{x^2}{\nu}\right)^{\frac{\nu+1}{2}}}, \quad \text{donde} \quad \nu = 1, 2, \ldots$$

Notación. $X \sim tstudent(\nu)$ significa que la variable aleatoria X tiene distribución t de student con ν grados de libertad.

La forma de la gráfica de la función de densidad de la distribución t de student, es una campana como la distribución normal estándar, pero con colas más «pesadas». Véase la siguiente gráfica donde se compara varias funciones de densidad con distribución t de student y la función de densidad de la normal estándar.

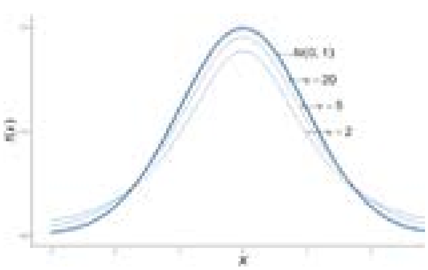


Figura 4.23. *Funciones de densidad de t-student y normal estándar*

Se puede observar que entre más grande sean los grados de libertad, la gráfica de la función de densidad de una distribución t de student se parece más a la gráfica de la función de densidad normal estándar.

Teorema 4.31. *Sea X una variable aleatoria con distribución t de student con ν grados de libertad, entonces,*

$$E[X] = 0, \quad si \quad \nu > 1,$$
$$V[X] = \frac{\nu}{\nu - 2}, \quad si \quad \nu > 2.$$

Observación. Para la distribución t de student no existe la función generadora de momentos.

Esta distribución es muy usada en inferencia estadística para construir intervalos de confianza y pruebas de hipótesis para un promedio o para la diferencia de dos promedios, entre otras aplicaciones.

A continuación, se presentará la definición de la distribución F.

Definición 4.15. Una variable aleatoria continua X tiene *distribución F* con ν_1 grados de libertad en el numerador y ν_2 grados de libertad en el denominador, si su función de densidad está dada por

$$f_X(x) = \frac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)}\left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}} \frac{(x)^{\frac{\nu_1}{2}-1}}{\left(1+\frac{\nu_1 x}{\nu_2}\right)^{\frac{\nu_1+\nu_2}{2}}} I_{(0,\infty)}(x),$$

donde $\nu_1, \nu_2 = 1, 2, \ldots$

Notación. $X \sim F(\nu_1, \nu_2)$ significa que la variable aleatoria X tiene distribución F con ν_1 grados de libertad en el numerador y ν_2 grados de libertad en el denominador.

La forma general de la gráfica de una función de densidad de una distribución F se da a continuación:

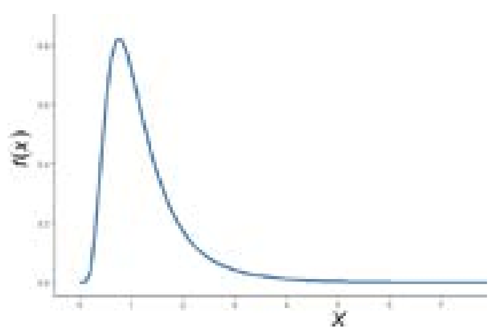


Figura 4.24. *Función de densidad. Distribución F*

Teorema 4.32. *Sea X una variable aleatoria con distribución F con ν_1 grados de libertad en el numerador y ν_2 grados de libertad en el denominador, entonces,*

$$E[X] = \frac{\nu_2}{\nu_2 - 2}, \quad si \quad \nu_2 > 2,$$

$$V[X] = \frac{2\nu_2^2(\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2 - 2)^2(\nu_2 - 4)}, \quad si \quad \nu_2 > 4.$$

Observación. Para la distribución F no existe la función generadora de momentos.

La distribución F tiene varias aplicaciones en inferencia estadística, en particular, cuando se comparan estadísticamente dos varianzas de poblaciones diferentes. También, en problemas de análisis de varianza, donde se comparan mutuamente varios promedios de diferentes poblaciones.

Tanto el cálculo de probabilidades como el cálculo de cuantiles de estas dos distribuciones se realizarán por medio de tablas especiales. Para la distribución t de student se puede encontrar una tabla en el apéndice A, donde se presentan para algunas probabilidades fijas diferentes cuantiles, considerando diversos grados de libertad. Para la distribución F, también en el apéndice A, se encontrarán para algunas probabilidades diferentes cuantiles, considerando diversos grados de libertad tanto en el numerador como en el denominador.

Ejercicios del capítulo 4

1.- Una pelota se lanza aleatoriamente entre los puntos **A** y **B**. Calcular la probabilidad de que

a) La pelota quede más cerca de **B** que de **A**.
b) La distancia con respecto a **A** sea más de cuatro veces la distancia con respecto a **B**.

2.- Se sabe que la tasa de interés anual de un crédito es una variable aleatoria con distribución normal con media 15 y varianza 3.2. Para un año determinado, calcular

a) La probabilidad de que la tasa sea mayor al 18 %.
b) La probabilidad de que la tasa sea menor al 14 %.
c) La probabilidad de que la tasa sea mayor que el 15 %, dado que se sabe que la tasa es menor al 17 %.
d) El 0.4−ésimo cuantil.

3.- Una pieza fabricada no es útil si tiene un diámetro menor a 90 milímetros o es superior a los 105 milímetros. Se sabe que el diámetro en milímetros de este tipo de piezas se distribuye normal con media y varianza 100 y 4.5 respectivamente. Calcular la probabilidad de que

a) Una pieza no sea útil.
b) Todas las piezas satisfagan las especificaciones de calidad, de 5 piezas seleccionadas al azar.
c) Al menos dos piezas no satisfagan las especificaciones, de 10 piezas seleccionadas al azar.

4.- Para la fabricación de un artículo se requiere de cierta cantidad de un producto. La cantidad del producto utilizado en un día es una variable aleatoria con distribución exponencial con media igual a 2.5 toneladas.

a) Calcular la probabilidad de que la empresa utilice más de 3 toneladas en un día determinado.
b) Calcular la mediana de la cantidad de producto usado en un día.
c) ¿Qué cantidad del producto se tendrá que almacenar, para que no se agote en un día dado, con una probabilidad del 0.96?

5.- Sea X una variable aleatoria con distribución $Gamma(r, \lambda)$, encontrar la media y la varianza de X.

6.- La variación en minutos del momento en que llega un alumno a clases es una variable aleatoria X, con función de densidad $f_X(x) = cI_{[-5,5]}(x)$. Obtener

a) El valor de c.
b) La función de distribución de X.
c) $E[X]$ y $V[X]$.
d) $P[X < -2|X > -4]$.
e) El 0.80-ésimo cuantil.

7.- El líquido en mililitros que arroja una máquina de refrescos es una variable aleatoria con distribución normal μ y varianza 9. La máquina se puede ajustar para lograr el valor de μ deseado.

a) Encontrar el valor de μ apropiado, si se desea que en vasos de 250 mililitros, solo se derrame el 1 % de las veces.
b) Con el valor de μ encontrado en a), calcular la probabilidad de que el líquido arrojado para un vaso sea menor a 243 mililitros.

8.- Sea X una variable aleatoria con distribución normal con parámetros μ y σ^2, encontrar la media y la varianza de X.

9.- El porcentaje de impurezas de cierto producto líquido por envase, es una variable aleatoria X con distribución $Beta(a = 2, b = 4)$. Un envase con más de 40 % de líquido impuro no se puede vender. Calcular

a) La media y la varianza de X.
b) La moda de X.
c) El 0.2-ésimo cuantil.
d) La probabilidad de que un envase de líquido seleccionada al azar no se pueda vender.
e) La probabilidad de que a lo más un envase de líquido se pueda vender, de 3 envases que se seleccionaron al azar.

10.- Si X es una variable aleatoria con distribución normal con media y varianza igual a 100 cada una, calcular $P[|X - 100| < 20]$.

11.- Sea X una variable aleatoria con distribución de Pareto con parámetros a y b, encontrar la media y la varianza de X.

12.- Si X es una variable aleatoria con distribución $Gamma(r = 3, \lambda = 2)$. Usando la función de distribución, calcular

a) $P[X > 1]$.
b) $P[]2 \leq X \leq 5]$.

13.- La proporción de gasto de mantenimiento mensual para una máquina, de acuerdo a lo presupuestado, es una variable aleatoria X, la cual se distribuye beta con parámetros $a = 4$ y $b = 1$.

a) Calcular la probabilidad de que la proporción de gasto en reparación en un mes sea mayor al 85 %.
b) Encontrar el 0.90-cuantil de la proporción de gasto mensual.

14.- La duración en horas de cierto dispositivo electrónico es una variable aleatoria X con distribución exponencial con media igual a 120. Cuatro de estos dispositivos trabajan en forma independientemente en un equipo. El equipo falla si al menos dos de los componentes fallan. Calcular

a) La probabilidad de que el equipo funcione menos de 100 horas.
b) La probabilidad de que el equipo funcione al menos durante 200 horas.
c) La media del número de componentes que duran al menos 200 horas.

15.- Si X es una variable aleatoria que se distribuye uniforme sobre el intervalo $(-3, 2)$, encontrar c tal que $P[X > c - E(X)] = 0.15$.

16.- La vida en horas de un dispositivo es una variable aleatoria X, con distribución $N(200, \sigma^2)$. Si un cliente requiere que el 90 % de los dispositivos tenga una vida de por lo menos 180 horas, ¿cuál es de la desviación estándar σ que satisface las necesidades del cliente?

17.- El número de automóviles que pasan por un crucero durante un intervalo de un minuto, tiene una distribución Poisson con media igual 2.

a) Sea Y el tiempo necesario para que el primer automóvil pase por el crucero. Calcular la probabilidad de que Y sea menor a 2 minutos.
b) Sea Y el tiempo necesario para que el cuarto auto pase por el crucero. Calcular la probabilidad de que Y sea superior a 5 minutos.

18.- Si X es una variable aleatoria con distribución beta, ¿es posible que $E\left[\frac{1}{X}\right] = 1$? Justificar la respuesta.

19.- Contestar las siguientes preguntas:

a) Si X es una variable aleatoria con distribución $U(5, b)$, donde $b > 5$ y se cumple $E[X] = 8V[X]$, calcular el valor de b.
b) Si X es una variable aleatoria con distribución $Exp(\lambda = 5)$, calcular la mediana.
c) Si X es una variable aleatoria con distribución $N(\mu, \sigma^2)$, y que cumple $P[X < 80] = P[X > 120] = 0.25$, calcular los valores de μ y σ.

20.- Si X es una variable aleatoria con distribución normal con media y varianza 0.5 y 1 respectivamente, calcular $P[4 < X^2 < 9]$.

21.- Sea X una variable aleatoria con distribución lognormal con parámetros μ y σ^2, demostrar que $E[X^r] = e^{r\mu + r^2\sigma^2/2}$.

22.- Supongamos que la cantidad reclamada en un seguro de automóvil es una variable aleatoria X con distribución lognormal con parámetros $\mu = 4$ y $\sigma = 1$. Calcular

a) La media y la desviación estándar de X.
b) La probabilidad de que se reclame menos de 150.
c) La probabilidad de que se reclame entre 100 y 250.
d) La probabilidad de que se reclame más de 200 dado que se sabe que se reclamaron más de 150.

23.- Suponga que el tiempo de espera de un cliente en un restaurante para que le sirvan lo que ordenó es una variable aleatoria con distribución exponencial con media igual a 15 minutos. Se supone que los tiempos de espera entre diferentes clientes son independientes.

a) Calcular la probabilidad de que un cliente sea atendido después de media hora.
b) A los clientes que esperan un tiempo muy prolongado, el consumo será gratis. El dueño del restaurante está dispuesto a pagar el 3 % de todos los consumos por exceso de espera. ¿A partir de qué tiempo de espera un cliente no pagará su consumo?
c) Para 5 clientes seleccionados al azar, calcular la probabilidad de que al menos uno de ellos no pague la cuenta.

24.- Sea X una variable aleatoria con recorrido los números reales y con función de densidad dada por $f_X(x) = k2^{-x^2}$. Encontrar k. *Recomendación: Expresar* $2 = e^{ln 2}$.

25.- Los rendimientos de una acción es una variable aleatoria X con distribución $N(\mu, \sigma^2)$. Si el 0.46-ésimo cuantil es igual 12 y $P[X > 15] = 0.32$. Encontrar la media y desviación estándar de los rendimientos.

26.- Demostrar el teorema 4.12.

27.- Sean X_1, X_2 y X_3 variables aleatorias independientes con distribuciones $N(0.5, 1)$, $N(1, 4)$ y $N(-0.5, 1)$, respectivamente. Calcular la probabilidad de que

a) Exactamente dos de estas variables sean menores que cero.
b) Por lo menos una de estas variables sea menor que cero.

28.- Las ganancias logradas en una empresa en un día, tiene distribución de Pareto, con media 80 y desviación estándar 90.

a) Calcular la probabilidad de que la empresa tenga una ganancia en un día, por más de 200 y menos de 300.
b) Encontrar el valor c, tal que la probabilidad de que las ganancias diarias de la empresa sean menores a c sea igual a 0.95.

29.- Si X es una variable aleatoria con distribución uniforme sobre el intervalo (a, b), tal que $E[X] = 0.5$ y $V[X] = 25/12$, encontrar a y b.

30.- La proporción de tiempo que una máquina está detenida en un mes es una variable aleatoria X con distribución $Beta(1, 3)$.

a) Calcular la media y la varianza de X.
b) Calcular la probabilidad de que la proporción de tiempo que la máquina esté parada en un mes dado sea superior al 50 %.
c) Encontrar τ tal que $P[X \leq \tau] = 0.85$.
d) El costo C de una interrupción por reparación está dado por $C = 20 + 8X + X^2$. Encontrar la media y la desviación estándar de C.

31.- El número de clientes que llegan a un servicio se distribuye Poisson con media igual a 5 clientes por hora. Calcular la probabilidad de que

a) Llegue el primer cliente antes de media hora.
b) Llegue el décimo cliente antes de los 90 minutos.

32.- Una caja es fabricada con una altura de 10 cm y con una base de X por $1.5X$ cm cuadrados. Si X se distribuye $U(19.5, 20.5)$, entonces, ¿cuál es el volumen esperado de la caja en cm cúbicos?

33.- Si $X \sim N(6, 16)$, calcular la probabilidad $P[X^2 - 12X \leq -16]$.

34.- Por un seguro médico, un asegurado puede reclamar hasta 4 siniestros en un año. La cantidad reclamada por un asegurado al año tiene distribuciones uniformes, sobre los intervalos $(50, 100)$, $(80, 150)$, $(120, 200)$ y $(150, 200)$, si el número de reclamos al año son $1, 2, 3, 4$, respectivamente. Por otro lado, las probabilidades de que un asegurado realice $0, 1, 2, 3$ y 4 reclamos al año son, respectivamente, $0.6, 0.15, 0.12, 0.1$ y 0.03. Calcular

a) La probabilidad de que un asegurado reclame más de 110 pero menos de 160 en un año.
b) La media y la varianza de la cantidad reclamada por un asegurado.

35.- Supongamos que X es una variable aleatoria con distribución beta con parámetros $a = k+2$ y $b = k+1$, calcular la moda de X y qué condición debe cumplir k.

36.- Si X es una variable aleatoria con distribución exponencial con parámetro λ, demostrar que $E[X^k] = \frac{k!}{\lambda^k}$, para $k = 1, 2, ...$

37.- La cantidad vendida en una tienda en un día es una variable aleatoria X con distribución de Weibull con parámetros $a = 50$ y $b = 1/3$. Para un día dado, calcular

a) La probabilidad de que se venda más de 400.
b) La probabilidad de que se venda entre 300 y 350.
c) La probabilidad de que se venda menos de 500, dado que se sabe que se vendió más de 300.
d) La media y la desviación estándar de la cantidad vendida en un día.

38.- La cantidad reclamada mensualmente por una póliza de seguro es una variable aleatoria con distribución de Pareto con parámetro $a = 2$ y media igual a 200. Encontrar la mediana de la distribución.

39.- Sea X una variable aleatoria con distribución uniforme sobre el intervalo (a, b), encontrar el k-ésimo momento con respecto al origen.

40.- Sea X una variable aleatoria con distribución $N(\mu, \sigma^2)$, tal que cumple $P[X < 49] = 0.80$ y $P[X < 64] = 0.9$. Encontrar μ y σ^2.

41.- Sea X una variable aleatoria con distribución lognormal con parámetros μ y σ^2. Demostrar que la moda es $M = e^{\mu-\sigma^2}$.

42.- Si X se distribuye beta con parámetros $a = 6$ y $b = 7$, ¿cuál es el valor esperado de $(1 - X)^{-5}$?, ¿cuál es el valor esperado de X^4?

43.- Demostrar el teorema 4.4.

44.- Sea X una variable aleatoria, encontrar la función de tasa de fallos e interpretarla, si

a) X se distribuye exponencial con parámetro λ.
b) X se distribuye Weibull con parámetros a y b.

45.- Se sabe que el precio de una acción es una variable aleatoria con distribución gamma con parámetros $r = 3$ y $\lambda = \frac{1}{\beta}$. Si la moda del precio de la acción es 4, calcular

a) El valor de β.
b) La media y la varianza del precio de la acción.
c) La probabilidad de que el precio de la acción sea menor a 8.5.
d) El 0.9-ésimo cuantil.

46.- Sea X una variable aleatoria la cual se distribuye $Beta(2, 2)$, calcular $P[|X - 0.3| > 0.2]$.

47.- La garantía de un artículo es de la siguiente manera: si el artículo falla en los primeros 3 meses, se reembolsa el 100 % del valor, si la falla es entre el tercer y sexto mes, se regresará el 80 %, en caso de que falle durante el segundo semestre, el reembolso es del 50 % y si el artículo deja de funcionar durante el segundo año, se pagará el 20 %. De otra manera, no hay reembolso. Se sabe que la duración del artículo tiene distribución exponencial con promedio de 3 años. Calcular el dinero esperado de la garantía por 100 artículos adquiridos, siendo el costo unitario de 50.

48.- Sea X una variable aleatoria tal que sus momentos con respecto al origen están dados de la siguiente manera: $E[X^{2m}] = \frac{(2m)!}{2^m m!}$ y $E[X^{2m-1}] = 0$, para $m = 1, 2, 3, \ldots$ Encontrar la función generadora de momentos y la función de densidad de X.

49.- Sea X una variable aleatoria con distribución uniforme sobre un intervalo de longitud $\frac{1}{2}$, la cual cumple con $P[0.2 < X < 0.4] = 0.4$. ¿Cuál es el rango de valores que puede tomar el límite inferior del intervalo de la distribución?, ¿y el rango de valores del límite superior?

50.- Un inversionista crea un portafolio con 4 instrumentos de inversión: 1, 2, 3 y 4, cada uno con una tasa variable. El 40 % de la inversión lo coloca en el instrumento 1, el 20 % en el 2, el 15 % en el 3 y el resto en el instrumento 4. La tasa de interés anual en porcentajes de cada instrumento se distribuye normal, con medias 6.2, 10.5, 15.2 y 12.5, respectivamente, y con varianzas 1.2, 2,4, 8.4 y 3, respectivamente. Calcular

a) La tasa de interés promedio del portafolio.
b) La varianza del portafolio.
c) La probabilidad de que la tasa anual del portafolio para un determinado año sea menor a 10.

51.- Calcular e interpretar la función de tasa de fallos para una distribución de Pareto con parámetros a y b.

52.- Si $m_X(t) = (1 - \frac{t}{3})^{-6}$ con $t < 3$ es la función generadora de momentos de una variable aleatoria X, encontrar $P[X < 4.5]$.

53.- El tiempo en horas que un trabajador tarda en terminar un trabajo, se distribuye beta con parámetros $a = 3$ y $b = 2$. El trabajador será compensado con un bono, si esta tarea la termina en forma rápida, esto es, si la termina antes de 15 minutos recibirá un bono de 100, si la termina entre 15 y 25 minutos recibirá un bono de 50 y si termina entre 25 y 30 minutos será compensado con 20, en cualquier otro caso, no recibirá bono alguno. Calcular el bono promedio que recibirá el trabajador.

54.- Una persona compra una lavadora con un costo de 42. La vida en años de esta lavadora se distribuye exponencial con media igual a 3. La empresa ofrece 3 tipos de garantías para que el cliente seleccione una a la hora de la compra. La garantía 1 le ofrece al cliente el reembolso del 50 % de la lavadora, si esta deja de funcionar durante el primer año de uso, y el 25 % si se descompone durante el segundo año, de otro modo, no hay reembolso. La garantía 2, consiste en pagar $42e^{-2x}$, donde x es el momento en que la lavadora deja de funcionar. Y la garantía 3 ofrece pagar el costo total, si esta se descompone durante los primeros 6 meses, de otra manera no hay reembolso. ¿Qué garantía le conviene más al cliente?

55.- Sea X una variable aleatoria con distribución normal estándar, encontrar la función de densidad truncada a la izquierda de 0.

56.- Sea X una variable aleatoria con distribución $Exp(\lambda)$, encontrar la función de densidad truncada a la izquierda de θ, donde $\theta > 0$.

57.- Sea X una variable aleatoria con distribución $U(a, b)$, encontrar la función de densidad truncada en el intervalo (c, d), donde $a < c < d < b$.

58.- Sea X una variable aleatoria con distribución $N(\mu, \sigma^2)$. Indicar cómo obtener la función de densidad truncada en el conjunto $(-\infty, 0) \cup (1, \infty)$.

59.- Sea X una variable aleatoria con distribución $Weibull(a, b)$, encontrar la función de densidad truncada en $(0, c) \cup (d, \infty)$, donde $0 < c < d$.

60.- Demostrar el teorema 4.25.

61.- El retorno de inversión mensual en porcentaje es una variable aleatoria con distribución $N(0.8, 0.09)$. Se sabe que las inversiones entre dos meses distintos son independientes. Se observa cada mes el retorno de inversión.

a) Para los 8 meses siguientes, calcular la probabilidad de observar al menos dos meses con retorno negativo.
b) Calcular la probabilidad de observar por tercera vez un mes con retorno negativo, después del sexto mes observado.
c) Hasta el quinto mes, se ha observado retornos de inversión positivos, calcular la probabilidad de observar por primera vez un retorno negativo antes del décimo mes.

62.- Para un determinado siniestro, es preocupante para una compañía de seguros, los reclamos muy grandes. El reclamo de un asegurado por este siniestro es una variable aleatoria X con distribución $Weibull(200, 0.8)$. Se supone independencia entre siniestro y siniestro.

a) Encontrar c tal que $X > c$ ocurre con una probabilidad de 0.05.
b) Calcular la probabilidad de que el primer siniestro con un reclamo superior a 1,000, sea antes del noveno siniestro.
c) Calcular la probabilidad de que el tercer siniestro con un reclamo superior a 1,000, sea antes del sexto siniestro.

Vectores aleatorios

Capítulo

Introducción

En este capítulo se estudiarán las distribuciones para vectores aleatorios, dicho de otra manera, las distribuciones conjuntas de dos o más variables aleatorias. Se obtendrán, a partir de la densidad conjunta, las funciones de densidad marginales y condicionales. Se presentará la definición de valor esperado de una función de un vector aleatorio, enfatizando en los casos particulares de medias, varianzas, y para dos variables, se darán las definiciones de covarianza y de coeficiente de correlación, con algunas interpretaciones prácticas. Posteriormente, se dará la definición de función generadora de momentos de un vector aleatorio y sus propiedades más importantes. Se tratará el concepto de variables aleatorias independientes, explicando su importancia en varias áreas de aplicación.

Un *vector* se dice que es *aleatorio*, si al menos uno de sus componentes es una variable aleatoria.

Las letras mayúsculas subrayadas representarán vectores y las letras minúsculas subrayadas representarán un punto específico del vector. Las letras no subrayadas seguirán representando variables aleatorias o algún valor específico, según sea el caso.

Esto es, un vector aleatorio se denotará como $\underline{X} = (X_1, X_2, \ldots, X_n)$.

Un punto específico del vector se representará como $\underline{x} = (x_1, x_2, \ldots, x_n)$. Siendo más específicos

$$\begin{aligned}\underline{x} = (x_1, x_2, \ldots, x_n) &= (X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)\\ &= ((X_1 = x_1) \cap (X_2 = x_2) \cap \cdots \cap (X_n = x_n)).\end{aligned}$$

También, será de interés el siguiente evento

$$\begin{aligned}&(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n)\\ &\qquad = ((X_1 \leq x_1) \cap (X_2 \leq x_2) \cap \cdots \cap (X_n \leq x_n)).\end{aligned}$$

En el caso particular de $\mathbb{R}^2$, el evento $(X_1 \leq a_1, X_2 \leq a_2)$ tiene la siguiente representación geométrica:

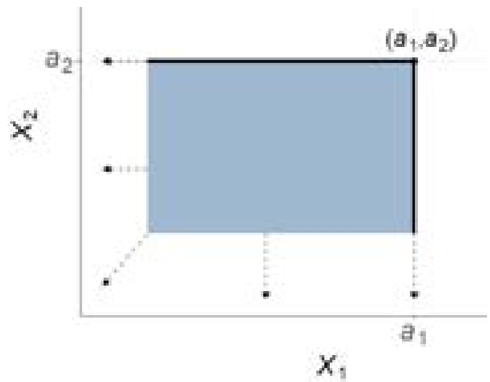


Figura 5.1. *Región de* $\{X_1 \leq a_1, X_2 \leq a_2\}$

El evento $(a_1 < X_1 \leq b_1, a_2 < X_2 \leq b_2, \ldots, a_n < X_n \leq b_n)$, también será de interés. En el caso particular de $(a_1 < X_1 \leq b_1, a_2 < X_2 \leq b_2)$ tiene la siguiente interpretación geométrica:

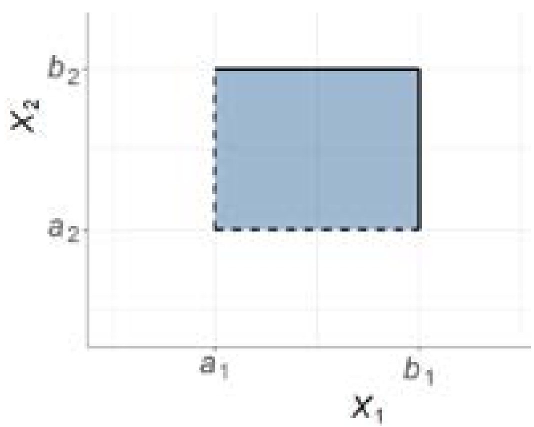


Figura 5.2. *Región de* $\{a_1 < X_1 \leq b_1, a_2 < X_2 \leq b_2\}$

Cabe mencionar que muchas de las definiciones para vectores aleatorios son generalizaciones de las definiciones que se presentaron para una variable aleatoria (véase capítulo 2).

5.1. Distribuciones conjuntas

A continuación, se presentará la definición de función de distribución de un vector aleatorio.

Definición 5.1. Sea $\underline{X} = (X_1, X_2, \ldots, X_n)$ un vector aleatorio, la *función de distribución* de $\underline{X}$ se denota como $F_{\underline{X}}(x_1, x_2, \ldots, x_n)$, y se define como

$$F_{\underline{X}}(x_1, \ldots, x_n) = P\left[X_1 \leq x_1, \ldots, X_n \leq x_n\right], \text{ para todo } (x_1, \ldots, x_n) \in \mathbb{R}^n.$$

A continuación, se presentará un ejemplo.

Ejemplo 5.1. Se lanzan 2 tetraedros y se definen las siguientes variables aleatorias: X_1 = Resultado del primer tetraedro, X_2 = El número más grande de ambos resultados. Obtener la función de distribución para valores del recorrido de $\underline{X}$.

Obsérvese que los recorridos individuales de X_1 y X_2 son, respectivamente, $x_1 = 1, 2, 3, 4$ y $x_2 = 1, 2, 3, 4$. Mas es necesario saber el recorrido conjunto de ambas variables para calcular la función de distribución del vector.

El recorrido de $\underline{X}$ es $\underline{x} = (1,1), (1,2), (1,3), (1,4), (2,2), (2,3), (2,4), (3,3), (3,4), (4,4)$.

Gráficamente el recorrido es

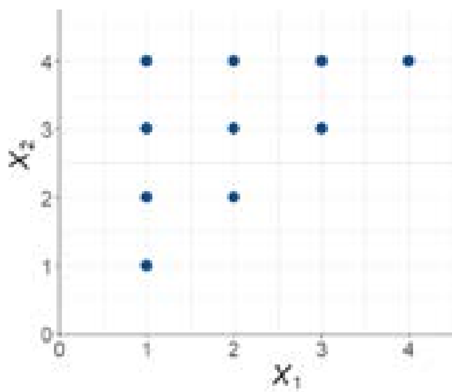


Figura 5.3. *Recorrido conjunto de X_1 y X_2, ejemplo 5.1*

Obsérvese que el espacio muestral es $\Omega = \{(1,1), (1,2), (1,3), (1,4), (2,1), (2,2), (2,3), (2,4), (3,1), (3,2), (3,3), (3,4), (4,1), (4,2), (4,3), (4,4)\}$.

Se calculará la función de distribución para algunos puntos del recorrido,

$F_{\underline{X}}(1,1) = P[X_1 = 1, X_2 = 1] = \frac{1}{16}$,

$F_{\underline{X}}(1,2) = P[X_1 \leq 1, X_2 \leq 2] = F_{\underline{X}}(1,1) + P[X_1 = 1, X_2 = 2]$
$= \frac{1}{16} + \frac{1}{16} = \frac{2}{16}$,

$F_{\underline{X}}(2,2) = P[X_1 \leq 2, X_2 \leq 2] = F_{\underline{X}}(1,2) + P[X_1 = 2, X_2 = 2]$
$= \frac{2}{16} + \frac{2}{16} = \frac{4}{16}$,

$F_{\underline{X}}(2,3) = F_{\underline{X}}(2,2) + P[X_1 = 1, X_2 = 3] + P[X_1 = 2, X_2 = 3]$
$= \frac{4}{16} + \frac{1}{16} + \frac{1}{16} = \frac{6}{16}$.

Obsérvese la función de distribución para diferentes puntos que no forman parte del recorrido.

$F_{\underline{X}}(5,1) = F_{\underline{X}}(4,1) = F_{\underline{X}}(3,1) = F_{\underline{X}}(2,1) = F_{\underline{X}}(1,1) = \frac{1}{16}$,

$F_{\underline{X}}(5,2) = F_{\underline{X}}(4,2) = F_{\underline{X}}(3,2) = F_{\underline{X}}(2,2) = \frac{4}{16}$,

$F_{\underline{X}}(5,3) = F_{\underline{X}}(4,3) = F_{\underline{X}}(3,3) = \frac{9}{16}$,

$F_{\underline{X}}(5,5) = F_{\underline{X}}(4,5) = F_{\underline{X}}(5,4) = F_{\underline{X}}(4,4) = 1$.

La siguiente tabla nos muestra valores de la función de distribución para puntos del recorrido del vector aleatorio, además, para algunos puntos fuera del recorrido.

x_1 \ x_2	**1**	**2**	**3**	**4**	**5**
1	1/16	2/16	3/16	4/16	4/16
2	1/16	4/16	6/16	8/16	8/16
3	1/16	4/16	9/16	12/16	12/16
4	1/16	4/16	9/16	1	1
5	1/16	4/16	9/16	1	1

Tabla 5.1. *Función de distribución, ejemplo 5.1*

Como en el caso univariado, la función de distribución tiene propiedades, estas se darán a continuación para el caso bivariado.

Propiedades de una función de distribución bivariada

Sea $\underline{X} = (X_1, X_2)$ un vector aleatorio y $F_{\underline{X}}(x_1, x_2)$ su función de distribución, entonces,

(I) $\lim\limits_{x_1 \to -\infty} F_{\underline{X}}(x_1, x_2) = 0.$

(II) $\lim\limits_{x_2 \to -\infty} F_{\underline{X}}(x_1, x_2) = 0.$

(III) $\lim\limits_{\substack{x_1 \to -\infty \\ x_2 \to -\infty}} F_{\underline{X}}(x_1, x_2) = 0.$

(IV) $\lim\limits_{\substack{x_1 \to \infty \\ x_2 \to \infty}} F_{\underline{X}}(x_1, x_2) = 1.$

(V) $\lim\limits_{0<h \to 0} F_{\underline{X}}(x_1 + h, x_2) = F_{\underline{X}}(x_1, x_2).$

(VI) $\lim\limits_{0<h \to 0} F_{\underline{X}}(x_1, x_2 + h) = F_{\underline{X}}(x_1, x_2).$

(VII) $\lim\limits_{0<h \to 0} F_{\underline{X}}(x_1 + h, x_2 + h) = F_{\underline{X}}(x_1, x_2).$

(VIII) $F_{\underline{X}}(x_1, x_2)$ es una función monótona no decreciente en $\mathbb{R}^2$. Esto es, para (a_1, b_1) y $(a_2, b_2) \in \mathbb{R}^2$, tal que $a_1 < a_2$ y $b_1 < b_2$, se cumple $F_{\underline{X}}(a_2, b_2) - F_{\underline{X}}(a_1, b_2) - F_{\underline{X}}(a_2, b_1) + F_{\underline{X}}(a_1, b_1) \geq 0.$

Observación. Cuando una de las variables tiende a $-\infty$, la función de distribución bivariada tiende a 0, veáse propiedades (I) y (II), no obstante, las dos variables deben tender a ∞, para que la función de distribución bivariada tienda a 1, veáse propiedad (IV). En el caso de que solo una variable tienda a ∞, se puede observar que la función de distribución conjunta tiende a una de las funciones de distribuciones univariadas, esto es,

$$\lim_{x_1 \to \infty} F_{\underline{X}}(x_1, x_2) = F_{X_2}(x_2),$$

$$\lim_{x_2 \to \infty} F_{\underline{X}}(x_1, x_2) = F_{X_1}(x_1).$$

Es importante mencionar que las propiedades de una función de distribución para un vector en $\mathbb{R}^n$, cuando $n \geq 3$, no son fáciles de expresar. También, para vectores en general, no es práctico el cálculo de probabilidades por medio de la función de distribución.

Definición 5 .2. Un vector aleatorio $\underline{X} = (X_1, X_2, \ldots, X_n)$ es *discreto*, si su recorrido en $\mathbb{R}^n$ es finito o infinito numerable.

Observación. Un vector aleatorio $\underline{X}$ es discreto si y solo si cada uno de sus componentes es una variable aleatoria discreta.

Definición 5 .3. Sea $\underline{X} = (X_1, X_2, \ldots, X_n)$ un vector aleatorio discreto. La *función de densidad* de $\underline{X}$ o la *función de densidad conjunta* de $X_1, \ldots, X_n$ es denotada como $f_{\underline{X}}(x_1, x_2, \ldots, x_n)$ y se define como

$$f_{\underline{X}}(x_1, x_2, \ldots, x_n) = P[X_1 = x_1, \ldots, X_n = x_n], \text{ para todo } (x_1, \ldots, x_n) \in \mathbb{R}^n.$$

Propiedades de una función de densidad conjunta discreta

Para una función de densidad $f_{\underline{X}}(x_1, x_2, \ldots, x_n)$, se cumplen las siguientes propiedades:

(I) $f_{\underline{X}}(x_1, x_2, \ldots, x_n) \geq 0$, para todo $(x_1, \ldots, x_n) \in \mathbb{R}^n$.

(II) $\displaystyle\sum_{x_n} \cdots \sum_{x_2} \sum_{x_1} f_{\underline{X}}(x_1, x_2, \ldots, x_n) = 1.$

Teorema 5.1. *Sea $\underline{X} = (X_1, X_2, \ldots, X_n)$ un vector aleatorio discreto, entonces, la función de densidad de $\underline{X}$ puede obtenerse a partir de la función de distribución de $\underline{X}$, y viceversa.*

Demostración. (Para $n = 2$)

La función de distribución en términos de la función de densidad es

$$F_{\underline{X}}(a_1, a_2) = \sum_{\{(x_1,x_2)|x_1 \leq a_1, x_2 \leq a_2\}} f_{\underline{X}}(x_1, x_2).$$

Ahora, la función de densidad puede expresarse en términos de la función de distribución de la siguiente manera:

$$\begin{aligned} f_{\underline{X}}(a_1, a_2) &= P[X_1 = a_1, X_2 = a_2] \\ &= F_{\underline{X}}(a_1, a_2) - \lim_{0<h\to 0} F_{\underline{X}}(a_1 - h, a_2) \\ &- \lim_{0<h\to 0} F_{\underline{X}}(a_1, a_2 - h) + \lim_{0<h\to 0} F_{\underline{X}}(a_1 - h, a_2 - h). \end{aligned}$$ □

Ejemplo 5.2. Obtener la función de densidad del vector aleatorio definido en el ejemplo 5.1.

La función de densidad de $\underline{X}$ para algunos puntos del recorrido
$f_{\underline{X}}(1,1) = P[X_1 = 1, X_2 = 1] = \frac{1}{16}$,
$f_{\underline{X}}(1,2) = P[X_1 = 1, X_2 = 2] = \frac{1}{16}$,
$f_{\underline{X}}(2,2) = P[X_1 = 2, X_2 = 2] = \frac{2}{16}$.

La siguiente tabla muestra valores de la función de densidad para los diferentes puntos del recorrido.

x_1 \ x_2	**1**	**2**	**3**	**4**
1	1/16	1/16	1/16	1/16
2	0	2/16	1/16	1/16
3	0	0	3/16	1/16
4	0	0	0	4/16

Tabla 5.2. *Función de densidad, ejemplo 5.1*

Queda como ejercicio para el estudiante confirmar los valores de la anterior tabla.

Observación. Si $\underline{X} = (X_1, \ldots, X_n)$ es un vector discreto y C es un subconjunto del recorrido de $\underline{X}$, entonces, $P[\underline{X} \in C]$ es la suma de $f_{\underline{X}}(x_1, \ldots, x_n)$ sobre la región C.

Ejemplo 5.3. En una urna hay 3 bolas rojas, 2 blancas y 3 negras. Se seleccionan al azar 3 bolas. Se definen las siguientes variables aleatorias: X_1 = Número de bolas rojas en la muestra, X_2 = Número de bolas blancas en la muestra. Encontrar la función de
a) Densidad del vector aleatorio $\underline{X} = (X_1, X_2)$.
b) Distribución para los puntos del recorrido del vector $\underline{X} = (X_1, X_2)$.

El recorrido del vector es
$(x_1, x_2) = (0,0), (0,1), (0,2), (0,3), (1,0), (1,1), (1,2), (2,0), (2,1)$.

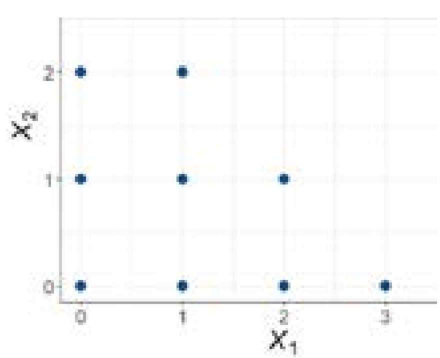


Figura 5.4. *Recorrido conjunto de X_1 y X_2, ejemplo 5.3*

a) La función de densidad para puntos del recorrido es

$$
\begin{aligned}
f_{\underline{X}}(x_1, x_2) &= P[X_1 = x_1, X_2 = x_2] \\
&= \frac{\binom{3}{x_1}\binom{2}{x_2}\binom{3}{3 - x_1 - x_2}}{\binom{8}{3}}.
\end{aligned}
$$

Obsérvese que los puntos del recorrido forman casi un triángulo formado por los puntos con componentes enteros, el cual está limitado por las rectas $X_2 = 0$, $X_1 = 0$ y $X_1 + X_2 = 3$, solo faltaría el punto $(0, 3)$. Si este último punto lo evaluamos en la anterior expresión de la función de densidad, el resultado es 0. Para facilitar la expresión de la función de densidad con funciones indicadoras, se incorporará este punto. De esta manera, podemos expresar la función de densidad con funciones indicadoras de la siguiente forma:

$$
f_{\underline{X}}(x_1, x_2) = \frac{\binom{3}{x_1}\binom{2}{x_2}\binom{3}{3 - x_1 - x_2}}{\binom{8}{3}} I_{\{0,1,\ldots 3-x_2\}}(x_1) I_{\{0,1,2,3\}}(x_2).
$$

Se puede observar que la función de densidad valuada en el punto $(0, 3)$ es cero, debido a que las combinaciones $\binom{2}{x_2}$ en este punto valen cero.

Habrá más ventajas de expresar la función de densidad conjunta como se acaba de hacer (puntos limitados por un triángulo), estas se apreciarán, cuando se obtengan más adelante, las funciones de densidad marginales y funciones de densidad condicionales, así como la obtención de la función generadora de momentos conjunta.

La función de densidad conjunta de X_1 y X_2 en formato de tabla, para puntos del recorrido, se da a continuación:

x_1 \ x_2	**0**	**1**	**2**
0	$1/56$	$6/56$	$3/56$
1	$9/56$	$18/56$	$3/56$
2	$9/56$	$6/56$	0
3	$1/56$	0	0

Tabla 5.3. *Función de densidad de X_1 y X_2, ejemplo 5.3*

Obsérvese que para cualquier otro punto en $\mathbb{R}^2$, la función de densidad vale cero, también se puede ver que la suma total de las anteriores probabilidades es igual a 1.

b) A continuación, se presentará la función de distribución para puntos del recorrido

x_1 \ x_2	**0**	**1**	**2**	**3**	**4**
0	$1/56$	$7/56$	$10/56$	$10/56$	$10/56$
1	$10/56$	$34/56$	$40/56$	$40/56$	$40/56$
2	$19/56$	$49/56$	$55/56$	$55/56$	$55/56$
3	$20/56$	$50/56$	1	1	1
4	$20/56$	$50/56$	1	1	1

Tabla 5.4. *Función de distribución de X_1 y X_2, ejemplo 5.3*

Observación. Las distribuciones de X_1 y de X_2 son, respectivamente, $H(N = 8, k = 3, n = 3)$ y $H(N = 8, k = 2, n = 3)$, esto es,

$$f_X(x_1) = \frac{\binom{3}{x_1}\binom{5}{3 - x_1}}{\binom{8}{3}} I_{\{0,1,2,3\}}(x_1),$$

$$f_X(x_2) = \frac{\binom{2}{x_2}\binom{6}{3 - x_2}}{\binom{8}{3}} I_{\{0,1,2\}}(x_2).$$

La función de distribución de X_1 para valores del recorrido

x_1	$F_{X_1}(x_1)$
0	$10/56$
1	$40/56$
2	$55/56$
3	1

Tabla 5.5. *Función de distribución de X_1, ejemplo 5.3*

La función de distribución de X_2 para valores del recorrido

x_2	$F_{X_2}(x_2)$
0	$20/56$
1	$50/56$
2	1

Tabla 5.6. *Función de distribución de X_2, ejemplo 5.3*

Obsérvese que los valores de la función de distribución de X_1 coinciden con los valores de la columna de X_2 cuando $x_2 = 2$ de la tabla de la función de distribución conjunta de X_1 y X_2. Y los valores de $F_{X_2}(x_2)$ coinciden con los valores del renglón de X_1 cuando $x_1 = 3$.

Definición 5.4. Un vector aleatorio $\underline{X}$ es *continuo* si su función de distribución $F_{\underline{X}}(x_1, \ldots, x_n)$ es una función continua para todo $(x_1, \ldots, x_n) \in \mathbb{R}^n$. El vector $\underline{X}$ es *absolutamente continuo*, si es continuo y si además existe una función $f_{\underline{X}}(x_1, \ldots, x_n)$ tal que

$$F_{\underline{X}}(x_1, \ldots, x_n) = \int_{-\infty}^{x_n} \cdots \int_{-\infty}^{x_2} \int_{-\infty}^{x_1} f_{\underline{X}}(t_1, \ldots, t_n) dt_1 dt_2 \cdots dt_n,$$

donde la función $f_{\underline{X}}(x_1, \ldots, x_n)$ es llamada *función de densidad* de $\underline{X}$.

Solo se tratarán vectores aleatorios absolutamente continuos y no vectores que sean continuos y que no sean absolutamente continuos. Entonces, cuando se mencione que el vector $\underline{X}$ es continuo, es porque también es absolutamente continuo.

Observación. Un vector es continuo, si y solo si cada componente del vector es una variable aleatoria continua.

La función de densidad del vector $\underline{X}$ o también llamada *función de densidad conjunta* de $X_1, X_2, \ldots, X_n$, es fundamental para el cálculo de probabilidades y valores esperados. A continuación, se presentarán las propiedades de esta función.

Propiedades de una función de densidad conjunta continua

(I) $f_{\underline{X}}(x_1, \ldots, x_n) \geq 0, \quad \text{para todo} \quad (x_1, \ldots, x_n) \in \mathbb{R}^n.$

(II) $\displaystyle\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{\underline{X}}(x_1, \ldots, x_n) dx_1 dx_2 \cdots dx_n = 1.$

Teorema 5.2. *Sea $\underline{X}$ un vector aleatorio continuo, la función de densidad de $\underline{X}$ se puede obtener a partir de la función de distribución de $\underline{X}$ y viceversa.*

Demostración. Por definición de vector absolutamente continuo,

$$F_{\underline{X}}(x_1, \ldots, x_n) = \int_{-\infty}^{x_n} \cdots \int_{-\infty}^{x_2} \int_{-\infty}^{x_1} f_{\underline{X}}(t_1, \ldots, t_n) dt_1 dt_2 \cdots dt_n.$$

Por el teorema fundamental de cálculo

$$f_{\underline{X}}(x_1, \ldots, x_n) = \frac{\partial^n}{\partial x_n \cdots \partial x_1} F_{\underline{X}}(x_1, \ldots, x_n). \qquad \square$$

Observación. Si $\underline{X} = (X_1, \ldots, X_n)$ es un vector continuo y C es un subconjunto del recorrido del vector $\underline{X}$, entonces, para calcular $P[\underline{X} \in C]$, se realizará la integral múltiple de la función de densidad $f_{\underline{X}}(x_1, \ldots, x_n)$ sobre la región C.

Ejemplo 5.4. Sea $\underline{X} = (X_1, X_2)$ un vector aleatorio con función de densidad dada por $f_{\underline{X}}(x_1, x_2) = I_{[0,1]}(x_1)I_{[0,1]}(x_2)$. Calcular
a) $P[X_1 \leq 0.5, X_2 \geq 0.5]$.
b) $P[X_1 \leq 0.25]$.
c) $P[X_1 \leq 0.5|X_2 \geq 0.5]$.
d) $F_{\underline{X}}(x_1, x_2)$.

El recorrido conjunto de X_1 y X_2 es

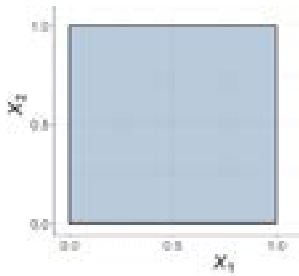


Figura 5.5. *Recorrido conjunto de X_1 y X_2, ejemplo 5.4*

a) El evento $(X_1 \leq 0.5, X_2 \geq 0.5)$ corresponde a la zona más obscura.

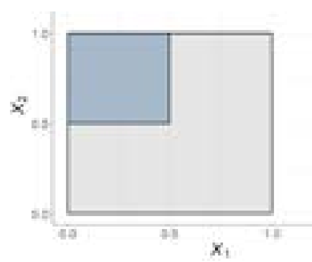


Figura 5.6. *Región del ejemplo 5.4 a)*

De esta manera,

$$P[X_1 \leq 0.5, X_2 \geq 0.5] = \int_{0.5}^{1} \int_{0}^{0.5} dx_1 dx_2 = 0.25.$$

b) La región más obscura corresponde al evento $(X_1 \leq 0.25)$.

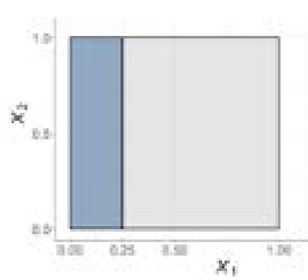


Figura 5.7. *Región del ejemplo 5.4 b)*

Por lo tanto,

$$P[X_1 \leq 0.25] = \int_0^1 \int_0^{0.25} dx_1 dx_2 = 0.25.$$

c) En la siguiente gráfica, la unión de los dos cuadrados más obscuros, corresponden al evento $(X_2 \geq 0.5)$, y el cuadrado más obscuro corresponde al evento $(X_1 \leq 0.5, X_2 \geq 0.5)$,

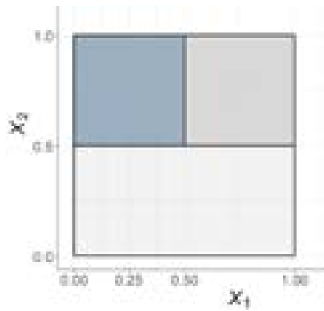


Figura 5.8. *Región del ejemplo 5.4 c)*

De esta manera,

$$P[X_2 \geq 0.5] = \int_0^1 \int_{0.5}^1 dx_2 dx_1 = 0.5.$$

Por lo tanto,

$$P[X_1 \leq 0.5 | X_2 \geq 0.5] = \frac{P[(X_1 \leq 0.5) \cap (X_2 \geq 0.5)]}{P[X_2 \geq 0.5]} = 0.5.$$

Observación. Como en el caso univariado, para distribuciones uniformes, el cálculo de probabilidades se puede lograr usando geometría y no realizando las integrales. Obsérvese que la función de densidad de este ejemplo es una constante, véase la siguiente gráfica.

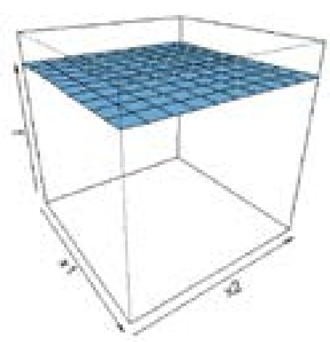


Figura 5.9. *Función de densidad de X_1 y X_2, ejemplo 5.4*

Las probabilidades se volverán a calcular, por medio de volúmenes

$$P[X_1 \leq 0.5, X_2 \geq 0.5] = 0.5 \times (1 - 0,5) \times 1 = 0.25.$$

$$P[X_1 \leq 0.25] = 0.25 \times (1 - 0) \times 1 = 0.25.$$

Para el cálculo de $P[X_1 \leq 0.5 | X_2 \geq 0.5]$, se puede ver en la figura 5.8, que el evento $(X_2 \geq 0.5)$ está representado por la unión de los dos cuadrados de arriba, y el evento $(X_1 \leq 0.5)$, condicionado a $(X_2 \geq 0.5)$, es representado por el cuadrado más oscuro, siendo la mitad de la unión de los dos cuadrados superiores. De esta manera, por uniformidad la probabilidad es igual a 0.5.

d) Para calcular la función de distribución, existen varios casos, véase la siguiente gráfica:

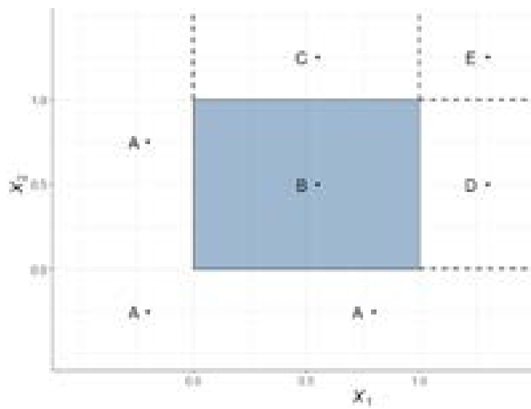


Figura 5.10. *Diferentes regiones para obtener la función de distribución, ejemplo 5.4*

Si $X_1 < 0$ o $X_2 < 0$ (región A), la función de distribución vale 0.

Si $0 \leq X_1 \leq 1$ y $0 \leq X_2 \leq 1$ (región B), la función de distribución se calcula como

$$F_{\underline{X}}(x_1, x_2) = \int_0^{x_2} \int_0^{x_1} dt_1 dt_2 = x_1 x_2.$$

Si $0 \leq X_1 \leq 1$ y $1 < X_2$ (región C), la función de distribución es

$$F_{\underline{X}}(x_1, x_2) = \int_0^{1} \int_0^{x_1} dt_1 dt_2 = x_1.$$

Si $1 < X_1$ y $0 \leq X_2 \leq 1$ (región D), entonces, la función de distribución se calcula de la siguiente manera:

$$F_{\underline{X}}(x_1, x_2) = \int_0^{x_2} \int_0^{1} dt_1 dt_2 = x_2.$$

Finalmente, si $1 < X_1$ y $1 < X_2$ (región E), la función de distribución es

$$F_{\underline{X}}(x_1, x_2) = \int_0^1 \int_0^1 dt_1 dt_2 = 1.$$

De esta manera,

$$\begin{aligned} F_{\underline{X}}(x_1, x_2) = x_1 x_2 I_{[0,1]}(x_1) I_{[0,1]}(x_2) + x_1 I_{[0,1]}(x_1) I_{(1,\infty)}(x_2) \\ + x_2 I_{(1,\infty)}(x_1) I_{[0,1]}(x_2) + I_{(1,\infty)}(x_1) I_{(1,\infty)}(x_2). \end{aligned}$$

Ejemplo 5.5. Sea $\underline{X} = (X_1, X_2)$ un vector aleatorio con función de densidad dada por

$$f_{\underline{X}}(x_1, x_2) = \begin{cases} k(1 - x_2), & \text{si } 0 \leq x_1 \leq x_2 \leq 1, \\ 0, & \text{en otro caso.} \end{cases}$$

a) Escribir la función de densidad con funciones indicadoras.
b) Encontrar k.
c) Calcular $P[X_1 \leq 0.75, X_2 \geq 0.5]$.
d) Calcular $P[X_1 \leq 0.5]$.
e) Calcular $P[X_1 \leq 0.5 | X_2 \leq 0.5]$.

El recorrido de $\underline{X} = (X_1, X_2)$ geométricamente es

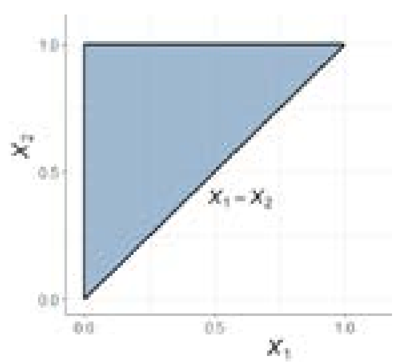


Figura 5.11. *Recorrido conjunto de X_1 y X_2, ejemplo 5.5*

a) La función de densidad con indicadoras

$$\begin{aligned} f_{\underline{X}}(x_1, x_2) &= k(1 - x_2) I_{[0,x_2]}(x_1) I_{[0,1]}(x_2) \\ &= k(1 - x_2) I_{[x_1,1]}(x_2) I_{[0,1]}(x_1). \end{aligned}$$

b) El cálculo de k

$$\int_0^1 \int_0^{x_2} k(1 - x_2) dx_1 dx_2 = 1.$$

Por lo tanto,

$$
\begin{aligned}
1 &= \int_0^1 k(1-x_2)x_2 dx_2 \\
&= k\left[\frac{x_2^2}{2} - \frac{x_2^3}{3}\right]_0^1 \\
&= k\left(\tfrac{1}{6}\right).
\end{aligned}
$$

Entonces, $k = 6$.
De esta manera,

$$
\begin{aligned}
f_{\underline{X}}(x_1, x_2) &= 6(1-x_2)I_{[0,x_2]}(x_1)I_{[0,1]}(x_2) \\
&= 6(1-x_2)I_{[x_1,1]}(x_2)I_{[0,1]}(x_1).
\end{aligned}
$$

c) La primera probabilidad

$$
\begin{aligned}
P[X_1 \leq 0.75, X_2 \geq 0.5] = 1 &- \int_0^{0.5}\int_0^{x_2} 6(1-x_2)dx_1dx_2 \\
&- \int_{0.75}^{1}\int_{0.75}^{x_2} 6(1-x_2)dx_1dx_2 = 0.4844.
\end{aligned}
$$

d) La segunda probabilidad

$$
P[X_1 \leq 0.5] = 1 - \int_{0.5}^{1}\int_{0.5}^{x_2} 6(1-x_2)dx_1dx_2 = 0.875.
$$

e) La probabilidad condicional

$$
\begin{aligned}
P[X_1 \leq 0.5 | X_2 \leq 0.5] &= \frac{P[X_1 \leq 0.5, X_2 \leq 0.5]}{P[X_2 \leq 0.5]} \\
&= \frac{\displaystyle\int_0^{0.5}\int_0^{x_2} 6(1-x_2)dx_1dx_2}{\displaystyle\int_0^{0.5}\int_0^{x_2} 6(1-x_2)dx_1dx_2} = 1.
\end{aligned}
$$

Ejemplo 5.6. Supóngase que X_1 y X_2 se distribuyen uniformemente sobre el siguiente triángulo

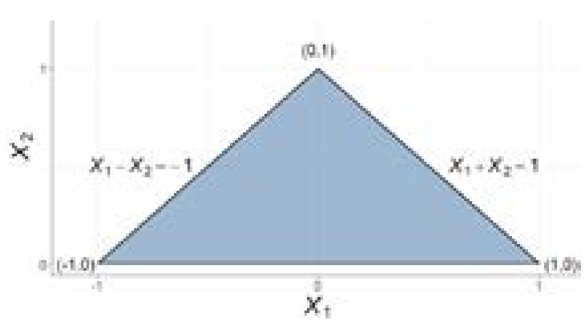


Figura 5.12. *Recorrido conjunto de X_1 y X_2, ejemplo 5.6*

a) Expresar la función de densidad conjunta de X_1 y X_2 con funciones indicadoras.
b) Calcular $P[X_1 \leq 0.75, X_2 \leq 0.75]$.
c) Calcular $P[X_1 - X_2 \geq 0]$.

a) Al ser uniforme la distribución, la altura de la función de densidad es igual a una constante, sea h la altura de la función de densidad. Entonces, (área del triángulo) $\times$ h = Volumen=1. De aquí se deduce que $h = 1$, de esta manera,

$$\begin{aligned} f_{\underline{X}}(x_1, x_2) &= I_{[x_2-1,1-x_2]}(x_1) I_{[0,1]}(x_2) \\ &= I_{[0,x_1+1]}(x_2) I_{[-1,0]}(x_1) + I_{[0,1-x_1]}(x_2) I_{(0,1]}(x_1). \end{aligned}$$

b) En la siguiente gráfica, la región de color obscuro corresponde al evento $\{X_1 \leq 0.75, X_2 \leq 0.75\}$.

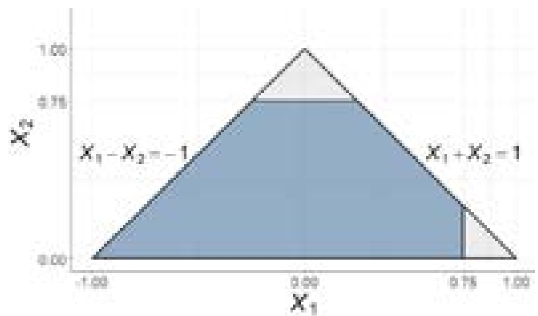


Figura 5.13. *Región del ejemplo 5.6 b)*

Por lo tanto,

$$\begin{aligned} P[X_1 \leq 0.75, X_2 \leq 0.75] &= 1 - (0.5)(0.25)(0.5)(1) - (0.25)(0.25)(0.5)(1) \\ &= \frac{29}{32}. \end{aligned}$$

c) Obsérvese en la siguiente gráfica que la región de color obscuro representa al evento $\{X_1 - X_2 \geq 0\}$.

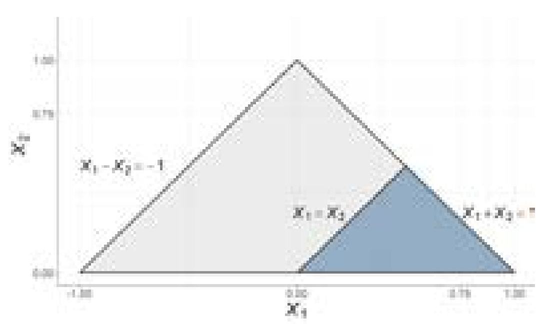


Figura 5.14. *Región del ejemplo 5.6 c)*

Por lo tanto,

$$\begin{aligned} P[X_1 - X_2 \geq 0] &= P[X_1 \geq X_2] \\ &= (1)(0.5)(0.5)(1) = 0.25. \end{aligned}$$

Observación. Es posible que las variables aleatorias que conforman un vector aleatorio $\underline{X}$, algunas de ellas sean discretas y otras continuas, véase el siguiente ejemplo, además, el ejemplo 5.27. Incluso es posible que algunas de las variables sea una mezcla de discreta y continua.

Ejemplo 5.7. Supóngase que X_1 es el número de reclamaciones realizadas en una compañía de seguros y X_2 es la cantidad de dinero reclamada, ambas en un mes. La función de densidad conjunta de X_1 y X_2 se da a continuación:

$$f_{\underline{X}}(x_1, x_2) = 0.003(0.7)^{x_1 - 1} e^{-\frac{x_2}{100}} I_{\{1,2,\ldots\}}(x_1) I_{(0,\infty)}(x_2).$$

Calcular la probabilidad $P[X_1 + X_2 < 3]$.

El recorrido conjunto de X_1 y X_2 está representado por las líneas azules de la siguiente gráfica:

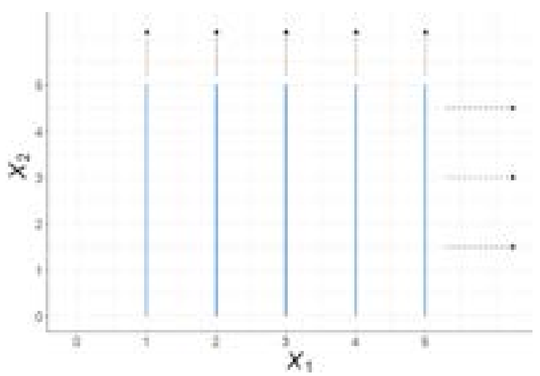


Figura 5.15. *Región del ejemplo 5.7*

Se puede observar que las líneas azules en la siguiente gráfica representan el evento $\{X_1 + X_2 < 3\}$.

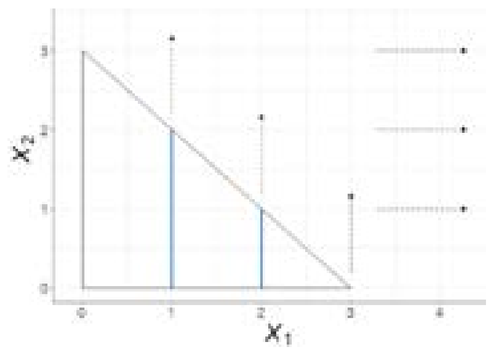


Figura 5.16. *Evento $X_1 + X_2 < 3$, ejemplo 5.7*

De esta manera, la probabilidad de que $X_1 + X_2 < 3$ se calcula como la suma de dos integrales, esto es,

$$\begin{aligned}
P[X_1 + X_2 < 3] &= \sum_{x_1=1}^{2} \int_0^{3-x_1} 0.003(0.7)^{x_1-1} e^{-\frac{x_2}{100}} dx_2 \\
&= \sum_{x_1=1}^{2} 0.3(0.7)^{x_1-1} \int_0^{3-x_1} \frac{1}{100} e^{-\frac{x_2}{100}} dx_2 \\
&= \sum_{x_1=1}^{2} 0.3(0.7)^{x_1-1} \left[1 - e^{-\frac{3-x_1}{100}}\right] \\
&= 0.00594 + 0.00209 = 0.00803.
\end{aligned}$$

5.2. Distribuciones marginales

A partir de la función de densidad conjunta de n variables, es posible obtener la función de densidad de una variable aleatoria específica de las que componen el vector aleatorio. Estas funciones de densidad de una variable son conocidas como las funciones de densidad marginales. Véase siguiente definición.

Definición 5.5. Sea $\underline{X} = (X_1, \ldots, X_n)$ un vector aleatorio con función de densidad $f_{\underline{X}}(x_1, \ldots, x_n)$. La *función de densidad marginal* de X_i es la función de densidad de X_i y se calcula de la siguiente manera:

$$f_{X_i}(x_i) = \sum_{x_n} \cdots \sum_{x_{i+1}} \sum_{x_{i-1}} \cdots \sum_{x_1} f_{\underline{X}}(x_1, \ldots, x_n),$$

si $\underline{X}$ es un vector discreto,

$$f_{X_i}(x_i) =$$

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\underline{X}}(x_1, \ldots, x_n) dx_1 \cdots dx_{i-1} dx_{i+1} \cdots dx_n,$$

si $\underline{X}$ es un vector continuo.

Obsérvese que la función marginal de una variable X_i se obtiene sumando o integrando sobre todos valores de las variables que componen el vector aleatorio, excepto sobre los valores de la variable X_i.

Ejemplo 5.8. Considerando el ejemplo 5.1 donde se lanzan de 2 tetraedros, encontrar las funciones de densidad marginales de X_1 y X_2.

Estas funciones de densidad marginales se obtendrán a partir de la función de densidad conjunta, esto es,

x_1 \ x_2	**1**	**2**	**3**	**4**	$f_{X_1}(x_1)$
1	1/16	1/16	1/16	1/16	1/4
2	0	2/16	1/16	1/16	1/4
3	0	0	3/16	1/16	1/4
4	0	0	0	4/16	1/4
$f_{X_2}(x_2)$	1/16	3/16	5/16	7/16	1

Tabla 5.7. *Funciones marginales, ejemplos 5.1 y 5.8*

Entonces, las funciones de densidad marginales de X_1 y X_2 quedan, respectivamente, como:

$$f_{X_1}(x_1) = \frac{1}{4} I_{\{1,2,3,4\}}(x_1),$$

$$f_{X_2}(x_2) = \left(\frac{2x_2 - 1}{16}\right) I_{\{1,2,3,4\}}(x_2).$$

Ejemplo 5.9. Obtener las funciones marginales de la función de densidad conjunta del ejemplo 5.5. Además, indicar cómo se distribuye cada variable.

$$\begin{aligned} f_{X_1}(x_1) &= \int_{x_1}^{1} 6(1 - x_2)dx_2 \\ &= \left[6x_2 - 6\left(\frac{x_2^2}{2}\right)\right]_{x_1}^{1} \\ &= 3 - 6x_1 + 3x_1^2. \end{aligned}$$

Por lo tanto,

$$f_{X_1}(x_1) = 3(1 - x_1)^2 I_{[0,1]}(x_1).$$

Obsérvese

$$X_1 \sim Beta(1, 3).$$

Por otro lado,

$$\begin{aligned} f_{X_2}(x_2) &= \int_{0}^{x_2} 6(1 - x_2)dx_1 \\ &= 6x_2(1 - x_2). \end{aligned}$$

Por lo tanto,

$$f_{X_2}(x_2) = 6x_2(1 - x_2)I_{[0,1]}(x_2).$$

En este caso,

$$X_2 \sim Beta(2, 2).$$

Ejemplo 5.10. Calcular las funciones de densidad marginales de la función de densidad conjunta del ejemplo 5.6.

Obsérvese que la función de densidad marginal de X_1 se calcula de la siguiente manera:

$$\begin{aligned} f_{X_1}(x_1) &= \int_0^{x_1+1} dx_2 \\ &= x_1 + 1, \quad \text{si } -1 \leq x_1 \leq 0. \end{aligned}$$

$$\begin{aligned} f_{X_1}(x_1) &= \int_0^{1-x_1} dx_2 \\ &= 1 - x_1, \quad \text{si } 0 < x_1 \leq 1. \end{aligned}$$

Por lo tanto,

$$f_{X_1}(x_1) = (x_1 + 1)I_{[-1,0]}(x_1) + (1 - x_1)I_{(0,1]}(x_1).$$

Y la función marginal de X_2 se calcula como:

$$\begin{aligned} f_{X_2}(x_2) &= \int_{x_2-1}^{1-x_2} dx_2 \\ &= 2 - 2x_2. \end{aligned}$$

Esto es,

$$f_{X_2}(x_2) = 2(1 - x_2)I_{[0,1]}(x_2).$$

Obsérvese que $X_2 \sim Beta(1, 2)$.

5.3. Distribuciones condicionales

Existen muchas aplicaciones tanto en temas de probabilidad, procesos estocásticos y estadística donde las distribuciones condicionales son utilizadas. Se presentará la definición de distribución condicional, algunas propiedades y ejemplos. A continuación, se presentará la definición de función de densidad condicional para un vector aleatorio bivariado y siendo la condición un valor específico de una de las variables.

Definición 5.6. Sea $\underline{X} = (X_1, X_2)$ un vector aleatorio que pertenece a $\mathbb{R}^2$. La *función de densidad condicional* de X_1 dado que $X_2 = x_2$ se denota como $f_{X_1|X_2}(x_1|x_2)$ y se define de la siguiente manera:

$$f_{X_1|X_2}(x_1|x_2) = \frac{f_{\underline{X}}(x_1, x_2)}{f_{X_2}(x_2)}, \quad \text{si} \quad f_{X_2}(x_2) > 0,$$

en el caso de $f_{X_2}(x_2) = 0$, se indefine la función de densidad condicional.

En forma análoga, la *función de densidad condicional* de X_2 dado que $X_1 = x_1$, se denota como $f_{X_2|X_1}(x_2|x_1)$ y se define como

$$f_{X_2|X_1}(x_2|x_1) = \frac{f_{\underline{X}}(x_1, x_2)}{f_{X_1}(x_1)}, \quad \text{si} \quad f_{X_1}(x_1) > 0,$$

en el caso de $f_{X_1}(x_1) = 0$, se indefine la función de densidad condicional.

La función de densidad condicional $f_{X_1|X_2}(x_1|x_2)$ es una función de densidad para la variable X_1, bajo la condición de que la variable X_2 toma el valor específico x_2, lo que significa que solo X_1 es la variable aleatoria.

Ejemplo 5.11. Se lanzan dos tetraedros y se definen las siguientes variables aleatorias: X_1 = resultado del primer tetraedro y X_2 = resultado del segundo tetraedro. Obtener las funciones de densidad: conjunta, marginales y condicionales de X_1 dado que $X_2 = x_2$ y X_2 dado que $X_1 = x_1$.

En la siguiente tabla se presenta, para valores del recorrido, la función de densidad conjunta de X_1 y X_2, y las funciones de densidad marginales.

x_1 \ x_2	**1**	**2**	**3**	**4**	$f_{X_1}(x_1)$
1	1/16	1/16	1/16	1/16	1/4
2	1/16	1/16	1/16	1/16	1/4
3	1/16	1/16	1/16	1/16	1/4
4	1/16	1/16	1/16	1/16	1/4
$f_{X_2}(x_2)$	1/4	1/4	1/4	1/4	1

Tabla 5.8. *Densidades conjunta y marginales, ejemplo 5.11*

Entonces, la función de densidad conjunta de X_1 y X_2 y las funciones de densidad marginales son, respectivamente,

$$f_{\underline{X}}(x_1, x_2) = \tfrac{1}{16} I_{\{1,2,3,4\}}(x_1) I_{\{1,2,3,4\}}(x_2),$$

$$f_{X_1}(x_1) = \tfrac{1}{4} I_{\{1,2,3,4\}}(x_1),$$

$$f_{X_2}(x_2) = \tfrac{1}{4} I_{\{1,2,3,4\}}(x_2).$$

La función de densidad condicional de X_1 dado que $X_2 = x_2$ es

$$\begin{aligned} f_{X_1|X_2}(x_1|x_2) &= \frac{f_{\underline{X}}(x_1, x_2)}{f_{X_2}(x_2)} \\ &= \frac{\frac{1}{16} I_{\{1,2,3,4\}}(x_1) I_{\{1,2,3,4\}}(x_2)}{\frac{1}{4} I_{\{1,2,3,4\}}(x_2)} \\ &= \tfrac{1}{4} I_{\{1,2,3,4\}}(x_1), \quad \text{si} \quad x_2 = 1, 2, 3, 4. \end{aligned}$$

Análogamente,

$$\begin{aligned} f_{X_2|X_1}(x_2|x_1) &= \frac{f_{\underline{X}}(x_1, x_2)}{f_{X_1}(x_1)} \\ &= \tfrac{1}{4} I_{\{1,2,3,4\}}(x_2), \quad \text{si} \quad x_1 = 1, 2, 3, 4. \end{aligned}$$

Ejemplo 5.12. Obtener las funciones de densidad condicionales de X_1 dado que $X_2 = x_2$ y de X_2 dado $X_1 = x_1$, del ejemplo 5.9 (y 5.5).

Obsérvese

$$
\begin{aligned}
f_{X_1|X_2}(x_1|x_2) &= \frac{f_{\underline{X}}(x_1, x_2)}{f_{X_2}(x_2)} \\
&= \frac{6(1-x_2)I_{[0,x_2]}(x_1)I_{(0,1)}(x_2)}{6x_2(1-x_2)I_{(0,1)}(x_2)} \\
&= \frac{1}{x_2}I_{[0,x_2]}(x_1), \qquad \text{si } 0 < x_2 < 1.
\end{aligned}
$$

Por lo tanto, $X_1|x_2 \sim U(0, x_2)$, donde $0 < x_2 < 1$.

Obsérvese que para cada valor de x_2 hay una función de densidad condicional de X_1.

Por otra parte,

$$
\begin{aligned}
f_{X_2|X_1}(x_2|x_1) &= \frac{f_{\underline{X}}(x_1, x_2)}{f_{X_1}(x_1)} \\
&= \frac{6(1-x_2)I_{[x_1,1]}(x_2)I_{[0,1)}(x_1)}{3(1-x_1)^2 I_{[0,1)}(x_1)} \\
&= \frac{2(1-x_2)}{(1-x_1)^2}I_{[x_1,1]}(x_2), \qquad \text{si } 0 \leq x_1 < 1.
\end{aligned}
$$

También obsérvese de este último resultado, que para cada valor de x_1, hay una función de densidad condicional de X_2.

Ejemplo 5.13. Considerando el ejemplo 5.1 (y 5.8), donde se lanzan dos tetraedros. Calcular las funciones de densidad condicionales.

En el ejemplo 5.8 se calcularon las funciones de densidad marginales.

Para el cálculo de la función de densidad condicional $f_{X_1|X_2}(x_1|x_2)$, se van a considerar diferentes casos para x_2.

Caso 1. Cuando $x_2 = 1$.

Obsérvese que el recorrido de X_1 es $x_1 = 1$.

De esta manera,

$$f_{X_1|X_2}(x_1|1) = \frac{f_{\underline{X}}(x_1, 1)}{f_{X_2}(1)}.$$

Esta función evaluada para el único valor del recorrido de X_1 es

$$f_{X_1|X_2}(1|1) = \frac{f_{\underline{X}}(1, 1)}{f_{X_2}(1)} = 1.$$

Por lo tanto,

$$f_{X_1|X_2}(x_1|1) = I_{\{1\}}(x_1).$$

Caso 2. Cuando $x_2 = 2$.

En este caso, el recorrido de X_1 es $x_1 = 1, 2$.

De esta manera,

$$f_{X_1|X_2}(x_1|2) = \frac{f_{\underline{X}}(x_1, 2)}{f_{X_2}(2)}.$$

Se evaluará esta función para los diferentes valores del recorrido de X_1

$$f_{X_1|X_2}(1|2) = \frac{f_{\underline{X}}(1, 2)}{f_{X_2}(2)} = \tfrac{1}{3},$$

$$f_{X_1|X_2}(2|2) = \frac{f_{\underline{X}}(2, 2)}{f_{X_2}(2)} = \tfrac{2}{3}.$$

Por lo tanto,

$$f_{X_1|X_2}(x_1|2) = \tfrac{1}{3}I_{\{1\}}(x_1) + \tfrac{2}{3}I_{\{2\}}(x_1).$$

Caso 3. Cuando $x_2 = 3$.
El recorrido de X_1 es $x_1 = 1, 2, 3$.
Por lo tanto,

$$f_{X_1|X_2}(x_1|3) = \tfrac{1}{5}I_{\{1,2\}}(x_1) + \tfrac{3}{5}I_{\{3\}}(x_1).$$

Caso 4. Cuando $x_2 = 4$.
El recorrido de X_1 es $x_1 = 1, 2, 3, 4$.
De esta manera,

$$f_{X_1|X_2}(x_1|4) = \tfrac{1}{7}I_{\{1,2,3\}}(x_1) + \tfrac{4}{7}I_{\{4\}}(x_1).$$

En resumen, la función de densidad condicional de X_1 dado que $X_2 = x_2$, queda de la siguiente manera:

$$f_{X_1|X_2}(x_1|x_2) = \begin{cases} I_{\{1\}}(x_1), & \text{si} \quad x_2 = 1, \\ \tfrac{1}{3}I_{\{1\}}(x_1) + \tfrac{2}{3}I_{\{2\}}(x_1), & \text{si} \quad x_2 = 2, \\ \tfrac{1}{5}I_{\{1,2\}}(x_1) + \tfrac{3}{5}I_{\{3\}}(x_1), & \text{si} \quad x_2 = 3, \\ \tfrac{1}{7}I_{\{1,2,3\}}(x_1) + \tfrac{4}{7}I_{\{4\}}(x_1), & \text{si} \quad x_2 = 4. \end{cases}$$

Queda como ejercicio para el estudiante, calcular la otra función de densidad condicional, $f_{X_2|X_1}(x_2|x_1)$.

Es importante observar en el ejemplo anterior que para cada valor de x_2 (4 valores), se obtuvo una función de densidad condicional de X_1 diferente.

Ejemplo 5.14. Calcular las funciones de densidad condicionales correspondientes al ejemplo 5.10 (y 5.6).

La función de densidad condicional de X_1 dado que $X_2 = x_2$ es

$$\begin{aligned} f_{X_1|X_2}(x_1|x_2) &= \frac{f_{\underline{X}}(x_1, x_2)}{f_{X_2}(x_2)} \\ &= \frac{I_{[x_2-1,1-x_2]}(x_1) I_{[0,1)}(x_2)}{2(1-x_2) I_{[0,1)}(x_2)} \\ &= \frac{1}{2(1-x_2)} I_{[x_2-1,1-x_2]}(x_1), \text{ si } 0 \leq x_2 < 1. \end{aligned}$$

Esto es, $X_1|x_2 \sim U(x_2 - 1, 1 - x_2), \text{ si } 0 \leq x_2 < 1.$

Para la función condicional de X_2 dado que $X_1 = x_1$, se deben considerar dos casos, ya que la función de densidad marginal de X_1 está dada en dos casos.

Caso 1. Cuando $-1 < x_1 \leq 0$.

$$\begin{aligned} f_{X_2|X_1}(x_2|x_1) &= \frac{f_{\underline{X}}(x_1, x_2)}{f_{X_1}(x_1)} \\ &= \frac{I_{[0,x_1+1]}(x_2) I_{(-1,0]}(x_1)}{(x_1+1) I_{(-1,0]}(x_1)} \\ &= \frac{1}{x_1+1} I_{[0,x_1+1]}(x_2). \end{aligned}$$

Caso 2. Cuando $0 < x_1 < 1$.

$$\begin{aligned} f_{X_2|X_1}(x_2|x_1) &= \frac{f_{\underline{X}}(x_1, x_2)}{f_{X_1}(x_1)} \\ &= \frac{I_{[0,1-x_1]}(x_2) I_{(0,1)}(x_1)}{(1-x_1) I_{(0,1)}(x_1)} \\ &= \frac{1}{1-x_1} I_{[0,1-x_1]}(x_2). \end{aligned}$$

De esta manera,

$$f_{X_2|X_1}(x_2|x_1) = \begin{cases} \dfrac{1}{x_1+1} I_{[0,x_1+1]}(x_2), & \text{si} \quad -1 < x_1 \leq 0, \\ \\ \dfrac{1}{1-x_1} I_{[0,1-x_1]}(x_2), & \text{si} \quad 0 < x_1 < 1. \end{cases}$$

Para el caso $-1 < x_1 \leq 0$, la distribución condicional de X_2 dado que $X_1 = x_1$ es una $U(0, x_1 + 1)$, y para el caso $0 < x_1 < 1$ es una $U(0, 1 - x_1)$.

5.4. Independencia de variables aleatorias

El concepto de variables aleatorias independientes es de mucha importancia para propósitos de inferencia estadística y algunas áreas como las matemáticas actuariales, por mencionar algunas aplicaciones.

Varias metodologías de inferencia estadística consideran en sus análisis, el concepto de muestra aleatoria, siendo esta, un conjunto de variables aleatorias independientes con una distribución común, véase definición 6.1.

En las matemáticas actuariales, en algunos modelos, también se puede suponer independencia entre las variables involucradas, debido a que las vidas de personas son sucesos que, por lo general, son eventos independientes.

Se presentarán dos definiciones de variables aleatorias independientes, la primera utiliza las funciones de distribución de las variables y en la segunda, las funciones de densidad. Cabe mencionar que ambas definiciones son equivalentes.

Definición 5.7. Las variables aleatorias $X_1, X_2, \ldots, X_n$ son *independientes* si y solo si $F_{\underline{X}}(x_1, x_2, \ldots, x_n) = F_{X_1}(x_1)F_{X_2}(x_2)\cdots F_{X_n}(x_n)$, para todo $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, donde $F_{\underline{X}}(x_1, x_2, \ldots, x_n), F_{X_1}(x_1), F_{X_2}(x_2), \ldots, F_{X_n}(x_n)$, son las funciones de distribución de $\underline{X}, X_1, X_2, \ldots, X_n$, respectivamente.

Definición 5.8. Las variables aleatorias $X_1, X_2, \ldots, X_n$ son *independientes* si y solo si $f_{\underline{X}}(x_1, x_2, \ldots, x_n) = f_{X_1}(x_1)f_{X_2}(x_2)\cdots f_{X_n}(x_n)$, para todo $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, donde $f_{\underline{X}}(x_1, x_2, \ldots, x_n), f_{X_1}(x_1), f_{X_2}(x_2), \ldots, f_{X_n}(x_n)$, son las funciones de densidad de $\underline{X}, X_1, X_2, \ldots, X_n$, respectivamente.

Ejemplo 5.15. Para el ejemplo 5.5, verificar si las variables X_1 y X_2 son independientes o no.

Las funciones marginales de X_1 y X_2 son $f_{X_1}(x_1) = 3(1-x_1)^2 I_{[0,1]}(x_1)$ y $f_{X_2}(x_2) = 6x_2(1-x_2)I_{[0,1]}(x_2)$, respectivamente.

Obsérvese

$$\begin{aligned} f_{X_1}(x_1)f_{X_2}(x_2) &= 3(1-x_1)^2 I_{[0,1]}(x_1) 6x_2(1-x_2)I_{[0,1]}(x_2) \\ &\neq 6(1-x_2)I_{[0,x_2]}(x_1)I_{(0,1]}(x_2) = f_{\underline{X}}(x_1, x_2). \end{aligned}$$

Por lo tanto, X_1 y X_2 son variables aleatorias dependientes.

Ejemplo 5.16. Considerando el ejemplo 5.11, donde se lanzan dos tetraedros, verificar si X_1 y X_2 son variables aleatorias independientes.

En el ejemplo 5.11 se obtuvieron las funciones de densidad conjunta de X_1 y X_2 y las funciones de densidad marginales de X_1 y X_2.

Se puede observar que

$$\begin{aligned} f_{X_1}(x_1)f_{X_2}(x_2) &= \tfrac{1}{4} I_{\{1,2,3,4\}}(x_1) \tfrac{1}{4} I_{\{1,2,3,4\}}(x_2) \\ &= \tfrac{1}{16} I_{\{1,2,3,4\}}(x_1) I_{\{1,2,3,4\}}(x_2) = f_{\underline{X}}(x_1, x_2). \end{aligned}$$

Por lo tanto, X_1 y X_2 son variables aleatorias independientes.

Se dejan como ejercicios para el estudiante, verificar en los ejemplos 5.8 y 5.10, si las variables aleatorias X_1 y X_2 son independientes.

Observación. Las variables aleatorias X_1 y X_2 son independientes si y solo si se cumple $f_{X_1|X_2}(x_1|x_2) = f_{X_1}(x_1)$ (o $f_{X_2|X_1}(x_2|x_1) = f_{X_2}(x_2)$).

Esto es, X_1 y X_2 son variables aleatorias independientes si y solo si se cumple

$$f_{X_1}(x_1)f_{X_2}(x_2) = f_{\underline{X}}(x_1, x_2).$$

La anterior igualdad se cumple, si y solo si

$$f_{X_2}(x_2) = \frac{f_{\underline{X}}(x_1, x_2)}{f_{X_1}(x_1)} = f_{X_2|X_1}(x_2|x_1).$$

O también,

$$f_{X_1}(x_1) = \frac{f_{\underline{X}}(x_1, x_2)}{f_{X_2}(x_2)} = f_{X_1|X_2}(x_1|x_2).$$

5.5. Valores esperados

En esta sección se tratarán diferentes valores esperados que se pueden obtener a partir de la función de densidad conjunta de n variables aleatorias. Entre los más importantes, las medias y las varianzas de las variables aleatorias que componen un vector aleatorio. También, se definirá un nuevo valor esperado, la covarianza de dos variables aleatorias, véase definición 5.10. Por otra parte, la función generadora de momentos de un vector aleatorio es también un valor esperado, véase definición 5.11.

Definición 5.9. Sea $\underline{X} = (X_1, \ldots, X_n)$ un vector aleatorio con función de densidad $f_{\underline{X}}(x_1, \ldots, x_n)$. Sea $g(X_1, \ldots, X_n) = g(\underline{X})$ una función del vector $\underline{X}$, de $\mathbb{R}^n$ a $\mathbb{R}$. El *valor esperado* de $g(\underline{X})$ se define como

$$E[g(\underline{X})] = \begin{cases} \displaystyle\sum_{x_n} \cdots \sum_{x_2} \sum_{x_1} g(\underline{x}) f_{\underline{X}}(\underline{x}), & \text{si } \underline{X} \text{ es discreto,} \\ \\ \displaystyle\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(\underline{x}) f_{\underline{X}}(\underline{x}) dx_1, \cdots, dx_n, & \text{si } \underline{X} \text{ es continuo.} \end{cases}$$

Algunos casos particulares de la función $g(\underline{X})$ pueden ser $g(\underline{X}) = X_i$, donde $E[g(\underline{X})] = E[X_i] = \mu_i$, también, $g(\underline{X}) = (X_i - \mu_i)^2$, donde $E[g(\underline{X})] = E[(X_i - \mu_i)^2] = \sigma_i^2$.

El siguiente teorema trata sobre las propiedades del valor esperado de un vector aleatorio.

Teorema 5.3. *Sea* $\underline{X} = (X_1, \ldots, X_n)$ *un vector aleatorio, y* $g_1(X_1, \ldots, X_n)$, $g_2(X_1, \ldots, X_n) \ldots, g_k(X_1, \ldots, X_n)$ *son funciones de* $\underline{X}$*, de* $\mathbb{R}^n$ *a* $\mathbb{R}$*, además sean* $c_1, \ldots, c_k$ *constantes, entonces,*

$$E\left[\sum_{i=1}^{k} c_i g_i(X_1, \ldots, X_n)\right] = \sum_{i=1}^{k} c_i E\left[g_i(X_1, \ldots, X_n)\right].$$

La demostración de este teorema se deja como ejercicio para el estudiante. Se sugiere suponer que el vector aleatorio sea continuo, siendo la comprobación para el caso discreto, un procedimiento análogo.

5.6. Covarianza y coeficiente de correlación

En esta sección se tratará con dos nuevos conceptos, la covarianza y el coeficiente de correlación de dos variables aleatorias. Se explicará las aplicaciones que tienen estas definiciones. Cabe mencionar que, en temas de estadística, en particular, en análisis de regresión, ambas son de suma importancia.

Definición 5.10. Sean X_1 y X_2 dos variables aleatorias tal que $E[X_1] = \mu_1$, $E[X_2] = \mu_2$, $V[X_1] = \sigma_1^2$ y $V[X_2] = \sigma_2^2$. La *covarianza* de X_1 y X_2 se denota como $Cov(X_1, X_2)$ y se define como

$$Cov(X_1, X_2) = E[(X_1 - \mu_1)(X_2 - \mu_2)].$$

El *coeficiente de correlación* de X_1 y X_2 se denota como ρ_{X_1,X_2} y se define como

$$\rho_{X_1,X_2} = \frac{Cov(X_1, X_2)}{\sigma_1 \sigma_2}.$$

Es importante mencionar que la covarianza de dos variables aleatorias mide la relación lineal entre las variables. Una covarianza «grande» positiva nos indica una fuerte relación lineal con tendencia positiva entre las variables. Una covarianza «grande» en valor absoluto, pero negativa, nos indica una relación lineal fuerte negativa entre las variables.

El coeficiente de correlación también evalúa la relación lineal de dos variables aleatorias. Más adelante, se demostrará que el valor absoluto del coeficiente de correlación es menor o igual a 1, véase corolario 5.2. Si el coeficiente de correlación es igual a 1, entonces, la dependencia lineal entre ambas variables es positiva y perfecta. Si el coeficiente de correlación es igual a -1, entonces, habrá una relación lineal perfecta negativa. Por otro lado, si el coeficiente de correlación es igual a 0, la dependencia lineal es nula.

El problema en la práctica, con la covarianza, es que será difícil de saber si esta es grande o no. Lo que no ocurre con el coeficiente de correlación, ya que carece de unidades.

En las siguientes gráficas se ilustra cómo el coeficiente de correlación es una medición de la relación lineal entre dos variables.

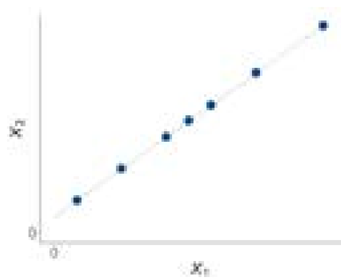


Figura 5.17. *Tendencia lineal,* $\rho_{X_1,X_2} = 1$

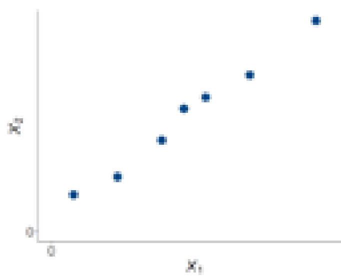


Figura 5.18. *Tendencia lineal positiva,* $\rho_{X_1,X_2} \approx 1$

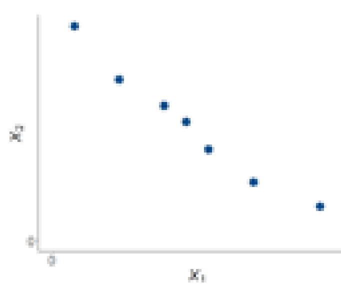


Figura 5.19. *Tendencia lineal negativa,* $\rho_{X_1,X_2} \approx -1$

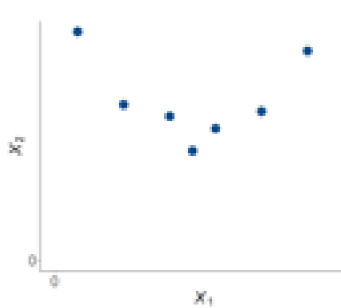


Figura 5.20. *Tendencia no lineal,* ρ_{X_1,X_2} *alejado de* $|1|$

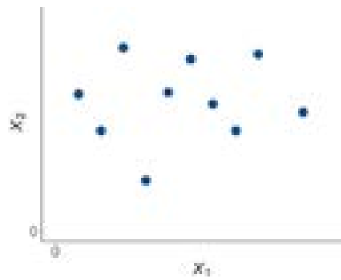


Figura 5.21. *Carencia de tendencia,* $\rho_{X_1,X_2} \approx 0$

Teorema 5.4. *Sean* X_1 *y* X_2 *dos variables aleatorias, entonces,*

$$Cov(X_1, X_2) = E[X_1X_2] - E[X_1]E[X_2].$$

Demostración.

$$\begin{aligned} Cov(X_1, X_2) &= E[X_1 - \mu_1)(X_2 - \mu_2)] \\ &= E[X_1X_2 - \mu_1X_2 - \mu_2X_1 + \mu_1\mu_2] \\ &= E[X_1X_2] - \mu_1E[X_2] - \mu_2E[X_1] + \mu_1\mu_2 \\ &= E[X_1X_2] - \mu_1\mu_2. \end{aligned}$$ □

Ejemplo 5.17. Considerando las variables aleatorias del ejemplo 5.1 (variables también consideradas en los ejemplos 5.2 y 5.8), calcular el coeficiente de correlación de X_1 y X_2.

Se usarán las funciones de densidad conjunta y funciones marginales de X_1 y X_2, estas se calcularon en los ejemplos 5.2 y 5.8, respectivamente. De esta manera,

$$\begin{aligned} E[X_1X_2] =&\sum_{x_2}\sum_{x_1} x_1x_2f_{\underline{X}}(x_1, x_2) \\ =&(1)(1)\left(\tfrac{1}{16}\right) + (2)(1)\left(\tfrac{1}{16}\right) + \cdots + (4)(4)\left(\tfrac{1}{4}\right) = 8.4375, \end{aligned}$$

$$E[X_1] = 1\left(\tfrac{1}{4}\right) + 2\left(\tfrac{1}{4}\right) + 3\left(\tfrac{1}{4}\right) + 4\left(\tfrac{1}{4}\right) = 2.5,$$

$$E[X_2] = 1\left(\tfrac{1}{16}\right) + 2\left(\tfrac{3}{16}\right) + 3\left(\tfrac{5}{16}\right) + 4\left(\tfrac{7}{16}\right) = 3.125.$$

Por lo tanto,

$$Cov(X_1, X_2) = 8.4375 - (2.5)(3.125) = 0.625.$$

Además,

$$E[X_1^2] = 1^2\left(\tfrac{1}{4}\right) + 2^2\left(\tfrac{1}{4}\right) + 3^2\left(\tfrac{1}{4}\right) + 4^2\left(\tfrac{1}{4}\right) = 7.5,$$

$$E[X_2^2] = 1^2\left(\tfrac{1}{16}\right) + 2^2\left(\tfrac{3}{16}\right) + 3^2\left(\tfrac{5}{16}\right) + 4^2\left(\tfrac{7}{16}\right) = 10.625.$$

De esta manera,
$V[X_1] = 1.25; V[X_2] = 0.859375; \sigma_1 = 1.11803$ y $\sigma_2 = 0.927024$.
Por lo tanto,

$$\rho_{X_1,X_2} = \frac{0.625}{(1.11803)(0.92702)} = 0.60302.$$

Ejemplo 5.18. Considerando las variables aleatorias del ejemplo 5.12 (variables también consideradas en los ejemplos 5.9 y 5.5), calcular el coeficiente de correlación de X_1 y X_2.

Se calculará primero $E[X_1X_2]$

$$\begin{aligned} E[X_1X_2] &= \int_0^1 \int_0^{x_2} x_1x_2 6(1-x_2)dx_1dx_2 \\ &= 3\int_0^1 x_2^3(1-x_2)dx_2 \\ &= 3\tfrac{\Gamma(4)\Gamma(2)}{\Gamma(6)} \int_0^1 \tfrac{\Gamma(6)}{\Gamma(4)\Gamma(2)} x_2^3(1-x_2)dx_2 \\ &= 3\tfrac{\Gamma(4)\Gamma(2)}{\Gamma(6)} = \tfrac{3}{20}. \end{aligned}$$

Por otro lado, como $X_1 \sim Beta(1,3)$ y $X_2 \sim Beta(2,2)$, entonces, $E[X_1] = \frac{1}{1+3} = \frac{1}{4}$; $V[X_1] = \frac{1(3)}{(1+3)^2(1+3+1)} = \frac{3}{80}$; $E[X_2] = \frac{2}{2+2} = \frac{1}{2}$ y $V[X_2] = \frac{2(2)}{(2+2)^2(2+2+1)} = \frac{1}{20}$.
Por lo tanto,

$$\begin{aligned} Cov(X_1, X_2) &= E[X_1X_2] - E[X_1]E[X_2] \\ &= \tfrac{3}{20} - \left(\tfrac{1}{4}\right)\left(\tfrac{1}{2}\right) = \tfrac{1}{40}. \end{aligned}$$

Siendo el coeficiente de correlación

$$\rho_{X_1,X_2} = \frac{\frac{1}{40}}{\sqrt{\frac{3}{80}}\sqrt{\frac{4}{80}}} = 0.5773.$$

5.7. Función generadora de momentos de un vector aleatorio

En esta sección se presentará la definición de función generadora de momentos de un vector aleatorio y también sus propiedades más importantes, las cuales serán de gran utilidad en los siguientes temas de probabilidad.

Definición 5.11. Sea $\underline{X} = (X_1, \ldots, X_n)$ un vector aleatorio con función de densidad dada por $f_{\underline{X}}(x_1, \ldots, x_n)$. La *función generadora de momentos (o función generatriz de momentos)* de $\underline{X}$ o *función generadora de momentos conjunta* de $X_1, \ldots, X_n$ es denotada como $m_{\underline{X}}(t_1, \ldots, t_n)$ y se define como

$$m_{\underline{X}}(t_1, \ldots, t_n) = E\left[e^{X_1t_1 + X_2t_2 + \cdots + X_nt_n}\right].$$

Una primera utilidad de esta función es obtener diferentes momentos y determinados valores esperados de las variables aleatorias involucradas.

Supongamos que para el vector aleatorio $\underline{X} = (X_1, \ldots, X_n)$ la función generadora de momentos $m_{\underline{X}}(t_1, \ldots, t_n)$ existe para toda t_i que pertenece al intervalo $(-h, h)$, donde $h > 0$, $i = 1, \ldots, n$. Entonces,

$$E\left[X_1^{r_1} X_2^{r_2} \cdots X_n^{r_n}\right] = \left.\frac{\partial^{r_1+r_2+\cdots+r_n} m_{\underline{X}}(\underline{t})}{\partial t_n^{r_n} \cdots \partial t_2^{r_2} \partial t_1^{r_1}}\right|_{\underline{t}=\underline{0}},$$

donde $\underline{t} = (t_1, \ldots, t_n)$ y $\underline{0} = (0, \ldots, 0)$.

La demostración de este resultado no es propósito de este tema.

Observación. La función generadora de X_i se puede obtener a partir de la función generadora de momentos de $\underline{X}$, de la siguiente manera:

$$m_{\underline{X}}(0, \ldots, 0, t_i, 0, \ldots, 0) = E[e^{t_i X_i}] = m_{X_i}(t_i).$$

Si la función generadora de X_i se obtiene a partir de la función generadora del vector $\underline{X}$, esta recibe el nombre de *función generadora de momentos marginal* de X_i.

Teorema 5.5. (Teorema de unicidad) *Sea $\underline{X} = (X_1, \ldots, X_n)$ un vector aleatorio con función generadora de momentos $m_{\underline{X}}(t_1, \ldots, t_n)$. La función generadora de momentos $m_{\underline{X}}(\underline{t})$ determina en forma única a la distribución de $\underline{X}$.*

Este teorema es de gran utilidad, en particular en la determinación de la distribución de funciones de variables aleatorias cuando se aplica el método de función generadora de momentos, véase el siguiente capítulo. La demostración de este teorema no es propósito de este material.

Ejemplo 5.19. Sea $\underline{X} = (X_1, X_2)$ un vector aleatorio con función de densidad dada por $f_{\underline{X}}(x_1, x_2) = e^{-(x_1+x_2)} I_{[0,\infty)}(x_1) I_{[0,\infty)}(x_2)$. Encontrar
a) La función generadora de momentos de $\underline{X}$.
b) La covarianza de X_1 y X_2.

a) La función generadora de momentos conjunta es

$$\begin{aligned}
m_{\underline{X}}(t_1, t_2) &= E\left[e^{X_1 t_1 + X_2 t_2}\right] \\
&= \int_0^\infty \int_0^\infty e^{x_1 t_1 + x_2 t_2} e^{-x_1 - x_2} dx_1 dx_2 \\
&= \int_0^\infty \int_0^\infty e^{-x_1(1-t_1)} e^{-x_2(1-t_2)} dx_1 dx_2 \\
&= \left[\int_0^\infty e^{-x_1(1-t_1)} dx_1\right] \left[\int_0^\infty e^{-x_2(1-t_2)} dx_2\right] \\
&= \left(-\frac{1}{1-t_1} e^{-x_1(1-t_1)}\Big|_0^\infty\right) \left(-\frac{1}{1-t_2} e^{-x_2(1-t_2)}\Big|_0^\infty\right) \\
&= (1-t_1)^{-1}(1-t_2)^{-1}, \quad \text{si } t_1 < 1 \text{ y } t_2 < 1.
\end{aligned}$$

b) El cálculo de $E[X_1]$, $E[X_2]$, $E[X_1X_2]$, $Cov(X_1, X_2)$ y ρ_{X_1,X_2} se da a continuación:

$$\begin{aligned} E[X_1] &= \left.\frac{\partial m_{\underline{X}}(t_1,t_2)}{\partial t_1}\right|_{\underline{t}=\underline{0}} \\ &= (-1)(1-t_2)^{-1}(-1)(1-t_1)^{-2}\Big|_{\underline{t}=\underline{0}} = 1, \end{aligned}$$

$$\begin{aligned} E[X_2] &= \left.\frac{\partial m_{\underline{X}}(t_1,t_2)}{\partial t_2}\right|_{\underline{t}=\underline{0}} \\ &= (-1)(1-t_1)^{-1}(-1)(1-t_2)^{-2}\Big|_{\underline{t}=\underline{0}} = 1, \end{aligned}$$

$$\begin{aligned} E[X_1X_2] &= \left.\frac{\partial^2 m_{\underline{X}}(t_1,t_2)}{\partial t_1 \partial t_2}\right|_{\underline{t}=\underline{0}} \\ &= (-1)(1-t_2)^{-2}(-1)(1-t_1)^{-2}\Big|_{\underline{t}=\underline{0}} = 1. \end{aligned}$$

Por lo tanto,

$$\begin{aligned} Cov(X_1, X_2) &= E[X_1X_2] - E[X_1]E[X_2] \\ &= 1 - (1)(1) = 0. \end{aligned}$$

De esta manera,

$$\rho_{X_1,X_2} = 0.$$

Cuando el coeficiente de correlación $\rho_{X_1,X_2} = 0$, se dice que las variables aleatorias X_1 y X_2 están *incorrelacionadas*.

5.8. Valores esperados e independencia

En esta sección se tratarán varios resultados donde se relacionan la independencia de variables aleatorias con algunas propiedades de valores esperados de las variables aleatorias involucradas. Estos resultados contribuirán de manera importante en algunos de temas de probabilidad, en particular ayudarán a verificar, con más facilidad, la independencia o dependencia de variables aleatorias.

Teorema 5.6. *Sean $X_1, \ldots, X_n$ variables aleatorias y $g_1(X_1), \ldots, g_n(X_n)$ funciones de estas variables aleatorias, respectivamente. Si $X_1, \ldots, X_n$ son variables aleatorias independientes, entonces,*

$$E[g_1(X_1) \cdots g_n(X_n)] = E[g_1(X_1)] \cdots E[g_n(X_n)].$$

Demostración. (Caso continuo)

$E[g_1(X_1)\cdots g_n(X_n)]$

$$\begin{aligned}
&= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} g_1(x_1)\cdots g_n(x_n) f_{\underline{X}}(x_1,\ldots,x_n)dx_1\cdots dx_n \\
&= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} g_1(x_1)\cdots g_n(x_n) f_{X_1}(x_1)\cdots f_{X_n}(x_n)dx_1\cdots dx_n \\
&= \left[\int_{-\infty}^{\infty} g_1(x_1) f_{X_1}(x_1)dx_1\right]\cdots\left[\int_{-\infty}^{\infty} g_n(x_n) f_{X_n}(x_n)dx_n\right] \\
&= E[g_1(X_1)]\cdots E[g_n(X_n)].
\end{aligned}$$

□

La demostración del caso discreto es análoga.

Corolario 5.1. *Si X_1 y X_2 son variables aleatorias independientes, entonces, X_1 y X_2 son variables incorrelacionadas.*

Demostración.

$$\begin{aligned}
Cov(X_1, X_2) &= E[X_1X_2] - E[X_1]E[X_2] \\
&= E[X_1]E[X_2] - E[X_1]E[X_2] = 0.
\end{aligned}$$

Por lo tanto, X_1 y X_2 están incorrelacionadas. □

Observación. El inverso no necesariamente es cierto, cuando dos variables aleatorias están incorrelacionadas significa que las variables son linealmente independientes, y este hecho no necesariamente implica que las variables sean independientes en general, véase el ejercicio 15 de este este capítulo.

Teorema 5.7. *Las variables aleatorias $X_1, \ldots, X_n$ son independientes, si y solo si*

$$m_{\underline{X}}(t_1, \ldots, t_n) = \prod_{i=1}^{n} m_{X_i}(t_i).$$

Demostración. Supongamos que $X_1, \ldots, X_n$ variables aleatorias independientes, entonces,

$$\begin{aligned}
m_{\underline{X}}(t_1, \ldots, t_n) &= E\left[e^{X_1t_1+\cdots+X_nt_n}\right] \\
&= E\left[e^{X_1t_1}\cdots e^{X_nt_n}\right] \\
&= E\left[e^{X_1t_1}\right]\cdots E\left[e^{X_nt_n}\right] \\
&= m_{X_1}(t_1)\cdots m_{X_n}(t_n) \\
&= \prod_{i=1}^{n} m_{X_i}(t_i).
\end{aligned}$$

Se considerará la demostración de regreso para el caso continuo, para el caso discreto es análogo. Supongamos que se cumple

$$m_{\underline{X}}(t_1, \ldots, t_n) = \prod_{i=1}^{n} m_{X_i}(t_i).$$

De esta manera,

$$
\begin{aligned}
&m_{\underline{X}}(t_1, \ldots, t_n) \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{x_1 t_1 + \cdots + x_n t_n} f_{\underline{X}}(x_1, \ldots, x_n) dx_1 \cdots dx_n \\
&= \left[\int_{-\infty}^{\infty} e^{x_1 t_1} f_{X_1}(x_1) dx_1 \right] \cdots \left[\int_{-\infty}^{\infty} e^{x_n t_n} f_{X_n}(x_n) dx_n \right] \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{x_1 t_1} f_{X_1}(x_1) \cdots e^{x_n t_n} f_{X_n}(x_n) dx_1 \cdots dx_n \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{x_1 t_1 + \cdots + x_n t_n} f_{X_1}(x_1) \cdots f_{X_n}(x_n) dx_1 \cdots dx_n.
\end{aligned}
$$

Por el teorema de unicidad, $f_{\underline{X}}(x_1, \ldots, x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n)$.
Por lo tanto, las variables $X_1, X_2, \ldots, X_n$ son independientes. □

Ejemplo 5.20. Considerando el ejemplo 5.19, verificar si las variables aleatorias X_1 y X_2 son independientes.

La función generadora de momentos ya fue obtenida

$$m_{\underline{X}}(t_1, t_2) = (1 - t_1)^{-1}(1 - t_2)^{-1}, \quad \text{si } t_1 < 1 \text{ y } t_2 < 1.$$

De esta función se obtendrán las funciones generadoras marginales

$$
\begin{aligned}
m_{X_1}(t_1) &= m_{\underline{X}}(t_1, 0) = (1 - t_1)^{-1}, \quad &\text{si } t_1 < 1, \\
m_{X_2}(t_2) &= m_{\underline{X}}(0, t_2) = (1 - t_2)^{-1}, \quad &\text{si } t_2 < 1.
\end{aligned}
$$

De esta manera,

$$
\begin{aligned}
m_{X_1}(t_1) m_{X_2}(t_2) &= (1 - t_1)^{-1}(1 - t_2)^{-1} \\
&= m_{\underline{X}}(t_1, t_2), \quad \text{si } t_1 < 1 \text{ y } t_2 < 1.
\end{aligned}
$$

Por lo tanto, X_1 y X_2 son variables aleatorias independientes.
El alumno también puede verificar la independencia de estas variables verificando que se cumple $f_{\underline{X}}(x_1, x_2) = f_{X_1}(x_1) f_{X_2}(x_2)$.

Ejemplo 5.21. Sea $\underline{X} = (X_1, X_2)$ un vector aleatorio con función de densidad dada por

$$f_{\underline{X}}(x_1, x_2) = c \frac{5^{x_1} 2^{x_2}}{x_1! x_2!} I_{\{0,1,2,\ldots\}}(x_1) I_{\{0,1,2,\ldots\}}(x_2).$$

a) Encontrar c.
b) Encontrar la función generadora de momentos conjunta de X_1 y X_2.
c) Calcular la covarianza de X_1 y X_2.
d) ¿Son X_1 y X_2 variables aleatorias independientes?

a) Se deberá cumplir

$$\begin{aligned}
1 &= \sum_{x_2=0}^{\infty}\sum_{x_1=0}^{\infty} c\frac{5^{x_1}2^{x_2}}{x_1!x_2!}\\
&= c\sum_{x_2=0}^{\infty}\frac{2^{x_2}}{x_2!}\sum_{x_1=0}^{\infty}\frac{5^{x_1}}{x_1!}\\
&= ce^2e^5 = ce^7.
\end{aligned}$$

De esta manera, $c = e^{-7}$.
Por lo tanto,

$$f_{\underline{X}}(x_1, x_2) = e^{-7}\frac{5^{x_1}2^{x_2}}{x_1!x_2!}I_{\{0,1,2,\ldots\}}(x_1)I_{\{0,1,2,\ldots\}}(x_2).$$

b) La función generadora de momentos

$$\begin{aligned}
m_{\underline{X}}(t_1, t_2) &= E\left[e^{X_1t_1+X_2t_2}\right]\\
&= \sum_{x_2=0}^{\infty}\sum_{x_1=0}^{\infty} e^{x_1t_1+x_2t_2}e^{-7}\frac{5^{x_1}2^{x_2}}{x_1!x_2!}\\
&= \left[\sum_{x_2=0}^{\infty} e^{x_2t_2}\frac{2^{x_2}}{x_2!}e^{-2}\right]\left[\sum_{x_1=0}^{\infty} e^{x_1t_1}\frac{5^{x_1}}{x_1!}e^{-5}\right]\\
&= e^{2(e^{t_2}-1)}e^{5(e^{t_1}-1)}.
\end{aligned}$$

Obsérvese que cada una de las dos últimas sumas representa la función generadora de momentos de una distribución Poisson.

c) Para la covarianza, son necesarios los siguientes valores esperados

$$\begin{aligned}
E[X_1] &= \left.\frac{\partial m_{\underline{X}}(t_1, t_2)}{\partial t_1}\right|_{\underline{t}=\underline{0}}\\
&= \left. e^{2(e^{t_2}-1)}e^{5(e^{t_1}-1)}5e^{t_1}\right|_{\underline{t}=\underline{0}} = 5,
\end{aligned}$$

$$\begin{aligned}
E[X_2] &= \left.\frac{\partial m_{\underline{X}}(t_1, t_2)}{\partial t_2}\right|_{\underline{t}=\underline{0}}\\
&= \left. e^{2(e^{t_2}-1)}e^{5(e^{t_1}-1)}2e^{t_2}\right|_{\underline{t}=\underline{0}} = 2,
\end{aligned}$$

$$\begin{aligned}
E[X_1X_2] &= \left.\frac{\partial^2 m_{\underline{X}}(t_1, t_2)}{\partial t_2\partial t_1}\right|_{\underline{t}=\underline{0}}\\
&= \left. e^{2(e^{t_2}-1)}e^{5(e^{t_1}-1)}2e^{t_2}5e^{t_1}\right|_{\underline{t}=\underline{0}} = 2(5) = 10.
\end{aligned}$$

Por lo tanto,

$$Cov(X_1, X_2) = E[X_1X_2] - E[X_1]E[X_2] = 0.$$

Entonces, X_1 y X_2 están incorrelacionadas.
d) Obsérvese

$$\begin{aligned} m_{X_1}(t_1) &= m_{\underline{X}}(t_1, 0) \\ &= e^{5(e^{t_1}-1)}, \end{aligned}$$

$$\begin{aligned} m_{X_2}(t_2) &= m_{\underline{X}}(0, t_2) \\ &= e^{2(e^{t_2}-1)}. \end{aligned}$$

De esta manera,

$$\begin{aligned} m_{X_1}(t_1)m_{X_2}(t_2) &= e^{5(e^{t_1}-1)}e^{2(e^{t_2}-1)} \\ &= m_{\underline{X}}(t_1, t_2). \end{aligned}$$

Por lo tanto, X_1 y X_2 son variables aleatorias independientes.

Teorema 5.8. *Sea $\underline{X} = (X_1, \ldots, X_n)$ un vector aleatorio con función de densidad $f_{\underline{X}}(x_1, \ldots, x_n)$. Sea A_i el recorrido de $X_i, i = 1, 2, \ldots, n$, entonces, $X_1, \ldots, X_n$ son variables aleatorias independientes si y solo si $f_{\underline{X}}(x_1, \ldots, x_n)$ puede expresarse como el producto $g_1(x_1)g_2(x_2)\cdots g_n(x_n)$, donde $g_i(x_i)$ es una función positiva en A_i, para $i = 1, \ldots, n$.*

Demostración. Supongamos que $X_1, \ldots, X_n$ son variables aleatorias independientes, entonces, $f_{\underline{X}}(x_1, \ldots, x_n) = f_{X_1}(x_1)\cdots f_{X_n}(x_n)$.

Sea

$$g_i(x_i) = f_{X_i}(x_i), \quad \text{para}\, i = 1, \ldots, n.$$

Por lo tanto,

$$\begin{aligned} f_{\underline{X}}(x_1, \ldots, x_n) &= f_{X_1}(x_1)\cdots f_{X_n}(x_n) \\ &= g_1(x_1)g_2(x_2)\cdots g_n(x_n), \text{ donde } g_i(x_i) > 0, \text{ para } x_i \in A_i. \end{aligned}$$

Para el regreso de la demostración, suponemos que $\underline{X}$ es continuo.
Suponemos que $f_{\underline{X}}(x_1, \ldots, x_n) = g_1(x_1)g_2(x_2)\cdots g_n(x_n)$, entonces, $f_{X_i}(x_i)$

$$\begin{aligned} &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_{\underline{X}}(x_1, \ldots, x_n)dx_1\cdots dx_{i-1}dx_{i+1}\cdots dx_n \\ &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} g_1(x_1)\cdots g_n(x_n)dx_1\cdots dx_{i-1}dx_{i+1}\cdots dx_n \\ &= g_i(x_i)\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} g_1(x_1)\cdots g_n(x_n)dx_1\cdots dx_{i-1}dx_{i+1}\cdots dx_n. \end{aligned}$$

El resultado de la anterior integral múltiple es una constante, sea c_i esta constante, de esta manera,

$$f_{X_i}(x_i) = c_i g_i(x_i), \quad \text{para}\, i = 1, 2, \ldots, n.$$

Por lo tanto,

$$\int_{-\infty}^{\infty} c_i g_i(x_i) dx_i = 1, \quad \text{para}\, i = 1, 2, \ldots, n.$$

De esta manera,

$$\begin{aligned}
1 &= \left[\int_{-\infty}^{\infty} c_1 g_1(x_1) dx_1\right]\left[\int_{-\infty}^{\infty} c_2 g_2(x_2) dx_2\right]\cdots\left[\int_{-\infty}^{\infty} c_n g_n(x_n) dx_n\right] \\
&= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} c_1\cdots c_n g_1(x_1)\cdots g_n(x_n) dx_1\cdots dx_n \\
&= c_1\cdots c_n \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} g_1(x_1)\cdots g_n(x_n) dx_1\cdots dx_n \\
&= c_1\cdots c_n \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_{\underline{X}}(x_1,\ldots,x_n) dx_1\cdots dx_n \\
&= c_1\cdots c_n.
\end{aligned}$$

Por lo tanto, $c_1\cdots c_n = 1$.
Entonces,

$$\begin{aligned}
f_{\underline{X}}(x_1,\ldots,x_n) &= g_1(x_1)\cdots g_n(x_n) \\
&= c_1\cdots c_n g_1(x_1)\cdots g_n(x_n) \\
&= c_1 g_1(x_1) c_2 g(x_2)\cdots c_n g_n(x_n) \\
&= f_{X_1}(x_1) f_{X_2}(x_2)\cdots f_{X_n}(x_n).
\end{aligned}$$

Por lo tanto, $X_1,\ldots,X_n$ son variables aleatorias independientes. □

La demostración para el caso discreto es análoga, en lugar de usar integrales se usan sumas.

Ejemplo 5.22. Verificar para las siguientes funciones de densidad si las variables son independientes o no,
a) $f_{\underline{X}}(x_1,x_2) = (x_1+x_2)I_{(0,1)}(x_1)I_{(0,1)}(x_2)$.
b) $f_{\underline{X}}(x_1,x_2) = 8x_1x_2I_{(0,x_2)}(x_1)I_{(0,1)}(x_2)$.
c) $f_{\underline{X}}(x_1,x_2) = 4x_1x_2I_{(0,1)}(x_1)I_{(0,1)}(x_2)$.
d) $f_{\underline{X}}(x_1,x_2) = \frac{1}{16}I_{\{1,2,3,4\}}(x_1)I_{\{1,2,3,4\}}(x_2)$.

a) Como no se puede expresar la función de densidad de $\underline{X}$ como el producto de dos funciones positivas, de X_1 y X_2, entonces, las variables aleatorias son dependientes.

b) Como la función indicadora $I_{(0,x_2)}(x_1)$ no se puede expresar como el producto dos funciones positivas, de X_1 y X_2, entonces, las variables aleatorias no son independientes.

c) La función de densidad de $\underline{X}$ se puede expresar como

$$f_{\underline{X}}(x_1, x_2) = [4x_1 I_{(0,1)}(x_1)][x_2 I_{(0,1)}(x_2)] = g_1(x_1)g_2(x_2).$$

Entonces, X_1 y X_2 son variables aleatorias independientes.

d) Obsérvese

$$f_{\underline{X}}(x_1, x_2) = \left[\tfrac{1}{16} I_{\{1,2,3,4\}}(x_1)\right] \left[I_{\{1,2,3,4\}}(x_2)\right] = g_1(x_1)g_2(x_2).$$

Por lo tanto, X_1 y X_2 son variables aleatorias independientes.

5.9. Valores esperados condicionales

A partir de una función de densidad condicional, es posible calcular diferentes valores esperados, estos reciben el nombre de *valores esperados condicionales*. En esta sección se tratarán diferentes valores esperados condicionales, como las medias y varianzas condicionales, también se tratarán sus principales propiedades.

Definición 5.12. Sea $\underline{X} = (X_1, X_2)$ un vector aleatorio y $g(X_1, X_2)$ una función real de $\underline{X}$, esto es, una función de $\mathbb{R}^2$ a $\mathbb{R}$. La *esperanza condicional* o *valor esperado condicional* de $g(X_1, X_2)$ dado que $X_2 = x_2$, se denota como $E[g(X_1, X_2)|X_2 = x_2]$ o $E[g(X_1, X_2)|x_2]$, y se define como

$$E[g(X_1, X_2)|x_2] = \begin{cases} \displaystyle\sum_{x_1} g(x_1, x_2) f_{X_1|X_2}(x_1|x_2), & \text{si } \underline{X} \text{ es discreto,} \\ \displaystyle\int_{-\infty}^{\infty} g(x_1, x_2) f_{X_1|X_2}(x_1|x_2)dx_1, & \text{si } \underline{X} \text{ es continuo,} \end{cases}$$

si la suma o la integral es finita.

Y la *esperanza condicional o valor esperando condicional* de $g(X_1, X_2)$ dado que $X_1 = x_1$, se denota como $E[g(X_1, X_2)|X_1 = x_1]$ o $E[g(X_1, X_2)|x_1]$, y se define como

$$E[g(X_1, X_2)|x_1] = \begin{cases} \displaystyle\sum_{x_2} g(x_1, x_2) f_{X_2|X_1}(x_2|x_1), & \text{si } \underline{X} \text{ es discreto,} \\ \displaystyle\int_{-\infty}^{\infty} g(x_1, x_2) f_{X_2|X_1}(x_2|x_1)dx_2, & \text{si } \underline{X} \text{ es continuo,} \end{cases}$$

si la suma o la integral es finita.

La esperanza condicional de $g(X_1, X_2)$ dado que $X_2 = x_2$ (o $X_1 = x_1$), también es conocida como *media o promedio condicional* de $g(X_1, X_2)$ dado que $X_2 = x_2$ (o $X_1 = x_1$).

En particular, si $g(X_1, X_2) = X_1$,

$$E[X_1|x_2] = \begin{cases} \displaystyle\sum_{x_1} x_1 f_{X_1|X_2}(x_1|x_2), & \text{si } \underline{X} \text{ es discreto,} \\ \displaystyle\int_{-\infty}^{\infty} x_1 f_{X_1|X_2}(x_1|x_2)dx_1, & \text{si } \underline{X} \text{ es continuo.} \end{cases}$$

Si $g(X_1, X_2) = X_2$,

$$E[X_2|x_1] = \begin{cases} \displaystyle\sum_{x_2} x_2 f_{X_2|X_1}(x_2|x_1), & \text{si } \underline{X} \text{ es discreto,} \\ \displaystyle\int_{-\infty}^{\infty} x_2 f_{X_2|X_1}(x_2|x_1)dx_2, & \text{si } \underline{X} \text{ es continuo.} \end{cases}$$

Ejemplo 5.23. Considerando las variables aleatorias del ejemplo 5.18 (también ejemplos 5.5, 5.9, 5.12), calcular los valores esperados condicionales $E[X_1|x_2]$ y $E[X_2|x_1]$.

Las funciones de densidad condicionales ya fueron calculadas en el ejemplo 5.12, esto es,

$$f_{X_1|X_2}(x_1|x_2) = \frac{1}{x_2} I_{[0,x_2]}(x_1), \text{ si } 0 < x_2 < 1,$$

$$f_{X_2|X_1}(x_2|x_1) = \frac{2(1-x_2)}{(1-x_1)^2} I_{[x_1,1]}(x_2), \text{ si } 0 \leq x_1 < 1.$$

De la primera función de densidad, se sabe

$$X_1|x_2 \sim U(0, x_2), \quad \text{si } 0 < x_2 < 1.$$

Por lo tanto,

$$E[X_1|x_2] = \frac{x_2}{2}, \quad \text{si } 0 < x_2 < 1.$$

Por otro lado,

$$\begin{aligned} E[X_2|x_1] &= \int_{x_1}^{1} \frac{2x_2(1-x_2)}{(1-x_1)^2} dx_2 \\ &= \frac{2}{(1-x_1)^2} \int_{x_1}^{1} (x_2 - x_2^2) dx_2 \\ &= \frac{2}{(1-x_1)^2} \left[\frac{x_2^2}{2} - \frac{x_2^3}{3} \right]_{x_1}^{1} \\ &= \frac{2}{(1-x_1)^2} \left[\frac{1}{2} - \frac{1}{3} - \frac{x_1^2}{2} + \frac{x_1^3}{3} \right]. \end{aligned}$$

Entonces,

$$E[X_2|x_1] = \frac{1 - 3x_1^2 + 2x_1^3}{3(1-x_1)^2}, \text{ si } 0 \leq x_1 < 1.$$

Obsérvese algunos casos particulares, $E[X_1|0.5] = 0.25$, $E[X_2|0.2] = 0.46667$.

Definición 5.13. Sea $\underline{X} = (X_1, X_2)$ un vector aleatorio, la *varianza condicional* de X_1 dado que $X_2 = x_2$, se denota como $V[X_1|x_2]$ y se define como

$$V[X_1|x_2] = E[X_1^2|x_2] - E^2[X_1|x_2].$$

Y la varianza condicional de X_2 dado que $X_1 = x_1$, se denota como $V[X_2|x_1]$ y se define como

$$V[X_2|x_1] = E[X_2^2|x_1] - E^2[X_2|x_1].$$

El valor esperado condicional $E[X_1|x_2]$ depende del valor particular que toma la variable X_2, esto es, de x_2, lo que significa que este valor esperado es una función de x_2. Dicho de otra manera, $E[X_1|x_2] = h(x_2)$. Esta función no es aleatoria, ya que depende de un valor específico de X_2.

Se puede definir una función de la variable aleatoria X_2, de la siguiente manera: $h(X_2) = E[X_1|X_2]$. En este caso, esta función si es una variable aleatoria, de manera análoga se puede definir también la función $h(X_1) = E[X_2|X_1]$. El siguiente teorema tiene que ver con este tipo de funciones.

Teorema 5.9. *Sea $\underline{X} = (X_1, X_2)$ un vector aleatorio, y $g(X_1)$ una función de X_1, entonces, $E[E[g(X_1)|X_2]] = E[g(X_1)]$.*

Demostración. (Caso continuo)

Definimos $h(X_2) = E[g(X_1)|X_2]$.

Entonces,

$$\begin{aligned}
E[E[g(X_1)|X_2]] &= E[h(X_2)] \\
&= \int_{-\infty}^{\infty} h(x_2) f_{X_2}(x_2) dx_2 \\
&= \int_{-\infty}^{\infty} E[g(X_1)|x_2] f_{X_2}(x_2) dx_2 \\
&= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} g(x_1) f_{X_1|X_2}(x_1|x_2) dx_1 \right] f_{X_2}(x_2) dx_2 \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x_1) f_{X_1|X_2}(x_1|x_2) f_{X_2}(x_2) dx_1 dx_2 \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x_1) f_{\underline{X}}(x_1, x_2) dx_1 dx_2 \\
&= E[g(X_1)].
\end{aligned}$$ □

Para el caso discreto, el procedimiento de la demostración es el mismo, solo que en lugar de las integrales, se realizarán sumas.

En particular, se cumplen las siguientes igualdades:

$$E[E[X_1|X_2]] = E[X_1],$$

$$E[E[X_2|X_1]] = E[X_2].$$

Ejemplo 5.24. Con la idea de ejemplificar el teorema anterior, considerar las variables aleatorias del ejemplo 5.23 (variables también usadas en los ejemplos 5.5, 5.9, 5.12, 5.18) y verificar que se cumple $E[E[X_1|X_2]] = E[X_1]$.

Para este caso, del ejemplo 5.23, véase que se cumple

$$E[X_1|X_2] = \frac{X_2}{2}.$$

Por lo tanto,

$$E[E[X_1|X_2]] = E\left[\frac{X_2}{2}\right] = \tfrac{1}{4}.$$

Por otro lado, se sabe que $X_1 \sim Beta(1,3)$. De esta manera, $E[X_1] = \frac{1}{4}$.

Para este mismo ejemplo, se deja como ejercicio para el estudiante verificar que se cumple $E[E[X_2|X_1]] = E[X_2]$.

Teorema 5.10. *Sea $\underline{X} = (X_1, X_2)$ un vector aleatorio, entonces,*

$$V[X_1] = V[E[X_1|X_2]] + E[V[X_1|X_2]],$$

$$V[X_2] = V[E[X_2|X_1]] + E[V[X_2|X_1]].$$

La demostración se queda como ejercicio para el estudiante.

Ejemplo 5.25. Con la idea de ejemplificar el teorema anterior, verificar que se cumple $V[X_1] = V[E[X_1|X_2]] + E[V[X_1|X_2]]$, con las variables aleatorias involucradas en el ejemplo 5.23 (también en los ejemplos 5.5, 5.9, 5.12, 5.18).

Se sabe que $X_1|x_2 \sim U(0, x_2)$, por lo tanto, $E[X_1|x_2] = \dfrac{x_2}{2}$ y $V[X_1|x_2] = \dfrac{x_2^2}{12}$, véase ejemplo 5.23.

De esta manera,

$$\begin{aligned} V[E[X_1|X_2]] + E[V[X_1|X_2]] &= V\left[\frac{X_2}{2}\right] + E\left[\frac{X_2^2}{12}\right] \\ &= \tfrac{1}{4}V[X_2] + \tfrac{1}{12}(V[X_2] + E^2[X_2]) \\ &= \left(\tfrac{1}{4}\right)\left(\tfrac{1}{20}\right) + \left(\tfrac{1}{12}\right)\left(\tfrac{1}{20} + \tfrac{1}{4}\right) = \tfrac{3}{80}. \end{aligned}$$

Ya se había calculado la varianza de X_1, esto es, $V[X_1] = \frac{3}{80}$.

También, ya se habían calculado $E[X_2]$ y $V[X_2]$, véase el ejemplo 5.18.

Teorema 5.11. *Sea $\underline{X} = (X_1, X_2)$ un vector aleatorio, entonces,*

$$E[h_1(X_2) + h_2(X_2)|x_1] = E[h_1(X_2)|x_1] + E[h_2(X_2)|x_1].$$

$$E[h_1(X_1)h_2(X_2)|x_1] = h_1(x_1)E[h_2(X_2)|x_1].$$

La demostración del anterior teorema se deja como ejercicio para el estudiante.

Ejemplo 5.26. La cantidad de dinero existente en una cuenta, al principio de cada mes, es una variable aleatoria X_1 con distribución $U(0, 2)$. Por otro lado, para una cantidad fija en la cuenta x_1, la cantidad gastada durante el mes es otra variable aleatoria X_2 con distribución $U(0, x_1)$. Obtener

a) La función de densidad conjunta de X_1 y X_2.

b) La probabilidad de que en un mes se gaste más de 0.5, si se sabe que en la cuenta había 1.

c) La probabilidad de haber tenido en la cuenta al inicio del mes más de 1, si se sabe que se gastaron más de 0.5.

d) $E[X_2]$.

a) Se sabe que $X_1 \sim U(0, 2)$ y $X_2|x_1 \sim U(0, x_1)$.
Gráficamente el recorrido conjunto de X_1 y X_2 es

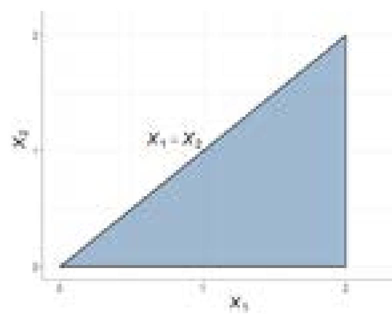


Figura 5.22. *Recorrido conjunto de X_1 y X_2, ejemplo 5.26*

La función de densidad conjunta es igual a

$$\begin{aligned} f_{\underline{X}}(x_1, x_2) &= f_{X_1}(x_1) f_{X_2|X_1}(x_2|x_1) \\ &= \frac{1}{2x_1} I_{(0,x_1)}(x_2) I_{(0,2)}(x_1) \\ &= \frac{1}{2x_1} I_{(x_2,2)}(x_1) I_{(0,2)}(x_2). \end{aligned}$$

b) Obsérvese

$$f_{X_2|X_1}(x_2|1) = I_{(0,1)}(x_2).$$

Por lo tanto,

$$P[X_2 > 0.5 | X_1 = 1] = 0.5.$$

c) La probabilidad que se pide se calcula de la siguiente manera:

$$P[X_1 > 1 | X_2 > 0.5] = \frac{P[(X_1 > 1) \cap (X_2 > 0.5)]}{P[X_2 > 0.5]}.$$

Donde

$$\begin{aligned} P[(X_1 > 1) \cap (X_2 > 0.5)] &= \int_1^2 \int_{0.5}^{x_1} \frac{1}{2x_1} dx_2 dx_1 \\ &= \int_1^2 \frac{1}{2x_1}(x_1 - 0.5) dx_1 \\ &= \int_1^2 \left(\frac{1}{2} - \frac{1}{4x_1} \right) dx_1 \\ &= [0.5x_1 - 0.25ln(x_1)]_1^2 = 0.32671, \end{aligned}$$

$$\begin{aligned} P[X_2 > 0.5] &= \int_{0.5}^2 \int_{0.5}^{x_1} \frac{1}{2x_1} dx_2 dx_1 \\ &= \int_{0.5}^2 \left(\frac{1}{2} - \frac{1}{4x_1} \right) dx_1 \\ &= [0.5x_1 - 0.25ln(x_1)]_{0.5}^2 = 0.40343. \end{aligned}$$

Por lo tanto,

$$P[X_1 > 1 | X_2 > 0.5] = \frac{0.3267}{0.4034} = 0.8098.$$

d) El cálculo del valor esperado

$$\begin{aligned} E[X_2] &= E[E[X_2|X_1]] \\ &= E\left[\frac{X_1}{2} \right] = 0.5. \end{aligned}$$

Ejemplo 5.27. El número de siniestros que ocurren por día es una variable aleatoria X, la cual se distribuye $Poisson(\lambda)$, a su vez λ es una variable aleatoria con distribución exponencial con media igual a 5. Encontrar el valor esperado de X.

a) Usando la función de densidad de X.

b) Sin usar la función de densidad de X.

a) X = Número de siniestros por día.
$X|\lambda \sim Poisson(\lambda)$ y $\lambda \sim Exp(\frac{1}{5})$.
De esta manera, la función de densidad conjunta de X y λ es

$$f_{X,\lambda}(x, \lambda) = \frac{\lambda^x}{x!} e^{-\lambda} I_{\{0,1,\ldots\}}(x) \frac{1}{5} e^{-\frac{\lambda}{5}} I_{[0,\infty)}(\lambda).$$

Por lo tanto,

$$\begin{aligned} f_X(x) &= \frac{1}{5} \int_0^\infty \frac{\lambda^x}{x!} e^{-\frac{6}{5}\lambda} d\lambda \\ &= \frac{1}{5} \left(\frac{5}{6} \right)^{x+1} \int_0^\infty \frac{\left(\frac{6}{5}\right)^{x+1}}{x!} \lambda^x e^{-\frac{6}{5}\lambda} d\lambda \\ &= \left(\frac{5}{6} \right)^{x+1} \frac{1}{5}. \end{aligned}$$

La última integral vale 1, ya que se integró la función de densidad de una $Gamma(x+1, \frac{6}{5})$. De esta manera,

$$f_X(x) = \left(\frac{5}{6}\right)^{x+1} \frac{1}{5} I_{\{0,1,\ldots\}}(x).$$

Ahora, obsérvese

$$f_X(x) = \left(\frac{5}{6}\right)^{x+1} \frac{1}{5} I_{\{0,1,\ldots\}}(x) = \left(\frac{1}{6}\right) \left(\frac{5}{6}\right)^{x} I_{\{0,1,\ldots\}}(x).$$

Por lo que concluimos, $X \sim$*Geométrica*$(p = \frac{1}{6})$. Entonces, $E[X] = 5$.

b) Sin usar la función de densidad de X, $E[X] = E[E[X|\lambda]]] = E[\lambda] = 5$.

Ejemplo 5.28. Considérese el ejemplo 5.3, el problema de la urna donde se sacaron 3 bolas al azar.

a) Encontrar las funciones de densidad marginales.
b) Calcular los valores esperados de X_1 y X_2.
c) Encontrar las funciones de densidad condicionales.
d) Calcular los valores esperados condicionales $E[X_1|x_2]$ y $E[X_2|x_1]$.
e) Verificar para este ejemplo que se cumple $E[E[X_1|X_2]] = E[X_1]$.

a) En el ejemplo 5.3 se obtuvo la función densidad conjunta,

$$\begin{aligned} f_{\underline{X}}(x_1, x_2) &= \frac{\binom{3}{x_1}\binom{2}{x_2}\binom{3}{3-x_1-x_2}}{\binom{8}{3}} I_{\{0,1,\ldots 3-x_2\}}(x_1) I_{\{0,1,2,3\}}(x_2) \\ &= \frac{\binom{3}{x_1}\binom{2}{x_2}\binom{3}{3-x_1-x_2}}{\binom{8}{3}} I_{\{0,1,\ldots 3-x_1\}}(x_2) I_{\{0,1,2,3\}}(x_1). \end{aligned}$$

Entonces,

$$\begin{aligned} f_{X_1}(x_1) &= \sum_{x_2=0}^{3-x_1} \frac{\binom{3}{x_1}\binom{2}{x_2}\binom{3}{3-x_1-x_2}}{\binom{8}{3}} \\ &= \frac{\binom{3}{x_1}}{\binom{8}{3}} \sum_{x_2=0}^{3-x_1} \binom{2}{x_2}\binom{3}{(3-x_1)-x_2} \\ &= \frac{\binom{3}{x_1}\binom{5}{3-x_1}}{\binom{8}{3}}, \text{ donde } x_1 = 0, 1, 2, 3. \end{aligned}$$

Por lo tanto, $X_1 \sim H(N = 8, K = 3, n = 3)$.

La función de densidad marginal de X_2

$$
\begin{aligned}
f_{X_2}(x_2) &= \sum_{x_1=0}^{3-x_2} \frac{\binom{3}{x_1}\binom{2}{x_2}\binom{3}{3-x_1-x_2}}{\binom{8}{3}} \\
&= \frac{\binom{2}{x_2}}{\binom{8}{3}} \sum_{x_1=0}^{3-x_2} \binom{3}{x_1}\binom{3}{(3-x_2)-x_1} \\
&= \frac{\binom{2}{x_2}\binom{6}{3-x_2}}{\binom{8}{3}}, \text{ donde } x_1 = 0, 1, 2.
\end{aligned}
$$

Por lo tanto, $X_2 \sim H(N = 8, K = 2, n = 3)$.

b) Por lo tanto, los valores esperados de X_1 y X_2 son, respectivamente, $E[X_1] = \frac{9}{8}$ y $E[X_2] = \frac{3}{4}$.

c) La función de densidad condicional de X_1 dado $X_2 = x_2$ se calcula de la siguiente manera:

$$
\begin{aligned}
f_{X_1|X_2}(x_1|x_2) &= \frac{f_{\underline{X}}(x_1, x_2)}{f_{X_2}(x_2)} \\
&= \frac{\dfrac{\binom{3}{x_1}\binom{2}{x_2}\binom{3}{3-x_1-x_2}}{\binom{8}{3}} I_{\{0,1,\ldots 3-x_2\}}(x_1) I_{\{0,1,2\}}(x_2)}{\dfrac{\binom{2}{x_2}\binom{6}{3-x_2}}{\binom{8}{3}} I_{\{0,1,2\}}(x_2)} \\
&= \frac{\binom{3}{x_1}\binom{3}{(3-x_2)-x_1}}{\binom{6}{3-x_2}} I_{\{0,1,\ldots,3-x_2\}}(x_1), \quad x_2 = 0, 1, 2.
\end{aligned}
$$

Se puede observar que la anterior función de densidad corresponde a la de una distribución hipergeométrica, siendo más específicos, se puede apreciar que $X_1|x_2 \sim H(N = 6, K = 3, n = 3 - x_2)$, donde $x_2 = 0, 1, 2$.

Ahora,

$$
\begin{aligned}
f_{X_2|X_1}(x_2|x_1) &= \frac{f_{\underline{X}}(x_1,x_2)}{f_{X_1}(x_1)}\\
&= \frac{\dfrac{\binom{3}{x_1}\binom{2}{x_2}\binom{3}{3-x_1-x_2}}{\binom{8}{3}} I_{\{0,1,\ldots 3-x_1\}}(x_2) I_{\{0,1,2,3\}}(x_1)}{\dfrac{\binom{3}{x_1}\binom{5}{3-x_1}}{\binom{8}{3}} I_{\{0,1,2,3\}}(x_1)}\\
&= \frac{\binom{2}{x_2}\binom{3}{(3-x_1)-x_2}}{\binom{5}{3-x_1}} I_{\{0,1,\ldots,3-x_1\}}(x_2), \quad x_1 = 0,1,2,3.
\end{aligned}
$$

En este caso, $X_2|x_1 \sim H(N=5, K=2, n=3-x_1)$, donde $x_1 = 0,1,2,3$.

d) Los valores esperados condicionales se podrán calcular a partir de los resultados anteriores, esto es,

$$E[X_1|x_2] = \frac{3-x_2}{2}, \text{ donde } x_2 = 0,1,2.$$

$$E[X_2|x_1] = \frac{2(3-x_1)}{5}, \text{ donde } x_1 = 0,1,2,3.$$

e) Ahora obsérvese

$$
\begin{aligned}
E[E[X_1|X_2]] &= E\left[\frac{3-X_2}{2}\right]\\
&= \tfrac{9}{8} = E[X_1].
\end{aligned}
$$

Ejemplo 5.29. Considerando los ejemplos 5.2, 5.8 y 5.13, verificar que se cumple $E[E[X_1|X_2]] = E[X_1]$.

La función de densidad condicional de X_1 dado $X_2 = x_2$ se encontró en el ejemplo 5.13,

$$
f_{X_1|X_2}(x_1|x_2) = \begin{cases}
I_{\{1\}}(x_1), & \text{si } x_2 = 1,\\[2ex]
\frac{1}{3}I_{\{1\}}(x_1) + \frac{2}{3}I_{\{2\}}(x_1), & \text{si } x_2 = 2,\\[2ex]
\frac{1}{5}I_{\{1,2\}}(x_1) + \frac{3}{5}I_{\{3\}}(x_1), & \text{si } x_2 = 3,\\[2ex]
\frac{1}{7}I_{\{1,2,3\}}(x_1) + \frac{4}{7}I_{\{4\}}(x_1), & \text{si } x_2 = 4.
\end{cases}
$$

Entonces, se puede obtener $E[X_1|x_2]$ de la siguiente manera:

$$E[X_1|x_2] = \begin{cases} 1, & \text{si } x_2 = 1, \\ \\ \left(\frac{1}{3}\right)1 + \left(\frac{2}{3}\right)2 = \frac{5}{3}, & \text{si } x_2 = 2, \\ \\ \left(\frac{1}{5}\right)(1+2) + \left(\frac{3}{5}\right)3 = \frac{12}{5}, & \text{si } x_2 = 3, \\ \\ \left(\frac{1}{7}\right)(1+2+3) + \left(\frac{4}{7}\right)4 = \frac{22}{7}, & \text{si } x_2 = 4. \end{cases}$$

Obsérvese que para cada valor de x_2 (4 valores), se encontró un valor de $E[X_1|x_2]$ diferente.

Por lo tanto,

$$\begin{aligned} E[E[X_1|X_2]] &= (1)f_{X_2}(1) + \left(\tfrac{5}{3}\right)f_{X_2}(2) + \left(\tfrac{12}{5}\right)f_{X_2}(3) + \left(\tfrac{22}{7}\right)f_{X_2}(4) \\ &= (1)\left(\tfrac{1}{16}\right) + \left(\tfrac{5}{3}\right)\left(\tfrac{3}{16}\right) + \left(\tfrac{12}{5}\right)\left(\tfrac{5}{16}\right) + \left(\tfrac{22}{7}\right)\left(\tfrac{7}{16}\right) \\ &= \tfrac{40}{16} = 2.5. \end{aligned}$$

Por otro lado, $E[X_1] = 1\left(\frac{1}{4}\right) + 2\left(\frac{1}{4}\right) + 3\left(\frac{1}{4}\right) + 4\left(\frac{1}{4}\right) = \frac{10}{4} = 2.5$.

Por lo tanto, $E[E[X_1|X_2]] = E[X_1]$.

Se deja como ejercicio para el estudiante, verificar $E[E[X_2|X_1]] = E[X_2]$.

Ejemplo 5.30. Considerando los ejemplos 5.6, 5.10 y 5.14, verificar que se cumple $E[E[X_1|X_2]] = E[X_1]$ y $E[E[X_2|X_1]] = E[X_2]$.

Las funciones de densidad marginales se obtuvieron en el ejemplo 5.10,

$$f_{X_1}(x_1) = (x_1+1)I_{(-1,0]}(x_1) + (1-x_1)I_{(0,1)}(x_1),$$

$$f_{X_2}(x_2) = 2(1-x_2)I_{(0,1)}(x_2).$$

De esta manera,

$$\begin{aligned} E[X_1] &= \int_{-1}^{0} x_1(x_1+1)dx_1 + \int_{0}^{1} x_1(1-x_1)dx_1 \\ &= \left[\frac{x_1^3}{3} + \frac{x_1^2}{2}\right]_{-1}^{0} + \left[\frac{x_1^2}{2} - \frac{x_1^3}{3}\right]_{0}^{1} = 0. \end{aligned}$$

Obsérvese que $X_2 \sim Beta(1,2)$, entonces, $E[X_2] = \frac{1}{3}$.

En el ejemplo 5.14 se obtuvieron las funciones condicionales, esto es,

$$X_1|x_2 \sim U(x_2-1, 1-x_2), \text{ si } 0 \leq x_2 < 1.$$

$$X_2|x_1 \sim \begin{cases} U(0, x_1+1), & \text{si } -1 < x_1 \leq 0, \\ \\ U(0, 1-x_1), & \text{si } 0 < x_1 < 1. \end{cases}$$

De esta manera,

$$E[X_1|x_2] = \frac{1 - x_2 + x_2 - 1}{2} = 0, \text{ si } 0 \leq x_2 < 1.$$

$$E[X_2|x_1] = \begin{cases} \dfrac{x_1+1}{2}, & \text{si } -1 < x_1 \leq 0, \\ \\ \dfrac{1-x_1}{2}, & \text{si } 0 < x_1 < 1. \end{cases}$$

Se puede observar $E[E[X_1|X_2]] = 0$, por lo tanto, $E[E[X_1|X_2]] = E[X_1]$. Sea $h(X_1) = E[X_2|X_1]$, entonces,

$$\begin{aligned} E[E[X_2|X_1]] &= E[h(X_1)] \\ &= \int_{-1}^{1} h(x_1) f_{X_1}(x_1) dx_1 \\ &= \int_{-1}^{0} h(x_1) f_{X_1}(x_1) dx_1 + \int_{0}^{1} h(x_1) f_{X_1}(x_1) dx_1 \\ &= \int_{-1}^{0} \left(\frac{1+x_1}{2}\right)(1+x_1)dx_1 + \int_{0}^{1} \left(\frac{1-x_1}{2}\right)(1-x_1)dx_1 \\ &= \frac{1}{2}\int_{-1}^{0} (1+2x_1+x_1^2)dx_1 + \frac{1}{2}\int_{0}^{1} (1-2x_1+x_1^2)dx_1 \\ &= \left(\frac{1}{2}\right)\left[x_1 + x_1^2 + \frac{x_1^3}{3}\right]_{-1}^{0} + \left(\frac{1}{2}\right)\left[x_1 - x_1^2 + \frac{x_1^3}{3}\right]_{0}^{1} = \tfrac{1}{3}. \end{aligned}$$

Por lo tanto, $E[E[X_2|X_1]] = E[X_2]$.

5.10. Desigualdad de Cauchy-Schwarz

En esta sección se presentará la desigualdad de Cauchy-Schwarz, este resultado tiene aplicaciones en análisis de regresión, específicamente contribuye para determinar una cota para el coeficiente de correlación.

Teorema 5.12. *(Desigualdad de Cauchy-Schwarz) Sean X_1 y X_2 variables aleatorias con segundos momentos finitos, entonces,*

$$E^2[X_1X_2] \leq E[X_1^2]E[X_2^2].$$

Cumpliéndose la igualdad si y solo si $P[cX_1 = X_2] = 1$, donde c es una constante.

Demostración. Si los segundos momentos $E[X_1^2]$ y $E[X_2^2]$ existen, entonces, también los primeros momentos $E[X_1]$ y $E[X_2]$, esto es, por la desigualdad de Jensen (teorema 2.9), $E^2[X_i] \leq E[X_i^2]$, $i = 1, 2$.

Definimos la siguiente función de t.

$$h(t) = E[(tX_1 - X_2)^2].$$

Se puede observar que $h(t) \geq 0$ para todo $t \in \mathbb{R}$, además, como función de t es un polinomio de segundo grado, esto es,

$$h(t) = t^2 E[X_1^2] - 2tE[X_1X_2] + E[X_2^2],$$

donde $\Delta(h(t)) = 4E^2[X_1X_2] - 4E[X_1^2]E[X_2^2]$ se le conoce como el discriminante del polinomio.

Se sabe que $h(t) > 0$ si y solo si $\Delta(h(t)) < 0$, esto es, si y solo si $E^2[X_1X_2] < E[X_1^2]E[X_2^2]$.

Por otro lado, $h(t) = 0$ si y solo si $\Delta(h(t)) = 0$, y esto ocurre si y solo si $E^2[X_1X_2] = E[X_1^2]E[X_2^2]$.

Además, no puede ser $\Delta(h(t)) > 0$, ya que tendría dos raíces, lo que significa que $h(t)$ es negativa para algunos valores de t, lo cual no es posible.

En conclusión,

$$E^2[X_1X_2] \leq E[X_1^2]E[X_2^2].$$

Ahora supongamos que se cumple la igualdad $E^2[X_1X_2] = E[X_1^2]E[X_2^2]$, entonces, se cumple $4E^2[X_1X_2] - 4E[X_1^2]E[X_2^2] = \Delta(h(t)) = 0$. Por lo tanto, existe una única raíz $t_0 \in \mathbb{R}$ tal que $h(t_0) = 0$. Si existe t_0 tal que $0 = h(t_0)$, entonces, $E[(t_0X_1 - X_2)^2] = 0$, lo cual implica $E[t_0X_1 - X_2] = 0$ (por la desigualdad de Jensen, teorema 2.9). De esta manera, también se cumple $V[t_0X_1 - X_2] = 0$, lo que significa que $t_0X_1 - X_2$ es una variable aleatoria degenerada en 0, esto es, $P[t_0X_1 - X_2 = 0] = 1$, esto es, $P[t_0X_1 = X_2] = 1$. En resumen, existe un número real c, en este caso $c = t_0$, tal que cumple con $P[cX_1 = X_2] = 1$.

De manera inversa, supongamos que existe un número real c, tal que cumple con $P[cX_1 = X_2] = 1$, entonces, $P[cX_1 - X_2 = 0] = 1$, lo que significa que $cX_1 - X_2$ es una variable aleatoria degenerada en 0, por lo cual se cumple que $E[cX_1 - X_2] = 0$ y $V[cX_1 - X_2] = 0$, lo anterior implica $E[(cX_1 - X_2)^2] = 0$. Esto es, existe c tal que $h(c) = E[(cX_1 - X_2)^2] = 0$, entonces, $\Delta(h(t)) = 0$, lo que significa $4E^2[X_1X - 2] - 4E[X_1^2]E[X_2^2] = 0$, esto es, $E^2[X_1X_2] = E[X_1^2]E[X_2^2]$. □

Corolario 5.2. *Sean X_1 y X_2 variables aleatorias, entonces,*

$$|\rho_{X_1,X_2}| \leq 1,$$

cumpliéndose la igualdad $|\rho_{X_1,X_2}| = 1$ si y solo si X_1 y X_2 son linealmente dependientes.

Demostración. Sean $U_1 = X_1 - \mu_1$ y $U_2 = X_2 - \mu_2$, donde $\mu_1 = E[X_1]$ y $\mu_2 = E[X_2]$.

Aplicando la desigualdad de Cauchy-Schwarz,

$$E^2[U_1U_2] \leq E[U_1^2]E[U_2^2].$$

Esto es,

$$E^2[(X_1-\mu_1)(X_2-\mu_2)] \leq E[(X_1-\mu_1)^2]E[(X_2-\mu_2)^2].$$

Dicho de otra manera,

$$\frac{E^2[(X_1-\mu_1)(X_2-\mu_2)]}{V[X_1]V[X_2]} \leq 1.$$

Por lo tanto,

$$(\rho_{X_1,X_2})^2 \leq 1.$$

Esto es,

$$|\rho_{X_1,X_2}| \leq 1.$$

Por otro lado, $(\rho_{X_1,X_2})^2 = 1$, si y solo si $\dfrac{E^2[(X_1-\mu_1)(X_2-\mu_2)]}{V[X_1]V[X_2]} = 1$, lo anterior, si y solo si $E^2[(X_1-\mu_1)(X_2-\mu_2)] = E[(X_1-\mu_1)^2]E[(X_2-\mu_2)^2]$. Esta última igualdad se cumple si y solo si existe un número c, tal que $P[c(X_1-\mu_1) = X_2-\mu_2] = 1$ (véase demostración del teorema anterior), esto es, si $P[X_2 = (\mu_2 - c\mu_1) + cX_1] = 1$. Dicho de otra manera, $|\rho_{X_1,X_2}| = 1$ si y solo si $P[X_2 = a + bX_1] = 1$, donde a y b son números reales. □

5.11. Distribución trinomial

La generalización de la distribución binomial es la distribución trinomial. En esta sección se va a presentar primero lo que es un experimento trinomial, posteriormente se va a definir la distribución trinomial, después, se presentarán las propiedades más importantes de esta distribución y algunos ejemplos que ilustrarán cómo puede aplicarse este modelo en problemas reales.

Definición 5.14. Un *experimento trinomial* consiste en realizar n pruebas idénticas e independientes. Cada prueba tiene la posibilidad de ocurrir en una de tres categorías posibles: categoría 1, categoría 2 o categoría 3. En cada prueba, la probabilidad de que ocurra la categoría 1 es p_1, la probabilidad de que ocurra la categoría 2 es p_2 y de que ocurra la categoría 3 es $p_3 = 1-p_1-p_2$. El objetivo del experimento es observar el número de pruebas que resultan en cada una de las categorías. El vector aleatorio $\underline{X} = (X_1, X_2)$ se define como el número de pruebas que resultan en la categoría 1 y el número de pruebas que resultan en la categoría 2.

Definición 5.15. Un vector aleatorio $\underline{X} = (X_1, X_2)$ tiene *distribución trinomial* con parámetros n, p_1 y p_2, si su función de densidad está dada por

$$f_{\underline{X}}(x_1, x_2) =$$

$$\frac{n!}{x_1!x_2!(n-x_1-x_2)!}p_1^{x_1}p_2^{x_2}(1-p_1-p_2)^{n-x_1-x_2}I_{\{0,1,\ldots,n-x_2\}}(x_1)I_{\{0,1,\ldots,n\}}(x_2),$$

donde n es un entero positivo, $0 < p_i < 1, i = 1, 2$ y además $0 < p_1 + p_2 < 1$.

Para un experimento que cumple con las características de un experimento trinomial, el vector aleatorio correspondiente podrá ser modelado por medio de la función de densidad trinomial.

Es importante mencionar que el experimento y la distribución multinomial es una generalización del experimento y distribución trinomial, respectivamente, donde cada ensayo tendrá la posibilidad de resultar en una de k categorías diferentes. El propósito en esta sección es solo considerar el modelo trinomial.

Las distribuciones trinomial y multinomial tiene varias aplicaciones en estadística, en particular, es un modelo que se utiliza para verificar la independencia de varias variables categóricas.

Notación. $\underline{X} \sim Trinomial(n, p_1, p_2)$ significa que el vector aleatorio $\underline{X}$ tiene distribución trinomial con parámetros n, p_1 y p_2.

A continuación, se presentará gráficamente el recorrido de un vector aleatorio con distribución trinomial. En esta gráfica $n = 10$.

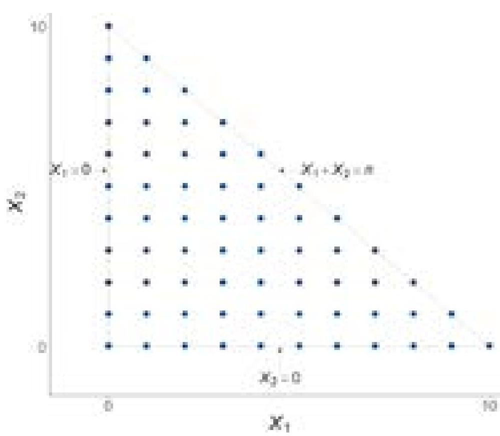


Figura 5.23. *Recorrido un vector aleatorio con distribución trinomial*

Obsérvese que la función de densidad trinomial también puede expresarse de la siguiente manera:

$$f_{\underline{X}}(x_1, x_2) =$$

$$\frac{n!}{x_1!x_2!(n-x_1-x_2)!}p_1^{x_1}p_2^{x_2}(1-p_1-p_2)^{n-x_1-x_2}I_{\{0,1,\ldots,n-x_1\}}(x_2)I_{\{0,1,\ldots,n\}}(x_1).$$

Se queda como ejercicio para el estudiante, demostrar que la función $f_{\underline{X}}(x_1, x_2)$ es una función de densidad.

Teorema 5.13. *Sea $\underline{X}$ un vector aleatorio con distribución trinomial con parámetros n, p_1 y p_2, entonces,*

$$E[X_1] = np_1; \quad E[X_2] = np_2,$$

$$V[X_1] = np_1(1-p_1); \quad V[X_2] = np_2(1-p_2),$$

$$\rho_{X_1,X_2} = -\sqrt{\frac{p_1p_2}{(1-p_1)(1-p_2)}},$$

$$m_{\underline{X}}(t_1, t_2) = (1 - p_1 - p_2 + p_1e^{t_1} + p_2e^{t_2})^n.$$

Demostración. Obsérvese

$$\begin{aligned}
&m_{\underline{X}}(t_1, t_2)\\
&= E[e^{X_1t_1+X_2t_2}]\\
&= \sum_{x_1=0}^{n}\sum_{x_2=0}^{n-x_1} e^{x_1t_1+x_2t_2}\frac{n!}{x_1!x_2!(n-x_1-x_2)!}p_1^{x_1}p_2^{x_2}(1-p_1-p_2)^{n-x_1-x_2}\\
&= \sum_{x_1=0}^{n}\frac{n!}{x_1!}e^{x_1t_1}p_1^{x_1}\sum_{x_2=0}^{n-x_1} e^{x_2t_2}\frac{1}{x_2!(n-x_1-x_2)!}p_2^{x_2}(1-p_1-p_2)^{n-x_1-x_2}\\
&= \sum_{x_1=0}^{n}\frac{n!}{x_1!(n-x_1)!}(e^{t_1}p_1)^{x_1}\\
&\qquad\times\sum_{x_2=0}^{n-x_1}\frac{(n-x_1)!}{x_2!((n-x_1)-x_2)!}(e^{t_2}p_2)^{x_2}(1-p_1-p_2)^{(n-x_1)-x_2}\\
&= \sum_{x_1=0}^{n}\binom{n}{x_1}(e^{t_1}p_1)^{x_1}\sum_{x_2=0}^{n-x_1}\binom{n-x_1}{x_2}(p_2e^{t_2})^{x_2}(1-p_1-p_2)^{(n-x_1)-x_2}.
\end{aligned}$$

Se aplica el teorema del binomio a la segunda suma (véase apéndice C), esto es,

$$\begin{aligned}
m_{\underline{X}}(t_1, t_2) &= \sum_{x_1=0}^{n}\binom{n}{x_1})(e^{t_1}p_1)^{x_1}(1-p_1-p_2+p_2e^{t_2})^{n-x_1}\\
&= (1-p_1-p_2+p_2e^{t_2}+p_1e^{t_1})^n.
\end{aligned}$$

Obsérvese que también a la última suma se le aplicó el teorema del binomio.

Se queda como ejercicios para el estudiante, calcular las medias, varianzas y el coeficiente de correlación. □

Teorema 5.14. *Sea $\underline{X}$ un vector aleatorio el cual se distribuye trinomial con parámetros n, p_1 y p_2, entonces,* $X_1 \sim B(n, p_1)$*;* $X_2 \sim B(n, p_2)$*;* $X_1|x_2 \sim B\left(n - x_2, \dfrac{p_1}{1-p_2}\right)$ *y* $X_2|x_1 \sim B\left(n - x_1, \dfrac{p_2}{1-p_1}\right)$.

Demostración. Obsérvese

$$
\begin{aligned}
f_{X_1}(x_1) &= \sum_{x_2=0}^{n-x_1} \frac{n!}{x_1!x_2!(n-x_1-x_2)!} p_1^{x_1} p_2^{x_2} (1-p_1-p_2)^{n-x_1-x_2} \\
&= \frac{n!}{(n-x_1)!x_1!} p_1^{x_1} \sum_{x_2=0}^{n-x_1} \frac{(n-x_1)!}{x_2!((n-x_1)-x_2)!} p_2^{x_2} (1-p_1-p_2)^{(n-x_1)-x_2} \\
&= \binom{n}{x_1} p_1^{x_1} (1-p_1)^{n-x_1}, \text{ donde } x_1 = 0, 1, \ldots, n.
\end{aligned}
$$

Por lo tanto, $X_1 \sim B(n, p_1)$.
Obsérvese que una vez más se aplicó el teorema del binomio.
Análogamente, $X_2 \sim B(n, p_2)$.
Ahora,

$$
f_{X_1|X_2}(x_1|x_2) = \frac{f_{\underline{X}}(x_1, x_2)}{f_{X_2}(x_2)}
$$

$$
= \frac{\dfrac{n!}{x_1!x_2!(n-x_1-x_2)!} p_1^{x_1} p_2^{x_2} (1-p_1-p_2)^{n-x_1-x_2} I_{\{0,\ldots,n-x_2\}}(x_1) I_{\{0,\ldots,n\}}(x_2)}{\dfrac{n!}{x_2!(n-x_2)!} p_2^{x_2} (1-p_2)^{n-x_2} I_{\{0,1,\ldots,n\}}(x_2)}
$$

$$
= \frac{(n-x_2)!}{x_1!((n-x_2)-x_1)!} \left[\frac{p_1}{1-p_2}\right]^{x_1} \left[\frac{1-p_1-p_2}{1-p_2}\right]^{n-x_2-x_1} I_{\{0,\ldots,n-x_2\}}(x_1).
$$

Se concluye que $X_1|x_2 \sim B\left(n - x_2, \dfrac{p_1}{1-p_2}\right)$.
Análogamente, $X_2|x_1 \sim B\left(n - x_1, \dfrac{p_2}{1-p_1}\right)$. □

Ejemplo 5.31. Para un artículo que se fabrica, el peso en gramos es una variable de calidad. El peso en gramos de este artículo tiene una distribución normal con media de 50 gramos y desviación estándar de 3 gramos. Las especificaciones del peso del producto deben encontrarse entre 45 y 53 gramos. Esto es, si un artículo pesa menos de 45 gramos, se le considera como producto con defecto tipo 1, si el producto pesa más de 53 gramos, este se le considera como artículo con defecto tipo 2, si el peso del artículo resulta dentro de las especificaciones, se considera un artículo de buena calidad. Se seleccionan 5 artículos al azar y se define el siguiente vector aleatorio

$\underline{X} = (X_1, X_2)$, donde X_1 = número de artículos en la muestra con defecto del tipo 1 y X_2 = número de artículos en la muestra con defectos del tipo 2. Encontrar

a) La función de densidad de $\underline{X}$.
b) La probabilidad de encontrar en la muestra más de 2 artículos con defecto del tipo 1 y a lo más dos artículos con defecto del tipo 2.
c) La probabilidad de no tener artículos con defecto del tipo 1 en la muestra, dado que se encontraron al menos 3 artículos con defecto del tipo 2.
d) La probabilidad de tener al menos dos artículos con defecto del tipo 2 en la muestra.

Para contestar cada uno de los incisos, el siguiente planteamiento y cálculos son importantes:

Y = Peso en gramos de un artículo.

$Y \sim N(50, 9)$.

Se calcularán las probabilidades de que un artículo resulte con defecto tipo 1 y con defecto tipo 2, respectivamente

$$\begin{aligned} P[Y < 45] &= P\left[Z < \frac{45-50}{3}\right] \\ &= P[Z < -1.67] = 0.0475. \end{aligned}$$

$$\begin{aligned} P[Y > 53] &= P\left[Z > \frac{53-50}{3}\right] \\ &= P[Z > 1] = 0.1587. \end{aligned}$$

a) Obsérvese

$$\underline{X} = (X_1, X_2) \sim Trinomial(n = 5, p_1 = 0.0475, p_2 = 0.1587).$$

Por lo tanto, la función de densidad de $\underline{X}$ es

$$\begin{aligned} f_{\underline{X}}(X_1, X_2) = &\frac{5!}{x_1!x_2!(5-x_1-x_2)!}(0.0475)^{x_1}(0.1587)^{x_2}(0.7938)^{5-x_1-x_2} \\ &\times I_{\{0,\ldots,5-x_1\}}(x_2) I_{\{0,\ldots,5\}}(x_1). \end{aligned}$$

b) La primera probabilidad es

$$\begin{aligned} P[X_1 > 2, X_2 \leq 2] = &f_{\underline{X}}(3,0) + f_{\underline{X}}(3,1) + f_{\underline{X}}(3,2) + f_{\underline{X}}(4,0) \\ &+ f_{\underline{X}}(4,1) + f_{\underline{X}}(5,0) = 0.000997. \end{aligned}$$

c) Obsérvese

$$P[X_1 = 0 | X_2 \geq 3] = \frac{P[(X_1 = 0) \cap (X_2 \geq 3)]}{P[X_2 \geq 3]},$$

$$P[(X_1 = 0) \cap (X_2 \geq 3)] = f_{\underline{X}}(0,3) + f_{\underline{X}}(0,4) + f_{\underline{X}}(0,5) = 0.027804,$$

$P[X_2 \geq 3] = f_{\underline{X}}(0,3) + f_{\underline{X}}(0,4) + f_{\underline{X}}(0,5) + f_{\underline{X}}(1,3) + f_{\underline{X}}(1,4) + f_{\underline{X}}(2,3) = 0.031059.$

Por lo tanto,

$$P[X_1 = 0 | X_2 \geq 3] = \frac{0.027804}{0.031059} = 0.8952.$$

d) Se sabe $X_2 \sim B(5, 0.1587)$, ya que, $X_2 =$ número de artículos con defecto del tipo 2 observados en la muestra, de este modo, la función de densidad de X_2 es

$$f_{X_2}(x_2) = \binom{5}{x_2}(0.1587)^{x_2}(0.8413)^{5-x_2} I_{\{0,1,2,3,4,5\}}(x_2).$$

Por lo tanto,

$$P[X_2 \geq 2] = 1 - [f_{X_2}(0) + f_{X_2}(1)]] = 0.18103.$$

Esta última probabilidad se puede calcular por medio de R. La instrucción es $1 - pbinom(1, 5, 0.1587)$, siendo 0.1810296 el resultado.

5.12. Distribución normal bivariada

A continuación, se va a definir la distribución normal bivariada, la cual tiene muchas aplicaciones en estadística, como análisis de regresión múltiple y en temas de estadística multivariada. Sus aplicaciones también se encuentran en otras áreas, como las finanzas. En esta sección, se presentará la definición de la distribución normal bivariada y varias de sus propiedades.

Definición 5.16. Un vector aleatorio $\underline{X} = (X_1, X_2)$ tiene una *distribución normal bivariada* con parámetros μ_1, μ_2, σ_1^2, σ_2^2 y ρ, si su función de densidad está dada por

$$f_{\underline{X}}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-c}, \quad \text{donde}$$

$$c = \frac{\left(\frac{x_1 - \mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1 - \mu_1}{\sigma_1}\right)\left(\frac{x_2 - \mu_2}{\sigma_2}\right) + \left(\frac{x_2 - \mu_2}{\sigma_2}\right)^2}{2(1-\rho^2)},$$

$\sigma_1 > 0, \quad \sigma_2 > 0$ y $-1 < \rho < 1$.

Un ejemplo de la gráfica de la función de densidad normal bivariada es

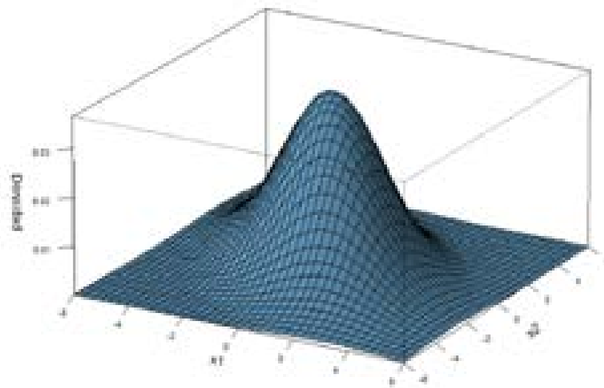


Figura 5.24. *Función de densidad, normal bivariada,* $\mu_1 = \mu_2 = 0, \sigma_1 = \sigma_2 = 5, \rho = 0.5$

Notación. $\underline{X} \sim NB(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, significa que el vector aleatorio $\underline{X}$ tiene distribución normal bivariada con parámetros μ_1, μ_2, σ_1^2, σ_2^2 y ρ.

El exponente c, considerado en la definición anterior, se va a expresar de otra manera. Esta nueva expresión para c, ayudará a encontrar varios resultados importantes de la distribución normal bivariada.

$$
\begin{aligned}
2(1-\rho^2)c &= \left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2 \\
&= \rho^2\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right) \\
&+ \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2 + \left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - \rho^2\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 \\
&= \left(\left(\frac{x_2-\mu_2}{\sigma_2}\right) - \rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\right)^2 + (1-\rho^2)\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 \\
&= \left(\frac{x_2 - \left(\mu_2 + \rho\dfrac{\sigma_2}{\sigma_1}(x_1-\mu_1)\right)}{\sigma_2}\right)^2 + (1-\rho^2)\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2.
\end{aligned}
$$

De esta manera,

$$
c = \left(\frac{x_2 - \left(\mu_2 + \rho\dfrac{\sigma_2}{\sigma_1}(x_1-\mu_1)\right)}{\sqrt{2(1-\rho^2)}\sigma_2}\right)^2 + \left(\frac{x_1-\mu_1}{\sqrt{2}\sigma_1}\right)^2.
$$

Sea $\mu_c = \mu_2 + \rho\dfrac{\sigma_2}{\sigma_1}(x_1-\mu_1)$.

Entonces,

$$c = \left(\frac{x_2 - \mu_c}{\sqrt{2(1-\rho^2)}\sigma_2} \right)^2 + \left(\frac{x_1 - \mu_1}{\sqrt{2}\sigma_1} \right)^2 .$$

Y de esta manera, la función de densidad conjunta de X_1 y X_2 de la normal bivariada se puede expresar de la siguiente manera:

$$\begin{aligned} f_{\underline{X}}(x_1, x_2) &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-c} \\ &= \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{1}{2}\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2} \frac{1}{\sqrt{2\pi(1-\rho^2)}\sigma_2} e^{-\frac{1}{2}\left(\frac{x_2-\mu_c}{\sqrt{1-\rho^2}\sigma_2}\right)^2}, \end{aligned}$$

donde $\mu_c = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x_1 - \mu_1)$.

Análogamente,

$$\begin{aligned} f_{\underline{X}}(x_1, x_2) &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-c} \\ &= \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{1}{2}\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2} \frac{1}{\sqrt{2\pi(1-\rho^2)}\sigma_1} e^{-\frac{1}{2}\left(\frac{x_1-\mu_d}{\sqrt{1-\rho^2}\sigma_1}\right)^2}, \end{aligned}$$

donde $\mu_d = \mu_1 + \rho\frac{\sigma_1}{\sigma_2}(x_2 - \mu_2)$.

Las anteriores factorizaciones de la función $f_{\underline{X}}(x_1, x_2)$ son útiles para demostrar varias propiedades de la distribución normal bivarida. Primero se demostrará que la función $f_{\underline{X}}(x_1, x_2)$ es de densidad.

Se puede ver

$$\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} > 0 \text{ y } e^{-c} > 0.$$

Por lo tanto,

$$f_{\underline{X}}(x_1, x_2) > 0.$$

Por otra parte,

$$\begin{aligned} &\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{\underline{X}}(x_1, x_2) dx_2 dx_1 \\ &\quad = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-c} dx_2 dx_1 \\ &\quad = \left[\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{1}{2}\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2} dx_1 \right] \\ &\quad \times \left[\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi(1-\rho^2)}\sigma_2} e^{-\frac{1}{2}\left(\frac{x_2-\mu_c}{\sqrt{1-\rho^2}\sigma_2}\right)^2} dx_2 \right] = 1 \times 1 = 1. \end{aligned}$$

Por lo tanto, $f_{\underline{X}}(x_1, x_2)$ es función de densidad.

Las funciones en las dos últimas integrales son funciones de densidad de distribuciones normales, la primera de una distribución $N(\mu_1, \sigma_1^2)$, y la segunda de una distribución $N(\mu_c, (1-\rho^2)\sigma_2^2)$. La media μ_c depende de x_1, pero no de x_2, entonces, μ_c es una constante para la segunda integral.

El siguiente resultado trata sobre las medias y las varianzas de X_1 y X_2, también sobre el coeficiente de correlación y la función generadora de momentos conjunta de X_1 y X_2.

Teorema 5.15. *Sea $\underline{X} = (X_1, X_2)$ un vector aleatorio con distribución normal bivariada con parámetros μ_1, μ_2, σ_1^2, σ_2^2 y ρ, entonces,*

$$\begin{gathered} E[X_1] = \mu_1; \quad E[X_2] = \mu_2, \\ V[X_1] = \sigma_1^2; \quad V[X_2] = \sigma_2^2, \\ \rho_{X_1, X_2} = \rho, \\ m_{\underline{X}}(t_1, t_2) = e^{\mu_1 t_1 + \mu_2 t_2 + \frac{1}{2}(\sigma_1^2 t_1^2 + 2\rho\sigma_1\sigma_2 t_1 t_2 + \sigma_2^2 t_2^2)}. \end{gathered}$$

Demostración.

$$\begin{aligned} m_{\underline{X}}(t_1, t_2) &= E[e^{X_1 t_1 + X_2 t_2}] \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{x_1 t_1 + x_2 t_2} \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-c} dx_2 dx_1 \\ &= \left[\int_{-\infty}^{\infty} e^{x_1 t_1} \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{1}{2}\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2}\right] \\ &\times \left[\int_{-\infty}^{\infty} e^{x_2 t_2} \frac{1}{\sqrt{2\pi(1-\rho^2)}\sigma_2} e^{-\frac{1}{2}\left(\frac{x_2-\mu_c}{\sqrt{1-\rho^2}\sigma_2}\right)^2} dx_2\right] dx_1. \end{aligned}$$

La segunda integral es la función generadora de momentos de la distribución normal univariada $N(\mu_c, (1-\rho^2)\sigma_2^2)$.

De esta manera,

$$\begin{aligned} m_{\underline{X}}(t_1, t_2) &= \int_{-\infty}^{\infty} e^{x_1 t_1} \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{1}{2}\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2} e^{\mu_c t_2 + \frac{(1-\rho^2)\sigma_1^2 t_2^2}{2}} dx_1 \\ &= \int_{-\infty}^{\infty} e^{x_1 t_1 + \mu_c t_2 + \frac{(1-\rho^2)\sigma_1^2 t_2^2}{2}} \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{1}{2}\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2} dx_1 \\ &= \left[e^{\mu_2 t_2 - \rho\frac{\sigma_2}{\sigma_1}\mu_1 t_2 + \frac{(1-\rho^2)\sigma_2^2 t_2^2}{2}}\right] \\ &\times \int_{-\infty}^{\infty} e^{x_1(\rho\frac{\sigma_2}{\sigma_1} t_2 + t_1)} \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{1}{2}\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2} dx_1. \end{aligned}$$

La última integral representa la función generadora de momentos de una distribución $N(\mu_1, \sigma_1^2)$ evaluada en $\rho\dfrac{\sigma_2}{\sigma_1}t_2 + t_1$, por lo tanto,

$$\begin{aligned} m_{\underline{X}}(t_1,t_2) &= \left[e^{\mu_2 t_2 - \rho\frac{\sigma_2}{\sigma_1}\mu_1 t_2 + \frac{(1-\rho^2)\sigma_2^2 t_2^2}{2}} \right] \\ &\quad \times \left[e^{\mu_1\left(\rho\frac{\sigma_2}{\sigma_1}t_2 + t_1\right) + \frac{\sigma_1^2}{2}\left(\rho^2\frac{\sigma_2^2}{\sigma_1^2}t_2^2 + t_1^2 + 2\rho\frac{\sigma_2}{\sigma_1}t_1 t_2\right)} \right] \\ &= e^{\mu_1 t_1 + \mu_2 t_2 + \frac{1}{2}(\sigma_1^2 t_1^2 + 2\rho\sigma_1\sigma_2 t_1 t_2 + \sigma_2^2 t_2^2)}. \end{aligned}$$ □

Se queda como ejercicio para el estudiante, calcular las medias, las varianzas de X_1 y X_2 y el coeficiente de correlación de X_1 y X_2.

El siguiente teorema trata de las distribuciones marginales y condicionales de una distribución normal bivariada.

Teorema 5.16. *Sea $\underline{X} = (X_1, X_2)$ un vector aleatorio con distribución normal bivariada con parámetros μ_1, μ_2, σ_1^2, σ_2^2 y ρ, entonces,*

$$X_1 \sim N(\mu_1, \sigma_1^2); \quad X_2 \sim N(\mu_2, \sigma_2^2),$$

$$X_1|x_2 \sim N(\mu_d, (1-\rho^2)\sigma_1^2), \text{ donde } \mu_d = \mu_1 + \rho\frac{\sigma_1}{\sigma_2}(x_2 - \mu_2),$$

$$X_2|x_1 \sim N(\mu_c, (1-\rho^2)\sigma_2^2), \text{ donde } \mu_c = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x_1 - \mu_1).$$

Demostración. Obsérvese

$$\begin{aligned} f_{X_1}(x_1) &= \int_{-\infty}^{\infty} f_{\underline{X}}(x_1,x_2)dx_2 \\ &= \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-c} dx_2 \\ &= \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{1}{2}\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2} \\ &\quad \times \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi(1-\rho^2)}\sigma_2} e^{-\frac{1}{2}\left(\frac{x_2-\mu_c}{\sqrt{1-\rho^2}\sigma_2}\right)^2} dx_2 \\ &= \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{1}{2}\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2}. \end{aligned}$$

La última integral vale 1, ya que es la integral de una función de densidad de una distribución normal univariada con media μ_c y varianza $(1-\rho^2)\sigma_2^2$.

Por lo tanto, $X_1 \sim N(\mu_1, \sigma_1^2)$.

Análogamente, se puede demostrar que $X_2 \sim N(\mu_2, \sigma_2^2)$.

Por otro lado,

$$
\begin{aligned}
f_{X_1|X_2}(x_1|x_2) &= \frac{f_{\underline{X}}(x_1,x_2)}{f_{X_2}(x_2)} \\
&= \frac{\frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{1}{2}\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2} \frac{1}{\sqrt{2\pi(1-\rho^2)}\sigma_1} e^{-\frac{1}{2}\left(\frac{x_1-\mu_d}{\sqrt{1-\rho^2}\sigma_1}\right)^2}}{\frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{1}{2}\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2}} \\
&= \frac{1}{\sqrt{2\pi(1-\rho^2)}\sigma_1} e^{-\frac{1}{2}\left(\frac{x_1-\mu_d}{\sqrt{1-\rho^2}\sigma_1}\right)^2}.
\end{aligned}
$$

Esta última expresión es la función de densidad de una normal univariada con media $\mu_d = \mu_1 + \rho\dfrac{\sigma_1}{\sigma_2}(x_2 - \mu_2)$ y varianza $(1-\rho^2)\sigma_1^2$.

Esto es,

$$
X_1|x_2 \sim N(\mu_d, (1-\rho^2)\sigma_1^2), \text{ donde } \mu_d = \mu_1 + \rho\frac{\sigma_1}{\sigma_2}(x_2 - \mu_2).
$$

Análogamente,

$$
X_2|x_1 \sim N(\mu_c, (1-\rho^2)\sigma_2^2), \text{ donde } \mu_c = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x_1 - \mu_1). \qquad \square
$$

En la normal bivariada, las variables aleatorias X_1 y X_2 son independientes si y solo si las variables están incorrelacionadas. El siguiente teorema trata con este resultado.

Teorema 5.17. *Sea $\underline{X} = (X_1, X_2)$ un vector aleatorio con distribución normal bivariada con parámetros μ_1, μ_2, σ_1^2, σ_2^2 y ρ, entonces, $\rho = 0$ si y solo si X_1 y X_2 son variables aleatorias independientes.*

Demostración. Si $\rho = 0$, entonces,

$$
\begin{aligned}
m_{\underline{X}}(t_1,t_2) &= e^{\mu_1 t_1 + \mu_2 t_2 + \frac{1}{2}(\sigma_1^2 t_1^2 + \sigma_2^2 t_2^2)} \\
&= e^{\mu_1 t_1 + \frac{1}{2}(\sigma_1^2 t_1^2)} e^{\mu_2 t_2 + \frac{1}{2}(\sigma_2^2 t_2^2)} \\
&= m_{X_1}(t_1) m_{X_2}(t_2).
\end{aligned}
$$

Por lo tanto, X_1 y X_2 son variables aleatorias independientes.

Si X_1 y X_2 son variables aleatorias independientes, entonces, X_1 y X_2 están incorrelacionadas (corolario 5.1). $\square$

En el siguiente resultado, se puede ver cómo a partir de un vector aleatorio con distribución normal bivariada, es posible construir otro vector aleatorio con distribución normal bivariada donde sus componentes son variables aleatorias independientes con distribución común, normal estándar.

Teorema 5.18. *Si* $\underline{X} = (X_1, X_2)$ *es un vector aleatorio con distribución normal bivariada con parámetros* μ_1, μ_2, σ_1^2, σ_2^2 *y* ρ, *entonces, las variables* $Z_1 = \dfrac{X_1 - \mu_1}{\sigma_1}$ *y* $Z_2 = \dfrac{X_2 - \mu_2}{\sqrt{1-\rho^2}\sigma_2} - \rho\left(\dfrac{X_1 - \mu_1}{\sqrt{1-\rho^2}\sigma_1}\right)$ *son variables aleatorias independientes, donde cada una tiene distribución normal estándar.*

La demostración es parte del siguiente tema, distribuciones de funciones de variables aleatorias. Se sugiere usar el método de función generadora de momentos, el cual será tratado en el siguiente capítulo.

Observación. De forma análoga, las variables aleatorias $Z_1 = \dfrac{X_2 - \mu_2}{\sigma_2}$ y $Z_2 = \dfrac{X_1 - \mu_1}{\sqrt{1-\rho^2}\sigma_1} - \rho\left(\dfrac{X_2 - \mu_2}{\sqrt{1-\rho^2}\sigma_2}\right)$ son también variables aleatorias independientes con distribución normal estándar.

El siguiente teorema indica, cómo a partir de dos variables aleatorias independientes con distribución $N(0,1)$, se puede construir un vector (X_1, X_2) con distribución normal bivariada con parámetros μ_1, μ_2, σ_1^2, σ_2^2 y ρ.

Teorema 5.19. *Si* Z_1 *y* Z_2 *son dos variables aleatorias independientes e idénticamente distribuidas con distribución normal estándar, entonces, el vector* $\underline{X} = (X_1, X_2)$, *donde* $X_1 = \mu_1 + \sigma_1 Z_1$ *y* $X_2 = \mu_2 + \rho\sigma_2 Z_1 + \sqrt{1-\rho^2}\sigma_2 Z_2$, *tiene una distribución normal bivariada con parámetros* μ_1, μ_2, σ_1^2, σ_2^2 *y* ρ.

La demostración se deja para el siguiente tema, distribuciones de funciones de variables aleatorias.

Una aplicación de este resultado es para simular valores de un vector aleatorio con distribución $NB(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, a partir de valores de dos variables independientes con distribución normal estándar. Los valores de una variable aleatoria con distribución $N(0,1)$, se pueden simular por el *método de Box-Müller*, véase en el siguiente capítulo el ejemplo 6.26.

Ejemplo 5.32. La longitud en centímetros y el peso en gramos de cierta especie tiene una distribución conjunta normal bivariada con parámetros $\mu_1 = 83$, $\mu_2 = 54$, $\sigma_1^2 = 9$, $\sigma_2^2 = 16$ y $\rho = 0.6$. Se selecciona al azar una especie, resultando un peso de 56 gramos. Calcular la probabilidad de que resulte una longitud por encima de 86 centímetros.

Obsérvese $X_1|x_2 \sim N(\mu_1 + \rho\dfrac{\sigma_1}{\sigma_2}(x_2 - \mu_2), (1-\rho^2)\sigma_1^2) = N(83.9, 5.76)$.

Por lo tanto,

$$\begin{aligned} P[X_1 > 86 | X_2 = 56] &= P\left[\frac{X_1 - 83.9}{2.4} > \frac{86 - 83.9}{2.4}\right] \\ &= P[Z > 0.875] = 0.19079. \end{aligned}$$

Por medio de R, la instrucción es $1 - pnorm(86, 83.9, 2.4)$, donde la respuesta es 0.190787.

Ejercicios del capítulo 5

1.- Terminar de calcular la función de densidad del ejemplo 5.2.

2.- En un grupo de diez profesionistas, cinco son actuarios, tres son matemáticos y dos son economistas. De este grupo, se seleccionan en forma aleatoria cuatro profesionistas. Sean X_1 y X_2 el número de actuarios y el número de economistas en la muestra, respectivamente. Obtener la función de densidad conjunta de X_1 y X_2.

3.- Supóngase que X_1 y X_2 son dos variables aleatorias con función de densidad conjunta dada por

$$f_{\underline{X}}(x_1, x_2) = \begin{cases} k, & 0 \leq x_1 \leq 2,\ 0 \leq x_2 \leq 2 \text{ y } 0 \leq x_1 + x_2 \leq 2, \\ \\ 0, & \text{en otro caso.} \end{cases}$$

Determinar

a) La función de densidad con funciones indicadoras.
b) El valor de k.
c) $P[X_1 \leq 3/2, X_2 \leq 3/2]$.
d) $P[X_1 \leq 1, X_2 \leq 1]$.
e) $P[X_1 \leq 1 | X_2 \leq 1\ ]$.

4.- La función de densidad conjunta de las variables X_1 y X_2 está dada por

$$f_{\underline{X}}(x_1, x_2) = \begin{cases} ke^{-x_1}, & 0 \leq x_2 \leq x_1 < \infty, \\ \\ 0, & \text{en otro caso.} \end{cases}$$

Determinar

a) La función de densidad con funciones indicadoras.
b) El valor de k.
c) $P[X_1 < 3, X_2 < 1]$.
d) $P[X_1 \leq X_2 + 2]$.
e) $P[X_1 + X_2 \geq 1]$.

5.- Sea (X_1, X_2) un punto seleccionado al azar de un círculo con radio r y centro en el origen. Obtener

a) La función de densidad conjunta de X_1 y X_2, expresarla con funciones indicadoras.
b) $P[X_1 \geq 0, X_2 \geq 0]$.
c) $P[X_1 \leq 2X_2]$.
d) $P[X_2 \geq 0 | X_1 \leq X_2]$.

6.- Sean X_1 y X_2 dos variables aleatorias con función de densidad conjunta dada por

$$f_{\underline{X}}(x_1, x_2) = \begin{cases} kx_1x_2e^{-\dfrac{x_1+x_2}{2}}, & x_1 > 0,\, x_2 > 0, \\ \\ 0, & \text{en otro caso.} \end{cases}$$

Determinar

a) La función de densidad con funciones indicadoras.
b) El valor de k.
c) $P[X_1 > 1, X_2 \leq 1]$.
d) $P[X_2 > 3]$.

7.- Considerando el ejercicio 2, calcular

a) La función densidad marginal de X_1.
b) La función de densidad condicional de X_2 dado que $X_1 = x_1$.
c) $P[X_2 = 1|X_1 = 2]$.
d) $P[X_3 = 1|X_2 = 1]$, donde $X_3 =$ número de profesionistas matemáticos en la muestra.

8.- Consideremos que (X_1, X_2) es un vector aleatorio que tiene una función de densidad dada por $f_{\underline{X}}(x_1, x_2) = kx_1^2x_2^3I_{(0,1)}(x_1)I_{(0,1)}(x_2)$. Obtener

a) El valor de k.
b) Las funciones de densidad marginales de X_1 y X_2.
c) $P[X_1 \leq 0.4|X_2 \geq 0.7]$.
d) Las funciones de densidad condicionales de X_1 dado que $X_2 = x_2$, y de X_2 dado que $X_1 = x_1$.
e) $P[X_1 \leq 0.7|X_2 = 0.4]$.

9.- Considerando el ejercicio 3, calcular

a) $P[X_1 \geq 1|X_2 \leq 0.5]$.
b) $P[X_1 \geq 1|X_2 = 0.5]$.

10.- Considerando el ejercicio 4, calcular

a) Las funciones de densidad marginales.
b) Las funciones de densidad condicionales.
c) $P[X_1 > 2|X_2 = 1]$.
d) $E[X_1]$ y $E[X_2]$.

11.- En cierta población se levanta cada día una encuesta para conocer el número de ciudadanos insatisfechos con el actual gobierno. Para cada día, se seleccionan al azar diez ciudadanos para observar el número de personas inconformes. Se sabe que la proporción p de ciudadanos inconformes varía entre un día a otro y tiene una distribución $U(0,1)$. Para determinado día, calcular

a) La función de densidad del número de ciudadanos inconformes en la muestra.
b) La probabilidad incondicional de que se observen menos de tres personas inconformes en la muestra.
c) El valor esperado del número de ciudadanos inconformes en la muestra, usando su función de densidad.
d) El valor esperado del número de ciudadanos inconformes en la muestra, sin usar su función de densidad.

12.- Cada cliente tiene tres opciones para ser atendido en una tienda: ventanilla 1, ventanilla 2 o ventanilla 3. Las probabilidades de ser atendido en cada una de las ventanillas son de 0.3, 0.5 y 0.2, respectivamente. Se seleccionan al azar 10 clientes y se observan las siguientes variables aleatorias: X_1 el número de clientes atendidos en la ventanilla 1 y X_2 el número de clientes atendidos en la ventanilla 2. Determinar

a) La función de densidad conjunta de X_1 y X_2 expresada con funciones indicadoras.
b) La probabilidad de que se atiendan más de 7 clientes en la ventanilla 1.
c) La probabilidad de que X_1 sea menor a dos y X_2 sea menor a 3.

13.- El tiempo de espera de un cliente para ser atendido tiene una distribución exponencial con media de 10 minutos. Sean X_1 y X_2 los tiempos de espera de dos clientes que se atienden en forma independiente. Encontrar

a) La función de densidad conjunta de X_1 y X_2, expresándola con funciones indicadoras.
b) La probabilidad de que $X_1 + X_2 \leq 10$.

14.- Considerando el ejercicio 8, calcular $COV(X_1, X_2)$.

15.- Las variables aleatorias X_1 y X_2 tienen la siguiente función de densidad conjunta: $f(x_1, x_2) = 1/4$ para $(x_1, x_2) = (-1, 0), (0, -1), (0, 1), (1, 0)$.

a) ¿Son X_1 y X_2 incorrelacionadas?
b) ¿Son X_1 y X_2 independientes?

16.- Se supone que el peso en kilogramos de un cerdo macho adulto, después de recibir una dieta desde su nacimiento, tiene una distribución normal con media de 180 kg y una desviación estándar de 9 kg. Se seleccionan al azar 5 cerdos adultos y se definen las siguientes variables aleatorias: X_1 = número de cerdos con peso menor a 180 kg y X_2 = número de cerdos con peso entre 180 y 190 kg. Determinar

a) La función de densidad conjunta de X_1 y X_2, expresándola con funciones indicadoras.
b) La probabilidad de que exactamente tres cerdos tengan un peso entre 180 y 190 kg y un cerdo tenga un peso superior a los 190 kg.
c) La probabilidad de que un cerdo pese más de 190 kg.
d) La probabilidad de tener a lo más un cerdo con peso menor a 180 kg y más de tres cerdos con peso mayor a 190 kg.

17.- Sean X_1 y X_2 dos variables aleatorias con funciones de distribución $F_{X_1}(x_1)$ y $F_{X_2}(x_2)$ respectivamente y función de distribución conjunta $F_{X_1,X_2}(x_1,x_2)$. Demostrar que se cumple la siguiente desigualdad: $F_{X_1}(x_1)+F_{X_2}(x_2)-1 \leq F_{X_1,X_2}(x_1,x_2) \leq \sqrt{F_{X_1}(x_1)F_{X_2}(x_2)}$.

18.- Sean X_1 y X_2 dos variables aleatorias independientes y con la misma distribución, dada por $P[X_i=1]=2/3$, $P[X_i=2]=1/3, i=1,2$. Sea $X_3=X_1X_2$.

a) Encontrar la función de densidad conjunta de X_1 y X_3.
b) Calcular $Cov(X_1,X_3)$.
c) ¿Son independientes X_1 y X_3?

19.- Se seleccionan al azar dos números sin reemplazo de una urna que contiene cuatro números -2,-1, 1 y 2. Sea X_1, el primer número obtenido y X_2, el número más grande de ambos números seleccionados.

a) Encontrar la función de densidad del vector (X_1,X_2).
b) Calcular $COV(X_1,X_2)$.
c) ¿Son X_1 y X_2 variables aleatorias independientes?

20.- Considerando, el ejemplo 5.13, calcular la función de densidad condicional de X_2 dado que $X_1=x_1$.

21.- Si X_1 es una variable aleatoria con función de densidad dada por $f_{X_1}(x_1)=0.5I_{\{1\}}(x_1)+(1/3)I_{\{2\}}(x_1)+(1/6)I_{\{3\}}(x_1)$ y X_2 otra variable aleatoria tal que $E[X_2|X_1=1]=1$, $E[X_2|X_1=2]=2$ y $E[X_2|X_1=3]=3$. Calcular $E[X_2]$.

22.- Cuatro dados son arrojados, X_1 denota el número de unos en los primeros tres dados y X_2 el número de unos en los últimos tres dados.

a) Encontrar la función de densidad conjunta de X_1 y X_2.
b) Encontrar las funciones de densidad marginales.
c) Encontrar la función de densidad condicional de X_2 dado que $X_1=x_1$.
d) ¿Son X_1 y X_2 independientes?

23.- Para una acción financiera, se sabe que baja para un mes determinado con una probabilidad de x_1 (esta probabilidad varía entre un mes y otro).

Consideremos los siguientes n meses, sea X_2 el número de periodos donde la acción baja. Se sabe que la distribución condicional de X_2 dado un valor específico de x_1 es $B(n, x_1)$ y la distribución de X_1 es $U(0,1)$. Encontrar

a) $E[X_2]$.
b) La distribución de X_2.
c) $E[X_2]$ usando la distribución encontrada en el inciso anterior.

24.- Sea (X_1, X_2) un vector aleatorio con función de densidad dada por $f_{\underline{X}}(x_1, x_2) = [1 - \alpha(1-2x_1)(1-2x_2)]I_{(0,1)}(x_1)I_{(0,1)}(x_2)$, donde el parámetro α satisface $-1 \leq \alpha \leq 1$.

a) Demostrar que $f_{\underline{X}}(x_1, x_2)$ es función de densidad.
b) Calcular las funciones de densidad marginales.
c) Calcular la covarianza de X_1 y X_2.
d) Demostrar que X_1 y X_2 son variables aleatorias independientes si y solo si X_1 y X_2 son variables que están incorrelacionadas.

25.- En una caja hay fichas, dos con el número 1, dos con el número 2 y dos con el número 3. Dos fichas son sacadas al azar sin reemplazo. Sea X_1 el número más pequeño de las fichas sacadas y X_2 el número más grande. Encontrar

a) La función de densidad conjunta de X_1 y X_2.
b) La función de densidad marginal de X_2.
c) La $Cov[X_1, X_2]$.

26.- Sean X_1 y X_2 dos variables aleatorias tal que la función de densidad de X_1 es $f_{X_1}(x_1) = (x_1/6)I_{\{1,2,3\}}(x_1)$ y la distribución condicional de X_2 dado $X_1 = x_1$ es una distribución $B(x_1, 0.1)$. Encontrar

a) $E[X_1]$ y $V[X_1]$.
b) $E[X_2]$.
c) La función de densidad conjunta de X_1 y X_2.

27.- Supongamos que $\underline{X} = (X_1, X_2)$ es un vector aleatorio con distribución trinomial con parámetros n, p_1 y p_2.

a) Demostrar que la función $f_{\underline{X}}(x_1, x_2)$ indicada en la definición 5.15, es de densidad.
b) Calcular $E[X_1]$ y $E[X_2]$.
c) Calcular $V[X_1]$ y $V[X_2]$.
d) Calcular $Cov[X_1, X_2]$.
e) Encontrar la distribución marginal de X_2.
f) Encontrar la distribución condicional de X_2 dado que $X_1 = x_1$.
g) Calcular $E[X_2|x_1]$.

28.- Si la función generadora de momentos de $\underline{X} = (X_1, X_2)$ está dada por $m_{\underline{X}}(t_1, t_2) = e^{{t_1}^2 + {t_2}^2}$.

a) ¿Cómo se distribuye $\underline{X}$?
b) Encontrar la distribución de X_2.
c) ¿Están X_1 y X_2 incorrelacionadas? ¿Por qué?
d) ¿Son independientes X_1 y X_2? ¿Por qué?

29.- Sean X_1 y X_2 dos variables aleatorias, con las siguientes funciones de densidad $f_{X_1|X_2}(x_1|x_2) = c_1 x_1^2/{x_2}^3 I_{(0,x_2)}(x_1)$, donde $0 < x_2 < 1$ y $f_{X_2}(x_2) = c_2 {x_2}^3 I_{(0,1)}(x_2)$. Encontrar

a) Las constantes c_1 y c_2.
b) La función de densidad conjunta de X_1 y X_2.
c) $P[1/4 < X_1 < 1/2 | X_2 = 5/8]$.
d) $P[1/4 < X_1 < 1/2]$.

30.- Sean X_1 y X_2 dos variables aleatorias con densidad conjunta dada por $f_{\underline{X}}(1,1) = 1/21$, $f_{\underline{X}}(1,2) = 1/42$, $f_{\underline{X}}(2,1) = 1/42$, $f_{\underline{X}}(0,0) = 1/21$, $f_{\underline{X}}(0,1) = 2/21$, $f_{\underline{X}}(0,2) = 1/7$, $f_{\underline{X}}(0,3) = 4/21$, $f_{\underline{X}}(1,0) = 2/21$, $f_{\underline{X}}(2,0) = 1/7$, $f_{\underline{X}}(3,0) = 4/21$, y cero en otro caso. Calcular el coeficiente de correlación.

31.- Sean X_1 y X_2 dos variables aleatorias independientes, y A y B eventos definidos de la siguiente manera: $A = \{(x_1, x_2) \in \mathbb{R}^2 | a < X_1 < b\}$ y $B = \{(x_1, x_2) \in \mathbb{R}^2 | c < X_2 < d\}$, tal que $P[A] = \frac{1}{6}$ y $P[B] = \frac{1}{5}$. Encontrar la probabilidad de

a) $A \cap B$.
b) $A - B$.
c) $A \cup B$.

32.- Se sabe que un ajustador de seguros de autos tardará en llegar al lugar del siniestro, entre 0 y 20 minutos, desde el momento en que se reportó el accidente. A partir de que el ajustador llega al lugar del siniestro, se tardará en gestionar el trámite del accidente, entre 20 y 30 minutos. Sea X_1 el tiempo que tarda en llegar el ajustador al lugar del accidente y X_2 el tiempo que tarda en gestionar el trámite. Ambas variables son independientes y se distribuyen uniformemente. Calcular la probabilidad de que en menos de 30 minutos se termine el trámite desde que se reportó el accidente.

33.- Sean X_1 y X_2 dos variables aleatorias que tienen función de densidad conjunta dada por $f_{\underline{X}}(x_1, x_2) = (3x_1 + 2x_2)/51$, donde el recorrido conjunto es $(x_1, x_2) = (1,1), (1,2), (1,3), (2,1), (2,2), (2,3)$. Determinar la media y varianza condicional de X_2, dado $X_1 = x_1$.

34.- Considerando la función de densidad conjunta del ejercicio 30.

a) Encontrar las dos funciones de densidad marginales.
b) Encontrar las dos funciones de densidad condicionales.
c) Calcular las dos medias condicionales.
d) Verificar que se cumple $E[E[X_2|X_1]] = E[X_2]$ sin usar el teorema 5.9.

35.- Sea $\Phi(t_1, t_2) = \ln(m_{\underline{X}}(t_1, t_2))$, donde $m_{\underline{X}}(t_1, t_2)$ es la función generadora de momentos conjunta de X_1 y X_2. Demostrar que la primera derivada de $\Phi(t_1, t_2)$, con respecto a t_i evaluada en $\underline{t} = (0, 0)$, proporciona la media de X_i, para $i = 1, 2$, que la segunda derivada de $\Phi(t_1, t_2)$, con respecto a t_i evaluada en $\underline{t} = (0, 0)$, proporciona la varianza de X_i, para $i = 1, 2$ y que la segunda derivada de $\Phi(t_1, t_2)$ con respecto a t_1 y con respecto a t_2, evaluada en $\underline{t} = (0, 0)$, nos proporciona la covarianza de X_1 y X_2.

36.- Demostrar el teorema 5.3.

37.- Sea $\underline{X} = (X_1, X_2)$ un vector aleatorio con función de densidad dada por

$$f_{\underline{X}}(x_1, x_2) = \binom{x_1}{x_2} (1/2)^{x_1} (x_1/10)\, I_{\{0,1,\ldots,x_1\}}(x_2) I_{\{1,\ldots,4\}}(x_1).$$

a) Demostrar que $f_{\underline{X}}(x_1, x_2)$ es función de densidad.
b) Encontrar la función de densidad marginal de X_2.
c) Calcular $E[X_2]$ usando la función de densidad de X_2.
d) Calcular $E[E[X_2|X_1]]$.

38.- En una compañía de seguros, cada siniestro es evaluado por un perito el cual dictamina, si se paga o no. Sea X_2 el número de siniestros reportados al mes y X_1 el número de siniestros que realmente se pagan en el periodo. La función de densidad conjunta de X_1 y X_2 está dada por

$$f_{\underline{X}}(x_1, x_2) = \frac{e^{-2}}{x_1!(x_2 - x_1)!} I_{\{0,1,\ldots,x_2\}}(x_1) I_{\{0,1,\ldots\}}(x_2).$$

a) Demostrar que $f_{\underline{X}}(x_1, x_2)$ es función de densidad.
b) Encontrar la función generadora de momentos de $\underline{X} = (X_1, X_2)$.
c) Calcular las medias y las varianzas de X_1 y X_2.
d) Calcular el coeficiente de correlación de X_1 y X_2.
e) Calcular $E[X_1|x_2]$.
f) Calcular $E[E[X_1|X_2]]$.

39.- Se sabe que el rendimiento de dos bonos financieros X_1 y X_2 son variables dependientes y que su distribución conjunta es una normal bivariada con parámetros $\mu_1 = 5$, $\mu_2 = 10$, $\sigma = 1$, $\sigma = 5$ y $\rho > 0$. Supongamos que se cumple $P[4 < X_2 < 16|x_1 = 5] = 0.954$, determinar el valor del coeficiente de correlación ρ.

40.- Sea $\underline{X} = (X_1, X_2)$ un vector aleatorio con función de densidad

$$f_{\underline{X}}(x_1, x_2) = \frac{1}{2\pi} e^{-1/2(x_1^2+x_2^2)} \left\{ 1 + x_1 x_2 e^{-1/2(x_1^2+x_2^2-2)} \right\}.$$

Demostrar que las funciones de densidad marginales de esta distribución conjunta son distribuciones normales. Este hecho demuestra que si las marginales son normales, esto no implica que la densidad conjunta sea forzosamente una normal bivariada.

41.- Sean X_1 y X_2 dos variables aleatorias con función de densidad conjunta $f_{\underline{X}}(x_1, x_2) = kI_{(0,x_1)}(x_2)I_{(0,2)}(x_1)$. Calcular

a) k.
b) $P[X_1 \leq 2X_2]$.
c) Las funciones de densidad marginales.
d) Las funciones de densidad condicionales.

42.- Sea (X_1, X_2) un vector aleatorio el cual se distribuye uniforme sobre la región $R = \{(x_1, x_2) | x_1 + x_2 < 3, x_1 > 0, x_2 > 0\}$. Sean $A = \{X_1 < 1\}$ y $B = \{X_2 > 1\}$ dos eventos. Calcular la probabilidad de que

a) Al menos uno de los eventos ocurra.
b) Ocurran los dos eventos.
c) Exactamente uno de los dos eventos ocurra.

43.- Suponga que X_1 y X_2 son dos variables aleatorias independientes, distribuidas de manera idéntica. Sea Z otra variable aleatoria definida como $Z = kX_1 + X_2$. Si el coeficiente de correlación de X_1 y Z es igual a 0.5, encontrar el valor de k.

44.- Sean X_1 y X_2 dos variables aleatorias con función de densidad conjunta $f_{\underline{X}}(x_1, x_2) = kx_1x_2I_{(0,x_2)}(x_1)I_{(0,2)}(x_2)$. Calcular

a) El valor de k.
b) $P[X_1 < 1/2 | X_2 = 1]$.

45.- Sean X_1 y X_2 dos variables aleatorias, tal que X_1 tiene una media igual a 16 y una varianza de 9. Si la media condicional de X_2 dado $X_1 = x_1$, es igual a x_1^2, calcular la media de X_2.

46.- Considerando el ejemplo 5.8, verificar si X_1 y X_2 son variables aleatorias independientes.

47.- Considerando el ejemplo 5.10, verificar si X_1 y X_2 son variables aleatorias independientes.

48.- Sean X_1 y X_2 dos variables aleatorias con función de densidad conjunta dada por $f_{\underline{X}}(x_1, x_2) = kx_1^2x_2^2I_{(0,x_2)}(x_1)I_{(0,1)}(x_2)$. Encontrar

a) El valor de k.
b) $E[X_1^2|X_2 = x_2]$.

49.- Sean X_1 y X_2 dos variables aleatorias independientes con distribución Poisson, tal que se cumple $4V[X_1] = V[X_2] = 8$. Calcular la probabilidad $P[X_1 + X_2 \geq 3]$.

50.- Sean X_1 y X_2 dos variables aleatorias, tal que la función de densidad de X_1 es $f_{X_1}(x_1) = 24x_1^2 I_{(0,1/2)}(x_1)$ y la función de densidad condicional de X_2 dado $X_1 = x_1$ es $f_{X_2|X_1}(x_2|x_1) = \dfrac{x_2}{2x_1^2} I_{(0,2x_1)}(x_2)$. Encontrar la distribución de X_2.

51.- Sean X_1 y X_2 dos variables aleatorias con función de densidad conjunta dada por $f_{\underline{X}}(x_1, x_2) = kI_{(x_2^2,1)}(x_1)I_{(0,1)}(x_2)$. Encontrar

a) El valor de k.
b) La función de densidad marginal de X_1.
c) La densidad condicional de X_2 dado que $X_1 = x_1$.

52.- Sea X_1 el importe de un siniestro reportado por un asegurado y X_2 la cantidad que la compañía de seguros realmente pagará al asegurado después de haber realizado el perito del siniestro. Supongamos que la función de densidad conjunta de X_1 y X_2 está dada por $f_{\underline{X}}(x_1, x_2) = ke^{-(x_1+x_2)} I_{(0,x_2)}(x_1)I_{(0,\infty)}(x_2)$.

a) Encontrar k.
b) ¿Cómo se distribuye X_1?
c) Encontrar la función de densidad condicional de X_2 dado que $X_1 = x_1$.
d) Encontrar el valor esperado condicional de X_2, dado que $X_1 = x_1$.

53.- Un punto (X_1, X_2) es seleccionado aleatoriamente de un disco circular de radio 1 con centro en $(1, 1)$. Determinar

a) La función de densidad para (X_1, X_2) con funciones indicadoras.
b) La función de densidad marginal de X_1.
c) La densidad condicional de X_2 dado que $X_1 = x_1$.
d) El valor esperado condicional de X_2 dado que $X_1 = x_1$.

54.- Sean X_1 y X_2 dos variables aleatorias, tal que la función de densidad condicional de X_1 dado $X_2 = x_2$ es $f_{X_1|X_2}(x_1|x_2) = \dfrac{3x_1^2}{1 - x_2^3} I_{(x_2,1)}(x_1)$ y la densidad marginal de X_2 es $f_{X_2}(x_2) = \dfrac{10}{3} x_2(1-x_2^3)I_{(0,1)}(x_2)$. Calcular $E[X_2|x_1]$.

55.- Sean X_1 y X_2 dos variables aleatorias independientes e idénticamente distribuidas como $U(-2a, 2a)$. Si $V[X_1X_2] = 256/9$, calcular a.

56.- Demostrar el teorema 5.10.

57.- Considerando el ejemplo 5.23 (5.5, 5.9, 5.12 y 5.18), verificar que se cumple $E[E[X_2|X_1]] = E[X_2]$ sin usar el teorema 5.9.

58.- Considerando el ejemplo 5.29 (5.1, 5.2, 5.8, 5.13), verificar que se cumple $E[E[X_2|X_1]] = E[X_2]$ sin usar el teorema 5.9.

59.- Sean X_1 y X_2 dos variables aleatorias, las cuales cumplen las siguientes propiedades: $E[X_2|X_1 = x_1] = 9x_1 - 3$ y $V[X_2|X_1 = x_1] = 81x_1^2$, además se sabe que X_1 tiene distribución uniforme en el intervalo $(1, 5)$. Calcular $V[X_2]$.

60.- Demostrar el teorema 5.11.

61.- Supongamos que el vector aleatorio $\underline{X} = (X_1, X_2)$ tiene una distribución nonrmal bivariada con parámetros $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho$. Calcular $E[X_1]$, $E[X_2]$, $V[X_1]$, $V[X_2]$ y ρ_{X_1,X_2}.

62.- Sea (X_1, X_2) un vector aleatorio con función de densidad $f_{\underline{X}}(x_1, x_2)$. Demostrar que se cumple $V[aX_1 + bX_2 + c] = a^2V[X_1] + b^2V[X_2] + 2abCov(X_1, X_2)$.

63.- Un actuario es asesor de una compañía de seguros. El asesor ha analizado que el tiempo de vida en años de Juan se distribuye uniforme sobre el intervalo $(0, 20)$ y además, que la vida de Pedro se distribuye exponencial con media 25 años. Se sabe que ambas vidas son independientes. Calcular las siguientes probabilidades

a) Juan viva más años que Pedro.
b) La suma de años entre Juan y Pedro sea menor a 20 años.

64.- Sea (X_1, X_2) un vector aleatorio que se distribuye uniformemente sobre la región $R = \{(x_1, x_2) \in \mathbb{R}^2 |\, 0 < x_2 < |x_1| \quad \text{y} -2 < x_1 < 2\}$. Encontrar

a) La función de densidad conjunta de (X_1 y X_2) con sus correspondientes funciones indicadoras.
b) $P[X_1 + X_2 < 2]$.
c) Las funciones de densidad marginales de X_1 y X_2.
d) La función de densidad condicional $f_{X_2|X_1}(x_2|x_1)$.

65.- Considerando un seguro de gastos médicos mayores, se sabe que el total de gastos hospitalarios por una intervención quirúrgica, es una variable aleatoria X_2, la cual se distribuye uniforme sobre el intervalo $(0, 5)$, además los gastos de hospedaje y atención médica en la habitación, es otra variable aleatoria X_1, que se distribuye uniforme sobre el intervalo $(0, x_2/4)$.

a) Encontrar la función de densidad de X_1.
b) Calcular $E[X_1]$.

Distribuciones de funciones de variables aleatorias

Capítulo

Introducción

En este capítulo se tratarán las distribuciones de funciones o transformaciones de variables y vectores aleatorios. Esto es, dada una variable o vector aleatorio con distribución conocida, el objetivo es encontrar la distribución de una nueva variable, la cual es función de la variable o del vector dado.

El tema de distribuciones de funciones de variables o vectores aleatorios tiene importantes aplicaciones en problemas de inferencia estadística, en simulación estocástica, y en problemas actuariales, entre otras áreas.

No existe mucha teoría sobre el tema de transformaciones de variables aleatorias, lo que existe son procedimientos y, por lo tanto, predominan más los ejemplos de diferente grado de dificultad. Se puede encontrar la distribución de la transformación de una variable o vector aleatorio por medio de alguno de cuatro métodos, estos se enlistan a continuación:

1. Método de la función de distribución.
2. Método discreto.
3. Técnica de cambio de variable.
4. Método de función generadora de momentos.

Es importante aclarar que en ocasiones existen problemas que se podrán resolver por más de un método, pero la recomendación es resolver el problema por el camino más simple, de ahí la importancia de que el estudiante domine cada una de estas metodologías.

6.1. Método de función de distribución

A continuación, se explicará en forma general en qué consiste el método de función de distribución para encontrar la distribución de la transformación de una variable o vector aleatorio. También, se indicará los alcances y limitaciones de este método.

Supongamos que conocemos la distribución de una variable aleatoria X, ya sea su función de densidad o su función de distribución. Sea $U = g(X)$ una transformación de esta variable, entonces, el objetivo del método de función de distribución es encontrar la función de distribución de la nueva variable aleatoria U, usando el conocimiento de la distribución de la variable X.

En la explicación anterior, X es una variable aleatoria y g es una transformación real. Los problemas con este método se pueden ampliar a transformaciones reales de un vector aleatorio $\underline{X} = (X_1, X_2)$, esto es, transformaciones de $\mathbb{R}^2$ a $\mathbb{R}$. Este método no es práctico para vectores $\underline{X}$ de mayor dimensión, ya que como se verá, los resultados por este método se recargan, casi siempre, en las gráficas de los recorridos, tanto de la variable o vector original, como de la nueva variable.

No queda más que realizar ejercicios con diferente grado de dificultad para asimilar y dominar este método. Se empezará con ejemplos donde la transformación va de $\mathbb{R}$ a $\mathbb{R}$.

Ejemplo 6.1. Sea X una variable aleatoria con función de densidad dada por $f_X(x) = 2xI_{(0,1)}(x)$, encontrar la función de densidad de $U = 3X - 1$.

Se recomienda realizar la gráfica de la transformación de la variable aleatoria, para tener claridad sobre el recorrido de la nueva variable U, además ayuda a confirmar características de la función como si es biyectiva o no, si es creciente o decreciente. Esta gráfica se da a continuación:

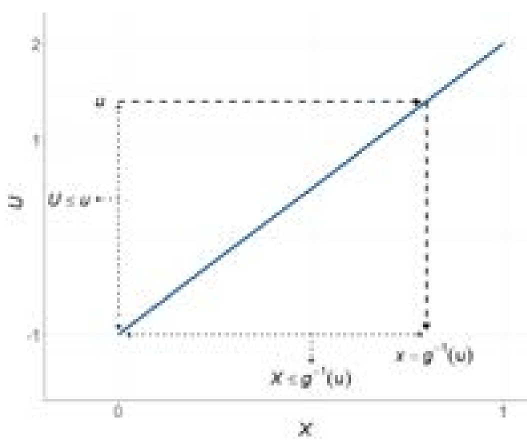


Figura 6.1. *Transformación del ejemplo 6.1*

Es notorio que el recorrido de la nueva variable es el intervalo $(-1, 2)$. La función de distribución de X para valores del recorrido es

$$\begin{aligned} F_X(x) &= P[X \leq x] \\ &= \int_0^x 2t dt \\ &= x^2. \end{aligned}$$

Ahora se obtendrá la función de distribución de U

$$
\begin{aligned}
F_U(u) &= P[U \leq u] \\
&= P[3X - 1 \leq u] \\
&= P\left[X \leq \frac{u+1}{3}\right] \\
&= F_X\left(\frac{u+1}{3}\right) \\
&= \left(\frac{u+1}{3}\right)^2, \text{ si } u \in (-1, 2).
\end{aligned}
$$

De esta manera,

$$F_U(u) = \left(\frac{u+1}{3}\right)^2 I_{(-1,2)}(u) + I_{[2,\infty)}(u).$$

Por lo tanto, la función de densidad está dada por

$$f_U(u) = \frac{2}{9}(u+1)I_{(-1,2)}(u).$$

Ejemplo 6.2. Sea X una variable aleatoria con función de densidad dada por $f_X(x) = 2(1-x)I_{(0,1)}(x)$, encontrar la función de densidad de $U = 4 - 7X$.

La gráfica de la transformación se da a continuación:

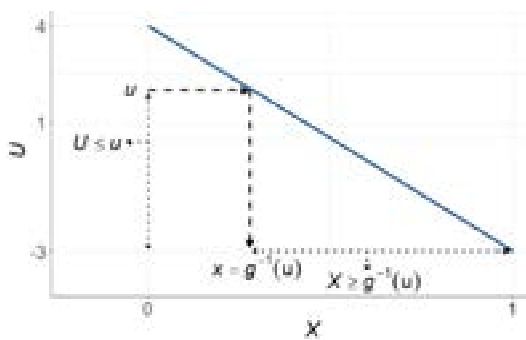


Figura 6.2. *Transformación del ejemplo 6.2*

Se puede observar que el recorrido de la nueva variable es $(-3, 4)$.
La función de distribución de X para valores de su recorrido es

$$
\begin{aligned}
F_X(x) &= P[X \leq x] \\
&= \int_0^x 2(1-t)dt \\
&= 2t - t^2 \Big|_0^x \\
&= 2x - x^2.
\end{aligned}
$$

La función de distribución de la nueva variable

$$\begin{aligned}
F_U(u) &= P[U \leq u] \\
&= P[4 - 7X \leq u] \\
&= P\left[X \geq \frac{4-u}{7}\right] \\
&= 1 - F_X\left(\frac{4-u}{7}\right) \\
&= 1 - 2\left(\frac{4-u}{7}\right) + \left(\frac{4-u}{7}\right)^2.
\end{aligned}$$

Por lo tanto,

$$F_U(u) = \left[1 - 2\left(\frac{4-u}{7}\right) + \left(\frac{4-u}{7}\right)^2\right] I_{(-3,4)}(u) + I_{[4,\infty)}(u).$$

De esta manera,

$$f_U(u) = \left(\frac{2}{49}\right)(3+u) I_{(-3,4)}(u).$$

Hasta el momento, los dos ejemplos tratados son transformaciones biyectivas. Por este método también se pueden tratar funciones que no son uno a uno, veamos el siguiente ejemplo.

Ejemplo 6.3. Sea X una variable aleatoria con distribución $U(-3,3)$. Sea $U = g(X) = 9 - X^2$. Encontrar la función de densidad de U.

La gráfica de la transformación se presenta a continuación:

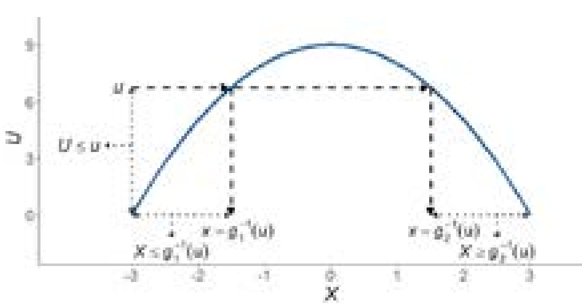


Figura 6.3. *Transformación del ejemplo 6.3*

Obsérvese que el recorrido de la nueva variable es el intervalo $(0,9)$. Además, es notorio que esta transformación no es uno a uno, es una función que tiene dos inversas para todos los valores de u.

Se puede observar que u proviene de dos preimágenes

$$x = g_1^{-1}(u) = -\sqrt{9-u} \text{ y } x = g_2^{-1}(u) = \sqrt{9-u}.$$

A continuación, encontraremos la función de distribución de U

$$\begin{aligned}F_U(u) &= P[U \leq u] \\ &= P[9 - X^2 \leq u].\end{aligned}$$

Se puede notar, véase la gráfica anterior, que la preimagen del evento $9 - X^2 \leq u$ es

$$\{x \in \mathbb{R} | -3 < X \leq -\sqrt{9-u} \quad \text{ó} \quad \sqrt{9-u} \leq X < 3\}.$$

De esta manera, la función de distribución es igual a

$$\begin{aligned}F_U(u) &= P[-3 < X \leq -\sqrt{9-u}] + P[\sqrt{9-u} \leq X < 3] \\ &= \frac{-\sqrt{9-u}+3}{6} + \frac{3-\sqrt{9-u}}{6} \\ &= \frac{6-2\sqrt{9-u}}{6}, \text{ donde } u \in (0,9).\end{aligned}$$

Por lo tanto,

$$F_U(u) = \frac{6-2\sqrt{9-u}}{6} I_{(0,9)}(u) + I_{[9,\infty)}(u).$$

Entonces, la función de densidad es

$$f_U(u) = \frac{1}{6\sqrt{9-u}} I_{(0,9)}(u).$$

Ejemplo 6.4. Encontrar la función de densidad de $U = g(X) = -|X-2|$, donde la variable aleatoria X tiene distribución uniforme sobre el intervalo $(0,4)$.

Igual que en los ejercicios anteriores, se presentará la gráfica de la transformación

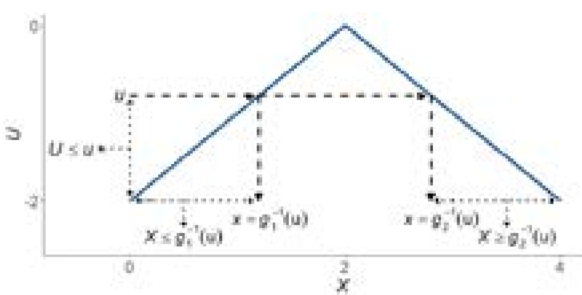


Figura 6.4. *Transformación del ejemplo 6.4*

Se puede ver en la gráfica que u tiene dos preimágenes,

$$x = g_1^{-1}(u) = u + 2 \quad \text{y} \quad x = g_2^{-1}(u) = 2 - u.$$

Se puede también notar que el recorrido de la variable aleatorio U es el intervalo $(-2, 0)$.

La función de distribución es

$$\begin{aligned}
F_U(u) &= P[U \le u] \\
&= P[-|X-2| \le u] \\
&= P[0 \le X \le u+2] + P[2-u \le X \le 4] \\
&= \frac{u+2}{4} + \frac{4-2+u}{4} \\
&= \frac{u+2}{2}.
\end{aligned}$$

De esta manera,

$$F_U(u) = \frac{u+2}{2} I_{(-2,0)}(u) + I_{[0,\infty)}(u).$$

Entonces,

$$f_U(u) = \frac{1}{2} I_{[-2,0]}(u).$$

Por lo tanto, $U \sim U(-2, 0)$.

Ejemplo 6.5. Sea X una variable aleatoria la cual se distribuye $U(2, 4)$. Encontrar la distribución de la siguiente transformación de X,

$$U = g(X) = \begin{cases} 3X - 6, & \text{si } X < 3, \\ \\ 12 - 3X, & \text{si } X \ge 3. \end{cases}$$

La gráfica de la transformación se da a continuación:

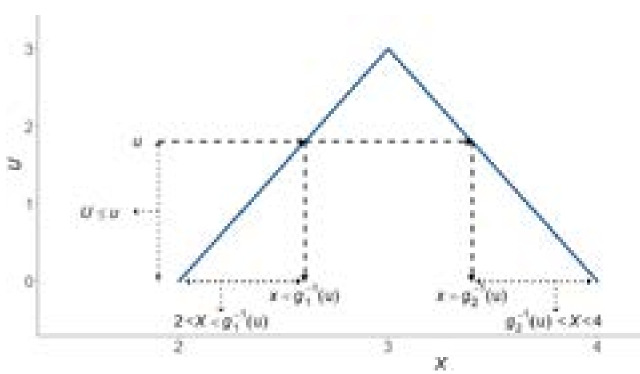


Figura 6.5. *Transformación del ejemplo 6.5*

Se puede observar que $u \in (0, 3)$.

En este caso,

$$g_1^{-1}(u) = \frac{u+6}{3} \text{ y } g_2^{-1}(u) = \frac{12-u}{3}.$$

Entonces, la función de distribución de U se determina de la siguiente manera:

$$\begin{aligned}
F_U(u) &= P[U \leq u] \\
&= P[g(X) \leq u] \\
&= P\left[2 < X < \frac{u+6}{3}\right] + \left[\frac{12-u}{3} < X < 4\right] \\
&= \frac{1}{2}\left[\frac{u+6}{3} - 2\right] + \frac{1}{2}\left[4 - \frac{12-u}{3}\right] \\
&= \frac{u}{3}.
\end{aligned}$$

Por lo tanto,

$$F_U(u) = \frac{u}{3} I_{(0,3)}(u) + I_{[3,\infty)}(u).$$

Y la función de densidad de U

$$f_U(u) = \frac{1}{3} I_{(0,3)}(u).$$

De esta manera, $U \sim U(0,3)$.

Ejemplo 6.6. Sea X una variable aleatoria la cual se distribuye uniforme sobre el intervalo $(0,5)$, y sea $U = g(X)$ una función de X dada de la siguiente manera: $g(x) = |x-2| I_{(0,4]}(x) + (6-x) I_{(4,5)}(x)$. Encontrar la función de densidad de $U = g(X)$.

Obsérvese las siguientes gráficas de la transformación de la variable aleatoria. En la primera, se muestra el caso cuando hay dos inversas y la segunda, se ilustra cuando hay tres inversas.

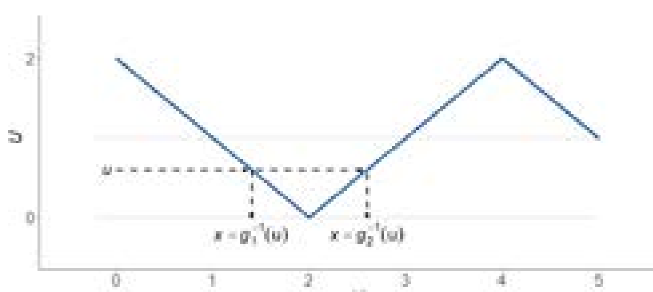


Figura 6.6. *Transformación del ejemplo 6.6, donde se ilustran dos inversas*

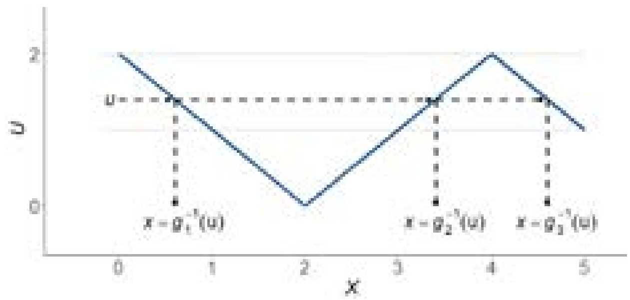


Figura 6.7. *Transformación del ejemplo 6.6, donde se ilustran tres inversas*

Obsérvese que el recorrido de la nueva variable U es el intervalo $(0, 2)$. Cuando $0 < u \leq 1$ hay dos inversas (primera gráfica) y cuando $1 < u \leq 2$ hay tres inversas (segunda gráfica), por esta razón, para obtener la función de distribución de U se deben de considerar dos casos, los cuales se presentarán a continuación.

Caso 1. Cuando $0 < u \leq 1$ (Cuando hay dos inversas).
Obsérvese la siguiente gráfica

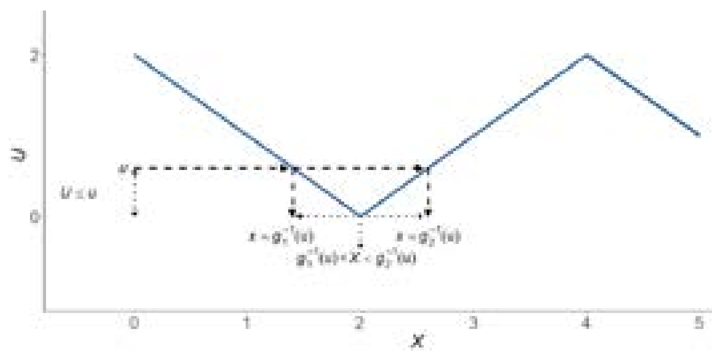


Figura 6.8. *Transformación del ejemplo 6.6, caso: dos inversas*

En este caso, $g_1^{-1}(u) = 2 - u$ y $g_2^{-1}(u) = 2 + u$.
Por lo tanto,

$$\begin{aligned} F_U(u) &= P[U \leq u] \\ &= P[g(X) \leq u] \\ &= P[2 - u \leq X \leq 2 + u] \\ &= \frac{2u}{5}. \end{aligned}$$

Caso 2. Cuando $1 < u \leq 2$ (Cuando hay tres inversas).

En la siguiente gráfica obsérvese con detalle, la importancia de las tres inversas en el cálculo de la función de distribución de U.

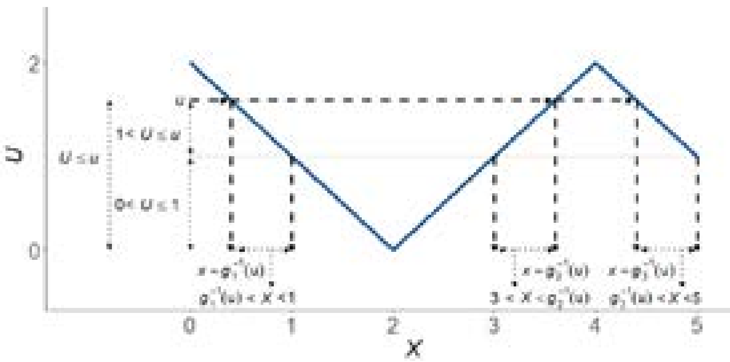


Figura 6.9. *Transformación del ejemplo 6.6, caso: tres inversas*

Obsérvese que, en este caso, las inversas son

$$g_1^{-1}(u) = 2 - u,\ g_2^{-1}(u) = 2 + u \ \text{ y } \ g_3^{-1}(u) = 6 - u.$$

Por lo tanto,

$$\begin{aligned}
F_U(u) &= P[U \leq u] \\
&= P[0 < U \leq 1] + P[1 < U \leq u] \\
&= F_U(1) + P[1 < U \leq u] \\
&= \frac{2}{5} + P[1 < g(X) \leq u] \\
&= \frac{2}{5} + P[2 - u \leq X < 1] + P[3 < X \leq u + 2] + P[6 - u \leq X < 5] \\
&= \frac{2}{5} + \frac{1 - 2 + u}{5} + \frac{u + 2 - 3}{5} + \frac{5 - 6 + u}{5} \\
&= \frac{3u - 1}{5}.
\end{aligned}$$

De esta manera,

$$F_U(u) = \frac{2}{5} u I_{(0,1]}(u) + \frac{3u - 1}{5} I_{(1,2)}(u) + I_{[2,\infty)}(u).$$

Por lo tanto, la función de densidad es

$$f_U(u) = \frac{2}{5} I_{(0,1]}(u) + \frac{3}{5} I_{(1,2)}(u).$$

Los ejemplos mostrados hasta este momento consideran transformaciones de los reales a los reales. En el siguiente ejemplo, la transformación que se considera es de un vector en $\mathbb{R}^2$ a los reales. Cabe mencionar que para este tipo de problemas, es de mucho apoyo tener claridad sobre el recorrido del vector aleatorio $\underline{X} = (X_1, X_2)$.

Ejemplo 6.7. Sea $\underline{X} = (X_1, X_2)$ un vector aleatorio con función de densidad dada por $f_{\underline{X}}(x_1, x_2) = 3x_1 I_{[0,x_1]}(x_2) I_{[0,1]}(x_1)$. Sea $U = X_1 - X_2$, una transformación de X_1 y X_2. Encontrar por el método de función de distribución, la función de densidad de U.

El recorrido del vector $\underline{X} = (X_1, X_2)$ está representado gráficamente por el triángulo de la siguiente gráfica

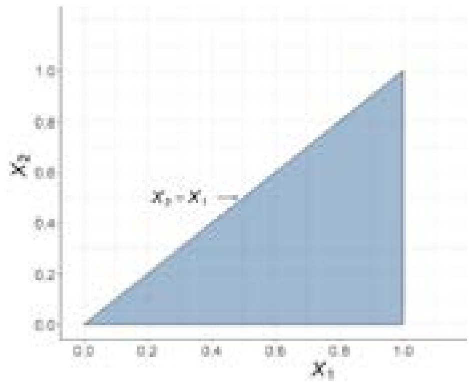


Figura 6.10. *Recorrido conjunto de* X_1 *y* X_2*, ejemplo 6.7*

Primero calcularemos la función de distribución de la nueva variable U, esto es,

$$\begin{aligned} F_U(u) &= P[U \leq u] \\ &= P[X_1 - X_2 \leq u]. \end{aligned}$$

Es importante tener claridad, geométricamente, sobre cuál es la región de $X_1 - X_2 \leq u$, para calcular con más facilidad la probabilidad anterior. Se puede notar que la desigualdad $X_1 - X_2 \leq u$ es la subregión del recorrido del vector $\underline{X}$, limitada por las siguientes rectas: $x_2 = x_1$; $x_2 = x_1 - u$; $x_2 = 0$ y $x_1 = 1$. En la siguiente gráfica esta región es la de color más obscuro.

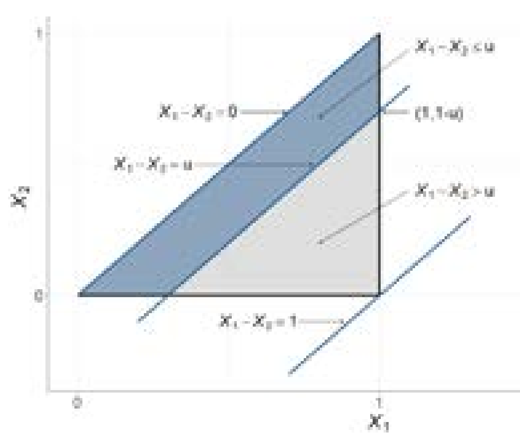


Figura 6.11. *Subregiones, ejemplo 6.7*

Puede notarse en la anterior gráfica, que los valores que toma la nueva variable U se encuentran en el intervalo $(0, 1)$. Esto es, obsérvese que de la familia de rectas $X_1 - X_2 = u$ que tocan la región del recorrido conjunto de X_1 y X_2 es para aquellas rectas donde los valores de u, se encuentran entre 0 y 1.

Entonces, la función de distribución de U se calcula de la siguiente manera:

$$\begin{aligned}
F_U(u) &= P[X_1 - X_2 \leq u] \\
&= 1 - P[X_1 - X_2 > u] \\
&= 1 - \int_0^{1-u} \int_{u+x_2}^{1} 3x_1 dx_1 dx_2 \\
&= 1 - \int_0^{1-u} \frac{3}{2}(1 - (u + x_2)^2) dx_2 \\
&= 1 - \int_0^{1-u} \frac{3}{2} dx_2 + \int_0^{1-u} \frac{3}{2}(u + x_2) dx_2 \\
&= 1 - \frac{3}{2}(1 - u) + \frac{3}{2}\left[\frac{(u + x_2)^3}{3}\right]_0^{1-u} \\
&= \frac{3}{2}u - \frac{u^3}{2}.
\end{aligned}$$

Por lo tanto, la función de distribución expresada con funciones indicadoras es

$$F_U(u) = \left(\frac{3}{2}u - \frac{u^3}{2}\right) I_{(0,1)}(u) + I_{[1,\infty)}(u).$$

De esta manera, derivando la anterior función, se obtiene la función de densidad, esto es,

$$f_U(u) = \left(\frac{3}{2} - \frac{3}{2}u^2\right) I_{(0,1)}(u) = \frac{3}{2}(1 - u^2) I_{(0,1)}(u).$$

Ejemplo 6.8. Sea $\underline{X} = (X_1, X_2)$ un vector aleatorio el cual se distribuye uniformemente sobre el cuadro unitario. Encontrar la función de densidad de $U = X_1 + X_2$.

Para este caso, la función de densidad del vector $\underline{X} = (X_1, X_2)$ está dada por

$$f_{\underline{X}}(x_1, x_2) = I_{(0,1)}(x_1) I_{(0,1)}(x_2).$$

Obsérvese que el recorrido del vector $\underline{X}$ está acotado por las siguientes rectas: $x_2 = 0; x_2 = 1$, $x_1 = 0$ y $x_1 = 1$, esta región se se puede apreciar en la siguiente gráfica.

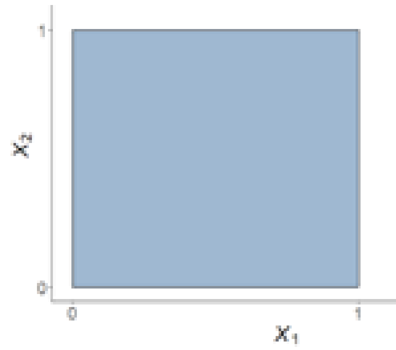


Figura 6.12. *Recorrido conjunto de X_1 y X_2, ejemplo 6.8*

La función de distribución de U, se calcula como

$$F_U(u) = P[X_1 + X_2 \leq u].$$

La desigualdad $X_1 + X_2 \leq u$ está representada por la subregión del recorrido de $\underline{X}$ que está por debajo de la recta $X_1 + X_2 = u$, véase la siguiente gráfica.

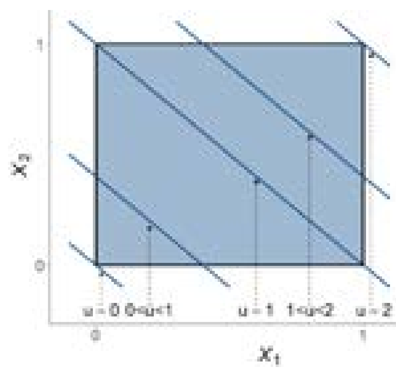


Figura 6.13. *Recorrido de U, ejemplo 6.8*

De esta última gráfica, se pueden apreciar las rectas $X_1+X_2 = u$ que pasan por el recorrido del vector $\underline{X}$, de aquí se puede afirmar que el recorrido de la nueva variable U es el intervalo $(0, 2)$.

Además, se puede observar que la subregión que representa la desigualdad $X_1 + X_2 \leq u$ siempre estará por debajo de la recta $X_1 + X_2 = u$. Esta subregión es un triángulo, mientras $0 < u \leq 1$ (región con color más obscuro), pero cuando $1 < u < 2$, la forma de esta subregión deja de ser un triángulo (región que no tiene color blanco). Estas dos situaciones se pueden apreciar en las siguientes dos gráficas. Lo que nos indica que el cálculo de la función de distribución se deberá realizar en dos casos diferentes. Este cálculo se presentará a continuación:

Caso 1. Cuando $0 < u \leq 1$.

Obsérvese la subregión de la desigualdad $X_1 + X_2 \leq u$ en la siguiente gráfica

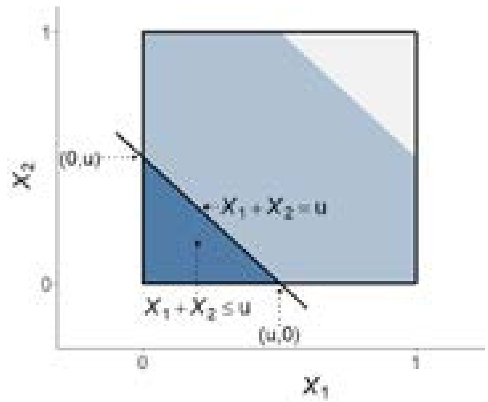


Figura 6.14. *Caso 1.* $0 < U \leq 1$, *ejemplo 6.8*

Entonces, la función de distribución queda como

$$
\begin{aligned}
F_U(u) &= P[X_1 + X_2 \leq u] \\
&= \frac{u^2}{2}.
\end{aligned}
$$

Caso 2. Cuando $1 < u < 2$.

Se puede observar cómo la subregión de $X_1 + X_2 \leq u$ cambió de forma

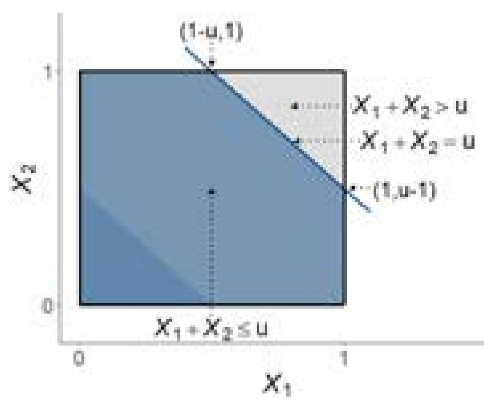


Figura 6.15. *Caso 2.* $1 < U < 2$, *ejemplo 6.8*

Entonces, la función de distribución se calculará de la siguiente manera:

$$
\begin{aligned}
F_U(u) &= P[X_1 + X_2 \leq u] \\
&= 1 - P[X_1 + X_2 > u] \\
&= 1 - \frac{(2-u)^2}{2}.
\end{aligned}
$$

De esta manera, la función de distribución es

$$F_U(u) = \frac{u^2}{2} I_{(0,1]}(u) + \left[1 - \frac{(2-u)^2}{2}\right] I_{(1,2)}(u) + I_{[2,\infty)}(u).$$

Y la función de densidad queda de la siguiente manera:

$$f_u(u) = uI_{(0,1]}(u) + (2-u)I_{(1,2)}(u).$$

El siguiente ejemplo, se deja para que lo resuelva el estudiante

Ejemplo 6.9. Sea $\underline{X} = (X_1, X_2)$ un vector aleatorio el cual se distribuye uniformemente sobre el cuadro unitario. Encontrar la función de densidad de U, donde
a) $U = X_1 - X_2$.
b) $U = \dfrac{X_1}{X_2}$.

El método de función de distribución puede ser usado para encontrar la fórmula de convolución para variables aleatorias continuas, la cual es una fórmula para encontrar la función de densidad de una variable aleatoria que es la suma de dos variables aleatorias, véase la siguiente observación.

Observación. Sean X_1 y X_2 variables aleatorias independientes continuas con funciones de densidad $f_{X_1}(x_1)$ y $f_{X_2}(x_2)$, respectivamente, sea $U = X_1 + X_2$, entonces,

$$f_U(u) = \int_{-\infty}^{\infty} f_{X_1}(x_1) f_{X_2}(u - x_1) dx_1.$$

A la anterior fórmula se le conoce como *fórmula de convolución para variables continuas.*

A continuación, realizaremos la demostración

$$\begin{aligned}
F_U(u) &= P[X_1 + X_2 \leq u] \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{u-x_1} f_{X_1}(x_1) f_{X_2}(x_2) dx_2 dx_1 \\
&= \int_{-\infty}^{\infty} f_{X_1}(x_1) \left[\int_{-\infty}^{u-x_1} f_{X_2}(x_2) dx_2\right] dx_1 \\
&= \int_{-\infty}^{\infty} f_{X_1}(x_1) F_{X_2}(u - x_1) dx_1.
\end{aligned}$$

Entonces, la función de densidad de U es

$$\begin{aligned}
f_U(u) &= \frac{d}{du} \int_{-\infty}^{\infty} f_{X_1}(x_1) F_{X_2}(u - x_1) dx_1 \\
&= \int_{-\infty}^{\infty} \frac{d}{du} f_{X_1}(x_1) F_{X_2}(u - x_1) dx_1 \\
&= \int_{-\infty}^{\infty} f_{X_1}(x_1) f_{X_2}(u - x_1) dx_1.
\end{aligned}$$

Análogamente, se puede demostrar que

$$f_U(u) = \int_{-\infty}^{\infty} f_{X_2}(x_2) f_{X_1}(u - x_2) dx_2.$$

Ejemplo 6.10. Considerando el ejemplo 6.8, encontrar de nuevo la función de densidad de $U = X_1 + X_2$, pero usando la fórmula de convolución.

El recorrido de X_i es el intervalo $(0, 1)$, por lo que el recorrido de U es el intervalo $(0, 2)$.

La función de densidad de $U = X_1 + X_2$ se calcula como

$$\begin{aligned} f_U(u) &= \int_{-\infty}^{\infty} f_{X_2}(x_2) f_{X_1}(u - x_2) dx_2 \\ &= \int_{-\infty}^{\infty} f_{X_1}(u - x_2) dx_2. \end{aligned}$$

La anterior integral se realizará sobre todos los valores de x_2 donde la función de densidad $f_{X_1}(u-x_2)$ es positiva. Esto ocurre cuando $0 < u-x_2 < 1$, además se sabe $0 < x_2 < 1$. Lo anterior ocurre si y solo si $u - 1 < x_2 < u$ y $0 < x_2 < 1$, véase la región conjunta de X_2 y u en la siguiente gráfica.

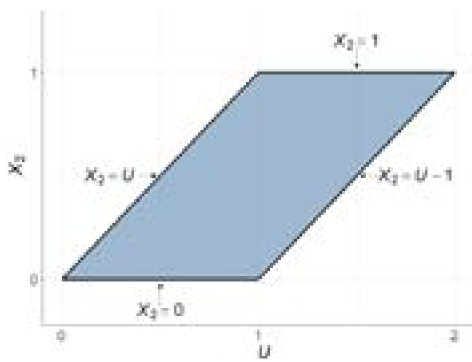


Figura 6.16. *Región x_2 y u, ejemplo 6.10*

De esta manera, existen dos casos para obtener $f_U(u)$.

Caso 1. Cuando $0 < u \leq 1$.

$$f_U(u) = \int_0^u dx_2 = u.$$

Caso 2. Cuando $1 < u < 2$.

$$f_U(u) = \int_{u-1}^{1} dx_2 = 1 - (u - 1) = 2 - u.$$

De este modo,

$$f_U(u) = uI_{(0,1]}(u) + (2 - u)I_{(1,2)}(u).$$

Observaciones

1. Este método está limitado a resolver problemas donde la transformación va de $\mathbb{R}$ a $\mathbb{R}$ o de $\mathbb{R}^2$ a $\mathbb{R}$, ya que la solución depende, en muchas ocasiones, de las gráficas de las subregiones de los eventos de la nueva variable.
2. El método puede ser aplicado para casos discretos con transformaciones de $\mathbb{R}$ a $\mathbb{R}$, por ejemplo, si X tiene distribución uniforme discreta o geométrica, en estos casos, la función de distribución se puede obtener en forma explícita. Mas siempre será más fácil resolver el problema por medio del método discreto.

6.2. Método discreto

En esta sección se tratará el método para encontrar distribuciones de funciones de variables o vectores aleatorios que son discretos. Se explicará en qué consiste el método y se tratarán algunos ejemplos.

Se tratarán primero transformaciones de $\mathbb{R}$ a $\mathbb{R}$. Esto es, supongamos que conocemos la función de densidad (o distribución) de una variable aleatoria discreta X y $g(X) = U$ es una función de X (de $\mathbb{R}$ a $\mathbb{R}$), entonces, el objetivo es encontrar la función de densidad de la nueva variable U de la siguiente manera:

$$f_U(u) = P[U = u] = P[g(X) = u].$$

La última probabilidad depende de la variable X, y al lograr despejar la variable X de la igualdad $g(X) = u$, será posible usar la función de densidad de la variable X, la cual es conocida.

Primero consideraremos el caso donde $U = g(X)$ es una transformación biyectiva, esto es, tiene solo una inversa. En este caso, la probabilidad para un valor específico de U, depende de la probabilidad de un valor específico de X. Esto es, la función de densidad de U se puede calcular de la siguiente manera:

$$\begin{aligned} f_U(u) &= P[U = u] \\ &= P[g(X) = u] \\ &= P[X = g^{-1}(u)] \\ &= f_X(g^{-1}(u)). \end{aligned}$$

Ejemplo 6.11. Sea X una variable aleatoria, la cual tiene distribución geométrica con parámetro p, sea $U = 2X - 1$. Encontrar la función de densidad de U.

Obsérvese que el recorrido de la nueva variable U es $u = -1, 1, 3, 5, \ldots$

La función de densidad de U se calcula como

$$\begin{aligned}
f_U(u) &= P[U = u] \\
&= P[g(X) = u] \\
&= P[2X - 1 = u] \\
&= P\left[X = \frac{u+1}{2}\right] \\
&= p(1-p)^{\frac{u+1}{2}}, \text{ para } u = -1, 1, 3, 5, \ldots
\end{aligned}$$

Por lo tanto,

$$f_U(u) = p(1-p)^{\frac{u+1}{2}} I_{\{-1,1,3,5,\ldots\}}(u).$$

El siguiente ejemplo considera una transformación que no es uno a uno.

Ejemplo 6.12. Sea X una variable aleatoria con distribución hipergeométrica con parámetros $N = 11$, $k = 4$ y $n = 4$ y sea $U = g(X) = (X-2)^2$ una transformación de X. Encontrar la función de densidad de U.

Obsérvese la gráfica de la transformación

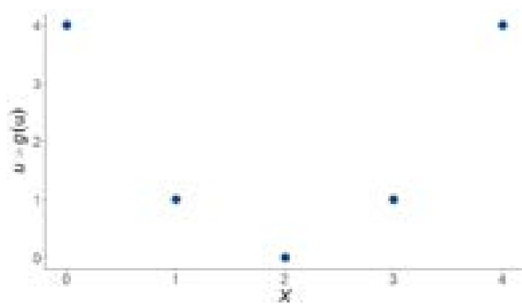


Figura 6.17. *Transformación del ejemplo 6.12*

De esta manera, se puede observar que el recorrido de la nueva variable es $u = 0, 1, 4$.

Para encontrar la función de densidad de U, existen dos casos, para cada caso el número de inversas es diferente.

Caso 1. Cuando $u = 0$ (existe solo una inversa).

$$\begin{aligned}
P[U = u] &= P[X = 2] \\
&= \frac{\binom{4}{2}\binom{7}{2}}{\binom{11}{4}}.
\end{aligned}$$

Caso 2. Cuando $u = 1, 4$ (existen dos inversas).

$$
\begin{aligned}
P[U = u] &= P[(X-2)^2 = u] \\
&= P[X = 2 \pm \sqrt{u}] \\
&= P[X = 2 - \sqrt{u}] + P[X = 2 + \sqrt{u}] \\
&= \frac{\binom{4}{2-\sqrt{u}}\binom{7}{4-(2-\sqrt{u})}}{\binom{11}{4}} + \frac{\binom{4}{2+\sqrt{u}}\binom{7}{4-(2+\sqrt{u})}}{\binom{11}{4}}.
\end{aligned}
$$

Por lo tanto,

$$
\begin{aligned}
f_U(u) = {} & \frac{\binom{4}{2}\binom{7}{2}}{\binom{11}{4}} I_{\{0\}}(u) \\
& + \left[\frac{\binom{4}{2-\sqrt{u}}\binom{7}{2+\sqrt{u}}}{\binom{11}{4}} + \frac{\binom{4}{2+\sqrt{u}}\binom{7}{2-\sqrt{u}}}{\binom{11}{4}} \right] I_{\{1,4\}}(u).
\end{aligned}
$$

Hasta el ejemplo anterior, las transformaciones consideradas van de $\mathbb{R}$ a $\mathbb{R}$. Este método también funciona para funciones que van de $\mathbb{R}^2$ a $\mathbb{R}^2$. Además, será posible resolver problemas donde la transformación va de $\mathbb{R}^2$ a $\mathbb{R}$, modificando la transformación, véase los siguientes dos problemas.

Ejemplo 6.13. Sean X_1 y X_2 variables aleatorias independientes tal que X_i tiene distribución Poisson con parámetro $\lambda_i, i = 1, 2$. Encontrar la distribución de $U = X_1 + X_2$.

La transformación de este problema va de $\mathbb{R}^2$ a $\mathbb{R}$. Se introducirá una nueva variable y se obtendrá de esta manera, una nueva transformación que va de $\mathbb{R}^2$ a $\mathbb{R}^2$. Esto es, la nueva transformación tendrá dos componentes, uno de ellos será precisamente la función de la cual se pretende obtener su distribución, es una función va de $\mathbb{R}^2$ a $\mathbb{R}$. El otro componente se define como otra transformación que va también de $\mathbb{R}^2$ a $\mathbb{R}$, no importando mucho cómo se define esta. Entonces, la idea es encontrar la función de densidad conjunta de las dos variables que componen el nuevo vector, y a partir de la función de densidad conjunta encontrada, se obtendrá la función de densidad marginal de la primera variable.

Para este ejemplo, la transformación inicial es $g_1(X_1, X_2) = X_1 + X_2 = U_1$ (de $\mathbb{R}^2$ a $\mathbb{R}$).

Definiremos una nueva transformación que va de $\mathbb{R}^2$ a $\mathbb{R}^2$ de la siguiente manera: $\underline{U} = (U_1, U_2) = (g_1(X_1, X_2), g_2(X_1, X_2)) = g(X_1, X_2)$. La definición de $U_2 = g_2(X_1, X_2)$ es arbitraria, finalmente no nos interesa obtener su función de densidad. En este sentido, conviene definirla como una función simple, fácil de trabajar, que nos facilite, primero, encontrar la función de densidad conjunta de U_1 y U_2, y segundo, a partir de la función de densidad conjunta, encontrar sin complicaciones la función de densidad de U_1.

Para este ejemplo, se propone $g_2(X_1, X_2) = X_2$, esto es, $\underline{U} = (U_1, U_2) = (X_1 + X_2, X_2)$.

La función de densidad conjunta de X_1 y X_2 es

$$f_{\underline{X}}(x_1, x_2) = \frac{\lambda_1^{x_1} \lambda_2^{x_2}}{x_1! x_2!} e^{-(\lambda_1+\lambda_2)} I_{\{0,1,2,\ldots\}}(x_1) I_{\{0,1,2,\ldots\}}(x_2).$$

El recorrido conjunto de X_1 y X_2 se presentará en la siguiente gráfica

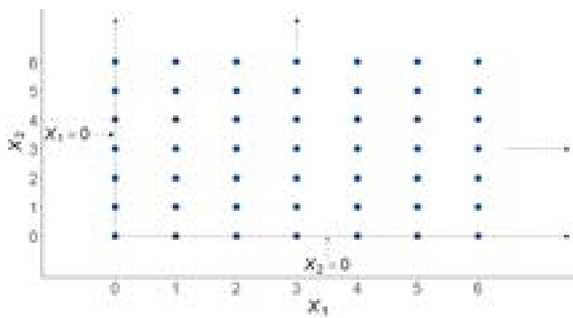


Figura 6.18. *Recorrido conjunto de X_1 y X_2, ejemplo 6.13*

Este recorrido está limitado por las rectas $X_1 = 0$ y $X_2 = 0$.

Considerando la transformación $g(X_1, X_2) = (X_1 + X_2, X_2)$, se puede observar que la recta $X_1 = 0$ se transforma en la recta $U_2 = U_1$ y la recta $X_2 = 0$ en la recta $U_2 = 0$, esto es, si $X_1 = 0$, entonces, $U_1 = X_2$ y $U_2 = X_2$, esto es, $U_1 = U_2$. Si $X_2 = 0$, entonces, $U_1 = X_1$ y $U_2 = 0$, esto es, $U_2 = 0$.

De esta manera, el recorrido conjunto de U_1 y U_2, gráficamente es

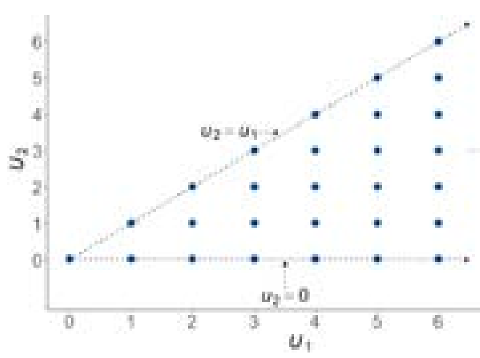


Figura 6.19. *Recorrido conjunto de U_1 y U_2, ejemplo 6.13*

La inversa de la transformación se obtendrá resolviendo el siguiente sistema de ecuaciones para x_1 y x_2

$$u_1 = x_1 + x_2,$$

$$u_2 = x_2.$$

La solución es

$$x_1 = u_1 - u_2 = k_1(u_1, u_2),$$

$$x_2 = u_2 = k_2(u_1, u_2),$$

donde $k_1(u_1, u_2)$ y $k_2(u_1, u_2)$ son los componentes de la función inversa.

De esta manera, se puede obtener la función de densidad de $\underline{U} = (U_1, U_2)$

$$\begin{aligned}
f_{\underline{U}}(u_1, u_2) &= P[U_1 = u_1, U_2 = u_2] \\
&= P[X_1 + X_2 = u_1, X_2 = u_2] \\
&= P[X_1 = u_1 - u_2, X_2 = u_2] \\
&= f_{\underline{X}}(u_1 - u_2, u_2) \\
&= \frac{\lambda_1^{u_1 - u_2} \lambda_2^{u_2}}{(u_1 - u_2)! u_2!} e^{-(\lambda_1 + \lambda_2)} I_{\{0,1,2,\ldots,u_1\}}(u_2) I_{\{0,1,2,\ldots\}}(u_1).
\end{aligned}$$

Ahora, se obtendrá la función de densidad de $U_1 = X_1 + X_2$

$$\begin{aligned}
f_{U_1}(u_1) &= \sum_{u_2=0}^{u_1} \frac{\lambda_1^{u_1 - u_2} \lambda_2^{u_2}}{(u_1 - u_2)! u_2!} e^{-(\lambda_1 + \lambda_2)} \\
&= e^{-(\lambda_1 + \lambda_2)} \sum_{u_2=0}^{u_1} \frac{\lambda_1^{u_1 - u_2} \lambda_2^{u_2}}{(u_1 - u_2)! u_2!} \\
&= e^{-(\lambda_1 + \lambda_2)} \frac{1}{u_1!} \sum_{u_2=0}^{u_1} \frac{u_1!}{(u_1 - u_2)! u_2!} \lambda_1^{u_1 - u_2} \lambda_2^{u_2} \\
&= e^{-(\lambda_1 + \lambda_2)} \frac{1}{u_1!} \sum_{u_2=0}^{u_1} \binom{u_1}{u_2} \lambda_1^{u_1 - u_2} \lambda_2^{u_2} \\
&= e^{-(\lambda_1 + \lambda_2)} \frac{1}{u_1!} (\lambda_1 + \lambda_2)^{u_1}.
\end{aligned}$$

Obsérvese que para el cálculo de la última suma, se aplicó el teorema del binomio (véase apéndice C).

Por lo tanto,

$$f_{U_1}(u_1) = \frac{(\lambda_1 + \lambda_2)^{u_1}}{u_1!} e^{-(\lambda_1 + \lambda_2)} I_{\{0,1,2,\ldots\}}(u_1).$$

Se concluye que $U_1 \sim Poisson(\lambda_1 + \lambda_2)$.

Ejemplo 6.14. Sea $\underline{X} = (X_1, X_2)$ un vector aleatorio el cual tiene distribución trinomial con parámetros n, p_1 y p_2. Encontrar la distribución de $U = X_1 + X_2$.

La función de densidad conjunta de X_1 y X_2 es

$$f_{\underline{X}}(x_1, x_2) =$$

$$\frac{n!}{x_1!x_2!(n - x_1 - x_2)!} p_1^{x_1} p_2^{x_2} (1 - p_1 - p_2)^{n - x_1 - x_2} I_{\{0,1,\ldots,n-x_2\}}(x_1) I_{\{0,1,\ldots,n\}}(x_2).$$

La gráfica del recorrido conjunto de las variables X_1 y X_2 se da a continuación:

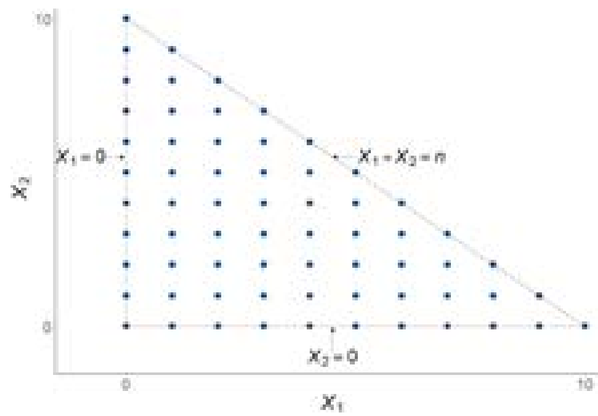


Figura 6.20. *Recorrido conjunto de X_1 y X_2 ($n = 10$), ejemplo 6.14*

Como se había comentado, las dimensiones del dominio y contradominio deben ser iguales, para lograrlo definimos la transformación $g(X_1, X_2)$ de la siguiente manera:

$$\begin{aligned} g(X_1, X_2) &= (g_1(X_1, X_2), g_2(X_1, X_2)) \\ &= (X_1 + X_2, X_1) \\ &= (U_1, U_2). \end{aligned}$$

En la anterior gráfica, se puede observar que el recorrido del vector $\underline{X} = (X_1, X_2)$ está limitado por las siguientes rectas: $x_2 = 0$, $x_1 = 0$ y $x_1 + x_2 = n$. Cada una de estas rectas, bajo la transformación $g(X_1, X_2)$, definida previamente, se transforman, respectivamente, en las siguientes rectas: $u_1 = u_2$, $u_2 = 0$ y $u_1 = n$. Esto es, si $x_2 = 0$, entonces, $u_1 = x_1$ y $u_2 = x_1$, y de esta manera, $u_1 = u_2$. Por otro lado, si $x_1 = 0$, entonces, $u_1 = x_2$ y $u_2 = 0$, por lo que $u_2 = 0$. Finalmente, si $x_1 + x_2 = n$, entonces, $u_1 = n$ y $u_2 = n - x_2$, de esta manera se concluye que $u_1 = n$.

Entonces, el recorrido conjunto de U_1 y U_2 gráficamente es

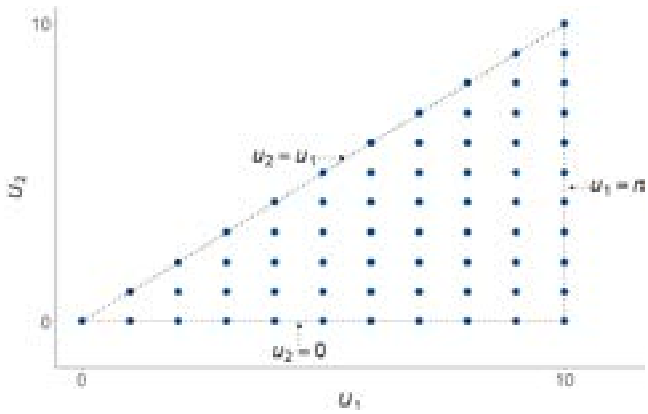


Figura 6.21. *Recorrido conjunto de las variables U_1 y U_2 ($n = 10$), ejemplo 6.14*

Ahora, se obtendrá la inversa de la transformación, resolviendo el siguiente sistema de ecuaciones:

$$u_1 = x_1 + x_2,$$

$$u_2 = x_1.$$

La solución es

$$x_1 = u_2 = k_1(u_1, u_2),$$

$$x_2 = u_1 - u_2 = k_2(u_1, u_2),$$

donde $k_1(u_1, u_2)$ y $k_2(u_1, u_2)$ son los componentes de la función inversa.

De este modo, se procederá a calcular la función de densidad del vector $\underline{U} = (U_1, U_2)$

$$\begin{aligned} f_{\underline{U}}(u_1, u_2) &= P[U_1 = u_1, U_2 = u_2] \\ &= P[X_1 + X_2 = u_1, X_2 = u_2] \\ &= P[X_1 = u_2, X_2 = u_1 - u_2] \\ &= f_{\underline{X}}(u_2, u_1 - u_2) \\ &= \frac{n!}{u_2!(u_1 - u_2)!(n - u_1)!} p_1^{u_2} p_2^{u_1 - u_2} (1 - p_1 - p_2)^{n - u_1}, \end{aligned}$$

donde el recorrido de $\underline{U}$ se puede observar en la gráfica anterior.

La función de densidad de $\underline{U}$ usando funciones indicadoras se da a continuación

$$\begin{aligned} f_{\underline{U}}(u_1, u_2) = &\left[\frac{n!}{u_2!(u_1 - u_2)!(n - u_1)!} p_1^{u_2} p_2^{u_1 - u_2} (1 - p_1 - p_2)^{n - u_1}\right] \\ &\times I_{\{0,1,\ldots,u_1\}}(u_2) I_{\{0,1,\ldots,n\}}(u_1). \end{aligned}$$

Ahora, se obtendrá la función de densidad de $U_1 = X_1 + X_2$,

$$\begin{aligned}
f_{U_1}(u_1) &= \sum_{u_2=0}^{u_1} \frac{n!}{u_2!(u_1-u_2)!(n-u_1)!} p_1^{u_2} p_2^{u_1-u_2} (1-p_1-p_2)^{n-u_1} \\
&= \frac{n!}{(n-u_1)!}(1-p_1-p_2)^{n-u_1} \sum_{u_2=0}^{u_1} \frac{1}{u_2!(u_1-u_2)!} p_1^{u_2} p_2^{u_1-u_2} \\
&= \frac{n!}{(n-u_1)!u_1!}(1-p_1-p_2)^{n-u_1} \sum_{u_2=0}^{u_1} \frac{u_1!}{u_2!(u_1-u_2)!} p_1^{u_2} p_2^{u_1-u_2} \\
&= \binom{n}{u_1}(1-p_1-p_2)^{n-u_1} \sum_{u_2=0}^{u_1} \binom{u_1}{u_2} p_1^{u_2} p_2^{u_1-u_2} \\
&= \binom{n}{u_1}(1-p_1-p_2)^{n-u_1}(p_1+p_2)^{u_1}, \text{ donde } u_1 = 0, 1, \ldots, n.
\end{aligned}$$

En la última suma se aplicó el teorema del binomio (ver apéndice C). Por lo tanto,

$$f_{U_1}(u_1) = \binom{n}{u_1}(p_1+p_2)^{u_1}(1-p_1-p_2)^{n-u_1} I_{\{0,1,\ldots,n\}}.$$

Se concluye que $U_1 = X_1 + X_2 \sim B(n, p_1 + p_2)$.

A continuación, se presenta la fórmula de convolución para variables aleatorias discretas.

Observación. Sean X_1 y X_2 variables aleatorias discretas independientes con valores enteros no negativos, con funciones de densidad $f_{X_1}(x_1)$ y $f_{X_2}(x_2)$ respectivamente, sea $U = X_1 + X_2$, entonces,

$$f_U(u) = \sum_{x_1=0}^{u} f_{X_1}(x_1) f_{X_2}(u - x_1). \tag{6.1}$$

Esta suma considera todos los casos donde la suma $X_1 + X_2$ es igual a u.

A la ecuación 6.1, se le conoce como *fórmula de convolución para variables discretas.*

Ejemplo 6.15. Considerando el ejemplo 3.10, donde el número de defectos en una pieza tiene distribución Poisson con promedio de 2.5, volver a resolver el inciso b usando la fórmula de convolución, esto es, la probabilidad de que el número de defectos en dos piezas sea igual a 2.

Definimos las siguientes variables

$X_i =$ Número de defectos en la pieza i, $i = 1, 2$.

$U = X_1 + X_2$.

De acuerdo con la fórmula de convolución 6.1,

$$\begin{aligned} f_U(2) &= \sum_{x_1=0}^{2} f_{X_1}(x_1) f_{X_2}(2-x_1) \\ &= f_{X_1}(0) f_{X_2}(2) + f_{X_1}(1) f_{X_2}(1) + f_{X_1}(2) f_{X_2}(0) = 0.08422. \end{aligned}$$

Las probabilidades anteriores fueron calculadas en el ejemplo 3.10.

Ejemplo 6.16. Considerando el ejemplo 6.13, donde X_1 y X_2 son variables aleatorias independientes, y X_i tienen distribución Poisson con $\lambda_i, i = 1, 2$. Encontrar la distribución de $U = X_1 + X_2$ usando la fórmula de convolución.

Aplicando la fórmula de convolución

$$\begin{aligned} f_U(u) &= \sum_{x_1=0}^{u} f_{X_1}(x_1) f_{X_2}(u-x_1) \\ &= \sum_{x_1=0}^{u} \frac{{\lambda_1}^{x_1}}{x_1!} e^{-\lambda_1} \frac{{\lambda_2}^{u-x_1}}{(u-x_1)!} e^{-\lambda_2}. \end{aligned}$$

La anterior suma ya se resolvió en el ejemplo 6.13, siendo el resultado

$$f_U(u) = \frac{e^{-(\lambda_1+\lambda_2)}}{u!} (\lambda_1 + \lambda_2)^u.$$

Concluyéndose de nuevo que $U = X_1 + X_2 \sim Poisson(\lambda_1 + \lambda_2)$.

Observación. La ecuación 6.1 se puede generalizar para variables aleatorias no independientes, esto es,

$$f_U(u) = \sum_{x_1=0}^{u} f_{\underline{X}}(x_1, u - x_1). \tag{6.2}$$

Se le sugiere al estudiante resolver de nuevo el ejemplo 6.14, pero ahora usando la fórmula 6.2.

6.3. Técnica de cambio de variable

Es una metodología propia para transformaciones de variables o vectores aleatorios continuos. El objetivo es encontrar la función de densidad de la transformación de una variable o un vector aleatorio continuo. Este método se presentará considerando varios casos, los cuales se presentarán a continuación.

Caso 1. Transformaciones biyectivas de $\mathbb{R}$ a $\mathbb{R}$.

Sea X una variable aleatoria continua con función de densidad conocida $f_X(x)$ y $U = g(X)$ una transformación real de X, la cual es biyectiva al menos en el recorrido de X. El objetivo es encontrar la función de densidad de U, usando el conocimiento de la distribución de X.

Si la función $g(X)$ es biyectiva (en el recorrido de X), entonces, la función es continua y además monótona creciente o decreciente.

Obsérvese la gráfica de $g(X)$ cuando es monótona creciente.

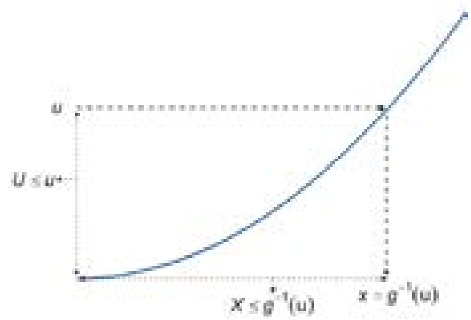


Figura 6.22. *Transformación creciente. Cambio de variable*

La función de distribución de U se calcula de la siguiente manera:

$$\begin{aligned} F_U(u) &= P[U \leq u] \\ &= P[g(X) \leq u] \\ &= P[X \leq g^{-1}(u)] \\ &= F_X(g^{-1}(u)). \end{aligned}$$

Por lo tanto, la función de densidad de U es

$$\begin{aligned} f_U(u) &= \frac{d}{du} F_U(u) \\ &= \frac{d}{du} F_X(g^{-1}(u)) \\ &= f_X(g^{-1}(u)) \left(\frac{d}{du} g^{-1}(u) \right). \end{aligned}$$

Ahora, obsérvese la gráfica de $g(X)$ cuando esta es monótona decreciente

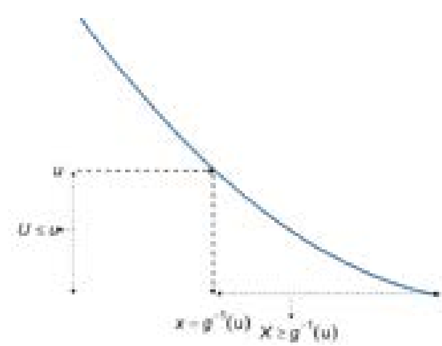


Figura 6.23. *Transformación decreciente. Cambio de variable*

La función de distribución de U queda de la siguiente manera:

$$\begin{aligned} F_U(u) &= P[U \leq u] \\ &= P[g(X) \leq u] \\ &= P[X \geq g^{-1}(u)] \\ &= 1 - P[X < g^{-1}(u)] \\ &= 1 - F_X(g^{-1}(u)). \end{aligned}$$

La función de densidad de U es

$$\begin{aligned} f_U(u) &= \frac{d}{du} F_U(u) \\ &= \frac{d}{du}[1 - F_X(g^{-1}(u))] \\ &= -f_X(g^{-1}(u)) \left(\frac{d}{du} g^{-1}(u) \right). \end{aligned}$$

Obsérvese que en este caso, la derivada de la función inversa es negativa.

En resumen, si $U = g(X)$ es una transformación biyectiva, entonces, la función de densidad de U es

$$f_U(u) = f_X(g^{-1}(u)) \left| \frac{d}{du} g^{-1}(u) \right|.$$

Ejemplo 6.17. Resolver el ejemplo 6.1 por el método de cambio de variable.

La función de densidad de X es $f_X(x) = 2xI_{(0,1)}(x)$ y la transformación $U = g(X) = 3X - 1$. En la solución de este ejercicio por el método de función de distribución se observó que el recorrido de U es el intervalo $(-1, 2)$.

La función inversa $g(x) = u$ es

$$x = \frac{1}{3}(u + 1) = g^{-1}(u).$$

La derivada de la función inversa es

$$\frac{d}{dx} g^{-1}(u) = \frac{1}{3}.$$

Por lo tanto, la función de densidad de U es

$$f_U(u) = \frac{2}{9}(u + 1) I_{(-1,2)}(u).$$

Ejemplo 6.18. Resolver el ejemplo 6.2 por el método de cambio de variable.

La función de densidad de X es $f_X(x) = 2(1 - x)I_{(0,1)}(x)$, y la transformación $U = 4 - 7X$. Este ejercicio se resolvió por el método de función de distribución y se observó que el recorrido de U es el intervalo $(-3, 4)$.

La función inversa es

$$x = \frac{1}{7}(4 - u) = g^{-1}(u).$$

La derivada de la inversa es

$$\frac{d}{dx} g^{-1}(u) = -\frac{1}{7}.$$

Por lo tanto, la función de densidad de U es

$$\begin{aligned} f_U(u) &= f_X\left(\frac{1}{7}(4 - u)\right)\left|-\frac{1}{7}\right| \\ &= \frac{2}{49}(3 + u). \end{aligned}$$

De esta manera,

$$f_U(u) = \frac{2}{49}(3 + u) I_{(-3,4)}(u).$$

Ejemplo 6.19. Sea X una variable aleatoria con distribución uniforme sobre el intervalo $(0, 1)$, sea $U = g(X) = -\alpha \ln(X)$ donde $\alpha > 0$, ¿cómo se distribuye U?

Obsérvese la gráfica de la transformación

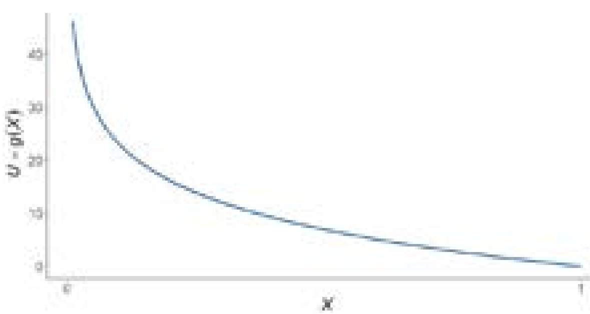


Figura 6.24. *Transformación del ejemplo 6.19*

Se puede observar en la gráfica, que el recorrido de la nueva variable U es el intervalo $(0, \infty)$, también se puede apreciar que la transformación es decreciente.

La inversa de la función es

$$x = e^{-\frac{u}{\alpha}} = g^{-1}(u).$$

La derivada de la función inversa es

$$\frac{d}{du} g^{-1}(u) = -\frac{1}{\alpha} e^{-\frac{u}{\alpha}}.$$

En este caso, la derivada resultó negativa, esto se debe a que la transformación $g(x)$ es decreciente.

Por lo tanto,

$$\left|\frac{d}{du}g^{-1}(u)\right| = \frac{1}{\alpha}e^{-\frac{u}{\alpha}}.$$

De esta manera, la función de densidad de U es

$$\begin{aligned} f_U(u) &= f_X\left(e^{-\frac{u}{\alpha}}\right)\frac{1}{\alpha}e^{-\frac{u}{\alpha}} \\ &= \frac{1}{\alpha}e^{-\frac{u}{\alpha}}, \text{ donde } 0 < u. \end{aligned}$$

Por lo tanto, $U \sim exp\left(\lambda = \frac{1}{\alpha}\right)$.

Uno de los propósitos de la simulación estocástica es generar valores de variables aleatorias con determinada distribución. Por ejemplo, este último resultado es útil para generar valores de una variable aleatoria con distribución exponencial. Con la idea de ilustrar lo anterior, se presentará una simulación específica que se realizó usando el *software* R.

Esto es, se calcularon 10,000 valores de una variable aleatoria X con distribución $U(0,1)$, por medio de la instrucción de R: $runif(10000, 0, 1)$. Posteriormente, para cada valor x generado de la distribución uniforme, se aplicó la transformación $u = -10ln(x)$, logrando así 10,000 valores de una variable aleatoria U con distribución exponencial con media igual a 10 (resultado obtenido del ejemplo anterior). Finalmente, con estos últimos 10,000 valores se construyó, también con R, un histograma.

En la siguiente gráfica, con la idea de comparar, se presenta el histograma junto con la gráfica de la función de densidad de una distribución exponencial con media igual a 10.

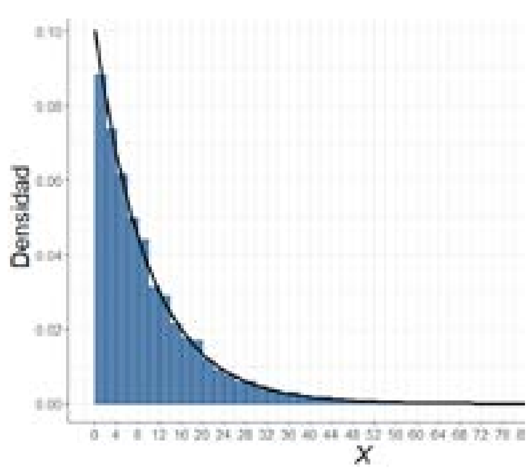


Figura 6.25. *Histograma vs. función de densidad*

Se puede apreciar que el histograma de los 10,000 valores y la gráfica de la función de densidad exponencial son muy parecidos.

Ejemplo 6.20. Sea X una variable aleatoria continua con función de densidad $f(x)$ y función de distribución $F(x)$. Sea $U = g(X) = F(X)$ una transformación de X. ¿Cómo se distribuye U?
a) Por el método de función de distribución.
b) Por el método de cambio de variable.

a) Es importante destacar que la gráfica de la función de distribución es continua, y además, para valores del recorrido, la función de distribución es una función monótona creciente. Dicho de otra manera, la función de distribución es una función biyectiva para valores del recorrido de la variable aleatoria X.

La gráfica típica de una función de distribución para una variable aleatoria continua es como la siguiente gráfica

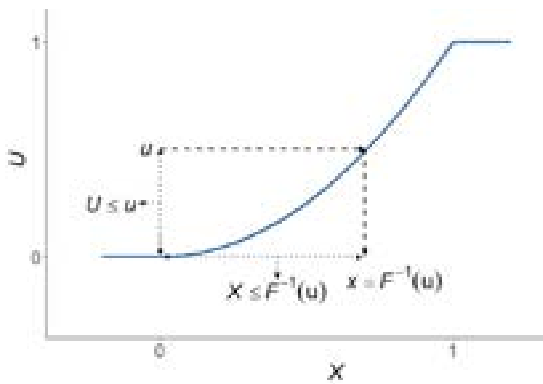


Figura 6.26. *Transformación del ejemplo 6.20*

Se puede observar que el recorrido de la variable U es el intervalo $(0,1)$. A continuación, se obtendrá la función de distribución de U.

$$
\begin{aligned}
F_U(u) &= P[U \leq u] \\
&= P[F(X) \leq u] \\
&= P[X \leq F^{-1}(u)] \\
&= F(F^{-1}(u)) = u, \text{ donde } u \in (0,1).
\end{aligned}
$$

Por lo tanto,

$$F_U(u) = uI_{(0,1)}(u) + I_{[1,\infty)}(u).$$

Entonces, la función de densidad de U es

$$f_U(u) = I_{(0,1)}(u).$$

En conclusión, $U \sim U(0,1)$.

b) Obsérvese que si $U = F(X)$, entonces, $X = F^{-1}(U)$.
De esta manera,

$$f_U(u) = f(F^{-1}(u)) \left| \frac{d}{du} F^{-1}(u) \right|.$$

Aplicando el resultado de la derivada de la inversa de una función

$$\begin{aligned} \frac{d}{du} F^{-1}(u) &= \frac{1}{F'(F^{-1}(u))} \\ &= \frac{1}{f(F^{-1}(u))}, \end{aligned}$$

donde F' es la derivada de la función F.
Por lo tanto,

$$\begin{aligned} f_U(u) &= f(F^{-1}(u)) \left| \frac{d}{du} F^{-1}(u) \right| \\ &= f(F^{-1}(u)) \left| \frac{1}{f(F^{-1}(u))} \right| \\ &= f(F^{-1}(u)) \frac{1}{f(F^{-1}(u))} = 1. \end{aligned}$$

De esta manera,

$$f_U(u) = I_{(0,1)}(u).$$

En conclusión, $U \sim U(0,1)$.

Este resultado sirve para simular números aleatorios con alguna distribución continua específica, en simulación la aplicación de este resultado se conoce como el *método de la función inversa* para generar números aleatorios con alguna distribución continua específica.

Caso 2. Transformaciones de $\mathbb{R}$ a $\mathbb{R}$, con k inversas.

Sea X una variable aleatoria continua con función de densidad conocida $f_X(x)$. Consideremos $U = g(X)$, una transformación real con k inversas en el recorrido de X. Sea A el recorrido de X y B el recorrido de U.

Si $g(X)$ tiene k inversas, significa que para cada elemento de B, este proviene de k elementos de A. Entonces, es posible particionar el conjunto A en k subconjuntos $A_1, A_2, \ldots A_k$, de tal manera, que la transformación $g(X)$ restringida a A_i (g: $A_i \longrightarrow B$), es una una función biyectiva con imagen B, esto es, la función $g(X)$ restingida a A_i, tiene solo una inversa, para $i = 1, 2, \ldots, k$.

Supongamos que las k inversas de $g(x)$, son denotadas como $g_i^{-1}(u), i = 1, 2, \ldots, k$. Obsérvese que la inversa $g_i^{-1}(u)$ tiene como dominio B y como imagen A_i.

De acuerdo al planteamiento anterior, la función de densidad de U se puede expresar de la siguiente manera:

$$f_U(u) = \sum_{i=1}^{k} f_X(g_i^{-1}(u)) \left| \frac{d}{du} g_i^{-1}(u) \right|.$$

Ejemplo 6.21. Sea X una variable aleatoria con distribución uniforme sobre el intervalo $(-3, 3)$ y $U = g(X) = 9 - X^2$ (ejemplo 6.3). Encontrar la función de densidad de U por la técnica cambio de variable.

La gráfica de la transformación se da a continuación:

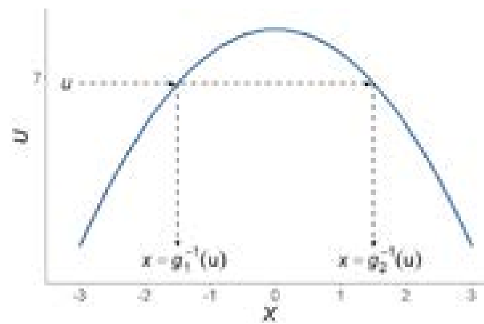


Figura 6.27. *Transformación del ejemplo 6.21*

Obsérvese que el recorrido de la nueva variable es el intervalo $(0, 9)$. Además, esta transformación no es uno a uno, es una función que tiene dos inversas para todos los valores de u.

La función de densidad de X es

$$f_X(x) = \frac{1}{6} I_{(-3,3)}(x).$$

Las dos inversas se obtendrán a continuación:

Si $u = 9 - x^2$, entonces, $x = \pm\sqrt{9-u}$. De esta manera, $g_1^{-1}(u) = -\sqrt{9-u}$ y $g_2^{-1}(u) = \sqrt{9-u}$.

Las derivadas de estas inversas son

$$\begin{aligned} \frac{d}{du} g_1^{-1}(u) &= -\frac{1}{2}(9-u)^{-\frac{1}{2}}(-1) \\ &= \frac{1}{2\sqrt{9-u}}, \end{aligned}$$

$$\begin{aligned} \frac{d}{du} g_2^{-1}(u) &= \frac{1}{2}(9-u)^{-\frac{1}{2}}(-1) \\ &= -\frac{1}{2\sqrt{9-u}}. \end{aligned}$$

Por lo tanto, la función de densidad de U es

$$\begin{aligned} f_U(u) &= f_X(g_1^{-1}(u)) \left| \frac{d}{du} g_1^{-1}(u) \right| + f_X(g_2^{-1}(u)) \left| \frac{d}{du} g_2^{-1}(u) \right| \\ &= \left(\frac{1}{6}\right) \frac{1}{2\sqrt{9-u}} + \left(\frac{1}{6}\right) \frac{1}{2\sqrt{9-u}} \\ &= \frac{1}{6\sqrt{9-u}}, \text{ para } u \in (0,9). \end{aligned}$$

De esta manera,

$$f_U(u) = \frac{1}{6\sqrt{9-u}} I_{(0,9)}(u).$$

Ejemplo 6.22. Sea Z una variable aleatoria con distribución normal estándar. Encontrar la distribución de $U = Z^2$.

La función de densidad de la variable Z es

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}.$$

Se puede apreciar en la gráfica que existen dos inversas

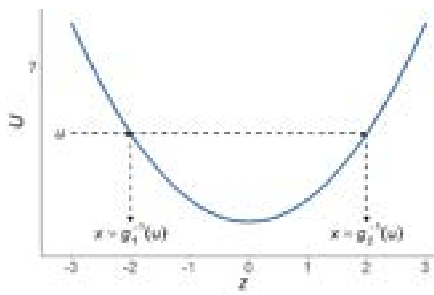


Figura 6.28. *Transformación del ejemplo 6.22*

El recorrido de la nueva variable es el intervalo $(0, \infty)$.
Si $u = z^2$, entonces, $z = \pm\sqrt{u}$. Esto es, $g_1^{-1}(u) = -\sqrt{u}$ y $g_2^{-1}(u) = \sqrt{u}$.
Por lo tanto, las derivadas de las inversas

$$\begin{aligned} \frac{d}{du} g_1^{-1}(u) &= -\frac{1}{2} u^{-\frac{1}{2}}, \\ \frac{d}{du} g_2^{-1}(u) &= \frac{1}{2} u^{-\frac{1}{2}}. \end{aligned}$$

Entonces, la función de densidad de U

$$\begin{aligned} f_U(u) &= f_Z(g_1^{-1}(u)) \left| \frac{d}{du} g_1^{-1}(u) \right| + f_Z(g_2^{-1}(u)) \left| \frac{d}{du} g_2^{-1}(u) \right| \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{u}{2}} \frac{1}{2} u^{-\frac{1}{2}} + \frac{1}{\sqrt{2\pi}} e^{-\frac{u}{2}} \frac{1}{2} u^{-\frac{1}{2}} \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{u}{2}} u^{-\frac{1}{2}}, \text{ donde } u \in (0, \infty). \end{aligned}$$

Esta función se puede expresar de la siguiente manera:

$$\begin{aligned} f_U(u) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{u}{2}} u^{-\frac{1}{2}} I_{(0,\infty)}(u) \\ &= \frac{\left(\frac{1}{2}\right)^{\frac{1}{2}}}{\Gamma\left(\frac{1}{2}\right)} e^{-\frac{u}{2}} u^{-\frac{1}{2}} I_{(0,\infty)}(u). \end{aligned}$$

Se puede observar que la anterior función de densidad corresponde a una distribución *Gamma* $\left(r = \frac{1}{2}, \lambda = \frac{1}{2}\right)$. Dicho de otra manera, es la función de densidad de una distribución ji cuadrada con un grado de libertad.

De esta manera, concluimos que $U = Z^2 \sim \chi^2(1)$.

Este último resultado es útil para generar valores de una variable aleatoria con distribución ji cuadrada con un grado de libertad, a partir de una variable con distribución normal estándar. Siendo más puntuales, por medio de R, se genera un valor de una variable aleatoria $N(0,1)$, puede ser por el método de Box-Müller, véase ejemplo 6.26, una vez generado este valor, se eleva al cuadrado, el resultado será un valor de una variable aleatoria con distribución ji cuadrada con un grado de libertad. Este resultado, además de ser útil en simulación de Monte Carlo, también contribuye a problemas de inferencia estadística.

Caso 3. Transformaciones biyectivas de $\mathbb{R}^n$ a $\mathbb{R}^n$.

Sea $\underline{X} = (X_1, X_2, \ldots, X_n)$ un vector aleatorio con función de densidad $f_{\underline{X}}(x_1, \ldots, x_n)$ la cual es conocida, consideremos la siguiente función biyectiva de $\underline{X}$, $g(X_1, X_2, \ldots, X_n) = (g_1(X_1, \ldots, X_n), \ldots, g_n(X_1, \ldots, X_n))$ $= (U_1, U_2, \ldots, U_n) = \underline{U}$. Entonces, la función de densidad de $\underline{U}$ se obtiene de la siguiente manera:

$$\begin{aligned} f_{\underline{U}}(u_1, \ldots, u_n) &= f_{\underline{X}}(k(u_1, \ldots, u_n))|J| \\ &= f_{\underline{X}}(k_1(u_1, \ldots, u_n), k_2(u_1, \ldots, u_n), \ldots, k_n(u_1, \ldots, u_n))|J|, \end{aligned}$$

donde $k(u_1, \ldots, u_n) = (k_1(u_1, \ldots, u_n), k_2(u_1, \ldots, u_n), \ldots, k_n(u_1, \ldots, u_n))$ es la función inversa de la función $g(x_1, \ldots, x_n)$ y $|J|$ es el valor absoluto del jacobiano el cual se calcula de la siguiente manera:

$$J = \begin{vmatrix} \dfrac{\partial k_1}{\partial u_1} & \dfrac{\partial k_1}{\partial u_2} & \cdots & \dfrac{\partial k_1}{\partial u_n} \\ \dfrac{\partial k_2}{\partial u_1} & \dfrac{\partial k_2}{\partial u_2} & \cdots & \dfrac{\partial k_2}{\partial u_n} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial k_n}{\partial u_1} & \dfrac{\partial k_n}{\partial u_2} & \cdots & \dfrac{\partial k_n}{\partial u_n} \end{vmatrix}.$$

Como se observa, los problemas que se pueden resolver con esta técnica son para transformaciones que van de $\mathbb{R}^n$ a $\mathbb{R}^n$, esto es, las dimensiones del dominio y contradominio deben de coincidir.

Pero también se pueden considerar problemas donde la transformación va de $\mathbb{R}^2$ a $\mathbb{R}$. En este caso, se sugiere construir una nueva transformación que va de $\mathbb{R}^2$ a $\mathbb{R}^2$, considerando como primer componente, la transformación de la cual existe interés en encontrar su función de densidad, y como segundo componente, una transformación arbitraria de $\mathbb{R}^2$ a $\mathbb{R}$. La idea es encontrar la función de densidad de la nueva transformación, la que va de $\mathbb{R}^2$ a $\mathbb{R}^2$, una vez lograda, a partir de esta función de densidad conjunta, se obtendrá la función marginal del componente que es la transformación de la cual existe interés en encontrar su función de densidad, véase el siguiente ejemplo.

Cabe mencionar, que esta idea de construir una nueva transformación que va de $\mathbb{R}^2$ a $\mathbb{R}^2$, ya se había sugerido en problemas discretos, donde también las dimensiones del dominio y contradominio, en primera instancia, no coincidían, se puede revisar los ejemplos discretos 6.13 y 6.14.

Ejemplo 6.23. Sea $\underline{X} = (X_1, X_2)$ un vector aleatorio con función de densidad $f_{\underline{X}}(x_1, x_2) = I_{[x_1, x_1+1]}(x_2) I_{[0,1]}(x_1)$. Sea $U = X_2 - 2X_1$, una función de X_1 y X_2. Encontrar la función de densidad de la variable aleatoria U.

Se puede observar que la transformación va de $\mathbb{R}^2$ a $\mathbb{R}$. Definiremos una nueva transformación que va de $\mathbb{R}^2$ a $\mathbb{R}^2$, de la siguiente manera:

$$\begin{aligned} g(X_1, X_2) &= (g_1(X_1, X_2), g_2(X_1, X_2)) \\ &= (X_2 - 2X_1, X_1) \\ &= (U_1, U_2) = \underline{U}. \end{aligned}$$

El primer componente de esta nueva transformación corresponde a la función en la que hay interés en encontrar su función de densidad. El segundo componente no importa como se define, no hay interés en encontrar su distribución, y por lo mismo se sugiere que se defina de la manera simple.

El recorrido gráficamente del vector $\underline{X} = (X_1, X_2)$ es

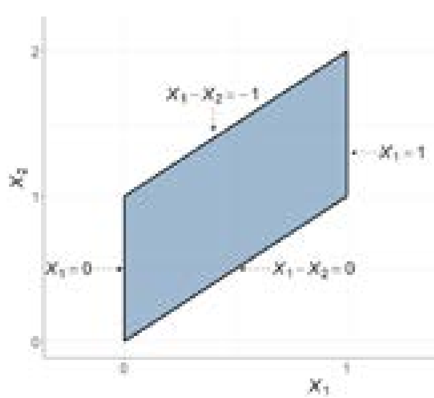


Figura 6.29. *Recorrido conjunto de X_1 y X_2, ejemplo 6.23*

Obsérvese que el recorrido del vector $\underline{X}$ está limitado por las rectas: $x_1 = 0$, $x_1 = 1$, $x_2 = x_1$ y $x_2 = x_1 + 1$. Estas rectas son importantes para encontrar el recorrido del nuevo vector $\underline{U}$. La idea es verificar cómo cada una de estas rectas cambia bajo la transformación y, de esta manera, tendremos las limitaciones del recorrido del vector $\underline{U}$.

A continuación, se encontrarán las limitaciones del recorrido del nuevo vector.

Si $x_1 = 0$, entonces, $u_1 = x_2$ y $u_2 = 0$. De aquí se deduce que la recta $x_1 = 0$ se transforma en la recta $u_2 = 0$.

Ahora, si $x_1 = 1$, entonces, $u_1 = x_2 - 2$ y $u_2 = 1$. Se concluye que la recta $x_1 = 1$, se transforma en la recta $u_2 = 1$.

En el caso, $x_1 = x_2$, implica $u_1 = -x_1$ y $u_2 = x_1$, esto es, $u_1 = -u_2$. Esto es, la recta $x_1 = x_2$ se transforma en $u_1 = -u_2$.

Si $x_2 = x_1 + 1$, entonces, $u_1 = 1 - x_1$ y $u_2 = x_1$, esto es, $u_1 = 1 - u_2$. La recta $x_2 = x_1 + 1$ se transforma en la recta $u_1 = 1 - u_2$.

Entonces, la nueva región queda determinada de la siguiente manera:

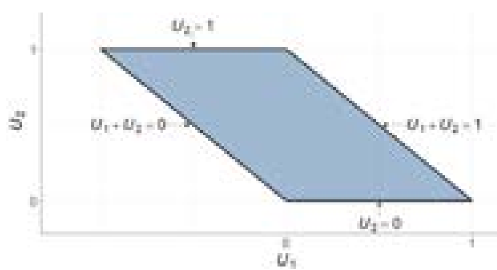


Figura 6.30. *Recorrido conjunto de U_1 y U_2, ejemplo 6.23*

Considerando los componentes de la transformación $g(x_1, x_2) = (u_1, u_2)$, se tiene el siguiente sistema de ecuaciones:

$$u_1 = x_2 - 2x_1,$$

$$u_2 = x_1.$$

Los componentes de la inversa se obtienen resolviendo el anterior sistema de ecuaciones, esto es,

$$x_2 = u_1 + 2u_2 = k_2(u_1, u_2),$$

$$x_1 = u_2 = k_1(u_1, u_2).$$

Se calculará el jacobiano

$$J = \begin{vmatrix} 0 & 1 \\ 1 & 2 \end{vmatrix} = -1.$$

Entonces,

$$\begin{aligned} f_{\underline{U}}(u_1, u_2) &= f_{\underline{X}}(k_1(u_1, u_2), k_2(u_1, u_2))|J| \\ &= f_{\underline{X}}(u_2, u_1 + 2u_2)|-1| = 1, \end{aligned}$$

donde el recorrido es indicado en la última gráfica.

Por lo tanto, la función de densidad de $\underline{U}$ con funciones indicadoras es

$$\begin{aligned} f_{\underline{U}}(u_1, u_2) &= I_{[-u_2, 1-u_2]}(u_1) I_{[0,1]}(u_2) \\ &= I_{[-u_1, 1]}(u_2) I_{[-1,0]}(u_1) + I_{[0, 1-u_1]}(u_2) I_{(0,1]}(u_1). \end{aligned}$$

El siguiente paso es encontrar la función de densidad marginal de U_1, esta se calculará por casos.

Caso 1. Cuando $-1 \leq u_1 \leq 0$.

$$\begin{aligned} f_{U_1}(u_1) &= \int_{-u_1}^{1} du_2 \\ &= 1 + u_1. \end{aligned}$$

Caso 2. Cuando $0 < u_1 \leq 1$.

$$\begin{aligned} f_{U_1}(u_1) &= \int_{0}^{1-u_1} du_2 \\ &= 1 - u_1. \end{aligned}$$

De esta manera,

$$f_{U_1}(u_1) = (1 + u_1) I_{[-1,0]}(u_1) + (1 - u_1) I_{(0,1]}(u_1).$$

Ejemplo 6.24. Sea $\underline{X} = (X_1, X_2)$ un vector aleatorio el cual se distribuye uniformemente sobre el cuadro unitario. Encontrar la función de densidad de $U = X_1 + X_2$ por medio de la técnica cambio de variable.

Este ejercicio ya se había resuelto por el método de función de distribución, véase el ejemplo 6.8, ahora se resolverá por la técnica de cambio de variable.

La transformación dada va de $\mathbb{R}^2$ a $\mathbb{R}$. Para igualar las dimensiones del dominio y contradominio, se definirá una nueva función de $\mathbb{R}^2$ a $\mathbb{R}^2$, de la siguiente manera:

$$\begin{aligned} g(X_1, X_2) &= (g_1(X_1, X_2), g_2(X_1, X_2)) \\ &= (X_1 + X_2, X_2) = (U_1, U_2) = \underline{U}. \end{aligned}$$

El recorrido conjunto de X_1 y X_2 es el cuadro unitario, el cual está limitado por las rectas: $x_1 = 0$, $x_1 = 1$, $x_2 = 0$ y $x_2 = 1$, este se puede apreciar en la siguiente gráfica.

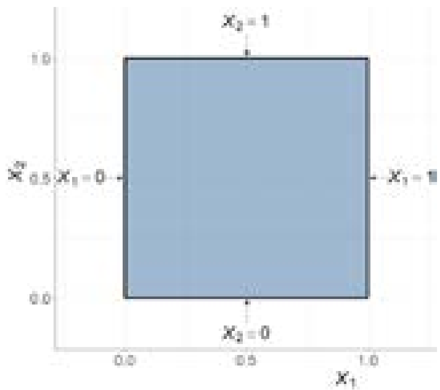


Figura 6.31. *Recorrido conjunto de X_1 y X_2, ejemplo 6.24*

La función de densidad conjunta de X_1 y X_2 es

$$f_{\underline{X}}(x_1, x_2) = I_{(0,1)}(x_1)I_{(0,1)}(x_2).$$

Obsérvese, si $x_1 = 0$, entonces, $u_1 = x_2$ y $u_2 = x_2$, por lo tanto, $u_1 = u_2$. En el caso de $x_1 = 1$, implica que $u_1 = 1 + x_2$ y $u_2 = x_2$, de esta manera, se cumple $u_2 = u_1 - 1$. Ahora, en el caso de $x_2 = 0$, se puede observar que $u_1 = x_1$ y $u_2 = 0$. Y si $x_2 = 1$, entonces, $u_1 = x_1 + 1$ y $u_2 = 1$. De esta manera, se obtienen cuatro líneas rectas que limitan el recorrido de $\underline{U}$. En conclusión, el recorrido de $\underline{U}$ queda limitado por las rectas: $u_1 = u_2$, $u_2 = u_1 - 1$, $u_2 = 0$ y $u_2 = 1$, véase la siguiente gráfica.

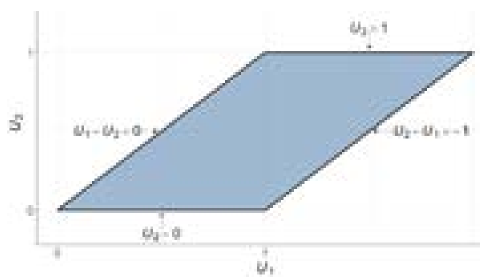


Figura 6.32. *Recorrido conjunto de U_1 y U_2, ejemplo 6.24*

Los componentes de la inversa de la transformación son

$$k_1(u_1, u_2) = u_1 - u_2,$$

$$k_2(u_1, u_2) = u_2.$$

El jacobiano es

$$J = \begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} = 1.$$

Entonces,

$$f_{\underline{U}}(u_1, u_2) = f_{\underline{X}}(k_1(u_1, u_2), k_2(u_1, u_2))|J| = 1,$$

donde el recorrido conjunto de U_1 y U_2 se muestra en la anterior gráfica.

De este modo,

$$f_{\underline{U}}(u_1, u_2) = I_{(0,u_1)}(u_2) I_{(0,1]}(u_1) + I_{(u_1-1,1)}(u_2) I_{(1,2)}(u_1).$$

Se obtendrá la función de densidad de U_1.

Caso 1. Cuando $0 < u_1 \leq 1$.

$$\begin{aligned} f_{U_1}(u_1) &= \int_0^{u_1} du_2 \\ &= u_1. \end{aligned}$$

Caso 2. Cuando $1 < u_1 < 2$.

$$\begin{aligned} f_{u_1}(u_1) &= \int_{u_1-1}^{1} du_2 \\ &= 2 - u_1. \end{aligned}$$

Por lo tanto,

$$f_{U_1}(u_1) = u_1 I_{(0,1]}(u_1) + (2 - u_1) I_{(1,2)}(u_1).$$

Ejemplo 6.25. Sean X_1 y X_2 dos variables aleatorias independientes tal que X_1 tiene distribución gamma con parámetros α y 1, y X_2 tiene distribución gamma con parámetros β y 1. Sean $U_1 = X_1 + X_2$ y $U_2 = \frac{X_1}{X_1+X_2}$.
a) ¿Son U_1 y U_2 variables aleatorias independientes?
b) Encontrar la distribución de U_1.
c) Encontrar la distribución de U_2.

En este caso, la transformación va de $\mathbb{R}^2$ a $\mathbb{R}^2$, esto es,

$$\begin{aligned} g(X_1, X_2) &= (g_1(X_1, X_2), g_2(X_1, X_2)) \\ &= \left(X_1 + X_2, \tfrac{X_1}{X_1+X_2}\right) \\ &= (U_1, U_2) = \underline{U}. \end{aligned}$$

La función de densidad conjunta de X_1 y X_2 es

$$\begin{aligned} f_{\underline{X}}(x_1, x_2) &= f_{X_1}(x_1) f_{X_2}(x_2) \\ &= \left[\frac{1}{\Gamma(\alpha)} x_1^{\alpha-1} e^{-x_1} I_{(0,\infty)}(x_1)\right] \left[\frac{1}{\Gamma(\beta)} x_2^{\beta-1} e^{-x_2} I_{(0,\infty)}(x_2)\right] \\ &= \frac{1}{\Gamma(\alpha)} \frac{1}{\Gamma(\beta)} x_1^{\alpha-1} x_2^{\beta-1} e^{-(x_1+x_2)} I_{(0,\infty)}(x_1) I_{(0,\infty)}(x_2). \end{aligned}$$

El recorrido conjunto de X_1 y X_2 es el primer cuadrante, esto es,

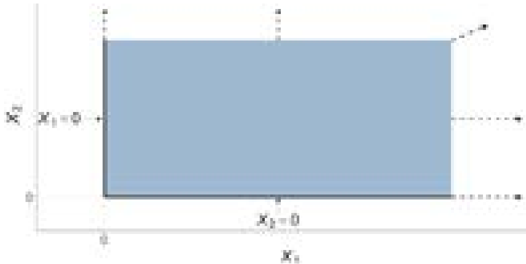


Figura 6.33. *Recorrido conjunto de X_1 y X_2, ejemplo 6.25*

Este recorrido está limitado por las rectas $x_1 = 0$ y $x_2 = 0$.

Si $x_1 = 0$, entonces, $u_1 = x_2$ y $u_2 = 0$. Si $x_2 = 0$, entonces, $u_1 = x_1$ y $u_2 = 1$. El nuevo recorrido está limitado por las rectas: $u_2 = 0$, $u_2 = 1$, y además por $u_1 = 0$, ya que $u_1 = x_1 + x_2 > 0$, véase siguiente gráfica.

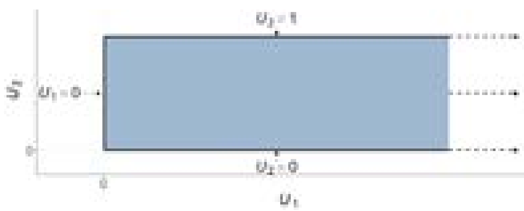


Figura 6.34. *Recorrido conjunto de U_1 y U_2, ejemplo 6.25*

La inversa de la transformación queda de la siguiente manera:

$$\begin{aligned} k(u_1, u_2) &= (k_1(u_1, u_2), k_2(u_1, u_2)) \\ &= (u_1 u_2, u_1 - u_1 u_2). \end{aligned}$$

El jacobiano es

$$\begin{aligned} J &= \begin{vmatrix} u_2 & u_1 \\ 1 - u_2 & -u_1 \end{vmatrix} \\ &= -u_1 u_2 - u_1(1 - u_2) = -u_1. \end{aligned}$$

Por lo tanto, $|J| = u_1$.

De esta manera, la función de densidad de $\underline{U}$ es

$$\begin{aligned} f_{\underline{U}}(u_1, u_2) &= \frac{1}{\Gamma(\alpha)} \frac{1}{\Gamma(\beta)} (u_1 u_2)^{\alpha-1} [u_1(1 - u_2)]^{\beta-1} e^{-u_1} (u_1) \\ &= \frac{1}{\Gamma(\alpha)} \frac{1}{\Gamma(\beta)} u_1^{\alpha+\beta-1} u_2^{\alpha-1} (1 - u_2)^{\beta-1} e^{-u_1}, \end{aligned}$$

donde el recorrido del vector $\underline{U}$ se muestra en la última gráfica.

Entonces,

$$f_{\underline{U}}(u_1, u_2) = \frac{1}{\Gamma(\alpha)} \frac{1}{\Gamma(\beta)} u_1^{\alpha+\beta-1} u_2^{\alpha-1} (1-u_2)^{\beta-1} e^{-u_1} I_{(0,\infty)}(u_1) I_{(0,1)}(u_2).$$

a) Obsérvese que la función de densidad encontrada se puede expresar de la siguiente manera:

$$\left[\frac{1}{\Gamma(\alpha+\beta)} u_1^{\alpha+\beta-1} e^{-u_1} I_{(0,\infty)}(u_1)\right] \left[\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} u_2^{\alpha-1} (1-u_2)^{\beta-1} I_{(0,1)}(u_2)\right].$$

Por el teorema 5.8, se puede afirmar que U_1 y U_2 son variables aleatorias independientes.

b) Se sacará la marginal de U_1

$$\begin{aligned}
f_{U_1}(u_1) &= \int_0^1 \left[\frac{1}{\Gamma(\alpha+\beta)} u_1^{\alpha+\beta-1} e^{-u_1}\right] \left[\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} u_2^{\alpha-1} (1-u_2)^{\beta-1}\right] du_2 \\
&= \frac{1}{\Gamma(\alpha+\beta)} u_1^{\alpha+\beta-1} e^{-u_1} \int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} u_2^{\alpha-1} (1-u_2)^{\beta-1} du_2 \\
&= \frac{1}{\Gamma(\alpha+\beta)} u_1^{\alpha+\beta-1} e^{-u_1}.
\end{aligned}$$

Obsérvese que la última integral se realiza sobre todos los valores de una función de densidad de una distribución $Beta(\alpha, \beta)$ y, por lo tanto, vale 1.

Por lo tanto,

$$f_{U_1}(u_1) = \frac{1}{\Gamma(\alpha+\beta)} u_1^{\alpha+\beta-1} e^{-u_1} I_{(0,\infty)}(u_1).$$

Esto es, $U_1 \sim Gamma(\alpha+\beta, 1)$.

c) La función de densidad de U_2

$$\begin{aligned}
f_{U_2}(u_2) &= \int_0^\infty \left[\frac{1}{\Gamma(\alpha+\beta)} u_1^{\alpha+\beta-1} e^{-u_1}\right] \left[\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} u_2^{\alpha-1} (1-u_2)^{\beta-1}\right] du_1 \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} u_2^{\alpha-1} (1-u_2)^{\beta-1} \int_0^\infty \frac{1}{\Gamma(\alpha+\beta)} u_1^{\alpha+\beta-1} e^{-u_1} du_1 \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} u_2^{\alpha-1} (1-u_2)^{\beta-1}.
\end{aligned}$$

La última integral se realiza sobre todos los valores de una función de densidad con distribución $Gamma(\alpha+\beta, 1)$, por lo que vale 1.

Por lo tanto,

$$f_{U_2}(u_2) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} u_2^{\alpha-1} (1-u_2)^{\beta-1} I_{(0,1)}(u_2).$$

Esto es, $U_2 \sim Beta(\alpha, \beta)$.

Cabe señalar que el resultado del inciso b sirve para generar valores de una variable aleatoria con distribución gamma, como suma de otras dos variables aleatorias independientes, cada una con distribución gamma. Este resultado se va a generalizar aplicando el método de función generadora de momentos, véase la siguiente sección.

El resultado del inciso c se aplica para simular valores de una variable aleatoria con distribución beta, a partir de valores de dos variables aleatorias independientes con distribución gamma cada una.

Ejemplo 6.26. Sea X_1 y X_2 una muestra aleatoria de tamaño 2 de una distribución uniforme sobre el intervalo $(0,1)$. Se define las siguientes variables aleatorias

$$U_1 = (-2\ln(X_1))^{1/2}\cos(2\pi X_2) \quad \text{y} \quad U_2 = (-2\ln(X_1))^{1/2}\operatorname{sen}(2\pi X_2).$$

a) Encontrar la distribución conjunta de U_1 y U_2.
b) ¿Son U_1 y U_2 variables aleatorias independientes?
c) ¿Cómo se distribuye $U_i, i = 1, 2$?

a) En este caso, la transformación es

$$\begin{aligned} g(X_1, X_2) &= (g_1(X_1, X_2), g_2(X_1, X_2)) \\ &= (-2\ln(X_1))^{1/2}\cos(2\pi X_2), (-2\ln(X_1))^{1/2}\operatorname{sen}(2\pi X_2)) \\ &= (U_1, U_2) = \underline{U}. \end{aligned}$$

La función de densidad conjunta de X_1 y X_2 es

$$f_{\underline{X}}(x_1, x_2) = I_{(0,1)}(x_1)I_{(0,1)}(x_2).$$

El recorrido conjunto de X_1 y X_2 es el cuadro unitario,

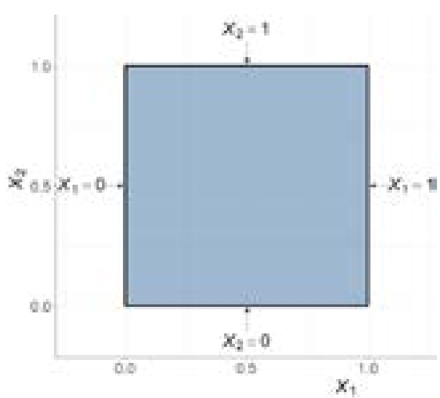


Figura 6.35. *Recorrido conjunto de X_1 y X_2, ejemplo 6.26*

En este caso, el procedimiento de considerar cada una de estas rectas que limitan el recorrido de X_1 y X_2, bajo la transformación, no nos lleva a una región concreta para U_1 y U_2, como en los ejemplos anteriores. Pero, analizando con detalle conjuntamente los componentes de la transformación $g(X_1, X_2)$, es posible observar que el recorrido conjunto de U_1 y U_2 es todo $\mathbb{R}^2$.

El siguiente paso es encontrar la función inversa de $g(X_1, X_2)$. Los componentes de esta transformación son

$$u_1 = (-2\ln(x_1))^{1/2}\cos(2\pi x_2) \quad \text{y} \quad u_2 = (-2\ln(x_1))^{1/2}\operatorname{sen}(2\pi x_2).$$

Elevando cada componente al cuadrado queda

$$\begin{aligned} u_1^2 &= -2\ln(x_1)\cos^2(2\pi x_2), \\ u_2^2 &= -2\ln(x_1)\operatorname{sen}^2(2\pi x_2). \end{aligned}$$

De esta manera,

$$u_1^2 + u_2^2 = -2\ln(x_1)\cos^2(2\pi x_2) - 2\ln(x_1)\operatorname{sen}^2(2\pi x_2) = -2\ln(x_1).$$

Por lo tanto,

$$\begin{aligned} x_1 &= e^{-\frac{u_1^2+u_2^2}{2}} \\ &= k_1(u_1, u_2). \end{aligned}$$

Por otro lado,

$$\frac{u_2}{u_1} = \tan(2\pi x_2).$$

Por lo tanto,

$$\begin{aligned} x_2 &= \frac{1}{2\pi}\tan^{-1}\left(\tfrac{u_2}{u_1}\right) \\ &= k_2(u_1, u_2). \end{aligned}$$

Ahora se calcularán los componentes del jacobiano

$$\frac{\partial k_1(u_1, u_2)}{\partial u_1} = e^{-\frac{u_1^2+u_2^2}{2}}(-u_1),$$

$$\frac{\partial k_1(u_1, u_2)}{\partial u_2} = e^{-\frac{u_1^2+u_2^2}{2}}(-u_2),$$

$$\begin{aligned} \frac{\partial k_2(u_1, u_2)}{\partial u_1} &= \frac{1}{2\pi}\left[\frac{1}{1+\left(\dfrac{u_2}{u_1}\right)^2}\left(-\frac{u_2}{u_1^2}\right)\right] \\ &= \frac{1}{2\pi}\left[\frac{-u_2}{u_1^2+u_2^2}\right], \end{aligned}$$

$$
\begin{aligned}
\frac{\partial k_2(u_1,u_2)}{\partial u_2} &= \frac{1}{2\pi}\left[\frac{1}{1+\left(\dfrac{u_2}{u_1}\right)^2}\left(\frac{1}{u_1}\right)\right] \\
&= \frac{1}{2\pi}\left[\frac{u_1}{u_1^2+u_2^2}\right].
\end{aligned}
$$

Entonces,

$$
\begin{aligned}
J &= \begin{vmatrix} e^{-\frac{u_1^2+u_2^2}{2}}(-u_1) & e^{-\frac{u_1^2+u_2^2}{2}}(-u_2) \\ \dfrac{1}{2\pi}\left[\dfrac{-u_2}{u_1^2+u_2^2}\right] & \dfrac{1}{2\pi}\left[\dfrac{u_1}{u_1^2+u_2^2}\right] \end{vmatrix} \\
&= -\frac{1}{2\pi}e^{-\frac{u_1^2+u_2^2}{2}}\left[\frac{u_1^2}{u_1^2+u_2^2}\right] - \frac{1}{2\pi}e^{-\frac{u_1^2+u_2^2}{2}}\left[\frac{u_2^2}{u_1^2+u_2^2}\right] \\
&= -\frac{1}{2\pi}e^{-\frac{u_1^2+u_2^2}{2}}\left[\frac{1}{u_1^2+u_2^2}\right](u_1^2+u_2^2) \\
&= -\frac{1}{2\pi}e^{-\frac{u_1^2+u_2^2}{2}}.
\end{aligned}
$$

Por lo tanto,

$$
|J| = \frac{1}{2\pi}e^{-\frac{u_1^2+u_2^2}{2}}.
$$

De esta manera,

$$
\begin{aligned}
f_{\underline{U}}(u_1,u_2) &= f_{\underline{X}}(k_1(u_1,u_2),k_2(u_1,u_2))|J| \\
&= (1)\frac{1}{2\pi}e^{-\frac{u_1^2+u_2^2}{2}},
\end{aligned}
$$

donde el recorrido del vector $\underline{U}$ es $\mathbb{R}^2$.

Esto es,

$$
f_{\underline{U}}(u_1,u_2) = \frac{1}{2\pi}e^{-\frac{u_1^2+u_2^2}{2}}.
$$

Esta es la función de densidad de una distribución normal bivariada con parámetros $\mu_1=0$, $\mu_2=0$, $\sigma_1^2=1$, $\sigma_2^2=1$ y $\rho=0$.

b) Al distribuirse $\underline{U}$ normal bivariada con $\rho=0$, entonces, U_1 y U_2 son variables aleatorias independientes (teorema 5.17).

c) Por el teorema 5.16, se puede afirmar $U_i \sim N(0,1), i=1,2$.

El resultado anterior es una forma de simular valores de una variable aleatoria con distribución normal estándar, por medio de valores de una variable aleatoria uniforme sobre el intervalo (0,1). A este método para generar valores de una distribución normal estándar, se le conoce en simulación como el *método de Box-Müller*.

Siendo más puntuales, por medio de R se simula el valor de una variable aleatoria con distribución $U(0,1)$, usando la instrucción $runif(1,0,1)$, supongamos que x es el resultado de este cálculo, entonces, se efectúa la siguiente transformación: $(-2ln(x))^{1/2}cos(2\pi x)$ (o $(-2ln(x))^{1/2}sen(2\pi x)$), el resultado será el valor de una variable aleatoria con distribución $N(0,1)$.

En seguida se presentará la simulación de 20,000 valores de una variable aleatoria con distribución normal estándar.

Por medio de R, se simularon 2 veces, 10,000 valores de una variable aleatoria con distribución $U(0,1)$, los primeros 10,000 se guardaron en un vector llamado $X1$, el resto en otro vector $X2$. Posteriormente, se generaron los vectores $U1$ y $U2$, también de tamaño 10,000 cada uno, donde los componentes que ocupa el lugar i se generaron de la siguiente manera: $U1[i] = ((-2*log(X1[i])**(1/2))*cos(2*pi*X2[i])$ y $U2[i] = ((-2*log(X1[i])**(1/2))*sin(2*pi*X2[i])$. Se apilaron los dos anteriores vectores en uno solo, quedando un vector de 20,000 componentes. De acuerdo al problema anterior, los componentes de este último vector forman una muestra aleatoria de una distribución $N(0,1)$. A continuación, se presenta el histograma de estos valores simulados junto con la gráfica de la función de densidad de una distribución $N(0,1)$.

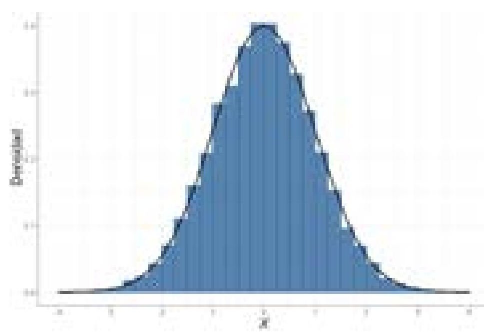


Figura 6.36. *Histograma y función de densidad*

El método de Box-Müller sirve para generar valores de una distribución normal estándar. Es muy fácil comprobar que si una variable $Z \sim N(0,1)$, entonces, la variable $X = \mu + \sigma Z \sim N(\mu, \sigma^2)$. Esto significa, que también se podrán generar fácilmente, valores de una variable aleatoria $N(\mu, \sigma^2)$.

Caso 4. Transformaciones de $\mathbb{R}^n$ a $\mathbb{R}^n$, con k inversas.

Sea $\underline{X} = (X_1, X_2, \ldots, X_n)$ un vector aleatorio continuo con función de densidad conocida $f_{\underline{X}}(x_1, \ldots, x_n)$. Consideremos $\underline{U} = g(X_1, X_2, \ldots, X_n)$, una transformación de $\mathbb{R}^n$ a $\mathbb{R}^n$ con k inversas en el recorrido de $\underline{X}$. Sea A el recorrido de $\underline{X}$ y B el recorrido de $\underline{U}$.

Si $g(\underline{X})$ tiene k inversas, significa que para cada elemento de B, este proviene de k elementos de A. Entonces, particionaremos el conjunto A en

k subconjuntos $A_1, A_2, \ldots, A_k$, de tal manera, que la transformación $g(\underline{X})$ restringida a A_i (g: $A_i \longrightarrow B$) es una una función biyectiva con imagen B, esto es, tiene solo una inversa, $i = 1, 2, \ldots, k$.

Supongamos que las k inversas de $g(\underline{x})$ son denotadas como $g_i^{-1}(\underline{u}) = (k_{i1}(u_1, \ldots, u_n), \ldots, k_{in}(u_1, \ldots, u_n))$, $i = 1, 2, \ldots, k$. Obsérvese que la inversa $g_i^{-1}(\underline{u})$ tiene como dominio B y como imagen A_i.

Entonces, la función de densidad conjunta de las variables $U_1, U_2, \ldots, U_n$ se puede obtener de la siguiente manera:

$$\begin{aligned} f_{\underline{U}}(u_1, u_2, \ldots, u_n) &= \sum_{i=1}^{k} f_{\underline{X}}(g_i^{-1}(\underline{u}))|J_i| \\ &= \sum_{i=1}^{k} f_{\underline{X}}(k_{i1}(u_1, \ldots, u_n), \ldots, k_{in}(u_1, \ldots, u_n))|J_i|, \text{ donde} \end{aligned}$$

$$J_i = \begin{vmatrix} \dfrac{\partial k_{i1}}{\partial u_1} & \dfrac{\partial k_{i1}}{\partial u_2} & \cdots & \dfrac{\partial k_{i1}}{\partial u_n} \\ \dfrac{\partial k_{i2}}{\partial u_1} & \dfrac{\partial k_{i2}}{\partial u_2} & \cdots & \dfrac{\partial k_{i2}}{\partial u_n} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial k_{in}}{\partial u_1} & \dfrac{\partial k_{in}}{\partial u_2} & \cdots & \dfrac{\partial k_{in}}{\partial u_n} \end{vmatrix}, \quad i = 1, 2, \ldots, k.$$

Por el momento no se presentará ejemplo alguno de la técnica de cambio de variable con una transformación $g(X_1, X_2, \ldots, X_n)$ de $\mathbb{R}^n$ a $\mathbb{R}^n$ con k inversas, más en el tema de estadísticas de orden, se podrá observar que la obtención de la función conjunta de n estadísticas de orden es un ejemplo donde se aplicará este resultado, véase teorema 6.14.

6.4. Método de función generadora de momentos

Como en los otros métodos ya tratados, se explicará en qué consiste el método de función generadora de momentos. Es importante resaltar que varios de los resultados obtenidos por este método tienen aplicaciones muy importantes, principalmente en temas relacionados con la estadística y simulación de Monte Carlo. A continuación, se explicará el método considerando primero transformaciones reales de una variable aleatoria, mas este método se puede extender a otro tipo de transformaciones de variables incluso funciones de vectores aleatorios.

Sea X una variable aleatoria con función generadora de momentos (f. g. m.) $m_X(t)$ conocida o que tiene una distribución conocida. Sea $U = g(X)$

una transformación de X, el objetivo de este método es, encontrar la f. g. m. de la variable U, esto es, $m_U(t)$, usando el conocimiento que se tiene de la distribución de la variable X.

Las siguientes observaciones nos indican algunos pros y contras de este método.

Observaciones

1. Este método se acaba de explicar para transformaciones $g\colon \mathbb{R} \longrightarrow \mathbb{R}$, pero el método se puede aplicar en forma eficiente y sin mucho grado de complejidad, para otras funciones con dimensiones diferentes, como $g\colon \mathbb{R}^n \longrightarrow \mathbb{R}^n$ o también para funciones $g\colon \mathbb{R}^n \longrightarrow \mathbb{R}$, lo cual es una ventaja de este método sobre los demás.
2. Este método no depende de la descripción gráfica de los recorridos, lo cual representa otra ventaja sobre los métodos anteriores.
3. El hecho de que la transformación no sea biyectiva, no representa complejidad adicional alguna, este método no hace distinción entre funciones biyectivas y las que no lo son.
4. Este método aportará resultados de mucha importancia para inferencia estadística, ya que las dimensiones del dominio y contradominio no afecta para encontrar la solución que se pretende.
5. No siempre es un método accesible, existen problemas donde es más sencillo aplicar los métodos previamente ya vistos. En particular, las variables con distribución uniforme o con función de densidad polinomial no son problemas sencillos para resolverse por medio de este método.

Obsérvese la demostración del teorema 4.19, es un primer ejemplo de la metodología de función generadora de momentos. Se recomienda al estudiante revisar paso a paso la demostración.

A continuación, se dará otro ejemplo.

Ejemplo 6.27. Sea Z una variable aleatoria con distribución normal estándar, encontrar la distribución de $U = Z^2$.

Obsérvese como se obtiene la función generadora de momentos de U

$$
\begin{aligned}
m_U(t) &= E[e^{Ut}] \\
&= E[e^{Z^2 t}] = \int_{-\infty}^{\infty} e^{z^2 t} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}\, dz \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2} z^2 (1-2t)} dz \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2(1-2t)^{-1}}}\, dz
\end{aligned}
$$

$$
\begin{aligned}
&= (1-2t)^{-\frac{1}{2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}(1-2t)^{-\frac{1}{2}}} e^{-\frac{z^2}{2(1-2t)^{-1}}} dz \\
&= (1-2t)^{-\frac{1}{2}}, \quad \text{si } (1-2t)^{-\frac{1}{2}} > 0.
\end{aligned}
$$

Esto es,

$$
m_U(t) = (1-2t)^{-\frac{1}{2}}, \quad \text{si} \quad t < \frac{1}{2}.
$$

Esta última es la función generadora de momentos de una distribución ji cuadrada con 1 grado de libertad, por lo tanto, se puede concluir que $U = Z^2 \sim \chi^2(1)$.

Obsérvese que la última integral se realiza sobre todos los valores de una función de densidad con distribución $N(0, (1-2t)^{-1})$.

Este último resultado ya se había resuelto por medio de la técnica cambio de variable (ver el ejemplo 6.22). También, se explicó cómo este resultado es útil para generar valores de una variable aleatoria ji cuadrada con 1 grado de libertad, dándose las indicaciones para hacer los cálculos por medio del *software* R.

Ejemplo 6.28. Sean $X_1, X_2, \ldots, X_n$ variables aleatorias independientes tal que $X_i \sim Poisson(\lambda_i)$, $i = 1, \ldots, n$. Encontrar la distribución de $U = X_1 + X_2 + \cdots + X_n$.

A continuación, se obtendrá la función generadora de momentos de U.

$$
\begin{aligned}
m_U(t) &= E[e^{Ut}] \\
&= E[e^{(X_1+X_2+\cdots+X_n)t}] \\
&= E[e^{X_1t}e^{X_2t}\cdots e^{X_nt}] \\
&= E[e^{X_1t}]E[e^{X_2t}]\cdots E[e^{X_nt}] \\
&= m_{X_1}(t)m_{X_2}(t)\cdots m_{X_n}(t) \\
&= e^{\lambda_1(e^t-1)}e^{\lambda_2(e^t-1)}\cdots e^{\lambda_n(e^t-1)} \\
&= e^{(\sum_{i=1}^n \lambda_i)(e^t-1)}.
\end{aligned}
$$

Obsérvese que esta última expresión es la función generadora de momentos de una distribución $Poisson\left(\sum_{i=1}^n \lambda_i\right)$. Por lo tanto, se puede concluir que $U = X_1 + X_2 + \cdots + X_n \sim Poisson\left(\sum_{i=1}^n \lambda_i\right)$.

Ejemplo 6.29. Sea X_1 y X_2 una muestra aleatoria de tamaño 2 de una distribución normal estándar.

a) ¿Cómo se distribuye $\underline{U} = (U_1, U_2) = (X_1 + X_2, X_1 - X_2)$?

b) ¿U_1 y U_2 son variables aleatorias independientes?

c) ¿Cómo se se distribuye U_i, $i = 1, 2$?

a) Se sacará la función generadora de momentos de $\underline{U}$,

$$\begin{aligned}
m_{\underline{U}}(t_1,t_2) &= E[e^{U_1t_1+U_2t_2}] \\
&= E[e^{(X_1+X_2)t_1+(X_1-X_2)t_2}] \\
&= E[e^{X_1(t_1+t_2)+X_2(t_1-t_2)}] \\
&= E[e^{X_1(t_1+t_2)}e^{X_2(t_1-t_2)}] \\
&= E[e^{X_1(t_1+t_2)}]E[e^{X_2(t_1-t_2)}] \\
&= m_{X_1}(t_1+t_2)m_{X_2}(t_1-t_2) \\
&= e^{\frac{(t_1+t_2)^2}{2}}e^{\frac{(t_1-t_2)^2}{2}} \\
&= e^{\frac{2t_1^2+2t_2^2}{2}}.
\end{aligned}$$

Esta es la función generadora de momentos de una distribución normal bivariada donde $\mu_1 = \mu_2 = 0$, $\sigma_1^2 = \sigma_2^2 = 2$ y $\rho = 0$. Por lo tanto, $\underline{U} = (U_1, U_2) \sim NB(0,0,2,2,0)$.

b) Como $\underline{U} = (U_1, U_2)$ tiene una distribución normal bivariada y $\rho = 0$, por el teorema 5.17, U_1 y U_2 son variables aleatorias independientes.

c) Por el teorema 5.16, $U_i \sim N(0,2), i = 1,2$.

Teorema 6.1. *Si $X_1, X_2, \ldots, X_n$ son variables aleatorias independientes con funciones generadoras de momentos $m_{X_1}(t), m_{X_2}(t), \ldots, m_{X_n}(t)$, respectivamente, entonces, la función generadora de momentos de la variable aleatoria $U = X_1 + X_2 + \cdots + X_n$ está dado por*

$$m_U(t) = \prod_{i=1}^{n} m_{X_i}(t).$$

Demostración. Obsérvese

$$\begin{aligned}
m_U(t) &= E[e^{ut}] \\
&= E[e^{(X_1+X_2+\cdots+X_n)t}] \\
&= E[e^{X_1t}e^{X_2t}\cdots e^{X_nt}] \\
&= E[e^{X_1t}]E[e^{X_2t}]\cdots E[e^{X_nt}] \\
&= m_{X_1}(t)m_{X_2}(t)\cdots m_{X_n}(t) \\
&= \prod_{i=1}^{n} m_{X_i}(t).
\end{aligned}$$

□

Teorema 6.2. *Si $X_1, X_2, \ldots, X_n$ son variables aleatorias independientes e idénticamente distribuidas con función generadora de momentos común $m_X(t)$, entonces, la función generadora de momentos de la variable $U = X_1 + X_2 + \cdots + X_n$ está dado por*

$$m_U(t) = [m_X(t)]^n.$$

Demostración. Aplicando el teorema anterior

$$\begin{aligned} m_U(t) &= \prod_{i=1}^{n} m_{X_i}(t) \\ &= \prod_{i=1}^{n} m_X(t) \\ &= [m_X(t)]^n. \end{aligned}$$ □

Los siguientes resultados son una aplicación de los dos teoremas anteriores, y en cada uno de ellos, la variable U se define como $U = \sum_{i=1}^{n} X_i$.

1. Sean $X_1, X_2, \ldots, X_n$ variables aleatorias independientes e idénticamente distribuidas como $Bernouilli(p)$, entonces, $U \sim B(n, p)$.
2. Sean $X_1, X_2, \ldots, X_n$ variables aleatorias independientes tal que $X_i \sim B(m_i, p), i = 1, \ldots, n$, entonces, $U \sim B\big(\sum_{i=1}^{n} m_i, p\big)$.
3. Sean $X_1, X_2, \ldots, X_n$ variables aleatorias independientes tal que $X_i \sim Poisson(\lambda_i), i = 1, \ldots, n$, entonces, $U \sim Poisson\big(\sum_{i=1}^{n} \lambda_i\big)$.
4. Sean $X_1, X_2, \ldots, X_n$ variables aleatorias independientes e idénticamente distribuidas, como $Geométrica(p)$, entonces, $U \sim BN(r = n, p)$.
5. Sean $X_1, X_2, \ldots, X_n$ variables aleatorias independientes tal que $X_i \sim BN(r_i, p), i = 1, \ldots, n$, entonces, $U \sim BN\big(\sum_{i=1}^{n} r_i, p\big)$.
6. Sean $X_1, X_2, \ldots, X_n$ variables aleatorias independientes e idénticamente distribuidas como $Exp(\lambda)$, entonces, $U \sim Gamma(r = n, \lambda)$.
7. Sean $X_1, X_2, \ldots, X_n$ variables aleatorias independientes tal que $X_i \sim Gamma(r_i, \lambda), i = 1, \ldots, n$, entonces, $U \sim Gamma\big(\sum_{i=1}^{n} r_i, \lambda\big)$.
8. Sean $X_1, X_2, \ldots, X_n$ variables aleatorias independientes tal que $X_i \sim \chi^2(\nu_i), i = 1, \ldots, n$, entonces, $U \sim \chi^2\big(\sum_{i=1}^{n} \nu_i\big)$.

Los anteriores resultados son útiles en materia de simulación, esto es, se pueden generar valores de una variable aleatoria, a partir de sumas de valores de otras variables aleatorias independientes.

Teorema 6.3. *Sean $X_1, X_2, \ldots, X_n$ variables aleatorias independientes e idénticamente distribuidas con función generadora de momentos común $m_X(t)$, entonces, la función generadora de momentos de la variable aleatoria $\bar{X} = \frac{X_1+X_2+\cdots+X_n}{n}$ está dada por*

$$m_{\bar{X}}(t) = \left[m_X\left(\tfrac{t}{n}\right)\right]^n.$$

Demostración. Obsérvese

$$\begin{aligned} m_{\bar{X}}(t) &= E[e^{\bar{x}t}] \\ &= E\left[e^{(\frac{X_1+X_2+\cdots+X_n}{n})t}\right] \\ &= E\left[e^{(X_1+X_2+\cdots+X_n)(\frac{t}{n})}\right]. \end{aligned}$$

La anterior es la función generadora de momentos de $X_1 + X_2 + \cdots + X_n$ valuada en $\frac{t}{n}$. Entonces, aplicando el teorema anterior

$$m_{\bar{X}}(t) = \left[m_X\left(\tfrac{t}{n}\right)\right]^n. \qquad \square$$

A la variable aleatoria $\bar{X}$ del teorema anterior se le conoce como la media muestral, véase definición 6.3.

Teorema 6.4. *Sean $X_1, X_2, \ldots, X_n$ variables aleatorias independientes y $Y_1 = g_1(X_1), Y_2 = g_2(X_2), \ldots, Y_n = g_n(X_n)$ son funciones de estas variables. Entonces, las variables $Y_1, Y_2, \ldots, Y_n$ también son independientes.*

Demostración. Obsérvese que la función generadora conjunta de $Y_1, Y_2, \ldots, Y_n$ se puede expresar de la siguiente manera:

$$\begin{aligned}
m_{\underline{Y}}(t_1, t_2, \ldots, t_n) &= E[e^{Y_1 t_1 + Y_2 t_2 + \cdots + Y_n t_n}] \\
&= E[e^{g_1(X_1)t_1 + g_2(X_2)t_2 + \cdots + g_n(X_n)t_n}] \\
&= E[e^{g_1(X_1)t_1} e^{g_2(X_2)t_2} \cdots e^{g_n(X_n)t_n}] \\
&= E[e^{g_1(X_1)t_1}]E[e^{g_2(X_2)t_2}] \cdots E[e^{g_n(X_n)t_n}] \text{ (por teorema 5.6)} \\
&= m_{Y_1}(t_1) m_{Y_2}(t_2) \cdots m_{Y_n}(t_n).
\end{aligned}$$

Por el teorema 5.7, las variables $Y_1, Y_2, \ldots, Y_n$ son independientes. $\square$

Existen resultados relacionados con la distribución normal, pero por la importancia que tienen en varias aplicaciones, principalmente en temas de inferencia estadística, estos se tratarán en la siguiente sección.

6.5. Aplicaciones a la estadística

En esta sección se tratarán resultados de distribuciones de variables aleatorias que contribuyen en temas de inferencia estadística. Estos resultados ayudarán a entender el vínculo que hay entre la probabilidad y la estadística.

Los primeros dos resultados tratan de la construcción de variables aleatorias, con distribuciones t de student y F. Estos resultados son importantes en la construcción de intervalos de confianza y pruebas de hipótesis, temas que son tratados en inferencia estadística.

Teorema 6.5. *Sea Z una variable aleatoria con distribución normal estándar y W una variable aleatoria con distribución ji cuadrada con ν grados de libertad, además Z y W son variables aleatorias independientes, entonces, la variable $\dfrac{Z}{\sqrt{W/\nu}}$ tiene distribución t de student con ν grados de libertad.*

Demostración. Sean Z y W dos variables aleatorias independientes, tales que $Z \sim N(0,1)$ y $W \sim \chi^2(\nu)$.

Por lo tanto, la función de densidad conjunta de las variables aleatorias Z y W es

$$\begin{aligned} f_{Z,W}(z,w) &= f_Z(z) f_W(w) \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \frac{1}{\Gamma\left(\frac{\nu}{2}\right) 2^{\frac{\nu}{2}}} w^{\frac{\nu}{2}-1} e^{-\frac{w}{2}} I_{(0,\infty)}(w). \end{aligned}$$

Definimos las siguientes variables aleatorias:

$$T = \frac{Z}{\sqrt{\frac{W}{\nu}}} \quad \text{y} \quad U = W,$$

las cuales están en función de Z y W.

Se puede observar que

$$-\infty < T < \infty \quad \text{y} \quad U > 0.$$

Obsérvese que la inversa de esta transformación es

$$Z = T\sqrt{\frac{W}{\nu}} = \tfrac{1}{\sqrt{\nu}} T U^{\frac{1}{2}} \quad \text{y} \quad W = U.$$

El jacobiano es

$$J = \begin{vmatrix} \frac{1}{\sqrt{\nu}} u^{\frac{1}{2}} & \frac{1}{\sqrt{\nu}} t \frac{1}{2} u^{-\frac{1}{2}} \\ 0 & 1 \end{vmatrix} = \frac{1}{\sqrt{\nu}} u^{\frac{1}{2}}.$$

De esta manera, se puede observar que el valor absoluto del jacobiano queda igual, esto es,

$$|J| = \left| \frac{1}{\sqrt{\nu}} u^{\frac{1}{2}} \right| = \frac{1}{\sqrt{\nu}} u^{\frac{1}{2}}.$$

La función de densidad conjunta de T y U, se calcula de la siguiente manera:

$$\begin{aligned} f_{T,U}(t,u) &= f_{Z,W}\left(t\sqrt{\frac{u}{\nu}}, u\right) |J| \\ &= \frac{1}{\sqrt{2\pi}} e^{-t^2\left(\frac{u}{\nu}\right)} \frac{1}{\Gamma\left(\frac{\nu}{2}\right) 2^{\frac{\nu}{2}}} u^{\frac{\nu}{2}-1} e^{-\frac{u}{2}} \frac{\sqrt{u}}{\sqrt{\nu}} \\ &= \frac{1}{\sqrt{2\pi\nu}\,\Gamma\left(\frac{\nu}{2}\right) 2^{\frac{\nu}{2}}} u^{\frac{\nu+1}{2}-1} e^{-\left(\frac{u}{2}\right)\left(1+\frac{t^2}{\nu}\right)}. \end{aligned}$$

De esta manera,

$$
\begin{aligned}
f_T(t) &= \int_0^{\infty} \frac{1}{\sqrt{2\pi\nu}\Gamma\left(\frac{\nu}{2}\right) 2^{\frac{\nu}{2}}} u^{\frac{\nu+1}{2}-1} e^{-(\frac{u}{2})(1+\frac{t^2}{\nu})} du \\
&= \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\Gamma\left(\frac{\nu}{2}\right) 2^{\left(\frac{\nu}{2}+\frac{1}{2}\right)} [(1+\frac{t^2}{\nu})(\frac{1}{2})]^{\frac{\nu+1}{2}}} \\
&\times \int_0^{\infty} \frac{[(1+\frac{t^2}{\nu})(\frac{1}{2})]^{\frac{\nu+1}{2}}}{\Gamma\left(\frac{\nu+1}{2}\right)} u^{\frac{\nu+1}{2}-1} e^{-u\left(1+\frac{t^2}{\nu}\right)\left(\frac{1}{2}\right)} du \\
&= \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\Gamma\left(\frac{\nu}{2}\right)\left(1+\frac{t^2}{\nu}\right)^{\frac{\nu+1}{2}}}.
\end{aligned}
$$

Obsérvese que la última integral se realiza sobre todos los valores de una función de densidad con distribución gamma con parámetros $r = \frac{\nu+1}{2}$ y $\lambda = (1+\frac{t^2}{\nu})(\frac{1}{2})$, por lo que el valor de esta integral es igual a uno.

Por lo tanto,

$$
f_T(t) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\,\Gamma\left(\frac{\nu}{2}\right)\left(1+\frac{t^2}{\nu}\right)^{\frac{\nu+1}{2}}}.
$$

La anterior es la función de densidad de una distribución $tStudent(\nu)$.

De esta manera, $T \sim tstudent(\nu)$. $\square$

Teorema 6.6. *Sean W_1 y W_2 dos variables aleatorias independientes con distribuciones ji cuadrada con ν_1 grados de libertad y ji cuadrada con ν_2 grados de libertad, respectivamente, entonces, $\dfrac{W_1/\nu_1}{W_2/\nu_2}$ se distribuye F con ν_1 y ν_2 grados de libertad en el numerador y denominador, respectivamente.*

Demostración. Sean $W_1 \sim \chi^2(\nu_1)$ y $W_2 \sim \chi^2(\nu_2)$, además, suponemos que W_1 y W_2 son variables independientes.

Por lo tanto,

$$
\begin{aligned}
f_{W_1,W_2}(w_1, w_2) &= f_{W_1}(w_1) f_{W_2}(w_2) \\
&= \frac{\left(\frac{1}{2}\right)^{\frac{\nu_1}{2}}\left(\frac{1}{2}\right)^{\frac{\nu_2}{2}}}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)} w_1^{\frac{\nu_1}{2}-1} w_2^{\frac{\nu_2}{2}-1} e^{-\frac{w_1+w_2}{2}} \\
&\quad \times I_{(0,\infty)}(w_1) I_{(0,\infty)}(w_2).
\end{aligned}
$$

Definimos las siguientes variables aleatorias

$$
F = \frac{\frac{W_1}{\nu_1}}{\frac{W_2}{\nu_2}} \quad \text{y} \quad U = W_2.
$$

Se pueden observar que los recorridos de estas variables son:

$$F > 0 \quad \text{y} \quad U > 0.$$

La inversa de la transformación es

$$W_1 = \left(\frac{\nu_1}{\nu_2}\right) UF \quad \text{y} \quad W_2 = U.$$

El jacobiano se calcula como

$$J = \begin{vmatrix} (\frac{\nu_1}{\nu_2})U & (\frac{\nu_1}{\nu_2})F \\ 0 & 1 \end{vmatrix} = \left(\frac{\nu_1}{\nu_2}\right) U.$$

Por lo tanto, la función de densidad conjunta de F y U es

$$\begin{aligned} f_{F,U}(f,u) &= f_{W_1,W_2}\left(\left(\frac{\nu_1}{\nu_2}\right) uf, u\right) |J| \\ &= \frac{1}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right) 2^{\frac{\nu_1+\nu_2}{2}}} \left[(\frac{\nu_1}{\nu_2})uf\right]^{\frac{\nu_1}{2}-1} u^{\frac{\nu_2}{2}-1} \\ &\times e^{-\frac{\left(\frac{\nu_1}{\nu_2}\right)uf+u}{2}} \left(\frac{\nu_1}{\nu_2}\right) u \\ &= \frac{(\frac{\nu_1}{\nu_2})^{\frac{\nu_1}{2}}}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right) 2^{\frac{\nu_1+\nu_2}{2}}} f^{\frac{\nu_1}{2}-1} u^{\frac{\nu_1+\nu_2}{2}-1} \\ &\quad \times e^{-\frac{u}{2}\left[\left(\frac{\nu_1}{\nu_2}\right)f+1\right]}, \quad \text{donde } f > 0 \quad \text{y} \quad u > 0. \end{aligned}$$

De esta manera, la función de densidad de F se calcula de la siguiente manera:

$$\begin{aligned} f_F(f) &= \int_0^\infty \frac{(\frac{\nu_1}{\nu_2})^{\frac{\nu_1}{2}}}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right) 2^{\frac{\nu_1+\nu_2}{2}}} f^{\frac{\nu_1}{2}-1} u^{\frac{\nu_1+\nu_2}{2}-1} e^{-\frac{u}{2}\left[\left(\frac{\nu_1}{\nu_2}\right)f+1\right]} du \\ &= \frac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)\left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}}}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)} \frac{f^{\frac{\nu_1}{2}-1}}{\left[\left(\frac{\nu_1}{\nu_2}\right) f + 1\right]^{\frac{\nu_1+\nu_2}{2}}} \\ &\times \int_0^\infty \frac{(\frac{1}{2})^{\frac{\nu_1+\nu_2}{2}} \left[\left(\frac{\nu_1}{\nu_2}\right) f + 1\right]^{\frac{\nu_1+\nu_2}{2}}}{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)} u^{\frac{\nu_1+\nu_2}{2}-1} e^{-\frac{u}{2}\left[\left(\frac{\nu_1}{\nu_2}\right)f+1\right]} du \end{aligned}$$

$$= \frac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)\left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}}}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)} \frac{f^{\frac{\nu_1}{2}-1}}{\left[\left(\frac{\nu_1}{\nu_2}\right)f+1\right]^{\frac{\nu_1+\nu_2}{2}}}$$

$$\times \int_0^\infty \frac{\left\{\left(\frac{1}{2}\right)\left[\left(\frac{\nu_1}{\nu_2}\right)f+1\right]\right\}^{\frac{\nu_1+\nu_2}{2}}}{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)} u^{\frac{\nu_1+\nu_2}{2}-1} e^{-\frac{1}{2}\left[\left(\frac{\nu_1}{\nu_2}\right)f+1\right]u} du$$

$$= \frac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)\left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}}}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)} \frac{f^{\frac{\nu_1}{2}-1}}{\left[1+\left(\frac{\nu_1}{\nu_2}\right)f\right]^{\frac{\nu_1+\nu_2}{2}}}, \quad \text{donde} \quad f > 0.$$

Por lo tanto,

$$f_F(f) = \frac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)\left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}}}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)} \frac{f^{\frac{\nu_1}{2}-1}}{\left[1+\left(\frac{\nu_1}{\nu_2}\right)f\right]^{\frac{\nu_1+\nu_2}{2}}} I_{(0,\infty)}(f).$$

La última integral se realizó sobre todos los valores de una función de densidad gamma con parámetros $r = \frac{\nu_1+\nu_2}{2}$ y $\lambda = \left(\frac{1}{2}\right)\left[1+\left(\frac{\nu_1}{\nu_2}\right)f\right]$, por lo que la integral vale uno. Por lo tanto, $F \sim F(\nu_1, \nu_2)$. □

Las siguientes definiciones y resultados están encaminados a aplicaciones en diferentes temas de estadística.

Definición 6.1. Sean $X_1, X_2, \ldots, X_n$ variables aleatorias independientes con una función de densidad común $f_{X_i}(x) = f_X(x)$, para $i = 1, 2, \ldots, n$, esto es, la función de densidad conjunta de las variables $X_1, X_2, \ldots, X_n$ es igual a $f_{\underline{X}}(x_1, x_2, \ldots, x_n) = f_X(x_1)f_X(x_2)\cdots f_X(x_n)$, entonces, se dice que $X_1, X_2, \ldots, X_n$ es una *muestra aleatoria* de una distribución con función de densidad $f_X(x)$.

Para que $X_1, X_2, \ldots, X_n$ sea una muestra aleatoria (m. a.) deberá cumplir dos propiedades, la primera es que las variables sean independientes y la segunda es que, además, tengan una distribución común. Esto es, $X_1, X_2, \ldots, X_n$ es una muestra aleatoria si y solo si las variables son independientes e idénticamente distribuidas (v. a. i. i. d.). Si alguna de las dos propiedades no se cumple, entonces, se puede considerar que es una muestra, pero no es aleatoria.

Definición 6.2. La distribución de una variable aleatoria que está en función de variables aleatorias que conforman una muestra es conocida como *distribución muestral*.

Definición 6.3. Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de alguna distribución, la *media muestral y la varianza muestral* se denotan, respectivamente, como $\bar{X}$ y S^2 y se definen de la siguiente manera:

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} \quad \text{y} \quad S^2 = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n-1}.$$

Observación. Algunos autores definen a la varianza muestral como

$$S^2 = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n}.$$

Cuando hagamos referencia a la varianza muestral, consideraremos la primera definición.

Los siguientes resultados tienen algo en común, las variables $X_1, X_2, \ldots, X_n$ tienen distribución normal. Estos resultados tienen una aportación importante en temas de inferencia estadística.

Teorema 6.7. *Sean $X_1, X_2, \ldots, X_n$ variables aleatorias independientes tal que X_i se distribuye normal con media μ_i y varianza σ_i^2, $i = 1, 2, \ldots, n$, y sean $k_1, k_2, \ldots, k_n$ constantes. Entonces,*

$$U = \sum_{i=1}^{n} k_i X_i \sim N\left(\sum_{i=1}^{n} k_i \mu_i, \sum_{i=1}^{n} k_i^2 \sigma_i^2 \right).$$

La demostración se deja como ejercicio para el estudiante.

Corolario 6.1. *Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una distribución normal con media μ y varianza σ^2, y $k_1, k_2, \ldots, k_n$ constantes. Entonces,*

$$U = \sum_{i=1}^{n} k_i X_i \sim N\left(\mu \sum_{i=1}^{n} k_i, \sigma^2 \sum_{i=1}^{n} k_i^2 \right).$$

Demostración. Por el teorema anterior

$$U = \sum_{i=1}^{n} k_i X_i \sim N\left(\sum_{i=1}^{n} k_i \mu, \sum_{i=1}^{n} k_i^2 \sigma^2 \right). \qquad \square$$

Observación. Supongamos que $X_1, X_2, \ldots, X_n$ es una muestra aleatoria de una distribución normal con media μ y varianza σ^2, aplicando el corolario anterior, en particular se cumple

1. $\bar{X} \sim N\left(\mu, \dfrac{\sigma^2}{n}\right)$.
2. $\sum_{i=1}^{n} X_i \sim N\left(n\mu, n\sigma^2\right)$.

Observación. Si $X_1, X_2, \ldots, X_n$ es una muestra aleatoria de una distribución normal con media μ y varianza σ^2, entonces,

$$\sqrt{n}\left(\frac{\bar{X}-\mu}{\sigma}\right) \sim N(0,1).$$

Este resultado se utiliza en estadística, en particular para la inferencia de un promedio, cuando la muestra aleatoria es obtenida de una población con distribución normal con media μ y varianza conocida σ^2.

Teorema 6.8. *Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una distribución normal con media μ y varianza σ^2. Entonces,*

$$U = \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2} \sim \chi^2(n).$$

Demostración. Se puede afirmar

$$\frac{X_i-\mu}{\sigma} \sim N(0,1), i = 1, 2, \ldots, n.$$

Entonces,

$$\left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \chi^2(1), i = 1, 2, \ldots, n.$$

Como $X_1, \ldots, X_n$ son variables independientes, entonces,

$$\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \chi^2(n).$$ □

Observación. Este último resultado es útil para hacer inferencia estadística para una varianza, cuando la muestra aleatoria es obtenida de una población la cual tiene una distribución normal con media μ conocida y varianza σ^2.

Teorema 6.9. *Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una distribución normal con media μ y varianza σ^2, entonces, $U = (n-1)\dfrac{S^2}{\sigma^2}$ tiene distribución ji cuadrada con $n-1$ grados de libertad, además, $\bar{X}$ y S^2 son variables aleatorias independientes.*

Demostración. Obsérvese que se cumple

$$\sum_{i=0}^{n}(x_i-\mu)^2 = \sum_{i=0}^{n}(x_i-\bar{x})^2 + n(\bar{x}-\mu)^2.$$

Si $X_1, \ldots, X_n$ es una muestra aleatoria de una distribución normal con media μ y varianza σ^2, entonces, la función de densidad conjunta de estas variables es igual a

$$\begin{aligned} f_{\underline{X}}(x_1, \ldots, x_n) &= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\dfrac{1}{2\sigma^2}\displaystyle\sum_{i=1}^{n}(x_i-\mu)^2} \\ &= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\dfrac{1}{2\sigma^2}\left[\displaystyle\sum_{i=1}^{n}(x_i-\bar{x})^2 + n(\bar{x}-\mu)^2\right]}. \end{aligned}$$

Por otro lado se sabe que

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right).$$

Por lo tanto, la función de densidad de $\underline{X}$ es

$$f_{\bar{X}}(\bar{x}) = \sqrt{n}\left(\frac{1}{\sqrt{2\pi}\sigma}\right) e^{-\dfrac{1}{2\sigma^2} n(\bar{x}-\mu)^2}.$$

Se va a considerar la siguiente transformación

$$Y_1 = \bar{X}, Y_2 = X_2, \ldots, Y_n = X_n.$$

Se puede observar que el recorrido conjunto de las variables $Y_1, Y_2, \ldots, Y_n$ es $\mathbb{R}^n$.

La inversa de esta transformación es

$$X_1 = nY_1 - Y_2 - \cdots - Y_n, X_2 = Y_2, \ldots, X_n = Y_n.$$

Se puede verificar fácilmente que el valor absoluto del jacobiano es igual $|J| = n$. Entonces, la densidad conjunta de $Y_1, Y_2, \ldots, Y_n$ es

$$\begin{aligned} &f_{\underline{Y}}(y_1, y_2, \ldots, y_n) \\ &\quad = f_{\underline{X}}(ny_1 - y_2 - \cdots - y_n, y_2, \ldots, y_n)n \\ &\quad = n\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \\ &\quad \times e^{-\dfrac{1}{2\sigma^2}[(ny_1 - y_2 - \cdots - y_n - y_1)^2 + \displaystyle\sum_{i=2}^{n}(y_i-y_1)^2 + n(y_1-\mu)^2]}. \end{aligned}$$

Por otro lado, se procederá a sacar la función de densidad condicional de $Y_2, \ldots, Y_n$ dado $Y_1 = y_1$.

$$
\begin{aligned}
&f_{Y_2,\ldots,Y_n|y_1}(y_2,\ldots,y_n|Y_1=y_1)\\
&=\frac{f_{\underline{Y}}(y_1,y_2,\ldots,y_n)}{f_{Y_1}(y_1)}\\
&=\frac{n\left[\dfrac{1}{\sqrt{2\pi}\sigma}\right]^n e^{-\frac{1}{2\sigma^2}\left[(ny_1-\cdots-y_n-y_1)^2+\sum_{i=2}^{n}(y_i-y_1)^2+n(y_1-\mu)^2\right]}}{\sqrt{n}\left[\dfrac{1}{\sqrt{2\pi}\sigma}\right]e^{-\frac{1}{2\sigma^2}n(\bar{x}-\mu)^2}}\\
&=\sqrt{n}\left[\frac{1}{\sqrt{2\pi}\sigma}\right]^{n-1} e^{-\frac{1}{2\sigma^2}\left[(ny_1-y_2-\cdots-y_n-y_1)^2+\sum_{i=2}^{n}(y_i-y_1)^2\right]}.
\end{aligned}
$$

La anterior función es una función de densidad conjunta para $Y_2,\ldots,Y_n$ condicionada a $Y_1=y_1$. El recorrido del vector $(Y_2,\ldots,Y_n)$ es $\mathbb{R}^{n-1}$. Si se integra esta función con respecto a cada una de las variables $Y_2,\ldots,Y_n$, sobre todos sus valores, el resultado es igual a uno, por ser una función de densidad. Lo anterior es válido para todo valor de $\sigma>0$.

Obsérvese que

$$
\begin{aligned}
(n-1)s^2&=\sum_{i=1}^{n}(x_i-\bar{x})^2\\
&=(ny_1-y_2-\cdots-y_n-y_1)^2+\sum_{i=2}^{n}(y_i-y_1)^2.
\end{aligned}
$$

Sea $R=(n-1)S^2$.

Por lo tanto,

$$
f_{Y_2,\ldots,Y_n|y_1}(y_2,\ldots,y_n|Y_1=y_1)=\sqrt{n}\left[\frac{1}{\sqrt{2\pi}\sigma}\right]^{n-1}e^{-\dfrac{1}{2\sigma^2}r}.
$$

Se procederá a obtener la función generadora de momentos condicional de $\dfrac{(n-1)S^2}{\sigma^2}=\dfrac{R}{\sigma^2}$ dado $Y_1=y_1$, esto es,

$$
\begin{aligned}
&m_{\frac{R}{\sigma^2}|Y_1}(t)\\
&\quad=E\left[e^{\left(\frac{R}{\sigma^2}\right)t}|Y_1=y_1\right]\\
&\quad=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}e^{\left(\frac{r}{\sigma^2}\right)t}\sqrt{n}\left[\frac{1}{\sqrt{2\pi}\sigma}\right]^{n-1}e^{-\left(\frac{1}{2\sigma^2}\right)r}dy_2\cdots dy_n\\
&\quad=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\sqrt{n}\left[\frac{1}{\sqrt{2\pi}\sigma}\right]^{n-1}e^{-\dfrac{r}{2\sigma^2(1-2t)^{-1}}}dy_2\cdots dy_n
\end{aligned}
$$

$$= (1-2t)^{-\frac{n-1}{2}}$$
$$\times \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\sqrt{n}\left[\frac{1}{\sqrt{2\pi}\sigma(1-2t)^{-\frac{1}{2}}}\right]^{n-1} e^{-\dfrac{r}{2\sigma^2(1-2t)^{-1}}}\, dy_2\cdots dy_n$$
$$= (1-2t)^{-\frac{n-1}{2}}, \quad \text{si } 1-2t>0.$$

Por lo tanto,

$$m_{\frac{R}{\sigma^2}|Y_1}(t) = (1-2t)^{-\frac{n-1}{2}}, \text{ si } t < \tfrac{1}{2}.$$

La última integral es sobre todos los valores de la función de densidad condicional $f_{Y_2,\ldots,Y_n|y_1}(y_2,\ldots,y_n|Y_1=y_1)$, con un nueva desviación estándar $(1-2t)^{-\frac{1}{2}}\sigma$, por lo que se debe cumplir que $t<\frac{1}{2}$.

Se puede observar que la función generadora de momentos condicional de $\dfrac{(n-1)S^2}{\sigma^2}$ dado $Y_1=y_1$ no depende de Y_1, esto es, no depende de $\bar{X}$. En conclusión, $\dfrac{(n-1)S^2}{\sigma^2}$ y $\bar{X}$ son variables aleatorias independientes (también lo son $\bar{X}$ y S^2). La función generadora de momentos obtenida es de una distribución ji cuadrada con $n-1$ grados de libertad, esta es la función generadora para $\dfrac{(n-1)S^2}{\sigma^2}$, ya que es independiente de $\bar{X}$, por lo tanto, $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$. □

En conclusión,

$$(n-1)\frac{S^2}{\sigma^2} = \frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{\sigma^2} \sim \chi^2(n-1),$$

además, $\bar{X}$ y S^2 son independientes.

Este último resultado es útil para construir intervalos de confianza y pruebas de hipótesis para una varianza o desviación estándar, cuando la muestra aleatoria proviene de una distribución normal y no es necesario conocer la media μ de la distribución.

Teorema 6.10. *Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una distribución normal con media μ y varianza σ^2, entonces,*

$$U = \sqrt{n}\left(\frac{\bar{X}-\mu}{S}\right) \sim tstudent(n-1).$$

Demostración. La muestra aleatoria $X_1, X_2, \ldots, X_n$ proviene de una distribución $N(\mu,\sigma^2)$, por lo tanto, $\bar{X}$ se distribuye $N(\mu,\sigma^2/n)$.

Entonces,

$$Z = \sqrt{n}\left(\frac{\bar{X}-\mu}{\sigma}\right) \sim N(0,1) \quad \text{y}$$

$$W = (n-1)\frac{S^2}{\sigma^2} \sim \chi^2(n-1),$$

además, las variables Z y W son independientes.

Por lo tanto,

$$\begin{aligned}\frac{Z}{\sqrt{\dfrac{W}{n-1}}} &= \frac{\sqrt{n}\left(\dfrac{\bar{X}-\mu}{\sigma}\right)}{\sqrt{\dfrac{(n-1)S^2}{\sigma^2(n-1)}}} \\ &= \sqrt{n}\left(\frac{\bar{X}-\mu}{S}\right) \sim tstudent(n-1).\end{aligned}$$

□

Este último resultado es útil para construir intervalos de confianza y pruebas de hipótesis para una media, cuando la muestra aleatoria proviene de una distribución normal y no es necesario conocer la varianza σ^2 de la distribución.

Teorema 6.11. *Sea $X_1, X_2, \ldots, X_{n_1}$ una muestra aleatoria de una distribución normal con media μ_1 y varianza σ^2, y $Y_1, Y_2, \ldots, Y_{n_2}$ otra muestra aleatoria, independiente de la anterior, de una distribución normal con media μ_2 y varianza σ^2, entonces,*

$$\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{S\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}} \sim tstudent(n_1+n_2-2),$$

donde $S = \sqrt{\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}$.

Demostración. Se sabe

$$\bar{X} \sim N\left(\mu_1, \frac{\sigma^2}{n_1}\right) \quad \text{y} \quad \bar{Y} \sim N\left(\mu_2, \frac{\sigma^2}{n_2}\right).$$

Por lo tanto,

$$\bar{X}-\bar{Y} \sim N\left(\mu_1-\mu_2, \frac{\sigma^2}{n_1}+\frac{\sigma^2}{n_2}\right).$$

De esta manera,

$$\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sigma\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}} \sim N(0,1).$$

Por otro lado, se sabe

$$(n_1 - 1)\frac{S_1^2}{\sigma^2} \sim \chi^2(n_1 - 1) \quad \text{y} \quad (n_2 - 1)\frac{S_2^2}{\sigma^2} \sim \chi^2(n_2 - 1).$$

Como las dos muestras son independientes entre sí, entonces, S_1^2 y S_2^2 también son variables aleatorias independientes, entonces,

$$(n_1 - 1)\frac{S_1^2}{\sigma^2} + (n_2 - 1)\frac{S_2^2}{\sigma^2} = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{\sigma^2} \sim \chi^2(n_1 + n_2 - 2).$$

Además,

$$\bar{X} - \bar{Y} \ \text{y} \ \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{\sigma^2} \quad \text{son variables independientes.}$$

Por lo tanto,

$$\frac{\dfrac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sigma\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}}{\sqrt{\dfrac{\dfrac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{\sigma^2}}{n_1 + n_2 - 2}}} \sim tstudent(n_1 + n_2 - 2).$$

De esta manera,

$$\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{S\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} \sim tstudent(n_1 + n_2 - 2). \qquad \square$$

Observación. El anterior resultado es útil para realizar inferencia estadística para la diferencia de dos medias, con la finalidad de compararlas, cuando se obtienen dos muestras aleatorias en forma independiente, cada una de distribución normal, con el supuesto de que las varianzas entre ambas poblaciones son iguales, se aclara que no es necesario conocer la varianza común de estas dos poblaciones.

Teorema 6.12. *Sea $X_1, X_2, \ldots, X_{n_1}$ una muestra aleatoria de una distribución normal con media μ_1 y varianza σ_1^2, y $Y_1, Y_2, \ldots, Y_{n_2}$ otra muestra aleatoria de una distribución normal con media μ_2 y varianza σ_2^2, independiente de la anterior, entonces,*

$$\frac{\dfrac{S_1^2\sigma_2^2}{S_2^2\sigma_1^2}}{} \sim F(n_1 - 1, n_2 - 1).$$

Demostración. Se sabe

$$W_1 = (n_1 - 1)\frac{S_1^2}{\sigma_1^2} \sim \chi^2(n_1 - 1)$$

Además,

$$W_2 = (n_2 - 1)\frac{S_2^2}{\sigma_2^2} \sim \chi^2(n_2 - 1).$$

Como las muestras son independientes entre sí, entonces S_1^2 y S_2^2 son también variables aleatorias independientes, por lo tanto,

$$\frac{\dfrac{W_1}{n_1 - 1}}{\dfrac{W_2}{n_2 - 1}} = \frac{S_1^2\sigma_2^2}{S_2^2\sigma_1^2} \sim F(n_1 - 1, n_2 - 1). \qquad \square$$

Observación. De manera análoga

$$\frac{S_2^2\sigma_1^2}{S_1^2\sigma_2^2} \sim F(n_2 - 1, n_1 - 1).$$

Observación. Los anteriores dos resultados son útiles para construir intervalos de confianza y pruebas de hipótesis para un cociente de dos varianzas, con la finalidad de comparar ambas varianzas, cuando se obtienen dos muestras aleatorias independientes entre sí, cada una obtenida de una población con distribución normal. Estos resultados también son usados en el tema de análisis de varianza, el cual tiene como objetivo comparar varios promedios a la vez.

6.6. Estadísticas de orden

Suponiendo que una muestra se ordena del menor al mayor elemento, en esta sección se van a tratar las distribuciones del mínimo y del máximo de la muestra ordenada. También, a partir de una muestra aleatoria ordenada, se presentará la distribución del elemento que ocupa el lugar k-ésimo, por otro lado, la distribución conjunta de dos elementos específicos y la distribución conjunta de la muestra ordenada. Cabe mencionar, que estas distribuciones tienen varias aplicaciones en temas de estadística, entre otros.

A continuación, se presentará una interpretación de la distribución del elemento mínimo de una muestra aleatoria.

Supongamos que se obtiene una muestra aleatoria de tamaño $n = 10$ de una distribución uniforme sobre el intervalo $(0, 1)$, sea $x_1, x_2, \ldots, x_{10}$ la muestra obtenida, entonces, la muestra se ordena del menor al mayor, sea $y_1, y_2, \ldots, y_{10}$ la muestra ordenada. De esta, se selecciona el elemento más pequeño, esto es, $y_1 = min(x_1, x_2, \ldots, x_{10})$. El anterior procedimiento lo repetimos 100,000 veces, entonces, tendremos 100,000 valores de $Y_1 = min(X_1, \ldots, X_{10})$.

Este procedimiento se simuló en R, después se construyó el histograma de los 1,000,000 valores de la variable X y el histograma de los 100,000 valores de la variable Y_1, véase siguientes gráficas.

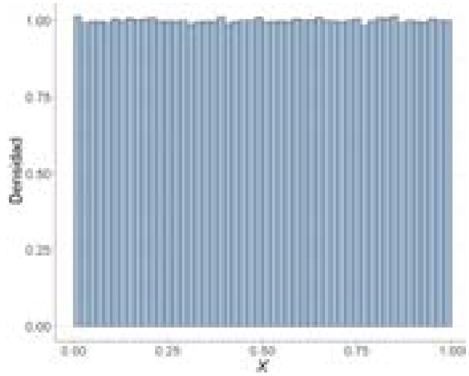


Figura 6.37. *Histograma de la variable X*

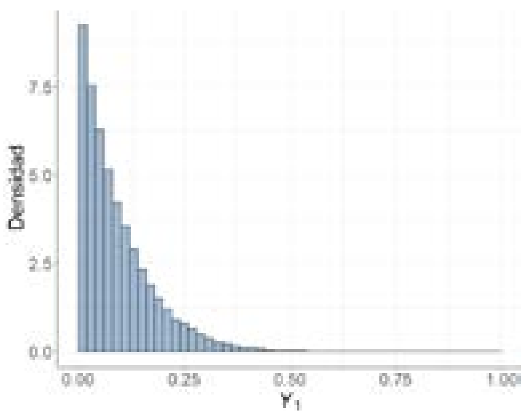


Figura 6.38. *Histograma de la variable Y_1*

Cabe señalar que la forma del histograma de los 100,000 valores de Y_1 (distribución de la variable Y_1), véase la gráfica 6.38, resultó ser muy diferente a la forma del histograma de los 1,000,000 valores de X (distribución de la variable X), obsérvese la gráfica 6.37, la cual es muy semejante a la función de densidad de la distribución uniforme sobre el intervalo $(0,1)$. Esta diferencia entre ambas distribuciones era de esperase, por lógica, los valores de Y_1 tienden a tomar valores más cerca de 0 (límite inferior del intervalo $(0,1)$).

Teorema 6.13. *Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una población con función de densidad $f_X(x)$, con función de distribución $F_X(x)$ y con recorrido el cual es denotado por A, entonces, la función de distribución de las variables $Y_1 = min(X_1, \ldots, X_n)$ y de $Y_n = max(X_1, \ldots, X_n)$ son, respectivamente,*

$$F_{Y_1}(y) = 1 - [1 - F_X(y)]^n, \quad \textit{donde } y \in A,$$

$$F_{Y_n}(y) = [F_X(y)]^n, \quad \textit{donde } y \in A.$$

Demostración. Primero la función de distribución de Y_1.

$$\begin{aligned}
F_{Y_1}(y) &= P[Y_1 \leq y] \\
&= 1 - P[Y_1 > y] \\
&= 1 - P[X_1 > y, \ldots, X_n > y] \\
&= 1 - P[X_1 > y] \cdots P[X_n > y] \\
&= 1 - \{(1 - P[X_1 \leq y]) \cdots (1 - P[X_n \leq y]\} \\
&= 1 - \{(1 - F_X(y)) \cdots (1 - F_X(y))\} \\
&= 1 - [1 - F_X(y)]^n, \text{ donde } y \in A.
\end{aligned}$$

En la anterior demostración, se consideró el siguiente razonamiento: el más pequeño de la muestra es mayor que y si y solo si todos los elementos de la muestra son mayores a y.

A continuación, se obtendrá la función de distribución de Y_n.

$$\begin{aligned}
F_{Y_n}(y) &= P[Y_n \leq y] \\
&= P[X_1 \leq y, \ldots, X_n \leq y] \\
&= P[X_1 \leq y] \cdots P[X_n \leq y] \\
&= F_{X_1}(y) \cdots F_{X_n}(y) \\
&= [F_X(y)]^n, \text{ donde } y \in A.
\end{aligned}$$ □

En el anterior procedimiento, se consideró el siguiente razonamiento: el más grande de la muestra es menor que y si y solo si todos los elementos de la muestra son menores a y.

También es importante aclarar que, en las dos demostraciones anteriores, la multiplicación de probabilidades se pudo realizar porque las variables $X_1, \ldots, X_n$ son independientes.

Observación. Si, además, las variables aleatorias son continuas, es posible obtener las funciones de densidad de Y_1 y Y_n, esto es,

$$f_{Y_1}(y) = n[1 - F_X(y)]^{n-1} f_X(y), \text{ donde } y \in A,$$
$$f_{Y_n}(y) = n[F_X(y)]^{n-1} f_X(y), \text{ donde } y \in A.$$

En caso de que las variables sean discretas, la función de densidad de Y_1 se podrá obtener de la siguiente manera: $f_{Y_1}(y) = F_{Y_1}(y)$ para el primer valor del recorrido de Y_1, para cualquier otro valor del recorrido, $f_{Y_1}(y) = F_{Y_1}(y) - F_{Y_1}(w)$, donde w es el anterior valor de y en el recorrido. De forma análoga, también se obtendrá la función de densidad de Y_n a partir de su función de distribución.

Ejemplo 6.30. La vida de cierto componente se distribuye exponencialmente con media de 100 horas. Si se instalan 10 componentes simultáneamente e independientemente, encontrar

a) La probabilidad de que todos los componentes estén descompuestos al término de 80 horas.
b) La vida promedio del primer componente que se descompone.

a) Para la primera probabilidad

$$\begin{aligned} P[Y_{10} \leq 80] &= F_{Y_{10}}(80) \\ &= [F_X(80)]^{10} \\ &= \left[1 - e^{-80/100}\right]^{10} = 0.002564. \end{aligned}$$

b) Se obtendrá la función de densidad de Y_1

$$\begin{aligned} f_{Y_1}(y) &= 10[1 - F_X(y)]^{10} f_X(y) \\ &= 10\left[1 - 1 + e^{-y/100}\right]^9 (1/100)e^{-y/100} \\ &= 1/10e^{-y/10}, \text{ donde } y > 0. \end{aligned}$$

Entonces, $Y_1 \sim Exp(\lambda = 1/10)$, por lo tanto, $E[Y_1] = 10$.

A continuación, se darán algunas definiciones de importancia, las cuales se estarán usando en el desarrollo de este tema.

Definición 6.4. Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una distribución, la muestra ordenada representada por $Y_1, Y_2, \ldots, Y_n$ es conocida como las *estadísticas de orden* de la muestra aleatoria $X_1, X_2, \ldots, X_n$.

Observación. Si la muestra $X_1, X_2, \ldots, X_n$ no es aleatoria, la muestra ordenada no se considera como las estadísticas de orden de la muestra, es un requisito que la muestra sea aleatoria.

Definición 6.5. Sean $Y_1, \ldots, Y_n$ las estadísticas de orden de una muestra aleatoria $X_1, X_2, \ldots, X_n$ de tamaño n. Si n es un número impar, la *mediana muestral* es el elemento que se encuentra exactamente a la mitad de la muestra ordenada. En caso de que n sea un número par, la *mediana muestral* es el promedio de los dos elementos que se encuentran exactamente a la mitad de la muestra ordenada.

Definición 6.6. Sean $Y_1, \ldots, Y_n$ las estadísticas de orden de una muestra aleatoria $X_1, X_2, \ldots, X_n$ de tamaño n. El *rango muestral*, denotado como R, es la diferencia entre el elemento mayor y el elemento menor, esto es, $R = Y_n - Y_1$.

Sean $Y_1, Y_2, \ldots, Y_n$ las estadísticas de orden de una muestra aleatoria $X_1, X_2, \ldots, X_n$ obtenida de una distribución continua, a continuación, se encontrarán:

1) La función de densidad conjunta de $Y_1, \ldots, Y_n$.
2) La función de densidad de la k-ésima estadística de orden, esto es, $f_{Y_k}(y)$.

3) La función de densidad conjunta de dos estadísticas de orden. Esto es, si Y_i y Y_j son dos estadísticas de orden, $i < j$, se encontrará $f_{Y_i,Y_j}(y_i, y_j)$.

Teorema 6.14. *Sean $Y_1, \ldots, Y_n$ las estadísticas de orden de una muestra aleatoria $X_1, \ldots, X_n$ de tamaño n de una distribución continua con función de densidad $f_X(x)$ y con recorrido $a < x < b$, entonces,*

$$f_{\underline{Y}}(y_1, \ldots, y_n) = n! \prod_{i=1}^{n} f_X(y_i), \quad \textit{donde} \quad a < y_1 < y_2 < \cdots < y_n < b.$$

Es importante aclarar que en la notación del recorrido de la variable X, a y b pueden ser números reales o puede ser $-\infty$ o ∞, respectivamente.

Demostración. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución continua con función de densidad $f_X(x)$ y función de distribución $F_X(x)$, y $Y_1, \ldots, Y_n$ las estadísticas de orden de esta muestra.

Primero se va a obtener la función de densidad conjunta de las estadísticas de orden de una muestra aleatoria de tamaño $n = 2$.

Definamos la siguiente transformación de $\mathbb{R}^2$ a $\mathbb{R}^2$

$$\begin{aligned} g(X_1, X_2) &= (g_1(X_1, X_2), g_2(X_1, X_2)) \\ &= (min(X_1, X_2), max(X_1, X_2)) \\ &= (Y_1, Y_2). \end{aligned}$$

Esta transformación tiene dos inversas. Sean $g_1^{-1}(y_1, y_2)$ y $g_2^{-1}(y_1, y_2)$, las cuales se dan a continuación:

$$g_1^{-1}(y_1, y_2) = (k_{11}(y_1, y_2), k_{12}(y_1, y_2)) = (y_1, y_2),$$

$$g_2^{-1}(y_1, y_2) = (k_{21}(y_1, y_2), k_{22}(y_1, y_2)) = (y_2, y_1).$$

Se puede observar que el número de inversas es igual al número de permutaciones de los elementos y_1, y_2, esto es, $2!$

Se aplicará el método de cambio de variable, considerando transformaciones con k inversas, véase caso 4 de este método. El método nos dice que la función de densidad conjunta de Y_1 y Y_2 se puede encontrar de la siguiente manera:

$$f_{\underline{Y}}(y_1, y_2) = f_{\underline{X}}(k_{11}(y_1, y_2), k_{12}(y_1, y_2))|J_1| + f_{\underline{X}}(k_{21}(y_1, y_2), k_{22}(y_1, y_2))|J_2|,$$

donde J_i es el i-ésimo jacobiano correspondiente a la i-ésima inversa.

Obsérvese que

$$f_{\underline{X}}(x_1, x_2) = f_X(x_1) f_X(x_2).$$

Además,

$$J_1 = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1 \quad \text{y} \quad J_2 = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = -1.$$

Por lo tanto, para $n = 2$

$$\begin{aligned} f_{\underline{Y}}(y_1, y_2) &= f_{\underline{X}}(y_1, y_2)|J_1| + f_{\underline{X}}(y_2, y_1)|J_2| \\ &= f_X(y_1) f_X(y_2)(1) + f_X(y_2) f_X(y_1)(1) \\ &= 2 f_X(y_1) f_X(y_2), \text{ donde } a < y_1 < y_2 < b. \end{aligned}$$

Ahora, para $n = 3$,
Definamos la siguiente transformación

$$\begin{aligned} g(X_1, X_2, X_3) &= (g_1(X_1, X_2, X_3), g_2(X_1, X_2, X_3), g_3(X_1, X_2, X_3)) \\ &= (Y_1, Y_2, Y_3), \end{aligned}$$

la cual tiene las siguientes inversas

$$\begin{aligned} g_1^{-1}(y_1, y_2, y_3) &= (y_1, y_2, y_3), \\ g_2^{-1}(y_1, y_2, y_3) &= (y_1, y_2, y_3), \\ g_3^{-1}(y_1, y_2, y_3) &= (y_2, y_1, y_3), \\ g_4^{-1}(y_1, y_2, y_3) &= (y_2, y_3, y_1), \\ g_5^{-1}(y_1, y_2, y_3) &= (y_3, y_1, y_2), \\ g_6^{-1}(y_1, y_2, y_3) &= (y_3, y_2, y_1). \end{aligned}$$

Se puede notar que el número de inversas es igual al número de permutaciones de los elementos y_1, y_2, y_3, en este caso $6 = 3!$

Obsérvese que el jacobiano para cada inversa será 1 o -1, ya que es igual al determinante de la matriz identidad o de una matriz con un número finito de renglones permutados de la matriz identidad. Lo que significa que $|J_i| = 1$, para $i = 1, \ldots, 6$.

La función de densidad de $\underline{X}$ es

$$f_{\underline{X}}(x_1, x_2, x_3) = f_X(x_1) f_X(x_2) f_X(x_3).$$

Entonces, la función de densidad conjunta de Y_1, Y_2 y Y_3 es

$$\begin{aligned} f_{\underline{Y}}(y_1, y_2, y_3) &= f_{\underline{X}}(y_1, y_2, y_3)|J_1| + f_{\underline{X}}(y_1, y_3, y_2)|J_2| + f_{\underline{X}}(y_2, y_1, y_3)|J_3| \\ &\quad + f_{\underline{X}}(y_2, y_3, y_1)|J_4| + f_{\underline{X}}(y_3, y_1, y_2)|J_5| + f_{\underline{X}}(y_3, y_2, y_1)|J_6| \\ &= 6 f_X(y_1) f_X(y_2) f_X(y_3) = 3! f_X(y_1) f_X(y_2) f_X(y_3). \end{aligned}$$

Por lo tanto,

$$f_{\underline{Y}}(y_1, y_2, y_3) = 3! f_X(y_1) f_X(y_2) f_X(y_3), \text{ donde } a < y_1 < y_2 < y_3 < b.$$

Considerando una muestra de tamaño n, el número de inversas que tiene la transformación $g(X_1, X_2, \ldots, X_n) = (Y_1, Y_2, \ldots, Y_n)$ es igual al número de permutaciones diferentes que se pueden realizar con los elementos $y_1, y_2, \ldots, y_n$, esto es, $n!$

Por otro lado, se puede observar con facilidad que $|J_i| = 1$, para todo $i = 1, 2, \ldots, n!$, ya que J_i es el determinante de la matriz identidad o de una matriz donde se han permutado algunos renglones o columnas de la matriz identidad.

Entonces, la función de densidad conjunta de $Y_1, Y_2, \ldots, Y_n$ es

$$\begin{aligned} f_{\underline{Y}}(y_1, \ldots, y_n) &= n! f_X(y_1) f_X(y_2) \ldots f_X(y_n) \\ &= n! \prod_{i=1}^{n} f_X(y_i), \text{ donde } a < y_1 < y_2 < \cdots < y_n < b. \quad \square \end{aligned}$$

Teorema 6.15. *Sean $Y_1, Y_2, \ldots, Y_n$ las estadísticas de orden de una muestra aleatoria de tamaño n, obtenida de una distribución continua con funciones de densidad y de distribución $f_X(x)$ y $F_X(x)$, respectivamente. Entonces, la función de densidad de la k-ésima estadística de orden, denotada como $f_{Y_k}(y)$, está dada por*

$$f_{Y_k}(y) = \frac{n!}{(k-1)!(n-k)!} [F_X(y)]^{k-1} [1 - F_X(y)]^{n-k} f_X(y),$$

donde $a < y < b$ y $k = 1, 2, \ldots, n$.

Demostración. Para obtener la función de densidad marginal de Y_k, se realizará la integral múltiple de la función de densidad conjunta de $Y_1, \ldots, Y_n$ con respecto a todas las variables excepto a la variable Y_k. Se integrará la función de densidad en el siguiente orden: primero con respecto a y_1, luego con respecto a $y_2, \ldots$, y así hasta integrar con respecto a y_{k-1}, después se integrará con respecto a y_n, seguido con respecto a $y_{n-1}, \ldots$, así hasta integrar con respecto a y_{k+1}, esto es,

$$\begin{aligned} & f_{Y_k}(y_k) \\ &= \int_{y_k}^{b} \cdots \int_{y_{n-1}}^{b} \int_{a}^{y_k} \cdots \int_{a}^{y_2} n! f_X(y_1) \cdots f_X(y_n) dy_1 \cdots dy_{k-1} dy_n \cdots dy_{k+1} \\ &= \int_{y_k}^{b} \cdots \int_{y_{n-1}}^{b} \int_{a}^{y_k} \cdots \int_{a}^{y_3} n! \prod_{i=2}^{n} f_X(y_i) \left[\int_{a}^{y_2} f_X(y_1) dy_1 \right] d_{y_2} \cdots dy_{k+1} \\ &= \int_{y_k}^{b} \cdots \int_{y_{n-1}}^{b} \int_{a}^{y_k} \cdots \int_{a}^{y_4} n! \prod_{i=3}^{n} f_X(y_i) \\ &\times \left[\int_{a}^{y_3} F_X(y_2) f_X(y_2) dy_2 \right] d_{y_3} \cdots dy_{k+1} \\ &= \int_{y_k}^{b} \cdots \int_{y_{n-1}}^{b} \int_{a}^{y_k} \cdots \int_{a}^{y_5} n! \prod_{i=4}^{n} f_X(y_i) \\ &\times \left[\int_{a}^{y_4} \frac{[F_X(y_3)]^2}{2} f_X(y_3) d_{y_3} \right] dy_4 \cdots dy_{k+1} \end{aligned}$$

$$
\begin{aligned}
&= \int_{y_k}^{b} \cdots \int_{y_{n-1}}^{b} \int_{a}^{y_k} \cdots \int_{a}^{y_6} n! \prod_{i=5}^{n} f_X(y_i) \\
&\times \left[\int_{a}^{y_5} \frac{[F_X(y_4)]^3}{2(3)} f_X(y_4) d_{y_4} \right] dy_5 \cdots dy_{k+1} \\
&\qquad\vdots \\
&= \int_{y_k}^{b} \cdots \int_{y_{n-1}}^{b} n! \prod_{i=k}^{n} f_X(y_i) \\
&\times \left[\int_{a}^{y_k} \frac{[F_X(y_{k-1})]^{k-2}}{2(3)\cdots(k-2)} f_X(y_{k-1}) d_{y_{k-1}} \right] dy_n \cdots dy_{k+1} \\
&= \int_{y_k}^{b} \cdots \int_{y_{n-1}}^{b} n! \prod_{i=k}^{n} f_X(y_i) \frac{[F_X(y_k)]^{k-1}}{(k-1)!} dy_n \cdots dy_{k+1}.
\end{aligned}
$$

Como se puede notar, en las integrales anteriores, se consideraron los siguientes cambios de variable: $u = F_X(y_i)$, para $i = 2, \ldots, k-1$, donde $du = f_X(y_i)dy_i$. Obsérvese que para el cambio de variable específico $u = F_X(y_i)$, la integral correspondiente se debe realizar de 0 a $F_X(y_{i+1})$.

A continuación, se resolverán las integrales restantes, en el siguiente orden: primero con respecto a y_n, luego con respecto y_{n-1}, …, así hasta integrar con respecto a y_{k+1}.

$$
\begin{aligned}
&f_{Y_k}(y_k) \\
&\quad = n! f_X(y_k) \frac{[F_X(y_k)]^{k-1}}{(k-1)!} \int_{y_k}^{b} \cdots \int_{y_{n-2}}^{b} \prod_{i=k+1}^{n-1} f_X(y_i) \\
&\quad \times \left[\int_{y_{n-1}}^{b} f_X(y_n) dy_n \right] dy_{n-1} \cdots dy_{k+1} \\
&\quad = n! f_X(y_k) \frac{[F_X(y_k)]^{k-1}}{(k-1)!} \int_{y_k}^{b} \cdots \int_{y_{n-3}}^{b} \prod_{i=k+1}^{n-2} f_X(y_i) \\
&\quad \times \left[\int_{y_{n-2}}^{b} [1 - F_X(y_{n-1})] f_X(y_{n-1}) dy_{n-1} \right] dy_{n-2} \cdots dy_{k+1} \\
&\quad = n! f_X(y_k) \frac{[F_X(y_k)]^{k-1}}{(k-1)!} \int_{y_k}^{b} \cdots \int_{y_{n-4}}^{b} \prod_{i=k+1}^{n-3} f_X(y_i) \\
&\quad \times \left[\int_{y_{n-3}}^{b} \frac{[1 - F_X(y_{n-2})]^2}{2} f_X(y_{n-2}) dy_{n-2} \right] dy_{n-3} \cdots dy_{k+1} \\
&\qquad\qquad \vdots \\
&\quad = n! f_X(y_k) \frac{[F_X(y_k)]^{k-1}}{(k-1)!} \int_{y_k}^{b} \frac{[1 - F_X(y_{k+1})]^{n-(k-1)}}{2(3)\cdots(n-(k-1))} f_X(y_{k+1}) dy_{k+1}
\end{aligned}
$$

$$= n! f_X(y_k) \frac{[F_X(y_k)]^{k-1}}{(k-1)!} \left[\frac{[1 - F_X(y_k)]^{n-k}}{2(3) \cdots (n-(k-1))(n-k)} \right].$$
$$= \frac{n!}{(k-1)!(n-k)!} [[1 - F_X(y_k)]^{n-k} [F_X(y_k)]^{k-1} f_X(y_k).$$

De esta manera, la función de densidad de Y_k es

$$f_{Y_k}(y_k) \frac{n!}{(k-1)!(n-k)!} [[1 - F_X(y_k)]^{n-k} [F_X(y_k)]^{k-1} f_X(y_k),$$

donde $a < y_k < b$.

En las anteriores integrales se consideraron los cambios de variable: $u = F_X(y_i)$, para $i = k+1, \ldots, n$, donde $du = f_X(y_i) dy_i$. Obsérvese que para el cambio de variable específico $u = F_X(y_i)$, la integral correspondiente va de $F_X(y_{i-1})$ a 1. □

Teorema 6.16. *Sean $Y_1, Y_2, \ldots, Y_n$ las estadísticas de orden de una muestra aleatoria de tamaño n obtenida de una distribución continua con función de densidad $f_X(x)$ y función de distribución $F_X(x)$. Entonces, la función de densidad conjunta de la i-ésima y j-ésima estadísticas de orden, donde $i < j$, es denotada como $f_{Y_i,Y_j}(y_i, y_j)$ y dada de la siguiente manera:*

$$\begin{aligned} f_{Y_i,Y_j}(y_i, y_j) = {} & \frac{n!}{(i-1)!(j-i-1)!(n-j)!} [F_X(y_i)]^{i-1} \\ & \times [F_X(y_j) - F_X(y_i)]^{j-i-1} [1 - F_X(y_j)]^{n-j} f_X(y_i) f_X(y_j), \end{aligned}$$

donde $a < y_i < y_j < b$ e $i, j = 1, 2, \ldots, n$.

La demostración del anterior resultado se logra realizando la integral múltiple de la función de densidad conjunta de $Y_1, Y_2, \ldots, Y_n$, con respecto a todas las variables, excepto y_i y y_j. Se sugiere integrar en el siguiente orden: primero con respecto a y_n, luego a $y_{n-1}, \ldots$, y así, hasta integrar con respecto a y_{j+1}. Luego con respecto a y_{j-1}, posteriormente, con a $y_{j-2}, \ldots$, y así, hasta integrar con respecto a y_{i+1}. Después, con respecto a y_1, luego a $y_2, \ldots$, y finalmente integrar con respecto a y_{i-1}. Considerando cambios de variable similares a los utilizados en la demostración del teorema anterior.

Ejemplo 6.31. En un banco, el tiempo en horas que tardan en tramitar la apertura de una nueva cuenta a un cliente es una variable aleatoria con distribución exponencial con media 1/2. Sean Y_1, Y_2, Y_3, Y_4, Y_5 las estadísticas de orden de una muestra aleatoria de 5 clientes seleccionados al azar. Encontrar

a) La función de densidad conjunta de Y_1, Y_2, Y_3, Y_4, Y_5.
b) La función de densidad conjunta de Y_1 y Y_5.
c) $P[Y_2 < 1/2]$.

a) La densidad conjunta se calcula de la siguiente manera:

$$\begin{aligned} f_{\underline{Y}}(y_1, y_2, y_3, y_4, y_5) &= 5!(2e^{-2y_1}) \cdots (2e^{-2y_5}) \\ &= 3840e^{-2(y_1+\cdots+y_n)}, \text{ donde } 0 < y_1 < \cdots < y_5 < \infty. \end{aligned}$$

b) La densidad conjunta de Y_1 y Y_5

$$\begin{aligned} f_{(Y_1,Y_5)}(y_1, y_5) &= \frac{5!}{(1-1)!(5-1-1)!(5-5)!} \\ &\times [1 - Fx(y_1)]^0 [F_X(y_5) - F_X(y_1)]^3 [F_X(y_5)]^0 f_X(y_1) f_X(y_5) \\ &= 20[e^{-2y_1} - e^{-2y_5}](2e^{-2y_1} 2e^{-2y_5}) \\ &= 80[e^{-2y_1} - e^{-2y_5}]e^{-2(y_1+y_5)}, \text{ donde } 0 < y_1 < y_5 < \infty. \end{aligned}$$

c) La marginal de Y_2

$$\begin{aligned} f_{y_2}(y) &= \frac{5!}{(2-1)!(5-2)!}[F_X(y)]^{2-1}[1 - F_X(y)]^{5-2} f_X(y_2) \\ &= 20[1 - e^{-2y}][e^{-2y}]^3[2e^{-2y}] \\ &= 40\{e^{-8y} - e^{-10y}\}, \text{ donde } 0 < y. \end{aligned}$$

Por lo tanto,

$$\begin{aligned} P[Y_2 < 1/2] &= \int_0^{1/2} f_{y_2}(y) dy_2 \\ &= 40 \left[\int_0^{1/2} (e^{-8y} - e^{-10y}) dy_2 \right] \\ &= 40 \left[-\frac{1}{8} e^{-8y} + \frac{1}{10} e^{-10y} \right]_0^{1/2} \\ &= -5e^{-4} + 4e^{-5} + 5 - 4 = 0.93537. \end{aligned}$$

Ejemplo 6.32. Sean $Y_1, Y_2, \ldots, Y_n$ las estadísticas de orden de una muestra aleatoria de tamaño $n = 2m + 1$ de una distribución continua con función de densidad $f_X(x)$ y función de distribución $F_X(x)$. Calcular la probabilidad de que la mediana muestral sea menor que la mediana poblacional.

La mediana muestral es la estadística de orden Y_{m+1}, y la mediana poblacional es aquel número M que cumple con $P[X < M] = 0.5$. El objetivo es calcular $P[Y_{m+1} < M]$.

Por lo que se obtendrá la función de densidad de Y_{m+1},

$$f_{Y_{m+1}}(y) = \frac{(2m+1)!}{(m!m!)}[F_X(y)]^m [1 - F_X(y)]^m f_X(y).$$

Entonces,

$$\begin{aligned} P[Y_{m+1} < M] &= \int_a^M f_{Y_{m+1}}(y)dy \\ &= \int_a^M \frac{(2m+1)!}{m!m!}[F_X(y)]^m[1-F_X(y)]^m f_X(y)dy. \end{aligned}$$

Se hace el siguiente cambio de variable $u = F_X(y)$, de esta manera, $du = f_X(y)dy$.

Obsérvese que el recorrido para la nueva variable U es el intervalo $(0,1)$. Entonces, la integral va de 0 a 0.5.

Esto es,

$$\begin{aligned} P[Y_{m+1} < M] &= \int_0^{0.5} \frac{(2m+1)!}{m!m!}u^m(1-u)^m du \\ &= \int_0^{0.5} \frac{\Gamma(2m+2)}{\Gamma(m+1)\Gamma(m+1)}u^m(1-u)^m du. \end{aligned}$$

La anterior integral, es la probabilidad acumulada hasta 0.5 de una distribución $Beta(m+1, m+1)$. Obsérvese que la función de densidad de esta distribución es simétrica con respecto a 0.5, lo que significa que la anterior integral vale 0.5. De este modo, se concluye que $P[Y_{m+1} < M] = 0.5$.

Ejemplo 6.33. Sean $Y_1, Y_2, \ldots, Y_n$ las estadísticas de orden de una muestra aleatoria de tamaño n de una distribución uniforme sobre el intervalo $(0,1)$, encontrar la distribución de la $k-$ésima estadística de orden.

Se sabe que las funciones de densidad y de distribución de X, son respectivamente, $f_X(x) = I_{(0,1)}(y)$ y $F_X(x) = xI_{(0,1)}(x) + I_{[1,\infty)}(x)$.

Ahora

$$\begin{aligned} f_{Y_k}(y) &= \frac{n!}{(k-1)!(n-k)!}[F_X(y)]^{k-1}[1-F_X(y)]^{n-k}f_X(y) \\ &= \frac{n!}{(k-1)!(n-k)!}y^{k-1}(1-y)^{n-k}(1), \text{ donde } 0 < y < 1. \end{aligned}$$

Se concluye, $Y_k \sim Beta(k, n-k+1)$.

Este resultado sirve para generar valores de una variable aleatoria con distribución beta con parámetros enteros.

Con la idea de explicar el ejemplo anterior, por medio de R se simularon 10,000 valores de una variable aleatoria con distribución $Beta(5,7)$. Esto es, se determina que los valores de k y n, en este caso, son respectivamente 5 y 11. Se generan por medio de R, 11 valores de una variable aleatoria con distribución $U(0,1)$, por medio de la instrucción $runif(11,0,1)$. Estos valores se ordenan de menor a mayor, el valor que ocupa el lugar 5 será un valor de la variable aleatoria con distribución $Beta(5,7)$. Este procedimiento se realiza 10,000 veces. A continuación, se presenta el histograma de los 10,000 valores de la distribución $Beta(5,7)$ junto con la densidad de una $Beta(5,7)$.

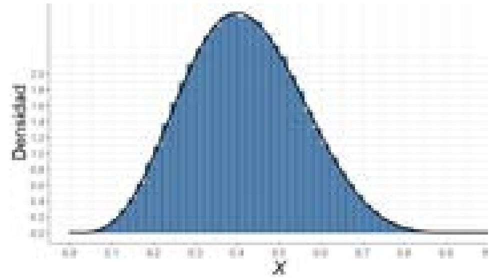


Figura 6.39. *Histograma y función de densidad*

Ejemplo 6.34. Sean Y_1, Y_2, Y_3, Y_4 las estadísticas de orden de una muestra aleatoria de una distribución uniforme sobre el intervalo $(0,1)$. Encontrar la distribución del rango muestral.

El rango muestral, denotado por R, está en función de las estadísticas de orden Y_1 y Y_4, entonces, primero se obtendrá la función de densidad conjunta de Y_1 y Y_4, de la siguiente manera:

$$\begin{aligned}
f_{Y_1,Y_4}(y_1,y_4) &= \frac{4!}{(1-1)!(4-1-1)!(4-4)!} \\
&= [F_X(y_1)]^{1-1}[F_X(y_4)-F_X(y_1)]^{4-1-1}[1-F_X(y_4)]^{4-4} \\
&\times f_X(y_1)f_X(y_4) \\
&= 12(y_4-y_1)^2, \text{ donde } 0 < y_1 < y_4 < 1.
\end{aligned}$$

Entonces, se obtendrá la distribución de la variable $R = Y_4 - Y_1$, en este caso se aplicará la técnica de cambio de variable, por lo que es necesario definir una nueva variable, sea $V = Y_1$.

Se obtendrá la función de densidad conjunta de $R = Y_4 - Y_1$ y $V = Y_1$.

El recorrido conjunto de Y_1 y Y_4 gráficamente es

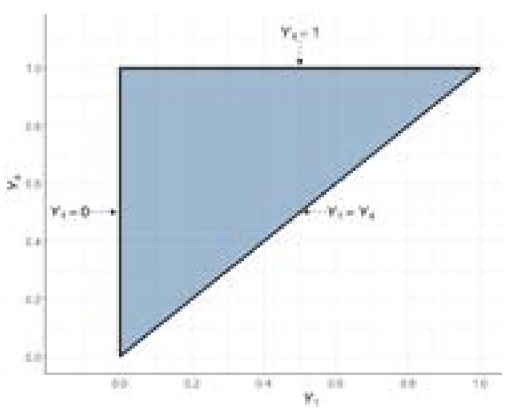


Figura 6.40. *Recorrido conjunto de Y_1 y Y_4, ejemplo 6.34*

Obsérvese que el recorrido conjunto de R y V gráficamente es

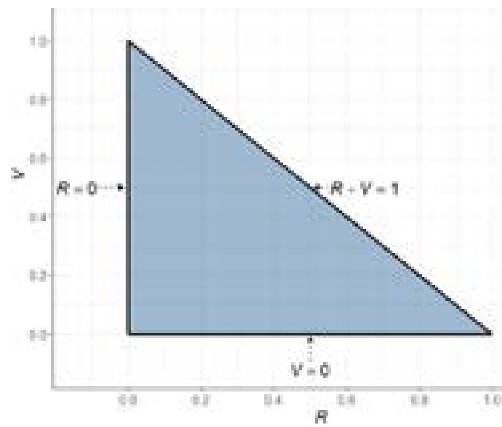


Figura 6.41. *Recorrido conjunto de R y V, ejemplo 6.34*

La función inversa de la transformación es

$$Y_1 = V \quad \text{y} \quad Y_4 = R + V.$$

El jacobiano es

$$J = \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} = 1.$$

La función de densidad conjunta de R y V es

$$\begin{aligned} f_{R,V}(r,v) &= f_{Y_1,Y_4}(v, r+v)|J| \\ &= 12r^2, \ \text{ donde } \ v > 0, r > 0 \text{ y } 0 < v + r < 1. \end{aligned}$$

Esto es,

$$f_{R,V}(r,v) = 12r^2 I_{(0,1-r)}(v) I_{(0,1)}(r).$$

Se obtendrá la marginal de R

$$\begin{aligned} f_R(r) &= \int_0^{1-r} 12r^2 dv \\ &= 12r^2(1-r). \end{aligned}$$

Esto es,

$$f_R(r) = 12r^2(1-r) I_{(0,1)}(r).$$

De esta manera, $R \sim Beta(3,2)$.

Se deja como ejercicio para el estudiante resolver este último ejemplo, pero por el método de función de distribución.

El resultado del anterior ejemplo se puede generalizar, considerando una muestra de tamaño n, véase el ejercicio 66.

Ejemplo 6.35. Sea $\underline{X} = (X_1, X_2)$ un vector aleatorio con función de densidad dada por $f_{\underline{X}}(x_1, x_2) = \dfrac{12}{7} x_1(x_1 + x_2) I_{(0,1)}(x_1) I_{(0,1)}(x_2)$.

a) Encontrar la función de densidad conjunta de $U = min(X_1, X_2)$ y $V = max(X_1, X_2)$.

b) ¿Son independientes U y V?

a) Es importante señalar que X_1 y X_2 no son una muestra aleatoria de una distribución, ya que las variables aleatorias X_1 y X_2 son dependientes, por lo que no podemos usar la fórmula para obtener la función de densidad conjunta de dos estadísticas de orden. Entonces, este problema se resolverá por medio de la técnica cambio de variable donde la transformación es

$$U = min(X_1, X_2), V = max(X_1, X_2).$$

El recorrido conjunto de X_1 y X_2 es el cuadro unitario, esto es,

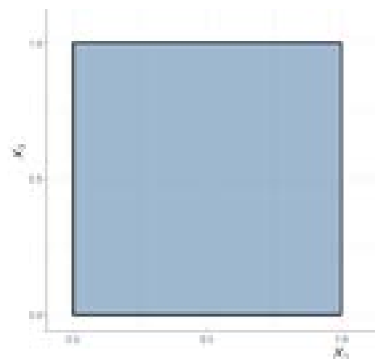


Figura 6.42. *Recorrido conjunto de X y Y, ejemplo 6.35*

De esta manera, el recorrido conjunto de U y V es

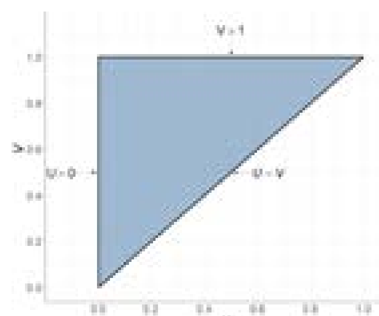


Figura 6.43. *Recorrido conjunto de U y V, ejemplo 6.35*

Es una transformación con dos inversas, esto es,

$$g_1^{-1}(u, v) = (u, v),$$

$$g_2^{-1}(u, v) = (v, u).$$

Se puede observar

$$|J_1| = 1 \quad \text{y} \quad |J_2| = 1.$$

De esta manera,

$$f_{U,V}(u,v) = \frac{12}{7}\left[u(u+v)(1) + v(v+u)(1)\right], \text{donde } 0 < u < v < 1.$$

Esto es,

$$f_{U,V}(u,v) = \frac{12}{7}(u+v)^2 I_{(0,v)}(u) I_{(0,1)}(v).$$

b) Como no es posible expresar la función de densidad conjunta $f_{U,V}(u,v)$ como el producto de dos funciones $g_1(u)$ y $g_2(v)$, entonces, U y V son variables aleatorias dependientes.

Ejemplo 6.36. Para un crédito determinado, la probabilidad de que un cliente se demore con su pago mensual es igual a p. Se seleccionan al azar n clientes. Para el cliente i, se observa X_i, el número de mensualidades sin demora, antes de tener por primera vez una mensualidad con atraso, para $i = 1, \ldots, n$. Encontrar la función de densidad de
a) $Y_n = max(X_1, \ldots, X_n)$.
b) $Y_1 = min(X_1, \ldots, X_n)$.
c) ¿Cómo se distribuye Y_1?

a) Obsérvese que X_i tiene distribución geométrica con parámetro p, y por ser la distribución de X_i del tipo discreto, no es posible aplicar la fórmula de la función de densidad del máximo. Entonces, se obtendrá primero, la función de distribución del máximo Y_n, y a partir de esta, se obtendrá su función de densidad.

Esto es,

$$F_{Y_n}(y) = [F_X(y)]^n.$$

La función de distribución de una geométrica con parámetro p, para valores del recorrido es

$$F_X(x) = 1 - (1-p)^{x+1}, \text{ para } x = 0, 1, \ldots$$

De esta manera,

$$F_{Y_n}(y) = [1 - (1-p)^{y+1}]^n, \text{ para } y = 0, 1, \ldots$$

Por lo tanto, para $y = 0$

$$\begin{aligned} f_{Y_n}(0) &= F_{Y_n}(0) \\ &= p^n. \end{aligned}$$

Y para $y = 1, 2, \ldots$

$$\begin{aligned} f_{Y_n}(y) &= F_{Y_n}(y) - F_{Y_n}(y-1) \\ &= [1 - (1-p)^{y+1}]^n - [1 - (1-p)^y]^n. \end{aligned}$$

La anterior igualdad, en principio se cumple para valores de y enteros positivos, pero si sustituimos $y = 0$, obsérvese que

$$[1 - (1 - p)]^n - [1 - (1 - p)^0]^n = p^n.$$

Por lo tanto, se cumple para todos los valores del recorrido de la variable Y_n, esto es,

$$f_{Y_n}(y) = \{[1 - (1 - p)^{y+1}]^n - [1 - (1 - p)^y]^n\} I_{\{0,1,2,\ldots\}}(y).$$

b) Se calculará primero la función de distribución de Y_1,

$$\begin{aligned} F_{Y_1}(y) &= 1 - [1 - F_X(y)]^n \\ &= 1 - [1 - 1 + (1 - p)^{y+1}]^n \\ &= 1 - (1 - p)^{(y+1)n}, \text{ para } y = 0, 1, \ldots \end{aligned}$$

Se obtendrá, primero, la función de densidad para $y = 0$

$$\begin{aligned} f_{Y_1}(0) &= F_{Y_1}(0) \\ &= 1 - (1 - p)^n. \end{aligned}$$

Ahora, para valores de $y = 1, 2, \ldots$

$$\begin{aligned} f_{Y_1}(y) &= F_{Y_1}(y) - F_{Y_1}(y - 1) \\ &= 1 - (1 - p)^{(y+1)n} - 1 + (1 - p)^{yn} \\ &= (1 - p)^{ny} - (1 - p)^{(y+1)n}. \end{aligned}$$

La anterior igualdad es válida, en principio, para valores de $y = 1, 2, \ldots$ Pero obsérvese que si esta expresión se evalúa en $y = 0$, se obtendrá el valor correspondiente a $f_{Y_1}(0)$. De esta manera, la función de densidad de Y_1 es

$$f_{Y_1}(y) = \{(1 - p)^{ny} - (1 - p)^{(y+1)n}\} I_{\{0,1,2,\ldots\}}(y).$$

c) Obsérvese que la función de densidad de Y_1 puede expresarse de la siguiente manera:

$$\begin{aligned} f_{Y_1}(y) &= \{(1 - p)^{ny} - (1 - p)^{(y+1)n}\} I_{\{0,1,2,\ldots\}}(y) \\ &= (1 - p)^{ny}(1 - (1 - p)^n) I_{\{0,1,2,\ldots\}}(y). \end{aligned}$$

Esta última expresión es la función de densidad de una distribución geométrica con probabilidad de «éxito» igual a $1 - (1 - p)^n$ y probabilidad de «fracaso» $(1 - p)^n$.

Por lo tanto, se concluye que Y_1 se distribuye *Geométrica*$(1 - (1 - p)^n)$.

Ejercicios del capítulo 6

1.- Sea X una variable aleatoria con función de densidad dada por $f_X(x) = |x|_{[-1,1]}(x)$. Encontrar la función de densidad de

a) $U = 5X$.
b) $U = 2 - X$.
c) $U = e^X$.
d) $U = X^2$.

2.- Sea X una variable aleatoria con función de densidad dada por $f_X(x) = 0.5I_{[0,1]}(x) + I_{(1,1.5]}(x)$. Sea $U = 5X - 2$. Encontrar

a) La función de densidad de U.
b) $E[U]$.

3.- Sean X_1 y X_2 dos variables aleatorias con función de densidad conjunta dada por $f_{\underline{X}}(x_1, x_2) = \frac{1}{2}I_{[0,2-x_2]}(x_1)I_{[0,2]}(x_2)$. Obtener la función de densidad de U si

a) $U = X_1 + X_2$.
b) $U = X_1 - X_2$.

4.- Sea X_1 y X_2 una muestra aleatoria de una distribución con función de densidad $f_X(x) = \frac{1}{2}xI_{(0,2)}(x)$. Calcular $P\left[X_1/X_2 \leq 1\right]$.

5.- Sea X una variable aleatoria con distribución geométrica con parámetro p, encontrar la función de densidad de $U = e^X$.

6.- Sea $U = X_1 + X_2$, donde X_1, X_2 es una muestra aleatoria de tamaño 2 de una distribución $U(0, 2)$. Encontrar la función de

a) Distribución de U.
b) Densidad de U.

7.- Sea X_1, X_2 una muestra aleatoria de tamaño 2 de una distribución uniforme sobre el intervalo $(0, 1)$. Usando el método de función de distribución, encontrar la función de densidad de U si

a) $U = X_2 - X_1$.
b) $U = X_2/X_1$.

8.- Realizar los incisos del ejercicio 7 por el método de cambio de variable.

9.- Sean X_1 y X_2 dos variables con función de densidad conjunta dada por $f_{\underline{X}}(x_1, x_2) = \left(\frac{1}{5}\right)^{x_1+x_2-2}\left(\frac{4}{5}\right)^{4-x_1-x_2} I_{\{1,2\}}(x_1)I_{\{1,2\}}(x_2)$, encontrar la función de densidad conjunta de las variables $U_1 = X_1 - X_2$ y $U_2 = X_1 + X_2$.

10.- Sean X_1 y X_2 dos variables aleatorias con función de densidad conjunta dada por $f_{\underline{X}}(x_1, x_2) = \dfrac{(x_1+1)(x_2+1)}{9} I_{\{0,1\}}(x_1) I_{\{0,1\}}(x_2)$. Encontrar la función de densidad de la variable aleatoria $U = X_1 - X_2$.

11.- Sea X una variable aleatoria con función de densidad dada por $f_X(x) = \frac{x^2}{9} I_{(0,3)}(x)$. ¿Cómo se distribuye la variable aleatoria $U = X^3$?

12.- Sea X una variable aleatoria con distribución de Weibull con parámetros $a = 2$ y $b = 2$. ¿Cómo se distribuye la variable aleatoria $U = X^2$?

13.- Sea X_1 y X_2 una muestra aleatoria de una distribución normal estándar. Sean $U_1 = X_1/X_2$ y $U_2 = X_2$.

a) Encontrar la función de densidad conjunta de U_1 y U_2.

b) Mostrar que la función de densidad de U_1 es $f_{U_1}(u_1) = \dfrac{1}{\pi(1+u_1^2)}$. Esto es, U_1 tiene una distribución Cauchy.

14.- Sea X una variable aleatoria con distribución uniforme sobre el intervalo $(-\pi/2, \pi/2)$. Mostrar que la variable $U = tan(X)$ tiene una distribución Cauchy.

15.- Sea X_1 y X_2 una muestra aleatoria de una distribución normal estándar. Muestre que $U_1 = \mu_1 + \sigma_1 X_1$ y $U_2 = \mu_2 + \rho\sigma_2 X_1 + \sigma_2\sqrt{1-\rho^2}X_2$, donde $\sigma_1 > 0$, $\sigma_2 > 0$, y $0 < \rho < 1$, tienen una distribución conjunta normal bivariada con parámetros μ_1, μ_2, σ_1^2, σ_2^2 y ρ.

16.- Sean X_1 y X_2 dos variables aleatorias independientes e idénticamente distribuidas con distribución común normal con media μ y varianza σ^2. Sean $U_1 = X_1 + X_2$ y $U_2 = X_1 - X_2$.

a) ¿Cómo se distribuye $\underline{U} = (U_1, U_2)$?
b) ¿Son independientes U_1 y U_2?, ¿por qué?
c) ¿Cómo se distribuye X_i, para $i = 1, 2$?

17.- Sea X una variable aleatoria y $U = g(X)$ una transformación de X definida de la siguiente manera: $g(x) = (0.5x+1)I_{[0,2)}(x) + xI_{[2,4)}(x) + (1/3)(28-4x)I_{[4,7]}(x)$. Encontrar la función de densidad de la variable U si

a) $X \sim U(0, 7)$.
b) $X \sim B(7, 0.5)$.

18.- Sea X_1 y X_2 una muestra aleatoria de tamaño 2 de una distribución normal con media μ y varianza σ^2. Sea $\underline{U} = (U_1, U_2)$, donde $U_1 = X_1 + X_2$ y $U_2 = X_1 + 2X_2$.

a) Encontrar la función de densidad de $\underline{U}$.
b) ¿Cómo se distribuye $\underline{U}$?

19.- Sea X una variable aleatoria con distribución uniforme sobre el intervalo $(-1, 1)$, encontrar la función de densidad de $U = X^2$.

20.- Sea X una variable aleatoria con distribución uniforme sobre el intervalo $(-1, 3)$, encontrar la función de densidad de $U = X^2$.

21.- Sea X_1 y X_2 una muestra aleatoria de tamaño 2 de una distribución normal estándar.

a) Encontrar la función de densidad conjunta de $U_1 = (X_1 + X_2)/\sqrt{2}$ y $U_2 = (X_1 - X_2)/\sqrt{2}$.
b) ¿Cómo se distribuye $\underline{U} = (U_1, U_2)$?
c) ¿Son independientes U_1 y U_2?
d) ¿Cómo se distribuye U_i, para $i = 1, 2$?

22.- Sea X una variable aleatoria con distribución uniforme sobre el intervalo $(0, 1)$. Encontrar la función de densidad de $U = 1/X$.

23.- Sea X la proporción de tiempo que una máquina está parada en un año, se sabe que X tiene una distribución beta con parámetros a y b. Encontrar la distribución de la proporción de tiempo que la máquina está funcionando al año.

24.- Sea X una variable aleatoria con distribución exponencial con media igual a 1. Encontrar la función de densidad de $U = X/(1 + X)$.

25.- Si X es una variable aleatoria con la siguiente función de densidad $f_X(x) = \dfrac{1}{\pi(1 + x^2)}$, encontrar la distribución de $U = 1/X$.

26.- Sea $U = X^2$, encontrar la función de densidad de U.

a) Si $X \sim U(0, \theta)$.
b) Si $X \sim U(-\theta, \theta)$.
c) Si $X \sim U(-2\theta, \theta)$.

27.- Sean X_1 y X_2 variables aleatorias independientes tal que tienen distribuciones $Beta(2, 1)$ y $Beta(1, 2)$, respectivamente. Encontrar la función de densidad de la variable $U = X_1 + X_2$ por medio de la técnica de cambio de variable.

28.- Sea X_1 y X_2 una muestra aleatoria de una distribución exponencial con parámetro λ.

a) Encontrar la función de densidad conjunta de $U_1 = X_1/X_2$ y $U_2 = X_1 + X_2$.
b) ¿Cómo se distribuye U_1?
c) ¿Cómo se distribuye U_2?

29.- Sean X_1 y X_2 variables aleatorias con función de densidad conjunta dada por $f_{\underline{X}}(x_1, x_2) = 2e^{-(x_1+x_2)}I_{[0,x_2)}(x_1)I_{[0,\infty)}(x_2)$.

a) Encontrar la función de densidad conjunta de $U_1 = X_1$ y $U_2 = X_1 + X_2$.
b) Encontrar la función de densidad marginal de U_2.

30.- Sean X_1 y X_2 variables aleatorias con función de densidad conjunta dada por $f_{\underline{X}}(x_1, x_2) = 3(x_1+x_2)I_{[0,1-x_2]}(x_1)I_{[0,1]}(x_2)$. Sean $U_1 = X_1+X_2$ y $U_2 = X_2 - X_1$.

a) Encontrar la función de densidad conjunta de U_1 y U_2.
b) Encontrar las funciones marginales de U_1 y U_2.

31.- Si X es una variable aleatoria con distribución exponencial con media igual a 1, encontrar la función de densidad de $U = e^{X/2}$.

32.- Sean X_1 y X_2 dos variables aleatorias con la siguiente función de densidad conjunta: $f_{\underline{X}}(x_1, x_2) = 8x_1x_2I_{(0,x_2)}(x_1)I_{(0,1)}(x_2)$. Sean $U_1 = X_1/X_2$ y $U_2 = X_2$.

a) Encontrar la función de densidad conjunta de U_1 y U_2.
b) ¿Cómo se distribuye U_i, para $i = 1, 2$?

33.- Sea X_1 y X_2 una muestra aleatoria de tamaño 2 de una distribución exponencial con parámetro λ.

a) Encontrar la función de densidad conjunta de U_1 y U_2 donde $U_1 = X_1$ y $U_2 = 2X_1 + 5X_2$.
b) A partir de la función de densidad conjunta de U_1 y U_2, encontrar las funciones marginales de U_1 y U_2.

34.- Sea X una variable aleatoria con función de densidad dada por $f_X(x) = 2x^{-2}I_{(1,2)}(x)$. Encontrar la función de densidad de $U = X^{1/2}$.

35.- Sea X_1 y X_2 una muestra aleatoria de tamaño 2 de una distribución exponencial con parámetro λ. Encontrar la función de densidad conjunta de U_1 y U_2 donde $U_1 = 3X_1 + 2X_2$ y $U_2 = 2X_1 + 3X_2$.

36.- Sean X_1 y X_2 variables aleatorias con función de densidad conjunta dada por $f_{\underline{X}}(x_1, x_2) = I_{[2x_2,2]}(x_1)I_{[0,1]}(x_2)$. Encontrar la función de densidad de $U = X_1 - X_2$.

a) Por medio del método de función de distribución.
b) Por medio de la técnica de cambio de variable.

37.- Sea X_1 y X_2 una muestra aleatoria de una distribución geométrica con parámetro p. Encontrar la distribución de $U = X_1 + X_2$, por medio del método discreto.

38.- Para los siguientes incisos se define a la variable aleatoria U de la siguiente manera: $U = \sum_{i=1}^{n} X_i$. Demostrar

a) Si $X_1, X_2, \ldots, X_n$ son variables aleatorias independientes e idénticamente distribuidas como *Bernouilli*(p), entonces, $U \sim B(n, p)$.
b) Si $X_1, X_2, \ldots, X_n$ son variables aleatorias independientes tal que $X_i \sim B(m_i, p), i = 1, \ldots, n$, entonces, $U \sim B\big(\sum_{i=1}^{n} m_i, p\big)$.
c) Si $X_1, X_2, \ldots, X_n$ son variables aleatorias independientes e idénticamente distribuidas, con distribución común *Geométrica*(p), entonces, $U \sim BN(r = n, p)$.
d) Si $X_1, X_2, \ldots, X_n$ son variables aleatorias independientes tal que $X_i \sim BN(r_i, p), i = 1, \ldots, n$, entonces, $U \sim BN\big(\sum_{i=1}^{n} r_i, p\big)$.
e) Si $X_1, X_2, \ldots, X_n$ son variables aleatorias independientes e idénticamente distribuidas como *Exp*(λ), entonces, $U \sim Gamma(r = n, \lambda)$.
f) Si $X_1, X_2, \ldots, X_n$ son variables aleatorias independientes tal que X_i tiene distribución $Gamma(r_i, \lambda)$, para $i = 1, \ldots, n$, entonces, $U \sim Gamma\big(\sum_{i=1}^{n} r_i, \lambda\big)$.
g) Si $X_1, X_2, \ldots, X_n$ son variables aleatorias independientes tal que $X_i \sim \chi^2(\nu_i)$, para $i = 1, \ldots, n$, entonces, $U \sim \chi^2\big(\sum_{i=1}^{n} \nu_i\big)$.

39.- Demostrar el teorema 6.7.

40.- La cantidad en kilogramos de materia prima **A** que se necesita para fabricar un artículo es una variable aleatoria X_1 la cual se distribuye normal con media 5 y varianza 0.25, la cantidad de materia prima **B** en litros que se necesita para la fabricación de este artículo es otra variable aleatoria X_2, la cual tiene una distribución normal con media 0.5 y varianza 0.01. La materia **A** cuesta \$8.00 por kilogramo y la materia **B** cuesta \$6.00 por cada litro. Además, existe un gasto fijo de manufactura de \$10.00. Entonces, el costo total de manufactura por artículo es $C = 10 + 8X_1 + 6X_2$. Suponiendo que las variables X_1 y X_2 son independientes

a) Calcular el valor esperado de C.
b) Encontrar la distribución de C.
c) Se van a fabricar 1,000 artículos, ¿cuánto dinero se debe presupuestar para que el costo real exceda la cantidad presupuestada un 0.5 % de las veces?

41.- Sea X_1 y X_2 una muestra aleatoria de tamaño 2 de una distribución normal con media μ y varianza σ^2. Sean $U_1 = X_1 + X_2$ y $U_2 = X_1 - X_2$. Demostrar por medio del método de función generadora de momentos que U_1 y U_2 son variables aleatorias independientes y además, se distribuyen cada una como normal. Indicar las medias y varianzas de estas distribuciones.

42.- Sean X_1 y X_2 variables aleatorias independientes, tal que X_1 tiene distribución binomial con parámetros n_1 y p y X_2 se distribuye binomial con parámetros n_2 y p.

   a) Demostrar que $n_2 - X_2 \sim B(n_2, 1-p)$.
   b) Si $p = 0.5$, demostrar que $X_1 - X_2 + n_2$ tiene una distribución $B(n_1 + n_2, 0.5)$.

43.- Sean X_1 y X_2 variables aleatorias con distribución conjunta normal bivariada con parámetros μ_1, μ_2, σ_1^2, σ_2^2 y ρ. Demostrar que $U_1 = \dfrac{X_1 - \mu_1}{\sigma_1}$ y $U_2 = \dfrac{X_2 - \mu_2}{\sigma_2\sqrt{1-\rho^2}} - \rho\left(\dfrac{X_1 - \mu_1}{\sigma_1\sqrt{1-\rho^2}}\right)$ son variables aleatorias independientes con distribución normal.

44.- Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de tamaño n de una distribución normal con media μ y varianza σ^2, sean $U_1 = \sum_{i=1}^n a_i X_i$ y $U_2 = \sum_{i=1}^n b_i X_i$, donde $a_1, \ldots, a_n$ y $b_1, \ldots, b_n$ son constantes.

   a) Encontrar la distribución conjunta de U_1 y U_2, por medio del método de función generadora de momentos.
   b) Observando la función generadora de U_1 y U_2, indicar cuando U_1 y U_2 son variables aleatorias independientes.

45.- Sea X la proporción de tiempo que una maquinaria está parada por fallo o mantenimiento en un mes, se sabe que esta variable tiene distribución beta con parámetros $a = 2$ y $b = 1$. Sea X_1, X_2, X_3 una muestra aleatoria de tamaño 3 de esta variable, calcular la probabilidad de que el más pequeño de la muestra exceda a la mediana de la distribución.

46.- Sea $X_1, \ldots, X_5$ una muestra aleatoria de tamaño 5 de una distribución uniforme discreta sobre el conjunto $\{1, \ldots, 6\}$. Demostrar que la función de densidad del más pequeño de la muestra, Y_1, es $f_{Y_1}(y_1) = \left[\left(\dfrac{7-y_1}{6}\right)^5 - \left(\dfrac{6-y_1}{6}\right)^5\right] I_{\{1,\ldots,6\}}(y_1)$.

47.- Sean $Y_1, \ldots, Y_5$ las estadísticas de orden de una muestra aleatoria de tamaño 5 de una distribución exponencial con media igual a 1. Demostrar que $U_1 = Y_2$ y $U_2 = Y_4 - Y_2$ son variables aleatorias independientes.

48.- Sean $Y_1, \ldots, Y_n$ las estadísticas de orden de una muestra aleatoria de tamaño n de una distribución de Weibull con parámetros a y b. Encontrar la distribución de Y_1.

49.- Sean Y_1, Y_2, Y_3 las estadísticas de orden de una muestra aleatoria de tamaño 3 de una distribución beta con parámetros $a = 2$ y $b = 1$. Sean $U_1 = Y_1/Y_2, U_2 = Y_2/Y_3$ y $U_3 = Y_3$.

a) Demostrar que U_1, U_2 y U_3 son variables aleatorias independientes.
b) ¿Cómo se distribuye cada variable $U_i, i = 1, 2, 3$? Justificar bien la respuesta.

50.- Considérese una muestra aleatoria de tamaño 2 de una distribución beta con parámetros $a = 1$ y $b = 2$. Obtener la probabilidad de que un elemento de la muestra es al menos tres veces más grande que el otro.

51.- Sean Y_1, Y_2, Y_3 las estadísticas de orden de una muestra aleatoria de tamaño 3 de una distribución uniforme sobre el intervalo $(0, 1)$.

a) Encontrar la función de densidad de $U = Y_1 + Y_3$.
b) Encontrar la función de densidad de $U = \dfrac{Y_1 + Y_3}{2}$.

52.- Sean X_1 y X_2 variables aleatorias independientes, tal que X_1 se distribuye beta con parámetros $a = 3$ y $b = 1$ y X_2 se distribuye beta con parámetros $a = 1$ y $b = 3$. Encontrar la función de densidad conjunta de U_1 y U_2, donde $U_1 = min(X_1, X_2)$ y $U_2 = max(X_1, X_2)$.

53.- Sean $Y_1, \ldots, Y_n$ las estadísticas de orden de una muestra aleatoria de tamaño n de una distribución exponencial con media igual a 1. Se definen las siguientes variables aleatorias: $U_1 = nY_1, U_2 = (n-1)(Y_2 - Y_1), U_3 = (n-2)(Y_3 - Y_2), \ldots, U_n = Y_n - Y_{n-1}$.

a) Demostrar para $n = 2$ que las variables aleatorias U_1, U_2 son independientes.
b) Demostrar que las variables aleatorias $U_1, U_2, \ldots, U_n$ son independientes.
c) ¿Cómo se distribuye $U_i, i = 1, \ldots, n$?
d) Demostrar que cualquier combinación lineal de $Y_1, \ldots, Y_n$, esto es, $\sum_{i=1}^{n} a_i Y_i$, puede expresarse como una función lineal de variables aleatorias independientes.

54.- Sea X_1, X_2 una muestra aleatoria de tamaño 2 de una distribución normal estándar, encontrar la función de densidad del rango muestral.

55.- Sea $\underline{X} = (X_1, X_2)$ un vector aleatorio con una distribución normal bivariada con parámetros μ_1, μ_2, σ_1^2, σ_2^2 y ρ.

a) Demostrar que la variable $U = aX_1 + bX_2 + c$ tiene distribución $N(a\mu_1 + b\mu_2 + c, a^2\sigma_1^2 + 2ab\rho\sigma_1\sigma_2 + b^2\sigma_2^2)$, donde a, b y c son constantes.
b) ¿Cómo se distribuye $X_1 + X_2$?
c) Si $\mu_1 = 10$, $\mu_2 = 15$, $\sigma_1^2 = 4$, $\sigma_2^2 = 16$, $\rho = 0.5$ y $U = -2X_1 + 5X_2$, calcular $P[25 < U < 50]$.
d) Si $\mu_1 = 10$, $\sigma_1^2 = 12$, $\mu_2 = -5$, $\sigma_2^2 = 5$ y $Cov(X_1, X_2) = 4$, calcular $P[X_1 + X_2 > 10]$.

56.- Sea $X_1, \ldots, X_{25}$ una muestra aleatoria de una distribución normal con media 0 y varianza 16, y $Y_1, \ldots, Y_{25}$ otra muestra aleatoria, independiente de la anterior, de una distribución normal con media 1 y varianza 9. Calcular $P[\bar{X} > \bar{Y}]$.

57.- Sea $X_1, \ldots, X_n$ una muestra aleatoria de tamaño n de una distribución normal con media μ y varianza σ^2, encontrar la media y la varianza de la desviación estándar muestral S, donde esta se define como $S = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}}$.

58.- Sea X_1, X_2 una muestra aleatoria de tamaño 2 de una distribución normal con media 0 y varianza σ^2. Encontrar la distribución U, donde

a) $U = \dfrac{X_1 - X_2}{\sqrt{2}\sigma}$.

b) $U = \dfrac{X_1 + X_2}{\sqrt{2}\sigma}$.

c) $U = \dfrac{(X_1 - X_2)^2}{(X_1 + X_2)^2}$.

d) $U = \dfrac{\sum_{i=1}^{2}(X_i - \bar{X})^2}{\sigma^2}$.

e) $U = \dfrac{X_1 - X_2}{\sqrt{(X_1 + X_2)^2}}$.

f) $U = \dfrac{X_1^2}{X_2^2}$.

59.- Sea $X_1, \ldots, X_n$ una muestra aleatoria de tamaño n de una distribución normal estándar. Se definen las siguientes variables aleatorias:

$$\bar{X}_k = \frac{1}{k}\sum_{i=1}^{k} X_i,$$

$$\bar{X}_{n-k} = \frac{1}{n-k}\sum_{i=k+1}^{n} X_i,$$

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

$$S_k^2 = \frac{1}{k-1}\sum_{i=1}^{k}(X_i - \bar{X}_k)^2,$$

$$S_{n-k}^2 = \frac{1}{n-k-1}\sum_{i=k+1}^{n}(X_i - \bar{X}_{n-k})^2,$$

$$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2.$$

Encontrar la distribución de U, donde

a) $U = \bar{X}_k + \bar{X}_{n-k}$.
b) $U = k\bar{X}_k^2 + (n-k)\bar{X}_{n-k}^2$.
c) $U = \dfrac{X_1^2}{X_n^2}$.
d) $U = \dfrac{X_1}{\sqrt{X_k^2}}$.
e) $U = (k-1)S_k^2 + (n-k-1)S_{n-k}^2$.
f) $U = \dfrac{S_k^2}{S_{n-k}^2}$.
g) $U = \sqrt{n}\left(\dfrac{\bar{X}_n}{S_n}\right)$.

60.- Sea $X_1, \ldots, X_{15}$ una muestra aleatoria de tamaño 15 de una distribución normal con media 0 y varianza σ^2. Calcular el

a) 0.90-cuántil de $\sqrt{15}\left(\dfrac{\bar{X}}{S}\right)$, donde $\bar{X}$ y S son la media y desviación estándar de la muestra aleatoria $X_1, \ldots, X_{15}$.
b) 0.95-ésimo cuantil de $\dfrac{\sum_{i=1}^{4} X_i^2}{\sum_{j=8}^{11} X_j^2}$.
c) 0.99-ésimo cuantil de $\dfrac{\sum_{i=1}^{6} X_i^2}{\sum_{j=7}^{15} X_j^2}$.
d) 0.90-ésimo cuantil de $\displaystyle\sum_{i=1}^{15}(X_i - \bar{X})^2$.
e) 0.05-ésimo cuantil de $\dfrac{\bar{X}}{S}$.

61.- Considerando una muestra aleatoria de una distribución continua, calcular la probabilidad de que la mayor de las observaciones exceda a la mediana. Si la muestra es de tamaño

a) $n = 2$.
b) n.

62.- Sean $Y_1, \ldots, Y_{10}$ las estadísticas de orden de una muestra aleatoria de tamaño 10 de una distribución uniforme sobre el intervalo $(0,1)$. Calcular $E\left[(1-Y_1)^5\right]$.

63.- Sean $Y_1, \ldots, Y_5$ las estadísticas de orden de una muestra aleatoria de tamaño 5 de una distribución beta con parámetros $a = 2$ y $b = 1$. Encontrar

a) La distribución de Y_5.
b) La media y la varianza de Y_5.

64.- Sean $Y_1, \ldots, Y_{17}$ las estadísticas de orden de una muestra aleatoria de tamaño 17 de una distribución uniforme sobre el intervalo $(0, b)$.

a) Para $b = 0.5$, encontrar la función de densidad de Y_1.
b) Para $b = 1$, encontrar la distribución de Y_1.

65.- Considerando el ejemplo 6.34, por medio del método de función de distribución, encontrar la distribución del rango.

66.- Sean $Y_1, \ldots, Y_n$ las estadísticas de orden de una muestra aleatoria de tamaño n de una distribución uniforme sobre el intervalo $(0, 1)$, encontrar la distribución del rango muestral. El resultado de este ejercicio es una generalización del ejemplo 6.34.

67.- Sean $Y_1, \ldots, Y_n$ las estadísticas de orden de una muestra aleatoria de tamaño n de una distribución exponencial con parámetro λ. Encontrar la distribución del mínimo Y_1.

68.- Se sabe que los tiempos que tardan 3 estudiantes E1, E2 y E3 en terminar un examen de probabilidad se distribuyen exponencialmente con medias de 50, 60 y 45 minutos, respectivamente. Suponiendo que estos tiempos entre estudiantes son independientes, calcular las siguientes probabilidades:

a) Que el primer estudiante que termine sea antes de una hora.
b) Que todos los estudiantes terminen antes de una hora.

69.- Resolver el ejemplo 6.14, pero ahora usando la fórmula 6.2.

70.- Resolver de nuevo el ejercicio 37, pero ahora usando la fórmula de convolución 6.1.

71.- Para cierta póliza de seguros de gastos médicos mayores, la cantidad reclamada se distribuye normal con media μ y varianza σ^2. La compañía, por cada reclamo, pagará el 90 % de la pérdida menos un deducible de \$30,000.

a) Encontrar la función generadora de momentos de la cantidad pagada por la aseguradora por un reclamo.
b) ¿Cómo se distribuye la cantidad pagada por la compañía por un reclamo?

72.- Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una distribución normal con media 0 y varianza σ^2. Sean $\bar{X}$ y S^2 la media y varianza muestrales, respectivamente. Encontrar

a) 0.05-ésimo cuantil de $\displaystyle\sum_{i=1}^{10} \frac{X_i^2}{\sigma^2}$.

b) 0.10-ésimo cuantil de $\sum_{i=1}^{10} X_i^2$.

c) 0.90-ésimo cuantil de $15\left(\frac{\bar{X}}{S}\right)$.

d) 0.10-ésimo cuantil de $\frac{\sum_{i=1}^{4} X_i^2}{\sum_{j=11}^{15} X_i^2}$.

73.- Sea X una variable aleatoria con distribución $Beta(a, b)$. Consideremos una transformación lineal de X, esto es, $Y = c + dX$. Encontrar la función de densidad de Y.

Convergencia de sucesiones de variables aleatorias

Capítulo

Introducción

En este capítulo se tratará el tema de convergencia de sucesiones de variables aleatorias. Muchos de los resultados de convergencia que se presentarán tienen aplicaciones muy importantes sobre todo en materia de inferencia estadística y simulación de Monte Carlo. Entre los resultados más importantes que se abordarán, se encuentran la ley de los grandes números y el teorema de límite central el cual repercute fuertemente en varias aplicaciones.

Siempre que existe un resultado límite de sucesiones de variables aleatorias, existe la posibilidad de aplicarlo, ya sea como una aproximación al cálculo de una probabilidad o a un resultado de estadística, entre otras aplicaciones. Cabe mencionar, varios de los resultados de convergencia que se presentarán se les podrá dar una interpretación estadística.

Es importante mencionar que la definición de convergencia de sucesiones de números reales no es apropiada para definir la convergencia de sucesiones de variables aleatorias, es necesario definir esta convergencia de otra manera. Algunas definiciones se van a presentar más adelante.

No obstante, antes de iniciar el tema de sucesiones de variables aleatorias, se recomienda al lector repasar los conceptos básicos de sucesiones de números reales, véase la sección C.2 del apéndice C, donde se trata la definición de convergencia de sucesiones de números reales y algunos ejemplos.

7.1. Sucesión de variables aleatorias

En esta sección se explicará qué es una sucesión de variables aleatorias, además se presentarán algunos ejemplos, con su correspondiente interpretación.

Si para una sucesión $\{U_n\}$, cada elemento $U_1, U_2, \ldots$ es una variable aleatoria, se dice que es una *sucesión de variables aleatorias.*

En una sucesión de números reales $\{U_n\}$, es posible saber el valor de cada elemento, esto es, para cada valor específico de n, se puede saber el valor U_n.

Pero en el caso de una sucesión $\{U_n\}$ de variables aleatorias, al conocer n, tendremos conocimiento de la variable aleatoria U_n, posiblemente también conocimiento de su recorrido y de su distribución, pero, por lo general, no tendremos conocimiento del valor de la variable.

Por lo anterior, la definición de convergencia de sucesión de números reales no puede aplicarse para definir la convergencia de sucesiones de variables aleatorias. Si una sucesión de variables aleatorias converge, existen dos posibilidades, que converja a una constante o a una variable aleatoria. En cualquier caso, no se puede garantizar para ninguna $m \in \mathbb{N}$, que la distancia entre la sucesión y el límite al que converge siempre sea inferior a un $\epsilon > 0$ dado, ya que la distancia es aleatoria.

Por ejemplo, sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución normal con media μ y varianza σ^2. Consideremos la sucesión $\{U_n\} = \{\bar{X}_n\}$, donde $\bar{X}_n$ es la media muestral. Se sabe que $\bar{X}_n \sim N(\mu, \sigma^2/n)$, esto es, para cada n la distribución de $\bar{X}_n$ es diferente (la varianza es diferente).

En la siguiente figura, se presentan diferentes gráficas de la función de densidad de $\bar{X}_n$, para $n = 1, 10, 20, 40, 200, 400$, $\mu = 0$ y $\sigma^2 = 1$.

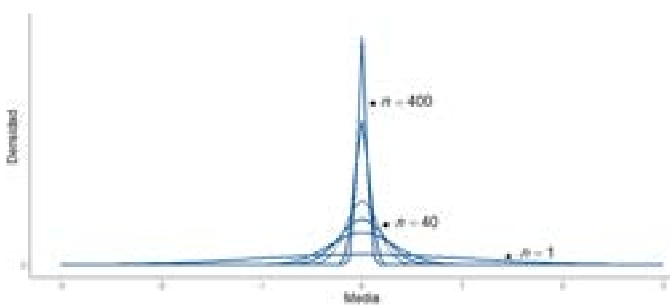


Figura 7.1. *Funciones de densidad de la media muestral*

Como complemento a estas gráficas, se presenta una tabla, donde se observa, para diferentes valores de n, el rango de valores para la media muestral, considerando siempre una probabilidad del 0.9974.

n	Rango de valores de $\bar{X}$ al 99.74 %
1	$(-3, 3)$
10	$(-0.94869, 0.94869)$
20	$(-0.67083, 0.67083)$
40	$(-0.47433, 0.47433)$
200	$(-0.21213, 0.21213)$
400	$(-0.13416, 0.13416)$

Tabla 7.1. *Rango de valores de la media muestral*

Algunas observaciones se dan a continuación: los valores de $\bar{X}_n$ siempre estarán alrededor de la media (de cero). Entre más grande sea el valor de n, más esbelta es la función de densidad y, por lo tanto, más estrecho el rango de valores de $\bar{X}_n$, lo que significa que la varianza va decreciendo. Es notorio que los valores de $\bar{X}_n$ se van acercando a la media (a cero), conforme el tamaño de la muestra se va incrementando, esta situación, con una probabilidad grande.

Por lo anterior, se puede intuir que $\bar{X}_n$ tiende a la media (a cero), conforme la n crece. En los ejemplos 7.1 y 7.2, se tratará formalmente con la convergencia de la anterior sucesión.

Cabe mencionar que la formalización y generalización del anterior ejemplo es conocido como la *ley de los grandes números*, la cual se presentará en el teorema 7.6.

Consideremos otro ejemplo, sean $Y_1, \ldots, Y_n$ las estadísticas de orden de una muestra aleatoria de tamaño n de una distribución uniforme sobre el intervalo $(0, 1)$. Se define la siguiente sucesión de variables aleatorias $\{U_n\} = \{Y_1\}$. Es importante resaltar que la distribución de Y_1 depende del tamaño de la muestra n, esto es, se puede comprobar que $Y_1 \sim Beta(1, n)$. Lo que significa que la distribución de Y_1 va cambiando conforme n crece. En la siguiente figura se presentan las funciones de densidad de Y_1 para los valores de $n = 10, 20, 50, 100$.

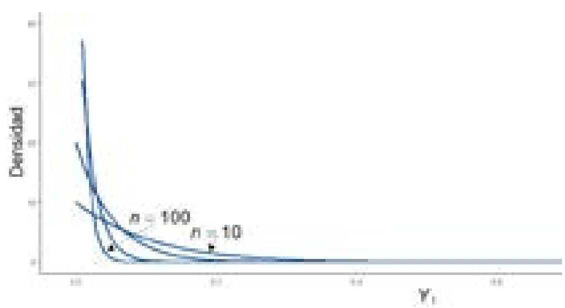


Figura 7.2. *Funciones de densidad del mínimo*

La gráfica de la función de densidad más alejada al eje vertical corresponde a la muestra de tamaño $n = 10$, y la gráfica más cercana al eje vertical corresponde a la muestra más grande, $n = 100$, lo que significa que entre más grande sea n, la probabilidad de que Y_1 tome valores más cercanos a cero crece. Por lo anterior, se puede intuir que la sucesión $\{U_n\} = \{Y_1\}$ converge a cero. El ejercicio 4 de este capítulo considera la convergencia de esta sucesión, se le sugiere al estudiante, más adelante, resolver este problema.

Es importante definir la convergencia de sucesiones de variables aleatorias de manera diferente a la definición tradicional de convergencia de sucesiones de números reales. En la siguiente sección, se tratarán las definiciones más importantes de convergencia de sucesiones de variables aleatorias.

7.2. Tipos de convergencia

La definición de convergencia de sucesiones de variables aleatorias tendrá que ser diferente a la definición tradicional de convergencia de sucesiones, cualquier forma de indicar que una sucesión se acerca a un límite, la distancia considerada será aleatoria, por lo que será necesario explicar el acercamiento por medio de una probabilidad o por medio de un valor esperado.

A continuación, se presentará la definición de variable aleatoria degenerada y sus propiedades, las cuales son muy importantes para algunos ejemplos de convergencia de sucesiones de variables aleatorias.

Definición 7.1. Se dice que una variable aleatoria X es *degenerada* en c, si toma este valor con probabilidad igual a 1, esto es, si $P[X = c] = 1$.

Teorema 7.1. *Una variable aleatoria degenerada en c tiene función de densidad, función de distribución, valor esperado, varianza y función generadora de momentos respectivamente, como*

$$f_X(x) = I_{\{c\}}(x),$$

$$F_X(X) = I_{[c,\infty)}(x),$$
$$E[X] = c,$$
$$V[X] = 0,$$
$$m_X(t) = e^{ct}.$$

El ejercicio 33 del capítulo 2, considera la demostración de este resultado.

Dicho de otra manera, una variable aleatoria degenerada es una variable aleatoria discreta que toma un solo valor.

En la siguiente gráfica se ilustra la función de distribución de una variable aleatoria degenerada en c, la cual es importante para algunos ejemplos de convergencia que se tratarán más adelante.

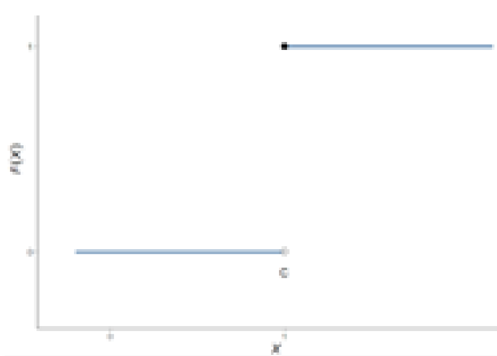


Figura 7.3. *Función de distribución de una variable aleatoria degenerada en c*

Por otro lado, para cualquier espacio de probabilidad $(\Omega, \mathcal{A}, P[A])$, se dice que el espacio muestral Ω es el *evento seguro*, y el conjunto vacío $\emptyset$ es el *evento imposible*.

Además, se dice que un evento $A \neq \Omega$ es un *evento casi seguro* si $P[A] = 1$, y un evento $A \neq \emptyset$ se considera un *evento casi imposible* si $P[A] = 0$.

Estos últimos conceptos son importantes en las definiciones de convergencia de sucesiones de variables aleatorias. A continuación, se presentarán cuatro modos de convergencia, así como algunos ejemplos.

Empezaremos con la definición de convergencia en probabilidad.

Definición 7.2. Convergencia en probabilidad. Sea $\{U_n\}$ una sucesión de variables aleatorias y U otra variable aleatoria, todas las variables $U_1, U_2, \ldots,$ y U en el mismo espacio de probabilidad $(\Omega, \mathcal{A}, P[A])$. Se dice que $\{U_n\}$ *converge en probabilidad* a U, si para todo $\epsilon > 0$.

$$P\left[|U_n - U| < \epsilon\right] \xrightarrow[n\to\infty]{} 1,$$

o en forma equivalente

$$P\left[|U_n - U| \geq \epsilon\right] \xrightarrow[n\to\infty]{} 0.$$

Notación. $\{U_n\} \xrightarrow{p} U$ denota, $\{U_n\}$ converge en probabilidad a U.

Obsérvese el siguiente ejemplo de convergencia en probabilidad.

Ejemplo 7.1. Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una distribución normal estándar, sea $U_n = \bar{X}_n$, demostrar que $\{U_n\}$ converge en probabilidad a 0.

Sea $\epsilon > 0$, obsérvese

$$\begin{aligned} P[|\bar{X}_n - 0| < \epsilon] &= P[-\epsilon < \bar{X}_n < \epsilon] \\ &= P\left[-\frac{\epsilon - 0}{1/\sqrt{n}} < \frac{\bar{X}_n - 0}{1/\sqrt{n}} < \frac{\epsilon - 0}{1/\sqrt{n}}\right] \\ &= P[-\sqrt{n}\epsilon < Z < \sqrt{n}\epsilon] \xrightarrow[n\to\infty]{} 1. \end{aligned}$$

Por lo tanto, $\{U_n\} = \{\bar{X}_n\} \xrightarrow{p} 0$.

A continuación, se dará la definición de convergencia en distribución.

Definición 7.3. Convergencia en distribución. Sea $\{U_n\}$ una sucesión de variables aleatorias, donde $F_{U_n}(u)$ es la función de distribución de U_n, $n = 1, 2, \ldots$, y sea U otra variable aleatoria con función de distribución $F_U(u)$, todas las variables aleatorias $U_1, U_2, \ldots,$ y U en el mismo espacio de probabilidad $(\Omega, \mathcal{A}, P[A])$. Se dice que $\{U_n\}$ *converge en distribución* a U, si para todo punto de continuidad de la función $F_U(u)$, se cumple

$$\lim_{n\to\infty} F_{U_n}(u) = F_U(u).$$

A la función de distribución $F_U(u)$, se le llama *función de distribución límite* de la sucesión $\{U_n\}$.

Notación. $\{U_n\} \xrightarrow{d} U$ denota, $\{U_n\}$ converge en distribución a U.

La definición de convergencia en distribución, se enfoca en la convergencia o no de $\{F_{U_n}(u)\}$, la cual tiene las siguientes posibilidades: que sea convergente a una función de distribución continua; que sea convergente a una función de distribución discreta en sus puntos de continuidad; que sea convergente a algo que no es función de distribución en sus puntos de continuidad, o que sea divergente.

Una aclaración, cuando se pida encontrar la distribución límite de una sucesión de variables aleatorias, se estará solicitando encontrar la convergencia en distribución de la sucesión.

En seguida, se presentará un ejemplo.

Ejemplo 7.2. Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una distribución normal estándar, encontrar la distribución límite de $\{U_n\}$ donde $U_n = \bar{X}_n$.

Se sabe que $\bar{X}_n \sim N(0, 1/n)$.

Entonces,

$$\begin{aligned} F_{\bar{X}_n}(u) &= P[\bar{X}_n \leq u] \\ &= P\left[\frac{\bar{X}_n - 0}{1/\sqrt{n}} \leq \frac{u-0}{1/\sqrt{n}}\right] \\ &= P[Z \leq \sqrt{n}u], \text{ donde } Z \sim N(0,1). \end{aligned}$$

Por lo tanto,

$$\lim_{n\to\infty} F_{U_n}(u) = \lim_{n\to\infty} P[Z \leq \sqrt{n}u].$$

El cálculo del anterior límite se debe realizar por casos, ya que depende del valor que toma u.

Caso 1. Cuando $u < 0$.

$$\lim_{n\to\infty} F_{u_n}(x) = \lim_{n\to\infty} P[Z \leq \sqrt{n}u] = 0.$$

Caso 2. Cuando $u = 0$.

$$\lim_{n\to\infty} F_{U_n}(u) = \lim_{n\to\infty} P[Z \leq \sqrt{n}u] = \lim_{n\to\infty} P[Z \leq 0] = 0.5.$$

Caso 3. Cuando $u > 0$.

$$\lim_{n\to\infty} F_{U_n}(u) = \lim_{n\to\infty} P[Z \leq \sqrt{n}u] = 1.$$

Esto es,

$$\lim_{n\to\infty} F_{U_n}(u) = 0.5 I_{\{0\}}(u) + I_{(0,\infty)}(u).$$

A continuación, se presenta la gráfica de este límite.

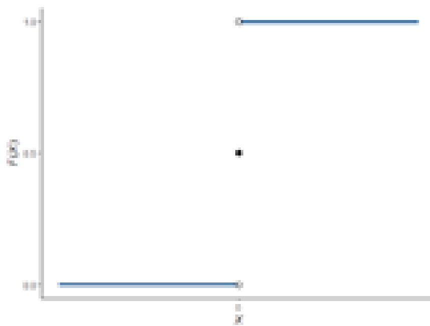


Figura 7.4. *Límite de la función de distribución de la media muestral, ejemplo 7.2*

La anterior función y la función de distribución $F_U(u) = I_{[0,\infty)}(u)$ son iguales para todo valor de $u \in (-\infty, 0) \cup (0, \infty)$, esto es, para todo punto de continuidad de la función $F_U(u)$. Al ser $F_U(u)$ la función de distribución de una variable aleatoria degenerada en 0, se concluye que $\{U_n\} \xrightarrow{d} U = 0$.

Se deja como ejercicio para el estudiante, demostrar $\{U_n\} = \{\bar{X}_n\} \xrightarrow{d} \mu$, cuando la muestra $X_1, \ldots, X_n$ proviene de una distribución $N(\mu, \sigma^2)$.

Las dos definiciones anteriores de convergencia son suficientes para la mayoría de las aplicaciones, no obstante, a continuación, se presentarán dos modos de convergencia más, con la idea de que el estudiante tenga conocimiento de la existencia de otras definiciones de convergencia.

A continuación, se presenta la definición de convergencia casi segura.

Definición 7.4. Convergencia casi segura. Sea $\{U_n\}$ una sucesión de variables aleatorias y U otra variable aleatoria, todas las variables $U_1, U_2, \ldots,$ y U en el mismo espacio de probabilidad $(\Omega, \mathcal{A}, P[A])$. Se dice que $\{U_n\}$ *converge casi seguramente* a U, si

$$P[\lim_{n\to\infty} U_n = U] = 1.$$

Notación. $\{U_n\} \xrightarrow{c.s.} U$ denota, $\{U_n\}$ converge casi seguramente a U.

Una aplicación de este tipo de convergencia es la ley de los grandes números, en su versión «ley fuerte», véase el teorema 7.6.

A continuación, se dará la definición de convergencia en media r-ésima.

Definición 7.5. Convergencia en media r-ésima. Sea $\{U_n\}$ una sucesión de variables aleatorias y U otra variable aleatoria, todas las variables $U_1, U_2, \ldots,$ y U en el mismo espacio de probabilidad $(\Omega, \mathcal{A}, P[A])$. Se dice que $\{U_n\}$ *converge en media r-ésima* a U, si

$$E\left[|U_n - U|^r\right] \xrightarrow[n\to\infty]{} 0.$$

Notación. $\{U_n\} \xrightarrow{r} U$ denota, $\{U_n\}$ converge en media r-ésima a U.

Muchos de los resultados que se tratarán en este tema de convergencia tienen importantes aplicaciones en diferentes temas de estadística, simulación de Monte Carlo y ciencias actuariales, entre otros. En este sentido, como ya se había comentado, los tipos de convergencia que principalmente contribuyen a las aplicaciones son la convergencia en probabilidad y la convergencia en distribución, razón por la cual, los ejemplos tratados en este tema solo consideran estos dos modos.

Observación. Una sucesión de números reales $\{U_n\}$ se puede considerar un caso particular de una sucesión de variables aleatorias, ya que las constantes son un caso particular de variables aleatorias degeneradas, véase el ejercicio 6 de este capítulo.

A continuación, se dará un ejemplo más.

Ejemplo 7.3. Sea $\{U_n\}$ una sucesión de variables aleatorias donde $U_n = n(\theta - Y_n)$ y Y_n es la $n-$ésima estadística de orden de una muestra aleatoria de tamaño n de una distribución $U(0,\theta)$. Encontrar la distribución límite de la sucesión $\{U_n\}$.

El recorrido de U_n es el intervalo $(0, n\theta)$, véase siguiente gráfica

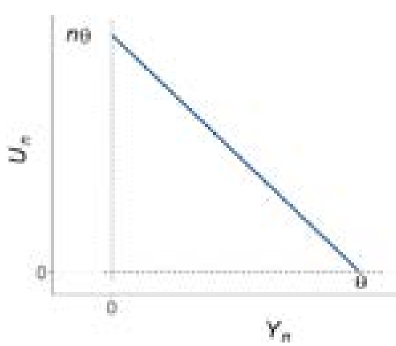


Figura 7.5. *Transformación U_n, ejemplo 7.3*

A continuación, se encontrarán las funciones de distribución de Y_n.

$$
\begin{aligned}
F_{Y_n}(y) &= [F_X(y)]^n \\
&= \left[\frac{y}{\theta}\right]^n, \text{ donde } 0 < y < \theta.
\end{aligned}
$$

Ahora la función de distribución de U_n.

$$
\begin{aligned}
F_{U_n}(u) &= P[U_n \leq u] \\
&= P[n(\theta - Y_n) \leq u] \\
&= P[Y_n \geq \theta - u/n]
\end{aligned}
$$

$$= 1 - F_{Y_n}(\theta - u/n)$$
$$= 1 - \left[\frac{\theta - u/n}{\theta}\right]^n, \text{ donde } 0 < u < n\theta.$$

Por lo tanto,

$$F_{U_n}(u) = \left\{1 - \left[\frac{\theta - u/n}{\theta}\right]^n\right\} I_{(0,n\theta)} + I_{[n\theta,\infty)}.$$

De esta manera,

$$\lim_{n\to\infty} F_{U_n}(u) = \lim_{n\to\infty}\left[1 - \left[1 - \frac{u}{n\theta}\right]^n\right] I_{(0,n\theta)}(u) + \lim_{n\to\infty} I_{[n\theta,\infty)}$$
$$= \left[1 - e^{-u/\theta}\right] I_{(0,\infty)}(u).$$

Se puede observar que en el penúltimo límite se aplicó el resultado límite C.2 del apéndice C.

Concluimos,

$$\{U_n\} \xrightarrow{d} U, \text{ donde } U \sim Exp(\lambda = 1/\theta).$$

7.3. Relaciones entre los tipos de convergencia

En esta parte se tratarán las relaciones que existen entre los cuatro modos de convergencia definidos en la anterior sección. Los siguientes resultados están encaminados a estas relaciones.

El siguiente teorema considera la relación que existe entre la convergencia casi segura y las convergencias en probabilidad y en distribución.

Teorema 7.2. *Si $\{U_n\}$ es una sucesión de variables aleatorias que converge casi seguramente a U, entonces, $\{U_n\}$ converge en probabilidad y también en distribución a U.*

El siguiente resultado trata con la relación existente entre la convergencia en media r-ésima y las convergencias en probabilidad y en distribución.

Teorema 7.3. *Si $\{U_n\}$ es una sucesión de variables aleatorias que converge en media r-ésima a U, entonces, $\{U_n\}$ converge en probabilidad y también en distribución a U.*

Los siguientes dos resultados consideran la relación que existe entre la convergencia en probabilidad y la convergencia en distribución.

Teorema 7.4. *Si $\{U_n\}$ es una sucesión de variables aleatorias que converge en probabilidad a U, entonces, $\{U_n\}$ converge en distribución a U.*

Las demostraciones de los anteriores resultados no son parte de este tema, corresponden a un curso de probabilidad más avanzado.

Teorema 7.5. *Sea $\{U_n\}$ una sucesión de variables aleatorias y c una constante. La sucesión $\{U_n\}$ converge en probabilidad a c si y solo si $\{U_n\}$ converge en distribución a c.*

Demostración. Primero, suponemos que la sucesión $\{U_n\}$ converge en probabilidad a c, esto es,

$$\lim_{n\to\infty} P[|U_n - c| < \epsilon] = 1, \text{ para todo } \epsilon > 0.$$

El propósito es demostrar $\lim_{n\to\infty} F_{U_n}(u) = 1$, cuando $u > c$ y $\lim_{n\to\infty} F_{U_n}(u) = 0$, cuando $u < c$.

Obsérvese

$$\begin{aligned} P[|U_n - c| < \epsilon] &= P[c - \epsilon < U_n < c + \epsilon] \\ &= F_{U_n}[(c+\epsilon)*] - F_{U_n}(c - \epsilon), \end{aligned}$$

donde $F_{U_n}[(c+\epsilon)*] = \lim_{0<h\to 0} F_{U_n}[(c + \epsilon - h)]$ (el límite de $F_{U_n}(u)$ cuando u tiende a $c + \epsilon$ por la izquierda).

Una manera práctica de interpretar $F_{U_n}[(c+\epsilon)*]$ es, $F_{U_n}[(c+\epsilon)*]$ es el valor de la función de distribución de U_n, un instante antes que $c + \epsilon$.

Ahora obsérvese

$$\begin{aligned} 1 &= \lim_{n\to\infty} P[|U_n - c| < \epsilon] \\ &= \lim_{n\to\infty} F_{U_n}[(c+\epsilon)*] - \lim_{n\to\infty} F_{U_n}(c - \epsilon). \end{aligned}$$

Como $F_{U_n}(u)$ toma valores en el intervalo $[0, 1]$ para todo u, la única forma como se puede cumplir que la anterior diferencia sea igual a 1 es que $\lim_{n\to\infty} F_{U_n}[(c+\epsilon)*] = 1$ y $\lim_{n\to\infty} F_{U_n}((c - \epsilon) = 0$, para todo $\epsilon > 0$.

Lo anterior significa

$$\lim_{n\to\infty} F_{U_n}(u) = 1, \text{para todo } u > c,$$

$$\lim_{n\to\infty} F_{U_n}(u) = 0, \text{para todo } u < c.$$

Por lo tanto, $\{U_n\}$ converge en distribución a c.

Ahora, suponemos que $\{U_n\}$ converge en distribución a c. Esto es,

$$\lim_{n\to\infty} F_{U_n}(u) = 1, \text{ si } u > c,$$

$$\lim_{n\to\infty} F_{U_n}(u) = 0, \text{ si } u < c.$$

Sea $\epsilon > 0$, entonces,

$$\begin{aligned} \lim_{n\to\infty} P[|U_n - c| < \epsilon] &= \lim_{n\to\infty} P[c - \epsilon < U_n < c + \epsilon] \\ &= \lim_{n\to\infty} F_{U_n}[(c+\epsilon)*] - \lim_{n\to\infty} F_{U_n}(c - \epsilon) \\ &= 1 - 0 = 1. \end{aligned}$$

Por lo tanto, $\{U_n\}$ converge en probabilidad a c. □

En resumen, las relaciones de los diferentes tipos de convergencia se presentan a continuación:

$$\{U_n\} \xrightarrow{c.s.} U \quad \Longrightarrow \begin{cases} \{U_n\} \xrightarrow{p} U, \\ \{U_n\} \xrightarrow{d} U. \end{cases}$$

$$\{U_n\} \xrightarrow{r} U \quad \Longrightarrow \begin{cases} \{U_n\} \xrightarrow{p} U, \\ \{U_n\} \xrightarrow{d} U. \end{cases}$$

$$\{U_n\} \xrightarrow{p} U \quad \Longrightarrow \{U_n\} \xrightarrow{d} U.$$

$$\{U_n\} \xrightarrow{p} U = c \quad \Longleftrightarrow \quad \{U_n\} \xrightarrow{d} U = c.$$

7.4. Ley de los grandes números

Uno de los teoremas más importantes en el tema de convergencia es la ley de los grandes números. Este resultado tiene aplicaciones importantes, tanto en probabilidad, como en estadística, entre otras áreas. En seguida se presenta este resultado.

Teorema 7.6. Ley de los grandes números. *Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución con una media μ y varianza σ^2, entonces,*

(I) $\{\bar{X}_n\} \xrightarrow{p} \mu$,

(II) $\{\bar{X}_n\} \xrightarrow{c.s.} \mu$.

La parte I del anterior teorema indica que la media muestral converge en probabilidad a la media de la distribución (o media poblacional), cuando el tamaño de la muestra crece. A este resultado se le conoce como la *ley débil de los grandes números.*

En el inciso II del teorema, también se afirma que la media muestral converge a la media poblacional, pero la convergencia es casi segura. A este resultado se le conoce como la *ley fuerte de los grandes números.*

En ambos resultados, la condición para que se cumpla la convergencia es que $X_1, \ldots, X_n$ sea una muestra aleatoria de una distribución con media y varianza finitas.

A continuación, se demostrará la ley débil de los grandes números.

Demostración. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución con media μ y varianza σ^2 finitas.

Se puede verificar fácilmente

$$E[\bar{X}_n] = \mu, \quad \text{y} \quad V[\bar{X}_n] = \frac{\sigma^2}{n}.$$

Por la desigualdad de Chebyshev (teorema 2.8), se puede afirmar

$$P\left[\left|\bar{X}_n - \mu\right| < k\frac{\sigma}{\sqrt{n}}\right] \geq 1 - \frac{1}{k^2}, \text{ donde } k > 0.$$

La desigualdad de Chebyshev es cierta para todo valor de $k > 0$, en particular para $k = \dfrac{\sqrt{n}}{\sigma}\epsilon$, donde $\epsilon > 0$.

Entonces,

$$P\left[\left|\bar{X}_n - \mu\right| < \epsilon\right] \geq 1 - \frac{\sigma^2}{n\epsilon^2}.$$

Además, toda probabilidad es menor o igual a 1, por lo tanto,

$$1 - \frac{\sigma^2}{n\epsilon^2} \leq P\left[\left|\bar{X}_n - \mu\right| < \epsilon\right] \leq 1.$$

De esta manera,

$$\lim_{n\to\infty}\left(1 - \frac{\sigma^2}{n\epsilon^2}\right) \leq \lim_{n\to\infty} P\left[\left|\bar{X}_n - \mu\right| < \epsilon\right] \leq \lim_{n\to\infty} 1.$$

Los límites de la partes izquierda y derecha de la anterior desigualdad son igual a 1, por lo que se concluye

$$\lim_{n\to\infty} P\left[\left|\bar{X}_n - \mu\right| < \epsilon\right] = 1, \text{ para } \epsilon > 0.$$

Por lo tanto,

$$\{\bar{X}_n\} \xrightarrow{p} \mu.$$

□

Hay otra manera de realizar la demostración de la ley débil de los grandes números, usando las funciones generadoras de momentos de la sucesión $\{\bar{X}_n\}$. En este caso, se debe aplicar el teorema 7.7 que se presentará en la siguiente sección. Es importante suponer que la función generadora de momentos de la distribución de X_i existe. Se deja como ejercicio para el estudiante, realizar la demostración por este camino, después de haber estudiado la siguiente sección.

7.5. Convergencia por función generadora de momentos

En esta sección se va a estudiar la convergencia en distribución de sucesiones de variables aleatorias, utilizando la funciones generadoras de momentos de los elementos de la sucesión, en lugar de las funciones de distribución, véase el siguiente teorema.

Teorema 7.7. *Sea $\{U_n\}$ una sucesión de variables aleatorias tal que U_n tiene función de distribución $F_{U_n}(u)$ y función generadora de momentos $m_{U_n}(t)$, la cual existe para toda $t \in (-h, h)$ para alguna $h > 0$. Si existe una variable aleatoria U con función de distribución $F_U(u)$ y función generadora de momentos $m_U(t)$, la cual existe para $t \in (-h_1, h_1)$ para alguna $0 < h_1 < h$, y se cumple $\lim_{n\to\infty} m_{U_n}(t) = m_U(t)$, entonces, $\{U_n\} \xrightarrow{d} U$.*

La demostración de este teorema no es propósito de este tema, es para un curso de probabilidad más avanzado. A continuación, se presentarán algunos ejemplos donde se aplicará este teorema.

Ejemplo 7.4. Sea $\{U_n\}$ una sucesión de variables aleatorias tal que U_n se distribuye binomial con parámetros n y p. Encontrar la distribución límite de $\{U_n\}$ cuando n tiende a ∞ y $np = \lambda$, donde λ es una constante que no depende de n.

Obsérvese

$$\begin{aligned}
\lim_{\substack{n\to\infty \\ np=\lambda}} m_{U_n}(t) &= \lim_{\substack{n\to\infty \\ np=\lambda}} (1 - p + pe^t)^n \\
&= \lim_{n\to\infty} \left(1 - \frac{\lambda}{n} + \frac{\lambda}{n}e^t\right)^n \\
&= \lim_{n\to\infty} \left(1 + \frac{\lambda(e^t - 1)}{n}\right)^n \\
&= e^{\lambda(e^t-1)}.
\end{aligned}$$

Una vez más se aplicó el resultado límite C.2 que se encuentra en el apéndice C.

La anterior es la función generadora de momentos de una variable aleatoria con distribución $Poisson(\lambda)$, por lo tanto, para $np = \lambda$,

$$\{U_n\} \xrightarrow{d} U, \text{ donde } U \sim Poisson(\lambda).$$

Ejemplo 7.5. Sea $\{U_n\}$ una sucesión de variables aleatorias tal que U_n se distribuye ji cuadrada con n grados de libertad. Encontrar la distribución límite de $\{Z_n\}$, donde $Z_n = \frac{U_n - n}{\sqrt{2n}}$.

Se obtendrá la función generadora de momentos de Z_n.

$$\begin{aligned}
m_{Z_n}(t) &= E\left[e^{Z_n t}\right] \\
&= E\left[e^{\left(\frac{U_n-n}{\sqrt{2n}}\right)t}\right] \\
&= E\left[e^{\frac{U_n t}{\sqrt{2n}} - \frac{nt}{\sqrt{2n}}}\right]
\end{aligned}$$

$$
\begin{aligned}
&= e^{-\frac{tn}{\sqrt{2n}}} E\left[e^{U_n \frac{t}{\sqrt{2n}}}\right] \\
&= e^{-\frac{tn}{\sqrt{2n}}} m_{U_n}\left(\frac{t}{\sqrt{2n}}\right) \\
&= e^{-\frac{tn}{\sqrt{2n}}} \left(1 - 2\frac{t}{\sqrt{2n}}\right)^{-\frac{n}{2}} \\
&= e^{-t\sqrt{\frac{2}{n}}\frac{n}{2}} \left(1 - \sqrt{\frac{2}{n}}t\right)^{-\frac{n}{2}} \\
&= \left[e^{t\sqrt{\frac{2}{n}}} \left(1 - \sqrt{\frac{2}{n}}t\right)\right]^{-\frac{n}{2}} \\
&= \left[e^{t\sqrt{\frac{2}{n}}} - t\sqrt{\frac{2}{n}}e^{t\sqrt{\frac{2}{n}}}\right]^{-\frac{n}{2}}.
\end{aligned}
$$

Se desarrollará con cuatro sumandos $e^{t\sqrt{\frac{2}{n}}}$,

$$
e^{t\sqrt{\frac{2}{n}}} = 1 + \sqrt{\frac{2}{n}}t + \left(\frac{2}{n}\right)\frac{t^2}{2!} + \left(\frac{2}{n}\right)^{3/2}\frac{t^3}{3!}e^{t\theta\sqrt{\frac{2}{n}}}, \text{ donde } 0 < \theta < 1.
$$

Se multiplica ambos lados de la igualdad anterior por $\sqrt{\frac{2}{n}}t$, de esta manera,

$$
\sqrt{\frac{2}{n}}te^{t\sqrt{\frac{2}{n}}} = \sqrt{\frac{2}{n}}t + \left(\frac{2}{n}\right)t^2 + \left(\frac{2}{n}\right)^{\frac{3}{2}}\frac{t^3}{2!} + \left(\frac{2}{n}\right)^2\frac{t^4}{3!}e^{t\theta\sqrt{\frac{2}{n}}}.
$$

Por lo tanto,

$$
\begin{aligned}
m_{Z_n}(t) &= \left[1 - \frac{t^2}{n} + \left(\frac{2}{n}\right)^{\frac{3}{2}}\frac{t^3}{3!}e^{t\theta\sqrt{\frac{2}{n}}} - \left(\frac{2}{n}\right)^{\frac{3}{2}}\frac{t^3}{2!} - \left(\frac{2}{n}\right)^2\frac{t^4}{3!}e^{t\theta\sqrt{\frac{2}{n}}}\right]^{-\frac{n}{2}} \\
&= \left[1 - \frac{t^2}{n} + \frac{\frac{2^{\frac{3}{2}}}{n^{\frac{1}{2}}}t^3\left[\frac{1}{3!}e^{t\theta\sqrt{\frac{2}{n}}} - \frac{1}{2!}\right] - \frac{4}{n}\frac{1}{3!}t^4e^{t\theta\sqrt{\frac{2}{n}}}}{n}\right]^{-\frac{n}{2}}.
\end{aligned}
$$

Se aplicará el resultado límite C.1 del apéndice C, considerando $\varphi(n)$, de la siguiente manera:

$$
\varphi(n) = \frac{2^{\frac{3}{2}}}{n^{\frac{1}{2}}}t^3\left[\frac{1}{3!}e^{t\theta\sqrt{\frac{2}{n}}} - \frac{1}{2!}\right] - \frac{4}{n}\frac{1}{3!}t^4e^{t\theta\sqrt{\frac{2}{n}}}.
$$

Obsérvese que

$$\lim_{n\to\infty} e^{t\theta\sqrt{\frac{2}{n}}} = 1.$$

Por lo tanto,

$$\lim_{n\to\infty} \varphi(n) = 0.$$

De esta manera,

$$\begin{aligned}\lim_{n\to\infty} m_{Z_n}(t) &= \lim_{n\to\infty} \left[1 - \frac{t^2}{n} + \frac{\varphi(n)}{n}\right]^{-\frac{n}{2}} \\ &= e^{(-t^2)(-\frac{1}{2})} = e^{\frac{t^2}{2}}.\end{aligned}$$

La anterior es la función generadora de momentos de una distribución normal estándar, por lo tanto,

$$Z_n \xrightarrow{d} Z, \text{ donde } Z \sim N(0,1).$$

Otra forma de explicar el anterior resultado es que la distribución de una variable aleatoria con distribución ji cuadrada tiende a una distribución nor-mal, si los grados de libertad tienden a crecer. Para acabar de entender este resultado límite, se puede apreciar en la figura 4.4, cómo la gráfica de una función de densidad de una distribución ji cuadrada tiende a una gráfica en forma de campana (distribución normal), conforme los grados de libertad son más grandes.

7.6. Teorema de límite central

En esta sección se presentará uno de los teoremas más importantes que existe en la teoría de la probabilidad. Su importancia se debe, principalmente, a sus diversas aplicaciones tanto en probabilidad como en estadística, entre otras áreas.

A continuación, se expondrá este resultado y se demostrará.

Teorema 7.8. Teorema de límite central. *Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución con media μ y varianza σ^2, entonces,*

$$Z_n = \sqrt{n}\left(\frac{\bar{X}_n - \mu}{\sigma}\right) \xrightarrow{d} Z \sim N(0,1).$$

Al anterior resultado se le conoce como el *teorema de límite central*. Como se verá más adelante, con algunos ejemplos, este resultado tiene muchas aplicaciones.

Obsérvese que Z_n es la media muestral estandarizada, ya que $E[\bar{X}_n] = \mu$ y $V[\bar{X}_n] = \frac{\sigma^2}{n}$.

También, se puede ver que

$$\begin{aligned}
Z_n &= \sqrt{n}\left(\frac{\bar{X}_n - \mu}{\sigma}\right) \\
&= n\left(\frac{\bar{X}_n - \mu}{\sqrt{n}\sigma}\right) \\
&= \frac{\displaystyle\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma}.
\end{aligned}$$

Esto es, Z_n también se puede expresar como el total muestral estandarizado, ya que $E\left[\displaystyle\sum_{i=1}^{n} X_i\right] = n\mu$ y $V\left[\displaystyle\sum_{i=1}^{n} X_i\right] = n\sigma^2$.

Entonces, el teorema de límite central también puede expresarse en términos del total muestral estandarizado, si las hipótesis del teorema se cumplen, esto es, $Z_n = \left(\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma}\right) \xrightarrow{d} Z \sim N(0,1)$.

En términos prácticos, se está afirmando que si $X_1, \ldots, X_n$ es una muestra aleatoria de una población que tiene distribución con media μ y varianza σ^2 y el tamaño de la muestra n es «suficientemente grande», entonces, $\bar{X}_n$ se distribuye aproximadamente normal con media μ y varianza σ^2/n, y, por otro lado, $\sum_{i=1}^{n} X_i$ también se distribuye aproximadamente normal, pero con media $n\mu$ y varianza $n\sigma^2$.

Siempre ha sido un tema de discusión qué significa un tamaño de muestra «suficientemente grande». Para algunos autores, un tamaño de muestra igual que 30 ya se considera suficientemente grande.

Lo anterior no siempre es cierto, si la muestra es obtenida de una población con distribución donde la función de densidad tiene un sesgo muy grande, como la distribución exponencial o la distribución geométrica, una $n = 30$ no es suficiente para apreciar una aproximación satisfactoria de la distribución de la media (o total muestral) con la distribución normal.

En cambio, si la muestra es obtenida de una población con una distribución donde la función de densidad es simétrica o casi simétrica, como la distribución uniforme o la distribución binomial con $p \approx 0.5$, un tamaño de muestra de 12, por ejemplo, en ocasiones, será suficiente para apreciar una buena aproximación.

En la siguiente gráfica, se aprecia cómo la distribución de la media muestral tiende a una distribución normal, cuando el tamaño de muestra se va incrementando. Para la construcción de esta gráfica, se simularon, por medio de R, muestras aleatorias obtenidas de una distribución exponencial con media igual a 10, por medio de la instrucción rexp(n,0.1), donde los valores de n se fueron variando.

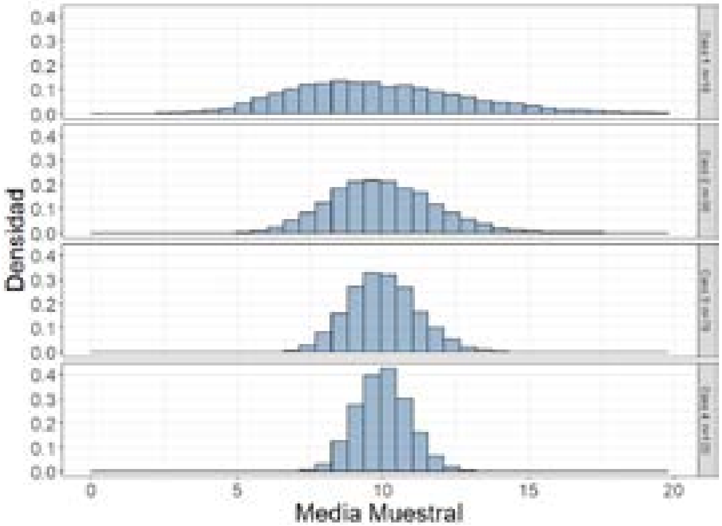


Figura 7.6. *Distribuciones de la media muestral con diferentes tamaño de muestra. Ilustración del teorema de límite central*

La primera gráfica, es el histograma de 10,000 medias muestrales, calculadas de muestras aleatorias de tamaño igual a 10. El segundo histograma fue construido también por 10,000 medias muestrales, pero calculadas de muestras de tamaño 30. El tercer histograma se hizo por medio de 10,000 medias muestrales, calculadas de muestras de tamaño 70, y la última gráfica se construyó por medio de 10,000 medias muestrales calculadas de muestras de tamaño 120.

Es notorio que tanto el sesgo por la derecha como la varianza van disminuyendo de una gráfica a otra, siendo la última gráfica, prácticamente, una campana con menos varianza.

A continuación, se realizará la demostración del teorema de límite central, donde se usará la función generadora de momentos del total muestral.

Demostración. (Teorema 7.8). Definimos la siguiente función de t.

$$\Phi(t) = E[e^{t(X-\mu)}] = e^{-t\mu} m_X(t).$$

Esta función se expresará por medio del polinomio de Maclaurin con tres sumandos, véase igualdad C.12 del apéndice C.

$$\Phi(t) = \Phi(0) + \Phi^{(1)}(0)t + \Phi^{(2)}(\xi)\frac{t^2}{2!}, \text{ donde } 0 < \xi < t.$$

Obsérvese

$$\Phi^{(1)}\big|_{t=0} = \left[e^{-t\mu} m_X^{(1)}(t) + m_X(t)[-\mu e^{-t\mu}]\right]_{t=0} = \mu - \mu = 0.$$

$$\Phi^{(2)}|_{t=0}$$

$$= \left[e^{-t\mu}m_X^{(2)}(t) - m_X^{(1)}(t)\mu e^{-t\mu} + \mu m_X(t)\mu e^{-t\mu} - \mu m_X^{(1)}(t)e^{-t\mu}\right]_{t=0}$$

$$= E[X^2] - \mu^2 + \mu^2 - \mu^2 = \sigma^2.$$

Entonces,

$$\Phi(t) = 1 + 0 + \Phi^{(2)}(\xi)\frac{t^2}{2!} = 1 + \frac{\sigma^2 t^2}{2} + \left[\Phi^{(2)}(\xi) - \sigma^2\right]\frac{t^2}{2},\ 0 < \xi < t.$$

Ahora, se desarrollará la función generadora de momento de Z_n

$$\begin{aligned}
m_{Z_n}(t) &= E\left[e^{Z_n t}\right] \\
&= E\left[e^{\left(\frac{\sum_{i=0}^{n} X_i - n\mu}{\sqrt{n}\sigma}\right)t}\right] \\
&= E\left[e^{\sum_{i=0}^{n}(X_i-\mu)\frac{t}{\sqrt{n}\sigma}}\right] \\
&= E\left[e^{(X_1-\mu)\frac{t}{\sqrt{n}\sigma}+\cdots+(X_n-\mu)\frac{t}{\sqrt{n}\sigma}}\right] \\
&= E\left[e^{(X_1-\mu)\frac{t}{\sqrt{n}\sigma}}\right]\cdots E\left[e^{(X_n-\mu)\frac{t}{\sqrt{n}\sigma}}\right] \\
&= \left[E\left[e^{(X-\mu)\frac{t}{\sqrt{n}\sigma}}\right]\right]^n \\
&= \left[\Phi\left(\tfrac{t}{\sqrt{n}\sigma}\right)\right]^n \\
&= \left[1 + \frac{t^2}{2n\sigma^2}\sigma^2 + \left(\Phi^{(2)}(\tau) - \sigma^2\right)\frac{t^2}{2n\sigma^2}\right]^n \\
&= \left[1 + \frac{\frac{t^2}{2}}{n} + \frac{\left(\Phi^{(2)}(\tau) - \sigma^2\right)\frac{t^2}{2\sigma^2}}{n}\right]^n, \ \text{donde}\ 0 < \tau < \frac{t}{\sigma\sqrt{n}}.
\end{aligned}$$

Sea

$$\varphi(n) = \left(\Phi^{(2)}(\tau) - \sigma^2\right)\frac{t^2}{2\sigma^2}.$$

Obsérvese

$$\lim_{n\to\infty} \Phi^{(2)}(\tau) = \sigma^2, \text{ ya que } \tau \to 0.$$

Por lo tanto,

$$\lim_{n\to\infty} \varphi(n) = \lim_{n\to\infty} \left(\Phi^{(2)}(\tau) - \sigma^2\right)\frac{t^2}{2\sigma^2} = 0.$$

De este modo,

$$\begin{aligned}
\lim_{n\to\infty} m_{Z_n}(t) &= \lim_{n\to\infty}\left[1+\frac{t^2/2}{n}+\frac{\left(\Phi^{(2)}(\tau)-\sigma^2\right)\dfrac{t^2}{2\sigma^2}}{n}\right]^n \\
&= \lim_{n\to\infty}\left[1+\frac{t^2/2}{n}+\frac{\varphi(n)}{n}\right]^n = e^{t^2/2}.
\end{aligned}$$

Esta última es la función generadora de momentos de una distribución normal estándar. Por lo tanto,

$$Z_n \xrightarrow{d} Z, \text{ donde } Z \sim N(0,1). \qquad \square$$

La anterior demostración se realizó suponiendo que la función generadora de momentos de X_i existe, y no es una condición del teorema de límite central. Existe la posibilidad de que la media y la varianza sean finitas y no exista la función generadora de momentos. La demostración de este teorema se puede generalizar; en lugar de usar la función generadora de momentos, se puede utilizar la función característica, véase definición 2.15, esta función existe para todas las variables aleatorias. Cabe señalar que el alumno necesita conocimientos de variable compleja para entender la demostración del teorema usando la función característica.

Es importante resaltar que la condición para que se cumpla el teorema de límite central es que las variables $X_1, X_2, \ldots, X_n$ sean independientes e idénticamente distribuidas, con media y varianza finitas.

El teorema de límite central nos indica que tanto la media muestral como el total muestral se distribuyen aproximadamente normal, si la muestra es «suficientemente grande». Lo interesante de este resultado es que la muestra puede ser obtenida de una distribución discreta, de una distribución continua o incluso de una mezcla de distribuciones. En la siguiente sección, se tratarán varios ejemplos.

7.7. Aplicaciones del teorema de límite central

En términos prácticos, como ya se mencionó, el teorema de límite central indica que tanto la media muestral como el total muestral, cuando son calculados de una muestra aleatoria obtenida de una distribución (población) con media y varianzas finitas, estas se distribuyen aproximadamente normal, cuando el tamaño de la muestra es «suficientemente grande». Entre varias aplicaciones que tiene este teorema, a continuación, se presentarán tres de las más importantes.

1.- Aproximación de la distribución binomial por medio de la distribución normal

A continuación, se presentará cómo las probabilidades de una variable aleatoria con distribución binomial se pueden aproximar por medio de una distribución normal.

Considérese la siguiente variable aleatoria, $Y =$ número de pruebas observadas con éxito al realizarse n pruebas idénticas e independientes. Supongamos que Y tiene distribución binomial con parámetros n y p.

Obsérvese que la variable Y se puede expresar como $Y = \sum_{i=1}^{n} X_i$, donde $X_1, \ldots, X_n$ es una muestra aleatoria de una distribución Bernoulli con parámetro p. En este caso, la variable X_i puede interpretarse como $X_i=$ número de éxitos en la $i-$ésima prueba.

De esta manera, el teorema de límite central se puede aplicar a la variable $Y = \sum_{i=1}^{n} X_i$, siendo la media y la varianza de X_i, respectivamente, $\mu = p$ y $\sigma^2 = p(1-p) = pq$.

Entonces, se puede afirmar, aplicando el teorema de límite central,

$$\frac{\displaystyle\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} = \frac{Y - np}{\sqrt{np(1-p)}} \xrightarrow{d} Z \sim N(0,1).$$

Dicho de otra manera, una variable aleatoria Y con distribución binomial con parámetros n y p tendrá aproximadamente una distribución normal, con media igual a $\mu = np$ y varianza $\sigma^2 = np(1-p)$, cuando n sea «suficientemente grande».

Es importante hacer la siguiente aclaración, cuando se quiera aproximar una probabilidad para una variable aleatoria discreta con valores enteros, por medio de una distribución continua, es recomendable utilizar lo que se llama el *factor de continuidad*, para lograr una mejor aproximación. Supongamos que X es una variable aleatoria discreta que toma valores enteros, y que es de interés calcular la siguiente probabilidad $P[a \leq X \leq b]$, donde a y b son enteros. Si este cálculo se quiere realizar en forma aproximada, por medio de una distribución continua, se sugiere considerar el factor de continuidad en cada lado de la desigualdad, el cual consiste en restar al número a, 0.5 y sumar al número b, 0.5. Entonces, se sugiere, usando la distribución continua, calcular la probabilidad $P[a - 0.5 \leq X \leq b + 0.5]$, en lugar de $P[a \leq X \leq b]$. A continuación, por medio de una gráfica, se ilustra la conveniencia de usar los factores de continuidad.

En la siguiente gráfica, el histograma representa la función de densidad de una variable discreta con valores enteros, y la gráfica continua representa la función de densidad de la distribución continua que supuestamente está aproximando la densidad discreta.

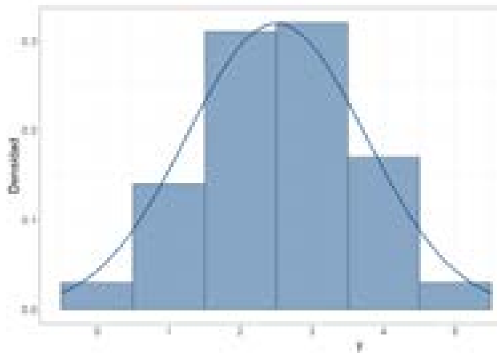


Figura 7.7. *Aproximación de la distribución binomial por la distribución normal*

Por ejemplo, el cálculo de la probabilidad $P[2 \leq Y \leq 3]$ por medio de la distribución discreta corresponde a la suma de las áreas de los dos rectángulos correspondientes a los valores del recorrido 2 y 3, la base del primero va de 1.5 a 2.5 y del segundo de 2.5 a 3.5. Si se pretende calcular esta probabilidad en forma aproximada por medio de la densidad continua, en lugar de considerar $P[2 \leq Y \leq 3]$, la aproximación resultará más exacta, si se calcula $P[1.5 \leq Y \leq 3.5]$, ya que de esta manera se estarían aproximando de mejor manera, las áreas de los rectángulos de $y = 2$ y $y = 3$.

Por ejemplo, en el caso del cálculo de $P[Y > 3]$, se recomienda calcular con la densidad continua, la probabilidad $P[Y \geq 3.5]$, ya que se estaría aproximando de mejor manera a las áreas de los rectángulos correspondientes de $y = 4$, $y = 5$, etc. Y en el caso de $P[Y = 3]$, se recomienda calcular con la distribución continua la probabilidad $P[2.5 \leq Y \leq 3.5]$, ya que se aproximaría de mejor manera el área del rectángulo correspondiente a $y = 3$.

Ejemplo 7.6. Para cierto tipo de seguro, se sabe que un siniestro tendrá un valor superior al deducible con una probabilidad de $p = 0.4$. Se seleccionan al azar 40 siniestros del último año. Sea Y el número de siniestros en la muestra que son superiores al deducible. Calcular en forma aproximada las siguientes probabilidades:
a) $P[Y > 16]$.
b) $P[10 < Y \leq 20]$.
c) $P[Y = 18]$.

a) La aproximación de la primera probabilidad es

$$\begin{aligned} P[Y > 16] &= P[Y \geq 16.5] \\ &= P\left[\frac{Y - 16}{\sqrt{40(0.4)(0.6)}} \geq \frac{16.5 - 16}{\sqrt{40(0.4)(0.6)}}\right] \\ &\approx P[Z \geq 0.16] = 0.4364. \end{aligned}$$

b) La segunda probabilidad se aproxima de la siguiente manera:

$$\begin{aligned} P[10 < Y \leq 20] &= P[10.5 \leq Y \leq 20.5] \\ &= P\left[\frac{10.5-16}{\sqrt{40(0.4)(0.6)}} \leq \frac{Y-16}{\sqrt{40(0.4)(0.6)}} \leq \frac{20.5-16}{\sqrt{40(0.4)(0.6)}}\right] \\ &\approx P[-1.78 \leq Z \leq 1.45] = 0.889. \end{aligned}$$

c) Y la tercera probabilidad

$$\begin{aligned} P[Y = 18] &= P[17.5 \leq Y \leq 18.5] \\ &= P\left[\frac{17.5-16}{\sqrt{40(0.4)(0.6)}} \leq \frac{Y-16}{\sqrt{40(0.4)(0.6)}} \leq \frac{18.5-16}{\sqrt{40(0.4)(0.6)}}\right] \\ &\approx P[0.48 \leq Z \leq 0.81] = 0.1066. \end{aligned}$$

Con la idea de comparar este último resultado, se calculará también por medio de la distribución binomial.

$$P[Y = 18] = \binom{40}{18}(0.40)^{18}(0.6)^{22} = 0.1026.$$

Se observa que ambas probabilidades son muy parecidas.

2.- Aproximación de la distribución de una proporción muestral

Otra aplicación del teorema del límite central es la aproximación de probabilidades de una proporción muestral por medio de la distribución normal. Cabe señalar, que estas aproximaciones pueden ser aplicadas en inferencia estadística.

Sea $X_1, \ldots, X_n$ una muestra aleatoria de tamaño n de una distribución Bernoulli con parámetro p. La proporción muestral se calcula como

$$\hat{p} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{Y}{n} = \bar{X}, \text{ donde } Y \sim B(n,p).$$

Entonces, el teorema de límite central se puede aplicar para $\bar{X} = \hat{p}$. Obsérvese

$$E[\hat{p}] = E\left[\frac{Y}{n}\right] = \frac{1}{n}np = p,$$

$$V[\hat{p}] = V\left[\frac{Y}{n}\right] = \frac{1}{n^2}np(1-p) = \frac{p(1-p)}{n}.$$

De esta manera, por el teorema de límite central,

$$\sqrt{n}\left(\frac{\bar{X}-\mu}{\sigma}\right) = \sqrt{n}\left(\frac{\hat{p}-p}{\sqrt{p(1-p)}}\right) \xrightarrow{d} Z \sim N(0,1).$$

Dicho de otra manera, la proporción muestral $\hat{p} = \frac{Y}{n}$, se distribuye aproximadamente normal con media p y varianza $\frac{p(1-p)}{n}$, cuando n es «suficientemente grande».

Esta aproximación tiene una aplicación muy importante en inferencia estadística, es utilizada para construir intervalos de confianza y pruebas de hipótesis para una proporción p.

A continuación, se presentarán un par de ejemplos.

Ejemplo 7.7. Considerando la misma distribución binomial del ejemplo anterior, calcular la probabilidad de que la proporción muestral $\hat{p}$ se desvíe de la proporción verdadera p, a lo más 0.05.

Se sabe que la proporción muestral $\hat{p}$ se distribuye aproximadamente $N\left(p, \frac{p(1-p)}{n}\right)$.

Por lo tanto,

$$\begin{aligned}
P[|\hat{p} - p| < 0.05] &= P[-0.05 \leq \hat{p} - p \leq 0.05] \\
&= P\left[\frac{-0.05}{\sqrt{\frac{p(1-p)}{n}}} \leq \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \leq \frac{0.05}{\sqrt{\frac{p(1-p)}{n}}}\right] \\
&\approx P[-0.65 \leq Z \leq 0.65] = 0.4844.
\end{aligned}$$

Ejemplo 7.8. Se saca una muestra aleatoria $X_1, \ldots, X_n$ de una distribución Bernoulli con parámetro $p = 0.2$. ¿De qué tamaño debe ser la muestra para que la proporción muestral se desvíe de la proporción verdadera, a lo más 0.03, con una probabilidad del 0.95?

Se sabe $\hat{p} \dot{\sim} N\left(0.2, \frac{0.2(0.8)}{n}\right)$.

De esta manera,

$$\begin{aligned}
&P[|\hat{p} - 0.2| \leq 0.03] \\
&= P[-0.03 \leq \hat{p} - 0.2 \leq 0.03] \\
&= P\left[\sqrt{n}\left(-\frac{0.03}{\sqrt{0.2(0.8)}}\right) \leq \sqrt{n}\left(\frac{\hat{p} - 0.2}{\sqrt{0.2(0.8)}}\right) \leq \sqrt{n}\left(\frac{0.03}{\sqrt{0.2(0.8)}}\right)\right] \\
&\approx P[-\sqrt{n}(0.075) \leq Z \leq \sqrt{n}(0.075)] = 0.95.
\end{aligned}$$

Entonces, $\sqrt{n}(0.075) = 1.96$, por lo tanto, $n = 682.95 \approx 683$.

Esto es, con una muestra aleatoria de tamaño $n = 683$, la proporción muestral no se alejará de la proporción poblacional (verdadera proporción), más de 0.03, aproximadamente un 95 % de las veces.

3.- Aproximación de la distribución de la media muestral (total muestral)

En general, por el teorema de límite central se pueden aproximar probabilidades de una media o total muestral por medio de una distribución normal. A continuación, se presentarán algunos ejemplos.

Supongamos que $X_1, \ldots, X_n$ representa una muestra aleatoria de tamaño n de una distribución, discreta o continua, con una media μ y varianza σ^2 finitas. Se sabe, por el teorema de límite central, que la media muestral y también el total muestral se distribuyen aproximadamente normal. Los siguientes ejercicios estarán encaminados con estos resultados.

Ejemplo 7.9. Supongamos que la cantidad reclamada en un mes, por un asegurado a una compañía de seguros, por cierto tipo de seguro, es una variable aleatoria con distribución uniforme sobre el intervalo $(0, 10)$. Además, supóngase que las cantidades reclamadas por diferentes asegurados se realizan en forma independiente. De 12 reclamaciones en un mes, calcular la probabilidad de que el total de reclamaciones sea entre
a) 70 y 85.
b) 40 y 70.

a) Sea X_i =cantidad reclamada por un asegurado en un mes, $i = 1, \ldots, 12$.
Como $X_i \sim U(0, 10)$, entonces, $\mu = E[X_i] = 5$, y $\sigma^2 = V[X_i] = \frac{100}{12}$.
Por lo tanto, $\sum_{i=1}^{12} X_i$ =total de reclamaciones en un mes.
Así, $E\left[\sum_{i=1}^{12} X_i\right] = 12(5) = 60$ y $V\left[\sum_{i=1}^{12} X_i\right] = 12(\frac{100}{12}) = 100$.
Aplicando el teorema de límite central, $\sum_{i=1}^{12} X_i \dot{\sim} N(60, 100)$.

$$\begin{aligned} P\left[70 \leq \sum_{i=1}^{12} X_i \leq 85\right] &= P\left[\frac{70-60}{10} \leq \frac{\sum_{i=1}^{12} X_i - 60}{10} \leq \frac{85-60}{10}\right] \\ &\approx P[1 \leq Z \leq 2.5] = 0.1525. \end{aligned}$$

b) La segunda probabilidad

$$\begin{aligned} P\left[40 \leq \sum_{i=1}^{12} X_i \leq 70\right] &= P\left[\frac{40-60}{10} \leq \frac{\sum_{i=1}^{12} X_i - 60}{10} \leq \frac{70-60}{10}\right] \\ &\approx P[-2 \leq Z \leq 1] = 0.8185. \end{aligned}$$

Como ya se había comentado, por ser la función de densidad de X_i simétrica, no es necesario que el tamaño de muestra sea muy grande (30 o más), para lograr una buena aproximación. Para algunos autores, en el caso de la distribución uniforme, con un tamaño de muestra entre 10 y 12 es suficiente.

Véase Rubinstein [27], donde se encuentra un procedimiento para simular valores de una variable aleatoria con distribución normal estándar, por medio de simulaciones de totales muestrales estandarizados, donde la muestra aleatoria proviene de una distribución $U(0,1)$.

Ejemplo 7.10. La proporción de impureza en una botella que contiene cierto líquido, se distribuye beta con parámetros $a = 2$ y $b = 4$. Se seleccionan aleatoriamente 50 botellas de la producción. Calcular la probabilidad de que el promedio muestral de las proporciones de impurezas.
a) Sea mayor a 0.40.
b) Se desvíe de la media verdadera de proporciones, a lo más 0.05.

Sea X_i = proporción de impurezas en la botella i, $i = 1, \ldots, 50$.
Se sabe que $X_i \sim Beta(2,4)$.
De este modo, $\mu = E[X_i] = \frac{1}{3}$ y $\sigma^2 = V[X_i] = \frac{2}{63}$.
En este caso, $\bar{X}$ = promedio muestral de proporciones de impurezas.
Se sabe que $E[\bar{X}] = \frac{1}{3}$ y $V\left[\bar{X}\right] = \frac{\sigma^2}{50} = \frac{1}{1575}$.
Por el teorema de límite central, se puede afirmar $\bar{X} \dot{\sim} N\left(\frac{1}{3}, \frac{1}{1575}\right)$.
a) La primera probabilidad se aproxima de la siguiente manera:

$$\begin{aligned} P[\bar{X} > 0.40] &= P\left[\frac{\bar{X} - 1/3}{\sqrt{1/1575}} > \frac{0.40 - 1/3}{\sqrt{1/1575}}\right] \\ &\approx P[Z > 2.64575] = 0.004. \end{aligned}$$

b) La segunda probabilidad,

$$\begin{aligned} P[|\bar{X} - 1/3| \leq 0.05] &= P[-0.05 \leq \bar{X} - 1/3 \leq 0.05] \\ &= P\left[-\frac{0.05}{\sqrt{1/1575}} \leq \frac{\bar{X} - 1/3}{\sqrt{1/1575}} \leq \frac{0.05}{\sqrt{1/1575}}\right] \\ &\approx P[-1.98431 \leq Z \leq 1.98431] = 0.9522. \end{aligned}$$

Ejemplo 7.11. Considerando el ejemplo anterior, ¿de qué tamaño debe ser la muestra, si se desea que la media muestral se desvíe de la verdadera media a lo más 0.04 con una probabilidad de 0.98?

Obsérvese
$P[|\bar{X} - 1/3| \leq 0.04]$

$$\begin{aligned} &= P[-0.04 \leq \bar{X} - 1/3 \leq 0.04] \\ &= P\left[-\sqrt{n}\left(\frac{0.04}{\sqrt{2/63}}\right) \leq \sqrt{n}\left(\frac{\bar{X} - 1/3}{\sqrt{2/63}}\right) \leq \sqrt{n}\left(\frac{0.04}{\sqrt{2/53}}\right)\right] \\ &\approx P[-0.2245\sqrt{n} \leq Z \leq 0.2245\sqrt{n}] = 0.98. \end{aligned}$$

Por lo tanto, se puede observar en la tabla de la normal estándar que $0.2245\sqrt{n} = 2.33$.

De esta manera se concluye que $n = 107.72 \approx 108$.

7.8. Teoremas de convergencia

En esta sección se tratarán diversos teoremas que tienen que ver con la convergencia en probabilidad y la convergencia en distribución. Estos teoremas ayudarán para lograr algunos resultados importantes, con aplicaciones en estimación puntual (estimadores consistentes), intervalos de confianza y pruebas de hipótesis, que son temas de suma importancia de inferencia estadística.

Teorema 7.9. *Sea $\{U_n\}$ una sucesión de variables aleatorias que converge en probabilidad a $c \neq 0$, entonces, la sucesión $\left\{\frac{U_n}{c}\right\}$ converge en probabilidad a 1.*

Demostración. Obsérvese

$$\begin{aligned}\lim_{n\to\infty} P\left[\left|\frac{U_n}{c} - 1\right| \geq \epsilon\right] &= \lim_{n\to\infty} P\left[\left|\frac{U_n - c}{c}\right| \geq \epsilon\right] \\ &= \lim_{n\to\infty} P\left[|U_n - c| \geq \epsilon|c|\right] = 0.\end{aligned}$$

Por lo tanto,

$$\left\{\frac{U_n}{c}\right\} \xrightarrow{p} 1. \qquad \square$$

Teorema 7.10. *Sea $\{U_n\}$ una sucesión de variables aleatorias que converge en probabilidad a $c > 0$, y $P[U_n < 0] = 0$ para toda n, entonces, la sucesión $\{\sqrt{U_n}\}$ converge en probabilidad a $\sqrt{c}$.*

Demostración. Obsérvese

$$\begin{aligned}P[|U_n - c| < \epsilon] &= P[|(\sqrt{U_n} - \sqrt{c})(\sqrt{U_n} + \sqrt{c})| < \epsilon] \\ &= P\left[\left|(\sqrt{U_n} - \sqrt{c})\right| < \frac{\epsilon}{\sqrt{U_n} + \sqrt{c}}\right] \\ &\leq P\left[\left|(\sqrt{U_n} - \sqrt{c})\right| < \frac{\epsilon}{\sqrt{c}}\right] \leq 1.\end{aligned}$$

Sea

$$\epsilon' = \frac{\epsilon}{\sqrt{c}}.$$

De esta manera,

$$P[|U_n - c| < \epsilon] \leq P\left[\left|(\sqrt{U_n} - \sqrt{c})\right| < \epsilon'\right] \leq 1.$$

Por lo tanto,

$$1 = \lim_{n\to\infty} P[|U_n - c| < \epsilon] \leq \lim_{n\to\infty} P\left[\left|(\sqrt{U_n} - \sqrt{c})\right| < \epsilon'\right] \leq \lim_{n\to\infty} 1 = 1.$$

Esto es,

$$\lim_{n\to\infty} P\left[\left|(\sqrt{U_n}-\sqrt{c})\right| < \epsilon'\right] = 1, \text{ para todo } \epsilon' > 0.$$

Por lo tanto,

$$\{\sqrt{U_n}\} \xrightarrow{p} \sqrt{c}. \qquad \square$$

El siguiente teorema trata sobre la suma, producto y cociente de dos sucesiones de variables aleatorias que convergen a una constante. Los resultados son útiles para la solución de la convergencia de algunas sucesiones, donde algunos ejemplos se presentarán más adelante. Este teorema, también es usado en inferencia estadística, en el tema de estimación puntual, para verificar si un estimador puntual de un parámetro es consistente.

Teorema 7.11. *Sea $\{U_n\}$ una sucesión de variables aleatorias que converge en probabilidad a c y $\{V_n\}$ una sucesión de variables aleatorias que converge en probabilidad a d, entonces, la sucesión*

a) $\{U_n + V_n\}$ converge en probabilidad a $c + d$.

b) $\{U_n V_n\}$ converge en probabilidad a cd.

c) $\left\{\frac{U_n}{V_n}\right\}$ converge en probabilidad a $\dfrac{c}{d}$, donde $d \neq 0$.

Se le sugiere al estudiante demostrar los resultados del anterior teorema.

Otro de los teoremas, que serán de gran utilidad en problemas de convergencia de variables aleatorias, se presentará a continuación:

Teorema 7.12. *Sea $\{U_n\}$ una sucesión de variables aleatorias que converge en distribución a una variable aleatoria X y $\{V_n\}$ una sucesión de variables aleatorias que converge en probabilidad a 1, entonces, la sucesión $\left\{\frac{U_n}{V_n}\right\}$ converge en distribución a la variable aleatoria X.*

Se le sugiere al estudiante realizar la demostración de este teorema.

A continuación, se presentarán algunos ejemplos donde se aplicarán varios de los resultados vistos durante este capítulo.

Ejemplo 7.12. Sea $\{Z_n\}$ una sucesión de variables aleatorias tal que $Z_n = \dfrac{U_n - np}{\sqrt{n\hat{p}(1-\hat{p})}}$, donde U_n se distribuye binomial con parámetros n y p, y $\hat{p} = \frac{U_n}{n}$, encontrar la distribución límite de la sucesión $\{Z_n\}$.

Se sabe, por el teorema de límite central

$$\left\{\frac{U_n - np}{\sqrt{np(1-p)}}\right\} \xrightarrow{d} Z \sim N(0,1).$$

Además, por la ley de los grandes números

$$\{\hat{p}\} \xrightarrow{p} p \quad \text{y} \quad \{1-\hat{p}\} \xrightarrow{p} 1-p.$$

Por el teorema 7.11, inciso b

$$\{\hat{p}(1-\hat{p})\} \xrightarrow{p} p(1-p).$$

Por el teorema 7.9

$$\left\{\frac{\hat{p}(1-\hat{p})}{p(1-p)}\right\} \xrightarrow{p} 1.$$

Por el teorema 7.10

$$\left\{\sqrt{\frac{\hat{p}(1-\hat{p})}{p(1-p)}}\right\} \xrightarrow{p} \sqrt{1} = 1.$$

De esta manera, por el teorema 7.12

$$\left\{\frac{\dfrac{U_n - np}{\sqrt{np(1-p)}}}{\sqrt{\dfrac{\hat{p}(1-\hat{p})}{p(1-p)}}}\right\} = \left\{\frac{U_n - np}{\sqrt{n\hat{p}(1-\hat{p})}}\right\} \xrightarrow{d} Z \sim N(0.1).$$

Ejemplo 7.13. Sea $\{Z_n\}$ una sucesión de variables aleatorias tal que $Z_n = \sqrt{Y_n}$, donde Y_n es la n-ésima estadística de orden de una muestra aleatoria de tamaño n de una distribución uniforme sobre el intervalo $(0,\theta)$. Demostrar que $\{Z_n\} \xrightarrow{p} \sqrt{\theta}$.

Primero se encontrará la función de distribución de Y_n

$$\begin{aligned} F_{Y_n}(y) &= [F_X(y)]^n \\ &= \left[\frac{y}{\theta}\right]^n, \qquad \text{donde } 0 < y < \theta. \end{aligned}$$

Ahora la función de distribución de Z_n

$$\begin{aligned} F_{Z_n}(z) &= P[Z_n \leq z] \\ &= P[\sqrt{Y_n} \leq z] \\ &= F_{Y_n}(z^2) \\ &= \left[\frac{z^2}{\theta}\right]^n, \qquad \text{donde } 0 < z < \sqrt{\theta}. \end{aligned}$$

Esto es,

$$F_{Z_n}(z) = \left[\frac{z^2}{\theta}\right]^n I_{(0,\sqrt{\theta})}(z) + I_{[\sqrt{\theta},\infty)}(z).$$

Por lo tanto,

$$\begin{aligned}
\lim_{n\to\infty} F_{Z_n}(z) &= \lim_{n\to\infty} \left\{ \left[\frac{z^2}{\theta}\right]^n I_{(0,\sqrt{\theta})}(z) + I_{[\sqrt{\theta},\infty)}(z) \right\} \\
&= 0 + I_{[\sqrt{\theta},\infty)}(z) \\
&= I_{[\sqrt{\theta},\infty)}(z).
\end{aligned}$$

El resultado de este límite es la función de distribución de una variable aleatoria degenerada en $\sqrt{\theta}$, entonces,

$$\{Z_n\} \xrightarrow{d} \sqrt{\theta}.$$

Aplicando el teorema 7.5,

$$\{Z_n\} \xrightarrow{p} \sqrt{\theta}.$$

Ejemplo 7.14. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución normal con parámetros μ y σ^2. Sea $\{U_n\}$ una sucesión de variables aleatorias tal que $U_n = S^2$, y donde S^2 es la varianza muestral, esto es, $S^2 = \dfrac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}$. Demostrar que $\{U_n\} \xrightarrow{p} \sigma^2$.

Se sabe que

$$(n-1)\frac{S^2}{\sigma^2} \sim \chi^2(n-1).$$

Por lo tanto,

$$\begin{aligned}
E[S^2] &= E\left[\frac{\sigma^2}{n-1}(n-1)\frac{S^2}{\sigma^2}\right] \\
&= \frac{\sigma^2}{n-1} E\left[(n-1)\frac{S^2}{\sigma^2}\right] \\
&= \frac{\sigma^2}{n-1}(n-1) = \sigma^2,
\end{aligned}$$

$$\begin{aligned}
V[S^2] &= V\left[\frac{\sigma^2}{n-1}(n-1)\frac{S^2}{\sigma^2}\right] \\
&= \frac{\sigma^4}{(n-1)^2} V\left[(n-1)\frac{S^2}{\sigma^4}\right] \\
&= \frac{\sigma^4}{(n-1)^2} 2(n-1) = \frac{2\sigma^4}{n-1}.
\end{aligned}$$

Se aplicará la desigualdad de Chebyshev para la varianza muestral S^2

$$P\left[\left|S^2 - \sigma^2\right| < k\frac{\sqrt{2}\sigma^2}{\sqrt{n-1}}\right] \geq 1 - \frac{1}{k^2}, \qquad \text{para toda } k > 0.$$

En particular, se selecciona k de la siguiente manera:

$$k = \frac{\sqrt{n-1}}{\sqrt{2}\sigma^2}\epsilon, \text{ donde } \epsilon > 0.$$

Por lo tanto,

$$P\left[\left|S^2 - \sigma^2\right| < \epsilon\right] \geq 1 - \frac{2\sigma^4}{\epsilon^2(n-1)}.$$

Además, toda probabilidad está acotada por el 1,

$$1 - \frac{2\sigma^4}{\epsilon^2(n-1)} \leq P\left[\left|S^2 - \sigma^2\right| < \epsilon\right] \leq 1.$$

Entonces,

$$\lim_{n\to\infty}\left\{1 - \frac{2\sigma^4}{\epsilon^2(n-1)}\right\} \leq \lim_{n\to\infty} P\left[\left|S^2 - \sigma^2\right| < \epsilon\right] \leq \lim_{n\to\infty} 1.$$

Los límites de las partes izquierda y derecha de la anterior desigualdad son igual a 1, de esta manera, se puede concluir que

$$\lim_{n\to\infty} P\left[\left|S^2 - \sigma^2\right| < \epsilon\right] = 1.$$

Por lo tanto,

$$\{U_n\} = \{S^2\} \xrightarrow{p} \sigma^2.$$

Cabe señalar, que fue posible aplicar la desigualdad de Chebyshev, ya que $E[S^2] = \sigma^2$. El siguiente ejemplo es muy parecido a este, pero el camino para encontrar la solución es diferente.

Ejemplo 7.15. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución normal con parámetros μ y σ^2. Sea $\{U_n\}$ una sucesión de variables aleatorias tal que $U_n = S_n^2$, donde S_n^2 está dada por $S_n^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}$. Demostrar que $\{U_n\} \xrightarrow{p} \sigma^2$.

Obsérvese

$$W_n = n\frac{S_n^2}{\sigma^2} \sim \chi^2(n-1).$$

Por lo tanto,

$$\begin{aligned} E[S_n^2] &= E\left[\frac{\sigma^2}{n} n \frac{S_n^2}{\sigma^2}\right] \\ &= \frac{\sigma^2}{n}(n-1) \neq \sigma^2. \end{aligned}$$

En este caso, se va a obtener la función generadora de momentos de S_n^2 para luego obtener su límite. Esto es,

$$\begin{aligned}
m_{S_n^2}(t) &= E\left[e^{S_n^2 t}\right] \\
&= E\left[e^{n\frac{S_n^2}{\sigma^2}\left(\frac{\sigma^2}{n}\right)t}\right] \\
&= m_{W_n}\left(\frac{\sigma^2 t}{n}\right) \\
&= \left[1-2\left(\frac{t\sigma^2}{n}\right)\right]^{-\frac{n-1}{2}} \\
&= \left[1+\frac{-2t\sigma^2}{n}\right]^{-\frac{n}{2}}\left[1+\frac{-2t\sigma^2}{n}\right]^{\frac{1}{2}}.
\end{aligned}$$

Aplicando el resultado C.2 del apéndice C tenemos,

$$\begin{aligned}
\lim_{n\to\infty} m_{S_n^2}(t) &= \lim_{n\to\infty}\left[1+\frac{-2t\sigma^2}{n}\right]^{-\frac{n}{2}}\left[1+\frac{-2t\sigma^2}{n}\right]^{\frac{1}{2}} \\
&= e^{-2t\sigma^2(-\frac{1}{2})}(1) = e^{\sigma^2 t}.
\end{aligned}$$

Esta es la función generadora de momentos de una variable aleatoria degenerada en σ^2.

Entonces,

$$\{U_n\} = \{S_n^2\} \xrightarrow{d} \sigma^2.$$

Por el teorema 7.5, se concluye

$$\{U_n\} = \{S_n^2\} \xrightarrow{p} \sigma^2.$$

Ejemplo 7.16. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución normal con parámetros μ y σ^2, encontrar la distribución límite de la sucesión de variables aleatorias $\{Z_n\}$, donde $Z_n = \sqrt{n}\left(\frac{\bar{X}_n - \mu}{S}\right)$, $\bar{X}_n$ y S son la media y desviación estándar muestrales respectivamente.

Se sabe, por el teorema de límite central,

$$\{Z_n\} = \left\{\sqrt{n}\left(\frac{\bar{X}_n - \mu}{\sigma}\right)\right\} \xrightarrow{d} Z \sim N(0,1).$$

Por otro lado, por el ejemplo 7.14,

$$\{S^2\} \xrightarrow{p} \sigma^2.$$

Por el teorema 7.9,

$$\left\{\frac{S^2}{\sigma^2}\right\} \xrightarrow{p} 1.$$

Por el teorema 7.10,

$$\left\{\frac{S}{\sigma}\right\} \xrightarrow{p} \sqrt{1} = 1.$$

Por lo tanto, por el teorema 7.12,

$$\left\{\frac{\sqrt{n}\left(\dfrac{\bar{X}_n - \mu}{\sigma}\right)}{\dfrac{S}{\sigma}}\right\} = \left\{\sqrt{n}\left(\frac{\bar{X}_n - \mu}{S}\right)\right\} \xrightarrow{d} Z \sim N(0,1).$$

Obsérvese que la muestra aleatoria $X_1, \ldots, X_n$ es obtenida de una distribución $N(\mu, \sigma^2)$, lo que implica que Z_n tiene una distribución t de student con $n-1$ grados de libertad. La interpretación del anterior resultado es, conforme el tamaño de la muestra crece (los grados de libertad también), la distribución de Z_n (t de student) converge a una distribución $N(0,1)$.

Cabe mencionar, que la variable aleatoria Z_n es utilizada en estadística, para construir intervalos de confianza y pruebas de hipótesis para un promedio μ, usándose la distribución t de student. Pero, si el tamaño de muestra es «grande» (mayor a 30), es posible utilizar la distribución normal estándar.

Obsérvese la figura 4.23, donde se presentan varias gráficas de la distribución t de student con diferentes grados de libertad y la función de densidad de una distribución normal estándar. Es notorio, cómo las gráficas de la distribución t de student se van pareciendo más a la gráfica de la distribución normal estándar, conforme los grados de libertad van aumentando. En general, la distribución t de student converge a la distribución normal estándar, cuando los grados de libertad crecen. La demostración de este resultado no es parte de este material.

Ejercicios del capítulo 7

1.- Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución con función de densidad $f(x) = e^{-(x-a)} I_{[a,\infty)}(x)$. Sea $Y_1 = \min\{X_1, \ldots, X_n\}$. Demostrar que $\{Y_1\} \xrightarrow{d} a$.

2.- Considerando la misma muestra aleatoria del ejercicio anterior, sea Y_1 la mínima estadística de orden, encontrar la distribución límite de $\{Z_n\}$, donde $Z_n = n[Y_1 - a]$.

3.- Sea $\bar{X}_n$ la media muestral de una muestra aleatoria de tamaño n de una distribución normal con media μ y varianza σ^2. Encontrar la distribución límite de $\{\bar{X}_n\}$.

4.- Sean $Y_1, \ldots, Y_n$ las estadísticas de orden de una muestra aleatoria de tamaño n de una distribución $U(0,1)$. Demostrar que $\{Y_1\}$ converge en distribución a 0.

5.- Sea U_n una variable aleatoria con media c y varianza $\frac{d}{n}$, donde, c y d son constantes que no dependen de n y $d > 0$. Demostrar que la sucesión $\{U_n\}$ converge en probabilidad a c.

6.- Sea $\{U_n\}$ una sucesión de números reales, tal que $\{U_n\} \to c$. Demostrar que $\{U_n\} \xrightarrow{p} c$.

7.- Sea Y_n la $n-$ésima estadística de orden de una muestra aleatoria de tamaño n de una distribución con función de densidad $f(x) = e^{-(x-a)} I_{[a,\infty)}(x)$. Encontrar la distribución límite de $\{Z_n\}$, donde $Z_n = n[1 - F(Y_n)]$.

8.- Sea Y_n la $n-$ésima estadística de orden de una muestra aleatoria de tamaño n de una distribución continua con función de distribución $F(x)$ y función de densidad $f(x) = F'(x)$. Encontrar la distribución límite de $\{Z_n\}$, donde $Z_n = n[1 - F(Y_n)]$.

9.- Sea Y_2 la segunda estadística de orden de una muestra aleatoria de tamaño n de una distribución uniforme sobre el intervalo $(0,1)$. Encontrar la distribución límite de $\{Z_n\}$, donde $Z_n = nY_2$.

10.- Sea Y_2 la segunda estadística de orden de una muestra aleatoria de tamaño n de una distribución continua con función de distribución $F(x)$ y función de densidad dada por $f(x) = F'(x)$. Encontrar la distribución límite de $\{Z_n\}$, donde $Z_n = nF(Y_2)$.

11.- Sea U_n una variable aleatoria degenerada en n. Demostrar que $\{U_n\}$ no tiene distribución límite.

12.- Sea Y_1 la primera estadística de orden de una muestra aleatoria de tamaño n de una distribución geométrica con parámetro p. Encontrar la distribución límite de $\{U_n\} = \{Y_1\}$.

13.- Sea Y_1 la primera estadística de orden de una muestra aleatoria de tamaño n de una distribución uniforme discreta sobre el conjunto $\{0, 1, \ldots, m\}$. Encontrar la distribución límite de $\{U_n\} = \{Y_1\}$.

14.- Sea Y_1 la primera estadística de orden de una muestra aleatoria de tamaño n de una distribución discreta con recorrido $x = 0, 1, \ldots, k$. Sea $\{U_n\} = \{Y_1\}$, demostrar que $\{U_n\}$ converge en probabilidad a 0.

15.- Sea Y_n la última estadística de orden de una muestra aleatoria de tamaño n de una distribución geométrica con parámetro p. Sea $\{U_n\} = \{Y_n\}$. ¿Tiene distribución límite $\{U_n\}$? Justifique su respuesta.

16.- Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de tamaño n de una distribución normal con media μ y varianza σ^2. Sea $Z_n = \sum_{i=1}^{n} X_i$, demostrar que $\{Z_n\}$ no tiene distribución límite.

17.- Sea U_n una variable aleatoria con distribución binomial con parámetros n y p. Demostrar que la sucesión $\left\{\frac{U_n}{n}\right\}$ converge en probabilidad a p.

18.- Sea U_n una variable aleatoria con distribución binomial con parámetros n y p. Demostrar que la sucesión $\left\{1 - \frac{U_n}{n}\right\}$ converge en distribución a $1 - p$.

19.- Sea $X_1, \ldots, X_n$ una muestra aleatoria obtenida de una distribución con media μ , varianza σ^2 y función generadora de momentos $m_X(t)$. Demostrar la ley débil de los grandes números, usando el teorema 7.7.

20.- Resolver el ejercicio 12 usando el teorema 7.7.

21.- Sea $S_n^2 = \sum_{i=1}^{n} \frac{(X_i - \bar{X_n})^2}{n}$ donde $X_1, \ldots, X_n$ es una muestra aleatoria de tamaño n de una distribución normal con media μ y varianza σ^2. Demostrar que la sucesión $\left\{\frac{nS_n^2}{(n-1)}\right\}$ converge en distribución a σ^2.

22.- Sea Y_n la última estadística de orden de una muestra aleatoria de tamaño n de una distribución uniforme sobre el intervalo $(0, \sqrt{a})$. Demostrar que $\{U_n\} = \{Y_n\}$ converge en distribución $\sqrt{a}$.

23.- Sea Y_1 la primera estadística de orden de una muestra aleatoria de tamaño n de una distribución uniforme sobre el intervalo $(0, \sqrt{a})$. Demostrar que $\{U_n\} = \{Y_1\}$ converge en distribución a 0.

24.- Sea X_n una variable aleatoria con distribución Erlang con parámetros n y λ y $U_n = \frac{X_n}{n}$. Encontrar la distribución límite de $\{U_n\}$.

25.- Sea X_n una variable aleatoria con distribución ji cuadrada con n grados de libertad. Encontrar la distribución límite de $\{U_n\}$, donde $U_n = \frac{X_n}{n^2}$.

26.- Sea X_n una variable aleatoria con distribución ji cuadrada con n grados de libertad. Encontrar la distribución límite de $\{U_n\}$, donde $U_n = \frac{X_n}{n}$.

27.- Sea X una variable aleatoria con distribución ji cuadrada con 50 grados de libertad. Aproximar la probabilidad $P\,(40 < X < 60)$.

28.- Para un bono, se sabe que el retorno de inversión en un día será positivo con una probabilidad de 0.9. Se seleccionan 50 días al azar. Encontrar

a) La probabilidad de que al menos 47 días tengan retornos positivos.
b) La probabilidad aproximada calculada en a) usando la distribución Poisson.

29.- Sea X_n una variable aleatoria con distribución Poisson con parámetro $\lambda = n$ y $U_n = \frac{X_n - n}{\sqrt{n}}$. Demostrar que la sucesión $\{U_n\}$ converge en distribución a una distribución normal estándar.

30.- Sea $X_1, X_2, \ldots X_n$ una muestra aleatoria de tamaño n de una distribución Poisson con parámetro λ. Sea $U_n = \sqrt{n}\left(\frac{\bar{X}_n - \lambda}{\sqrt{\lambda}}\right)$, donde $\bar{X}_n$ es la media muestral. Demostrar, sin usar el teorema de límite central, que la sucesión $\{U_n\}$ converge en distribución a una variable aleatoria Z la cual tiene distribución normal estándar.

31.- Sea $X_1, X_2, \ldots X_n$ una muestra aleatoria de tamaño n de una distribución exponencial con parámetro $\lambda = 1/\beta$. Sea $U_n = \sqrt{n}\left(\frac{\bar{X}_n - \beta}{\beta}\right)$, donde $\bar{X}_n$ es la media muestral. Demostrar sin usar el teorema de límite central, que la sucesión $\{U_n\}$ converge en distribución a una variable aleatoria Z la cual tiene distribución normal estándar.

32.- Sean $Y_1, Y_2, \ldots, Y_n$ las estadísticas de orden de una muestra aleatoria de tamaño n obtenida de una distribución con función de densidad dada por $f_X(x) = e^{-(x-a)} I_{[a,\infty)}(x)$. Sea $U_n = Y_n - \ln(n)$. ¿La función de distribución de U_n converge cuando n tiende a infinito?, en caso afirmativo, ¿a qué función de distribución converge?

33.- Sean $Y_1, Y_2, \ldots, Y_n$ las estadísticas de orden de una muestra aleatoria de tamaño n obtenida de una distribución beta con parámetros $a = m$ y $b = 1$. Sea $U_n = n^{1/m} Y_1$. Demostrar que la sucesión $\{U_n\}$ converge en distribución a una variable aleatoria con distribución de Weibull.

34.- Sea $X_1, X_2, \ldots, X_{100}$ una muestra aleatoria de tamaño $n = 100$ de una distribución continua. Aproximar las siguientes probabilidades:

a) $P[59.5 < \bar{X} < 60.5]$, si la distribución es ji cuadrada con 60 grados de libertad.
b) $P[11 < \bar{X} < 13]$, si la distribución es gamma con parámetros $r = 3$ y $\lambda = 1/4$.
c) $P[0.45 < \bar{X} < 0.55]$, si la distribución es beta con parámetros $a = 2$ y $b = 2$.

35.- Sea Y la suma de los elementos de una muestra aleatoria de tamaño 15 de una distribución $UD\{1, 2, \ldots, 10\}$. Calcular aproximadamente $P(80 \leq Y \leq 90)$.

36.- Supongamos que $X_1, X_2, \ldots, X_{12}$ es una muestra aleatoria de 12 reclamos a un compañía de seguros por determinado seguro. Se sabe que X_i se distribuye $U(0, 1)$, $i = 1, \ldots, 12$. Sea Y el total de los 12 reclamos, calcular aproximadamente $P[5 \leq Y \leq 8]$.

37.- Sea Y una variable aleatoria con distribución binomial con parámetros n y p. Calcular aproximadamente

a) $P(0.25 < \hat{p})$, donde $\hat{p} = Y/n$, para una $n = 200$ y $p = 1/5$.
b) $P(Y = 55)$, para una $n = 100$ y $p = 1/2$.
c) El valor más pequeño de n para el cual se cumple $P(\hat{p} > 0.5) \geq 0.90$, considerando $p = 0.55$.

38.- Sea X la pérdida por la ocurrencia de determinado siniestro. La función de densidad de X está dada por $f_X(x) = \frac{1}{x^2} I_{(1,\infty)}(x)$. Considérese una muestra aleatoria 100 pérdidas por la ocurrencia de este siniestro. Calcular la probabilidad aproximada de que más de 80 siniestros sean menores a 3.

39.- Uno de los procedimientos para evaluar la calidad de un producto es medir su diámetro en cm. En cada turno se seleccionan aleatoriamente 25 piezas y se mide su diámetro. El error de medición es una variable aleatoria con distribución uniforme sobre el intervalo $(-0.1, 0.1)$. Calcular la probabilidad de que la suma de las 25 mediciones se aleje de cero, menos de 1 cm.

40.- El tiempo en segundos para que un trabajador termine un tarea específica tiene un promedio y una desviación estándar de 45 y 5 segundos, respectivamente. Se registra la duración en segundos de 81 tiempos seleccionados al azar. Calcular la probabilidad de que el promedio muestral de los tiempos

a) Sea mayor a 46 segundos.
b) Se desvíe de la verdadera media a lo más 1 segundo.

41.- Se sabe que las estaturas de hombres en cierta población tienen una desviación estándar de 10 cm. Se obtiene una muestra aleatoria de 64 hombres de esta población.

a) Calcular la probabilidad de que la media muestral se desvíe de la verdadera media a lo más 2 cm.
b) Si se desea que la media muestral se desvíe de la verdadera media a lo más 1 cm con una probabilidad del 0.95, encontrar el tamaño de muestra que cumple esta condición.

42.- Sea Y una variable aleatoria con distribución binomial con parámetros n y p.

a) Demostrar que la varianza de $\frac{Y}{n}$ alcanza su valor máximo en $p = 0.5$ para una n fija.
b) Se seleccionan al azar una muestra aleatoria de n elementos de una distribución Bernoulli con parámetro p, se define a Y como el número de éxitos en la muestra, ¿qué valor de n garantiza que la

proporción muestral $\hat{p} = \frac{Y}{n}$ se desvíe de la proporción real p a lo más 0.1 con una probabilidad de 0.95?

43.- Se sabe que la cantidad promedio de pérdidas por un siniestro, cuando este ocurre, es de 240 con una desviación estándar de 36. Se obtiene una muestra aleatoria de 49 siniestros, calcular la probabilidad de que la media muestral se encuentre entre 238 y 239.

44.- Se ha observado que los asegurados muestran insatisfacción hacia la compañía de seguros, cuando en un accidente automovilístico ellos esperan al ajustador, en el lugar de los hechos por más de 20 minutos. Se sabe que el tiempo de espera de un asegurado para que llegue el ajustador, desde que reporta el siniestro, es una variable aleatoria con distribución gamma con parámetros $r = 2$ y $\lambda = 0.1$. Se seleccionan al azar 30 siniestros automovilísticos, calcular la probabilidad aproximada de que al menos 5 asegurados esperen al ajustador más de 20 minutos.

45.- A una conferencia invitaron a 120 personas. Se sabe que una persona no llegará a la conferencia con una probabilidad del 10 %, además la asistencia o no de una persona es independiente de cualquier otra persona. El cupo del auditorio es de 100 lugares. ¿Cuál es la probabilidad de que todas las personas que lleguen a la conferencia tengan su lugar seguro?

46.- Se sabe que, para una acción financiera, 1 de cada 100 días tiene un retorno de rendimiento diario negativo. De 250 días seleccionados al azar, calcular la probabilidad aproximada de tener

a) Más de cuatro días con retorno de inversión negativo

b) Menos de 2 días con retorno negativo.

47.- Se sabe que el 25 % de los habitantes en una ciudad están en desacuerdo con el gobierno actual de esa población. Se seleccionan al azar 81 ciudadanos. Calcular aproximadamente

a) La probabilidad de que la proporción muestral de ciudadanos que están en desacuerdo difiera de la proporción verdadera a los más 0.05.

b) El tamaño de muestra n para que la proporción muestral de los ciudadanos que están en desacuerdo se desvíe de la proporción verdadera a lo más 0.03, con una probabilidad del 0.98.

48.- Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de tamaño n de una distribución gamma con parámetros r y $\lambda = 2$. Sea $U_n = \frac{\sqrt{n}(2\bar{X}_n - r)}{\sqrt{2\bar{X}_n}}$. Demostrar que $\{U_n\}$ converge en distribución a una variable aleatoria Z con distribución normal estándar.

49.- Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una distribución normal con parámetros μ y σ^2 y sea U_n definida como $U_n = \frac{\bar{X}_n - \mu}{\sqrt{\frac{S_n^2}{(n-1)}}}$, donde $\bar{X}_n$ y $S_n^2 = \sum_{i=1}^n \frac{(X_i - \bar{X}_n)^2}{n}$ son la media y la varianza muestrales respectivamente. Demostrar que $\{U_n\}$ converge en distribución a una variable Z con distribución normal estándar.

50.- Sean $X_1, X_2, \ldots, X_n$ y $Y_1, Y_2, \ldots, Y_n$ dos muestras aleatorias de tamaño n, independientes entre sí, obtenidas de distribuciones con medias μ_1 y μ_2, respectivamente, y con varianza común σ^2. Encontrar la distribución límite de $\frac{(\bar{X}_n - \bar{Y}_n) - (\mu_1 - \mu_2)}{\sigma\sqrt{\frac{2}{n}}}$, donde $\bar{X}_n$ es la media muestral de la primera muestra y $\bar{Y}_n$ la media muestral de la segunda muestra.

51.- Un actuario muestra, en cierta región, que el reclamo promedio individual de las mujeres, en un seguro de auto, es de 45,600 con una varianza de 98,800. Por otro lado, que el reclamo promedio individual de los hombres es 47,000 con una varianza igual a la varianza de los reclamos de las mujeres. De cada grupo de mujeres y hombres, se seleccionan al azar, en forma independiente, 49 reclamos. Calcular la probabilidad de que el promedio anual de reclamaciones de los hombres supere al promedio anual de reclamaciones de las mujeres al menos 1,500.

Bibliografía

[1] Asmussen, S. *Ruin probabilities.* World Scientific, Advanced Series on Statistical Science and Applied Probability, volumen 2, 2000.

[2] Batanero, C. y Serrano, L. *La aleatoriedad, sus significados e implicaciones educativas.* UNO, 5,15-28, 1995.

[3] Bennett, D. *The development of the mathematical concept of randomness; educational implications*, tesis doctoral. New York University, 1993.

[4] Canavos, G. C. *Probabilidad y Estadística. Aplicaciones y Métodos.* Mc Graw Hill, 1988.

[5] Capinski, M. y Kopp, E. *Measure, Integral and Probability*, segunda edición. Springer, 2005.

[6] Casella, G. y Berger, R. L. *Statistical Inference*, segunda edición. Duxbury, 2002.

[7] Feller, W. *A introduction to probability theory and its applications*, tercera edición, volumen I. Nueva York: Wiley, 1968.

[8] Feller, W. *A introduction to probability theory and its applications*, tercera edición, volumen II. Nueva York: Wiley, 1971.

[9] Galaz, F. *Medida e integral de Lebesgue en* $\mathbb{R}^n$. Oxford University Press, 2002.

[10] Guerrero, V. *Análisis estadístico de series de tiempo económicas.* International Thomson Editores, 2003.

[11] Harris, B. *Theory of probability.* Massachusetts: Addison-Wesley, 1966.

[12] Hogg, R. V. y Craig, A. T, *Introduction to the mathematical statistics*, cuarta edición. Nueva York: Collier Macmillan, 1978.

[13] Hogg, R. V. Mckean, J. y Craig, A.T. *Introduction to the mathematical statistics*, séptima edición. Edimburgo: Person education, 2014.

[14] Hogg, R. V. y Tanis, E. A, *Probability and Statistical Inference*, sexta edición. Prentice-Hall, 2001.

[15] Hull J. C. *Options, Futures, and other Derivatives*, novena edición. Pearson, 2015.

[16] Johnson N. L., Kemp A. W. y Kotz S. *Univariate discrete distributions*, tercera edición. Nueva York: Wiley, 2005.

[17] Johnson N. L., Kotz S. y Balakrishnan N. *Continuous univariate distributions*, segunda edición, volume 1. Nueva York: Wiley, 1994.

[18] Johnson N. L., Kotz S. y Balakrishnan N. *Continuous univariate distributions*, segunda edición, volume 2. Nueva York: Wiley, 1995.

[19] Kiburg H. *The logical foundations of statistical inference*. Boston: Reidel, 1974.

[20] Meyer, P. L. *Probabilidad y aplicaciones estadísticas*, segunda edición. Addison-Weslwy Iberoamerica, 1992.

[21] Mood, A. M., Graybill, F. A. y Boes, D. C. *Introduction to the theory of statistics*, tercera edición. Nueva York: Mc Graw Hill, 1974.

[22] Parzen, E. *Modern probability theory and its aplications*. Nueva York: Wiley, 1992.

[23] Poincaré H. *El azar*. (Artículo publicado originalmente en lengua inglesa en *Journal of the American Statistical Association*, 31, 10-30), 1936.

[24] Reiss, R. D. y Thomas, M. *Statistical analysis of extreme values with applications to insurance, finance, hydrology and other fields*, tercera edición. Birkhäuser Verlang, 2007.

[25] Ross, S. *A first course in probability*, décima edición. Reino Unido: Pearson, 2019.

[26] Ross, S. *Introduction to probability models*, décima edición. Reino Unido: Pearson, 2010.

[27] Rubinstein, R. Y. *Simulation and the Monte Carlo Method*. Wiley, 1981.

[28] Wackerly, D. D., Mendenhall, W. y Scheaffer, R. L. *Estadística matemática con aplicaciones*, séptima edición. Cengage Learning, 2010.

[29] Yiu-Kuen Tse. *Nonlife Actuarial Models Theory, Methods and Evaluation*. Cambridge University Press, 2009.

Apéndice A

Tablas

Tabla 1. Normal estándar: función de distribución

$$\phi(z) = P[Z \leq z]$$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

Valores más usados:

z	1.282	1.645	1.96	2.326	2.576	3.09	3.291	3.891	4.417
$\phi(z)$	0.90	0.95	0.975	0.99	0.995	0.999	0.9995	0.99995	0.99999
$2[1-\phi(z)]$	0.20	0.10	0.0.05	0.0.02	0.01	0.002	0.001	0.0001	0.00001

Tabla 2. *T Student*

$(1-\alpha)$-ésimo cuantiles

α / g.l.	0.1	0.05	0.025	0.01	0.005	α / g.l.
1	3.078	6.314	12.706	31.821	63.657	1
2	1.886	2.920	4.303	6.965	9.925	2
3	1.638	2.353	3.182	4.541	5.841	3
4	1.533	2.132	2.776	3.747	4.604	4
5	1.476	2.015	2.571	3.365	4.032	5
6	1.440	1.943	2.447	3.143	3.707	6
7	1.415	1.895	2.365	2.998	3.499	7
8	1.397	1.860	2.306	2.896	3.355	8
9	1.383	1.833	2.262	2.821	3.250	9
10	1.372	1.812	2.228	2.764	3.169	10
11	1.363	1.796	2.201	2.718	3.106	11
12	1.356	1.782	2.179	2.681	3.055	12
13	1.350	1.771	2.160	2.650	3.012	13
14	1.345	1.761	2.145	2.624	2.977	14
15	1.341	1.753	2.131	2.602	2.947	15
16	1.337	1.746	2.120	2.583	2.921	16
17	1.333	1.740	2.110	2.567	2.898	17
18	1.330	1.734	2.101	2.552	2.878	18
19	1.328	1.729	2.093	2.539	2.861	19
20	1.325	1.725	2.086	2.528	2.845	20
21	1.323	1.721	2.080	2.518	2.831	21
22	1.321	1.717	2.074	2.508	2.819	22
23	1.319	1.714	2.069	2.500	2.807	23
24	1.318	1.711	2.064	2.492	2.797	24
25	1.316	1.708	2.060	2.485	2.787	25
26	1.315	1.706	2.056	2.479	2.779	26
27	1.314	1.703	2.052	2.473	2.771	27
28	1.313	1.701	2.048	2.467	2.763	28
29	1.311	1.699	2.045	2.462	2.756	29
30	1.310	1.697	2.042	2.457	2.750	30
31	1.309	1.696	2.040	2.453	2.744	31
32	1.309	1.694	2.037	2.449	2.738	32
33	1.308	1.692	2.035	2.445	2.733	33
34	1.307	1.691	2.032	2.441	2.728	34
35	1.306	1.690	2.030	2.438	2.724	35
36	1.306	1.688	2.028	2.434	2.719	36
37	1.305	1.687	2.026	2.431	2.715	37
38	1.304	1.686	2.024	2.429	2.712	38
39	1.304	1.685	2.023	2.426	2.708	39
40	1.303	1.684	2.021	2.423	2.704	40
g.l. / α	0.1	0.05	0.025	0.01	0.005	g.l. / α

Tabla 3. Ji cuadrada. Cola izquierda

$(1-\alpha)$-ésimo cuantiles

α g.l.	**0.995**	**0.99**	**0.975**	**0.95**	**0.9**	α g.l.
1	0.0000	0.0002	0.0010	0.0039	0.0158	**1**
2	0.0100	0.0201	0.0506	0.1026	0.2107	**2**
3	0.0717	0.1148	0.2158	0.3518	0.5844	**3**
4	0.2070	0.2971	0.4844	0.7107	1.0636	**4**
5	0.4117	0.5543	0.8312	1.1455	1.6103	**5**
6	0.6757	0.8721	1.2373	1.6354	2.2041	**6**
7	0.9893	1.2390	1.6899	2.1673	2.8331	**7**
8	1.3444	1.6465	2.1797	2.7326	3.4895	**8**
9	1.7349	2.0879	2.7004	3.3251	4.1682	**9**
10	2.1559	2.5582	3.2470	3.9403	4.8652	**10**
11	2.6032	3.0535	3.8157	4.5748	5.5778	**11**
12	3.0738	3.5706	4.4038	5.2260	6.3038	**12**
13	3.5650	4.1069	5.0088	5.8919	7.0415	**13**
14	4.0747	4.6604	5.6287	6.5706	7.7895	**14**
15	4.6009	5.2293	6.2621	7.2609	8.5468	**15**
16	5.1422	5.8122	6.9077	7.9616	9.3122	**16**
17	5.6972	6.4078	7.5642	8.6718	10.0852	**17**
18	6.2648	7.0149	8.2307	9.3905	10.8649	**18**
19	6.8440	7.6327	8.9065	10.1170	11.6509	**19**
20	7.4338	8.2604	9.5908	10.8508	12.4426	**20**
21	8.0337	8.8972	10.2829	11.5913	13.2396	**21**
22	8.6427	9.5425	10.9823	12.3380	14.0415	**22**
23	9.2604	10.1957	11.6886	13.0905	14.8480	**23**
24	9.8862	10.8564	12.4012	13.8484	15.6587	**24**
25	10.5197	11.5240	13.1197	14.6114	16.4734	**25**
26	11.1602	12.1981	13.8439	15.3792	17.2919	**26**
27	11.8076	12.8785	14.5734	16.1514	18.1139	**27**
28	12.4613	13.5647	15.3079	16.9279	18.9392	**28**
29	13.1211	14.2565	16.0471	17.7084	19.7677	**29**
30	13.7867	14.9535	16.7908	18.4927	20.5992	**30**
31	14.4578	15.6555	17.5387	19.2806	21.4336	**31**
32	15.1340	16.3622	18.2908	20.0719	22.2706	**32**
33	15.8153	17.0735	19.0467	20.8665	23.1102	**33**
34	16.5013	17.7891	19.8063	21.6643	23.9523	**34**
35	17.1918	18.5089	20.5694	22.4650	24.7967	**35**
36	17.8867	19.2327	21.3359	23.2686	25.6433	**36**
37	18.5858	19.9602	22.1056	24.0749	26.4921	**37**
38	19.2889	20.6914	22.8785	24.8839	27.3430	**38**
39	19.9959	21.4262	23.6543	25.6954	28.1958	**39**
40	20.7065	22.1643	24.4330	26.5093	29.0505	**40**
41	21.4208	22.9056	25.2145	27.3256	29.9071	**41**
42	22.1385	23.6501	25.9987	28.1440	30.7654	**42**
g.l. α	**0.995**	**0.99**	**0.0.975**	**0.95**	**0.9**	**g.l.** α

Tabla 4. Ji cuadrada. Cola derecha

$(1-\alpha)$-ésimo cuantiles

α g.l.	**0.1**	**0.05**	**0.025**	**0.01**	**0.005**	α g.l.
1	2.7055	3.8415	5.0239	6.6349	7.8794	**1**
2	4.6052	5.9915	7.3778	9.2103	10.5966	**2**
3	6.2514	7.8147	9.3484	11.3449	12.8382	**3**
4	7.7794	9.4877	11.1433	13.2767	14.8603	**4**
5	9.2364	11.0705	12.8325	15.0863	16.7496	**5**
6	10.6446	12.5916	14.4494	16.8119	18.5476	**6**
7	12.0170	14.0671	16.0128	18.4753	20.2777	**7**
8	13.3616	15.5073	17.5345	20.0902	21.9550	**8**
9	14.6837	16.9190	19.0228	21.6660	23.5894	**9**
10	15.9872	18.3070	20.4832	23.2093	25.1882	**10**
11	17.2750	19.6751	21.9200	24.7250	26.7568	**11**
12	18.5493	21.0261	23.3367	26.2170	28.2995	**12**
13	19.8119	22.3620	24.7356	27.6882	29.8195	**13**
14	21.0641	23.6848	26.1189	29.1412	31.3193	**14**
15	22.3071	24.9958	27.4884	30.5779	32.8013	**15**
16	23.5418	26.2962	28.8454	31.9999	34.2672	**16**
17	24.7690	27.5871	30.1910	33.4087	35.7185	**17**
18	25.9894	28.8693	31.5264	34.8053	37.1565	**18**
19	27.2036	30.1435	32.8523	36.1909	38.5823	**19**
20	28.4120	31.4104	34.1696	37.5662	39.9968	**20**
21	29.6151	32.6706	35.4789	38.9322	41.4011	**21**
22	30.8133	33.9244	36.7807	40.2894	42.7957	**22**
23	32.0069	35.1725	38.0756	41.6384	44.1813	**23**
24	33.1962	36.4150	39.3641	42.9798	45.5585	**24**
25	34.3816	37.6525	40.6465	44.3141	46.9279	**25**
26	35.5632	38.8851	41.9232	45.6417	48.2899	**26**
27	36.7412	40.1133	43.1945	46.9629	49.6449	**27**
28	37.9159	41.3371	44.4608	48.2782	50.9934	**28**
29	39.0875	42.5570	45.7223	49.5879	52.3356	**29**
30	40.2560	43.7730	46.9792	50.8922	53.6720	**30**
31	41.4217	44.9853	48.2319	52.1914	55.0027	**31**
32	42.5847	46.1943	49.4804	53.4858	56.3281	**32**
33	43.7452	47.3999	50.7251	54.7755	57.6484	**33**
34	44.9032	48.6024	51.9660	56.0609	58.9639	**34**
35	46.0588	49.8018	53.2033	57.3421	60.2748	**35**
36	47.2122	50.9985	54.4373	58.6192	61.5812	**36**
37	48.3634	52.1923	55.6680	59.8925	62.8833	**37**
38	49.5126	53.3835	56.8955	61.1621	64.1814	**38**
39	50.6598	54.5722	58.1201	62.4281	65.4756	**39**
40	51.8051	55.7585	59.3417	63.6907	66.7660	**40**
41	52.9485	56.9424	60.5606	64.9501	68.0527	**41**
42	54.0902	58.1240	61.7768	66.2062	69.3360	**42**
g.l. α	**0.1**	**0.05**	**0.025**	**0.01**	**0.005**	g.l. α

Tabla 5. Distribución F.

0.9-ésimo cuantiles

v_1 = Grados de libertad del numerador
v_2 = Grados de libertad del denominador

v_1 / v_2	1	2	3	4	5	6	7	8	9	10
1	39.864	49.500	53.593	55.833	57.240	58.204	58.906	59.439	59.857	60.195
2	8.526	9.000	9.162	9.243	9.293	9.326	9.349	9.367	9.381	9.392
3	5.538	5.462	5.391	5.343	5.309	5.285	5.266	5.252	5.240	5.230
4	4.545	4.325	4.191	4.107	4.051	4.010	3.979	3.955	3.936	3.920
5	4.060	3.780	3.619	3.520	3.453	3.405	3.368	3.339	3.316	3.297
6	3.776	3.463	3.289	3.181	3.108	3.055	3.014	2.983	2.958	2.937
7	3.589	3.257	3.074	2.961	2.883	2.827	2.785	2.752	2.725	2.703
8	3.458	3.113	2.924	2.806	2.726	2.668	2.624	2.589	2.561	2.538
9	3.360	3.006	2.813	2.693	2.611	2.551	2.505	2.469	2.440	2.416
10	3.285	2.924	2.728	2.605	2.522	2.461	2.414	2.377	2.347	2.323
11	3.225	2.860	2.660	2.536	2.451	2.389	2.342	2.304	2.274	2.248
12	3.177	2.807	2.606	2.480	2.394	2.331	2.283	2.245	2.214	2.188
13	3.136	2.763	2.560	2.434	2.347	2.283	2.234	2.195	2.164	2.138
14	3.102	2.726	2.522	2.395	2.307	2.243	2.193	2.154	2.122	2.095
15	3.073	2.695	2.490	2.361	2.273	2.208	2.158	2.119	2.086	2.059
16	3.048	2.668	2.462	2.333	2.244	2.178	2.128	2.088	2.055	2.028
17	3.026	2.645	2.437	2.308	2.218	2.152	2.102	2.061	2.028	2.001
18	3.007	2.624	2.416	2.286	2.196	2.130	2.079	2.038	2.005	1.977
19	2.990	2.606	2.397	2.266	2.176	2.109	2.058	2.017	1.984	1.956
20	2.975	2.589	2.380	2.249	2.158	2.091	2.040	1.999	1.965	1.937

0.95-ésimo cuantiles

v_1 / v_2	1	2	3	4	5	6	7	8	9	10
1	161.446	199.499	215.707	224.583	230.160	233.988	236.767	238.884	240.543	241.882
2	18.513	19.000	19.164	19.247	19.296	19.329	19.353	19.371	19.385	19.396
3	10.128	9.552	9.277	9.117	9.013	8.941	8.887	8.845	8.812	8.785
4	7.709	6.944	6.591	6.388	6.256	6.163	6.094	6.041	5.999	5.964
5	6.608	5.786	5.409	5.192	5.050	4.950	4.876	4.818	4.772	4.735
6	5.987	5.143	4.757	4.534	4.387	4.284	4.207	4.147	4.099	4.060
7	5.591	4.737	4.347	4.120	3.972	3.866	3.787	3.726	3.677	3.637
8	5.318	4.459	4.066	3.838	3.688	3.581	3.500	3.438	3.388	3.347
9	5.117	4.256	3.863	3.633	3.482	3.374	3.293	3.230	3.179	3.137
10	4.965	4.103	3.708	3.478	3.326	3.217	3.135	3.072	3.020	2.978
11	4.844	3.982	3.587	3.357	3.204	3.095	3.012	2.948	2.896	2.854
12	4.747	3.885	3.490	3.259	3.106	2.996	2.913	2.849	2.796	2.753
13	4.667	3.806	3.411	3.179	3.025	2.915	2.832	2.767	2.714	2.671
14	4.600	3.739	3.344	3.112	2.958	2.848	2.764	2.699	2.646	2.602
15	4.543	3.682	3.287	3.056	2.901	2.790	2.707	2.641	2.588	2.544
16	4.494	3.634	3.239	3.007	2.852	2.741	2.657	2.591	2.538	2.494
17	4.451	3.592	3.197	2.965	2.810	2.699	2.614	2.548	2.494	2.450
18	4.414	3.555	3.160	2.928	2.773	2.661	2.577	2.510	2.456	2.412
19	4.381	3.522	3.127	2.895	2.740	2.628	2.544	2.477	2.423	2.378
20	4.351	3.493	3.098	2.866	2.711	2.599	2.514	2.447	2.393	2.348

Tabla 5. Distribución F. Continuación

0.975-ésimo cuantiles

v_1 = Grados de libertad del numerador
v_2 = Grados de libertad del denominador

v_1 / v_2	1	2	3	4	5	6	7	8	9	10
1	647.793	799.482	864.151	899.599	921.835	937.114	948.203	956.643	963.279	968.634
2	38.506	39.000	39.166	39.248	39.298	39.331	39.356	39.373	39.387	39.398
3	17.443	16.044	15.439	15.101	14.885	14.735	14.624	14.540	14.473	14.419
4	12.218	10.649	9.979	9.604	9.364	9.197	9.074	8.980	8.905	8.844
5	10.007	8.434	7.764	7.388	7.146	6.978	6.853	6.757	6.681	6.619
6	8.813	7.260	6.599	6.227	5.988	5.820	5.695	5.600	5.523	5.461
7	8.073	6.542	5.890	5.523	5.285	5.119	4.995	4.899	4.823	4.761
8	7.571	6.059	5.416	5.053	4.817	4.652	4.529	4.433	4.357	4.295
9	7.209	5.715	5.078	4.718	4.484	4.320	4.197	4.102	4.026	3.964
10	6.937	5.456	4.826	4.468	4.236	4.072	3.950	3.855	3.779	3.717
11	6.724	5.256	4.630	4.275	4.044	3.881	3.759	3.664	3.588	3.526
12	6.554	5.096	4.474	4.121	3.891	3.728	3.607	3.512	3.436	3.374
13	6.414	4.965	4.347	3.996	3.767	3.604	3.483	3.388	3.312	3.250
14	6.298	4.857	4.242	3.892	3.663	3.501	3.380	3.285	3.209	3.147
15	6.200	4.765	4.153	3.804	3.576	3.415	3.293	3.199	3.123	3.060
16	6.115	4.687	4.077	3.729	3.502	3.341	3.219	3.125	3.049	2.986
17	6.042	4.619	4.011	3.665	3.438	3.277	3.156	3.061	2.985	2.922
18	5.978	4.560	3.954	3.608	3.382	3.221	3.100	3.005	2.929	2.866
19	5.922	4.508	3.903	3.559	3.333	3.172	3.051	2.956	2.880	2.817
20	5.871	4.461	3.859	3.515	3.289	3.128	3.007	2.913	2.837	2.774

0.99-ésimo cuantiles

v_1 / v_2	1	2	3	4	5	6	7	8	9	10
1	4052.185	4999.340	5403.534	5624.257	5763.955	5858.950	5928.334	5980.954	6022.397	6055.925
2	98.502	99.000	99.164	99.251	99.302	99.331	99.357	99.375	99.390	99.397
3	34.116	30.816	29.457	28.710	28.237	27.911	27.671	27.489	27.345	27.228
4	21.198	18.000	16.694	15.977	15.522	15.207	14.976	14.799	14.659	14.546
5	16.258	13.274	12.060	11.392	10.967	10.672	10.456	10.289	10.158	10.051
6	13.745	10.925	9.780	9.148	8.746	8.466	8.260	8.102	7.976	7.874
7	12.246	9.547	8.451	7.847	7.460	7.191	6.993	6.840	6.719	6.620
8	11.259	8.649	7.591	7.006	6.632	6.371	6.178	6.029	5.911	5.814
9	10.562	8.022	6.992	6.422	6.057	5.802	5.613	5.467	5.351	5.257
10	10.044	7.559	6.552	5.994	5.636	5.386	5.200	5.057	4.942	4.849
11	9.646	7.206	6.217	5.668	5.316	5.069	4.886	4.744	4.632	4.539
12	9.330	6.927	5.953	5.412	5.064	4.821	4.640	4.499	4.388	4.296
13	9.074	6.701	5.739	5.205	4.862	4.620	4.441	4.302	4.191	4.100
14	8.862	6.515	5.564	5.035	4.695	4.456	4.278	4.140	4.030	3.939
15	8.683	6.359	5.417	4.893	4.556	4.318	4.142	4.004	3.895	3.805
16	8.531	6.226	5.292	4.773	4.437	4.202	4.026	3.890	3.780	3.691
17	8.400	6.112	5.185	4.669	4.336	4.101	3.927	3.791	3.682	3.593
18	8.285	6.013	5.092	4.579	4.248	4.015	3.841	3.705	3.597	3.508
19	8.185	5.926	5.010	4.500	4.171	3.939	3.765	3.631	3.523	3.434
20	8.096	5.849	4.938	4.431	4.103	3.871	3.699	3.564	3.457	3.368

Apéndice B

Tablas de distribuciones

En este apéndice se presentarán las tablas de distribuciones discretas y continuas. Para cada distribución, se indicará la función de densidad correspondiente, también el valor esperado, la varianza y la función generadora de momentos, si esta existe. Para ampliar las propiedades de cada distribución se puede consultar el capítulo 3 para las distribuciones discretas y el capítulo 4 para las distribuciones continuas.

Tabla 1. Distribuciones discretas

Distribución	Función de densidad	Media	Varianza	Función generadora de momentos
Uniforme discreta	$\frac{1}{N} I_{\{1,\ldots,N\}}(x)$	$\frac{N+1}{2}$	$\frac{N^2-1}{12}$	$\frac{e^t(1-e^{tN})}{N(1-e^t)}$, si $t \neq 0$ 1 si $t = 0$
Bernouilli	$p^x q^{1-x} I_{\{0,1\}}(x)$	p	$p(1-p)$	$q + pe^t$
Binomial	$\binom{n}{x} p^x q^{n-x}$ $\times I_{\{0,\ldots,n\}}(x)$	np	npq	$(q + pe^t)^n$
Hipergeométrica	$\frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$ $\times I_{\{0,\ldots,n\}}(x)$	$n\left(\frac{K}{N}\right)$	$n\left(\frac{K}{N}\right)\left(1-\frac{K}{N}\right)$ $\times\left(\frac{N-n}{N-1}\right)$	No es útil
Geométrica	$pq^x I_{\{0,1,\ldots\}}(x)$	$\frac{q}{p}$	$\frac{q}{p^2}$	$\frac{p}{1-qe^t}$, si $t < -ln(q)$
Binomial negativa	$\binom{r+x-1}{x} p^r q^x$ $\times I_{\{0,1,\ldots\}}(x)$	$\frac{rq}{p}$	$\frac{rq}{p^2}$	$\left(\frac{p}{1-qe^t}\right)^r$, si $t < -ln(q)$
Poisson	$\frac{\lambda^x}{x!} e^{-\lambda} I_{\{0,1,\ldots\}}(x)$	λ	λ	$e^{\lambda(e^t-1)}$

Tabla 2. Distribuciones continuas

Distribución	Función de densidad	Media	Varianza	Función generadora de momentos
Uniforme continua	$\frac{1}{b-a} I_{[a,b]}(x)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{bt}-e^{at}}{t(b-a)}$, si $t \neq 0$ 1 si $t = 0$
Gamma	$\frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x} I_{(0,\infty)}(x)$	$\frac{r}{\lambda}$	$\frac{r}{\lambda^2}$	$\left(\frac{\lambda}{\lambda-t}\right)^r$, si $t < \lambda$
Exponencial	$\lambda e^{-\lambda x} I_{[0,\infty)}(x)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{\lambda}{\lambda-t}$, si $t < \lambda$
Beta	$\frac{1}{\beta(a,b)} x^{a-1}(1-x)^{b-1}$ $\times I_{(0,1)}(x)$	$\frac{a}{a+b}$	$\frac{ab}{(a+b)^2(a+b+1)}$	No es útil
Weibull	$\frac{b}{a^b} x^{b-1} e^{-(\frac{x}{a})^b}$ $\times I_{(0,\infty)}(x)$	$a\Gamma(1/b+1)$	$a^2\Gamma(2/b+1)$ $-a^2\Gamma^2(1/b+1)$	No existe para $0 < b < 1$
Normal	$\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2	$e^{\mu t + \frac{\sigma^2 t^2}{2}}$

Continuación Tabla 2. Distribuciones continuas

Distribución	Función de densidad	Media	Varianza	Función generadora de momentos
Lognormal	$\frac{1}{x\sqrt{2\pi}\sigma}e^{-\frac{1}{2\sigma^2}(ln(x)-\mu)^2}$ $\times I_{(0,\infty)}(x)$	$e^{\mu+\frac{\sigma^2}{2}}$	$e^{2\mu+2\sigma^2}$ $-e^{2\mu+\sigma^2}$	No existe para $t>0$
Pareto	$\frac{ab^a}{(b+x)^{a+1}}I_{(0,\infty)}(x)$	$\frac{b}{a-1}$ si $a>1$	$\frac{ab^2}{(a-1)(a-2)}$ si $a>2$	No existe
Ji cuadrada	$\frac{(1/2)^{\nu/2}}{\Gamma(\nu/2)}x^{\nu/2-1}e^{-x/2}$ $\times I_{(0,\infty)}(x)$	ν	2ν	$(1-2t)^{-\nu/2}$, si $t<1/2$
T de student	$\frac{\Gamma\left((\nu+1)/2\right)}{\Gamma\left(\nu/2\right)}\frac{1}{\sqrt{\pi\nu}}$ $\times\frac{1}{(1+x^2/\nu)^{(\nu+1)/2}}$	0, si $\nu>1$	$\frac{\nu}{\nu-2}$, si $\nu>2$	No existe
F	$\frac{\Gamma\left((\nu_1+\nu_2)/2\right)}{\Gamma\left(\nu_1/2\right)\Gamma\left(\nu_2/2\right)}\left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}}$ $\times\frac{(x)^{\nu_1/2-1}}{(1+\nu_1x/\nu_2)^{(\nu_1+\nu_2)/2}}$ $\times I_{(0,\infty)}(x)$	$\frac{\nu_2}{\nu_2-2}$, si $\nu_2>2$	$\frac{2\nu_2^2(\nu_1+\nu_2-2)}{\nu_1(\nu_2-2)^2(\nu_2-4)}$, si $\nu_2>4$	No existe

Apéndice C

Resultados de matemáticas

En este apéndice se presentarán diversos resultados de matemáticas que son usados en temas de probabilidad. En primer lugar, se mostrarán los números combinatorios y sus principales propiedades. Posteriormente, se darán los resultados de algunas sumas que son usadas en probabilidad, principalmente en tema de distribuciones discretas. En seguida se presentarán las definiciones de series y polinomios de Taylor, así como algunas propiedades. Finalmente, se mostrará la definición de convergencia de sucesiones de números reales, dando algunos ejemplos que facilitarán al lector su comprensión.

C.1. Números combinatorios

En esta sección se van a presentar los números combinatorios y sus propiedades más importantes. Estos números son fundamentales en las técnicas de conteo y, por lo tanto, en el cálculo de probabilidades. Los números combinatorios también aparecen en diferentes modelos de probabilidad.

El *factorial* de un número entero positivo n, se denota como $n!$, y es el producto desde el número 1 hasta el número n, esto es,

$$n! = (1)(2)\cdots(n).$$

En el caso particular del factorial de 0, por convención, se define como

$$0! = 1.$$

El número de maneras de seleccionar un subconjunto con k elementos de un conjunto con n elementos, se denota C_k^n o $\binom{n}{k}$, y se calcula como

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Los números $\binom{n}{k}$ son llamados *números combinatorios o coeficientes binomiales.*

Como se había comentado, los números combinatorios son muy importantes en el análisis combinatorio o técnicas de conteo y, por lo tanto, en el cálculo de probabilidades, véase la sección 1.10. Algunas propiedades de estos números se presentan a continuación:

1) $\binom{n}{0} = \binom{n}{n} = 1.$

2) $\binom{n}{1} = \binom{n}{n-1} = n.$

3) $\binom{n}{2} = \binom{n}{n-2} = \dfrac{n(n-1)}{2}.$

4) $\binom{n}{k} = \binom{n}{n-k}.$

La propiedad 1) indica que solo hay un subconjunto con cero elementos, siendo, en este caso, el conjunto vacío. También indica que solo hay un subconjunto con n elementos, siendo este conjunto, el conjunto universo.

La propiedad 2) nos indica que hay n subconjuntos con un elemento y n subconjuntos con $n-1$ elementos. Obsérvese que, por cada subconjunto con un elemento, existe un subconjunto complemento con $n-1$ elementos.

En la siguiente propiedad, se presenta una opción de calcular en forma más directa el número de subconjuntos de tamaño 2, y el número de subconjuntos de tamaño $n-2$.

Una manera de interpretar la propiedad 4) es que por cada subconjunto de tamaño k, existe un subconjunto complemento de tamaño $n-k$, por lo que el número de subconjuntos de tamaño k, coincide con el número de subconjuntos de tamaño $n-k$.

Los números combinatorios aparecen en el *triángulo de Pascal*, a continuación, se presentan los 7 primeros renglones de este triángulo.

$$
\begin{array}{ccccccccccccc}
 & & & & & & 1 & & & & & & \\
 & & & & & 1 & & 1 & & & & & \\
 & & & & 1 & & 2 & & 1 & & & & \\
 & & & 1 & & 3 & & 3 & & 1 & & & \\
 & & 1 & & 4 & & 6 & & 4 & & 1 & & \\
 & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & \\
1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1
\end{array}
$$

Tabla C.1. *Triángulo de Pascal, (7 renglones)*

Obsérvese que el renglón $k+1$ se puede obtener de la siguiente manera:

$$\binom{k}{0}, \binom{k}{1}, \ldots, \binom{k}{k}.$$

Se puede notar, que el primer renglón del triángulo está conformado por el 1. Para cada uno de los siguientes renglones, el primer y el último elemento son el número 1, y para cualquier otro elemento, este es la suma de los dos números más cercanos del renglón anterior.

Los números combinatorios aparecen también en el teorema del binomio, entre otros resultados, véase en este mismo apéndice, la igualdad C.7.

C.2. Sucesiones

En diferentes temas de probabilidad son utilizados algunos límites de sucesiones, principalmente en el último tema: convergencia de sucesiones de variables aleatorias. En esta sección se presentarán las definiciones básicas de sucesiones y algunos ejemplos que ayudarán a entender el concepto de convergencia.

Definición C.1. Una *sucesión de números reales* o simplemente una *sucesión* es una función real con dominio, los números naturales.

Dicho de otra manera, a cada número natural se le asigna un número real de acuerdo a una regla de correspondencia.

Una sucesión la denotaremos como $\{U_n\}$.

Ejemplo C.1. A continuación, se darán varios ejemplos de sucesiones

$$\{U_n\} = \left\{\frac{1}{2^n}\right\} = \left\{\frac{1}{2}, \frac{1}{4}, \ldots\right\},$$

$$\{U_n\} = \{2^n\} = \{2, 4, \ldots, \},$$

$$\{U_n\} = \left\{\frac{1}{n}\right\} = \left\{1, \frac{1}{2}, \frac{1}{3}, \ldots\right\},$$

$$\{U_n\} = \{n\} = \{1, 2, \ldots, \},$$

$$\{U_n\} = \{\sqrt[n]{2}\} = \{2, \sqrt{2}, \sqrt[3]{2}, \ldots, \}.$$

En seguida se presentará la definición de convergencia de una sucesión de números reales.

Definición C.2. Se dice que una sucesión $\{U_n\}$ converge a L, si para todo $\epsilon > 0$, existe una $m \in \mathbb{N}$, tal que para todo $n \geq m$, se cumple $|U_n - L| < \epsilon$.

Si la sucesión $\{U_n\}$ converge a L, se denotará como $\{U_n\} \to L$.

Si una sucesión $\{U_n\}$ no converge a un número, se dice que la sucesión $\{U_n\}$ *diverge.*

Obsérvese para cada sucesión del ejemplo anterior, si la sucesión converge o diverge.

$$\{U_n\} = \left\{\frac{1}{2^n}\right\} \to 0,$$
$$\{U_n\} = \{2^n\} \text{ diverge,}$$
$$\{U_n\} = \left\{\frac{1}{n}\right\} \to 0,$$
$$\{U_n\} = \{n\} \text{ diverge,}$$
$$\{U_n\} = \{\sqrt[n]{2}\} \to 1.$$

Ejemplo C.2. Demostrar que $\{U_n\} = \left\{\frac{1}{2^n}\right\} \to 0.$

Sea $\epsilon > 0$, busquemos $m \in \mathbb{N}$ tal que $|U_m - 0| < \epsilon$.
Obsérvese, $|U_m - 0| < \epsilon$ si y solo si $\left|\frac{1}{2^m}\right| < \epsilon$. Esto es, si y solo si $\frac{1}{2^m} < \epsilon$.
La anterior desigualdad se cumple, si y solo si $m > \frac{-ln(\epsilon)}{ln(2)}$.
En conclusión, para $m > \frac{-ln(\epsilon)}{ln(2)}$ se cumple $|U_m - 0| < \epsilon$.
Por ejemplo, para el caso particular de $\epsilon = 0.001$, $m > \frac{-ln(0.001)}{ln(2)} = 9.96578 \approx 10$.
Esto es, la desigualdad se va a cumplir para $n = 10, 11, \ldots$

Ejemplo C.3. Demostrar que $\{U_n\} = \{\sqrt[n]{2}\} \to 1.$

Sea $\epsilon > 0$, el objetivo es encontrar $m \in \mathbb{N}$ tal que $|U_m - 1| < \epsilon$, esto es, m que cumple con $|\sqrt[m]{2} - 1| < \epsilon$.

Obsérvese que $\sqrt[m]{2} - 1 > 0$ para cualquier valor de m. De esta manera, $|\sqrt[m]{2} - 1| < \epsilon$ se cumple, si y solo si $\sqrt[m]{2} - 1 < \epsilon$. La anterior desigualdad se cumple si y solo si $m > \frac{ln(2)}{ln(\epsilon+1)}$.

Consideremos el siguiente caso particular, si $\epsilon = 0.0001$, entonces, $m > \frac{ln(2)}{ln(1.0001)} = 6931.81837 \approx 6932$. En conclusión, la desigualdad se cumple para valores de $n = 6932, 6933, \ldots$

C.3. Resultados límites

En esta sección se presentarán dos resultados límites que son usados en temas de probabilidad.

Los siguientes dos resultados son generalizaciones de la definición del número e.

$$\lim_{n\to\infty} \left(1 + \frac{A}{n} + \frac{\varphi(n)}{n}\right)^{Bn} = e^{AB}, \text{ si } \varphi(n) \to 0. \tag{C.1}$$

En particular,

$$\lim_{n\to\infty} \left(1 + \frac{A}{n}\right)^{Bn} = e^{AB}. \tag{C.2}$$

Es importante aclarar, para el caso particular $a = b = 1$ y $\varphi(n) = 0$, se considera como la defınción del número e.

C.4. Sumas y series

A continuación, se presentarán diferentes resultados de sumas que son utilizadas en temas de probabilidad, algunas son sumas finitas y otras son infinitas, varios de estos resultados son sumas o series muy utilizadas en matemáticas.

C.4.1. Suma de los primeros n números naturales

La suma de los primeros n números naturales es igual a

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}. \tag{C.3}$$

C.4.2. Suma de los cuadrados de los primeros n números naturales

La suma de los cuadrados de los primeros n números naturales es igual a

$$\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}. \tag{C.4}$$

Ejemplo C.4. Obtener la suma de los 10 primeros números naturales y la suma de los cuadrados de los primeros 5 números naturales.

Las sumas son, respectivamente,

$$\frac{10(11)}{2} = 55.$$

$$\frac{(5)(6)(11)}{6} = 55.$$

C.4.3. Sumas geométricas

Una *serie geométrica* se define como la siguiente suma

$$a + ar + ar^2 + \cdots = a\sum_{i=0}^{\infty} r^i.$$

Obsérvese que, para cualquier sumando, a partir del segundo, si este se divide entre su predecesor, el resultado siempre dará r. Entonces, para una serie geométrica, a es el *coeficiente* de cada término y r es la *razón* de dos términos sucesivos.

La suma de los primeros $n+1$ elementos de una serie geométrica se calcula como

$$\sum_{i=0}^{n} ar^i = a\left(\frac{1-r^{n+1}}{1-r}\right). \tag{C.5}$$

Por otro lado, la suma de los elementos de una serie geométrica se calcula de la siguiente manera:

$$\sum_{i=0}^{\infty} ar^i = \frac{a}{1-r}, \quad \text{si} \quad |r|<1. \tag{C.6}$$

El anterior resultado, se obtiene sacando el límite de la ecuación C.5 cuando n tiende a ∞. Este límite será cierto siempre y cuando $|r|<1$.

Ejemplo C.5. Consideremos la serie $5+5/3+5/9+5/27+\ldots$, calcular
a) El valor de la serie.
b) La suma de los primeros 10 elementos de la serie.

Obsérvese que la serie es geométrica, donde $a=5$ y $r=1/3$.
a) El valor de la serie es igual a

$$\frac{a}{1-r} = \frac{5}{1-1/3} = \frac{15}{2} = 7.5.$$

b) La suma de los primeros 10 elementos de la serie es

$$\sum_{i=0}^{9} 5\left(\frac{1}{3}\right)^i = 5\left(\frac{1-(1/3)^{10}}{1-1/3}\right) = 7.49987.$$

Las sumas geométricas son utilizadas en varios resultados de probabilidad, en particular, son usadas para comprobar algunas propiedades de la distribución geométrica, véase sección 3.5.

C.4.4. Sumas con números combinatorios

Existen sumas de mucha utilidad en las matemáticas donde están involucrados algunos números combinatorios. Algunas de estas sumas se utilizan en temas de probabilidad.

La primera suma es conocida como el *teorema del binomio* y esta se da a continuación:

$$(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i}. \tag{C.7}$$

Esta suma calcula la potencia de un binomio, y en cada sumando aparece un coeficiente binomial.

También se puede ver que para una n determinada, los coeficientes binomiales que aparecen en la suma corresponden a los elementos del renglón $n+1$ del triángulo de Pascal, véase tabla C.1.

El teorema del binomio es usado en varias ocasiones en probabilidad, en particular, en la distribución discreta binomial, véase la sección 3.3.

La segunda suma donde aparecen coeficientes binomiales se presenta a continuación:

$$\sum_{i=0}^{n}\binom{A}{i}\binom{N-A}{n-i}=\binom{N}{n}. \tag{C.8}$$

Esta igualdad tiene una interpretación práctica. Primero, el número de subconjuntos de tamaño n que se pueden formar de un conjunto de tamaño N es

$$\binom{N}{n}.$$

Por otra parte, pensemos que el conjunto de tamaño N, se particiona en dos subconjuntos, uno de tamaño A, el cual llamaremos S_1 y el otro de tamaño $N-A$, el cual llamaremos S_2. Entonces, cada subconjunto de tamaño n formado del conjunto de tamaño N, se puede formar de varias maneras, considerando 0 elementos del subconjunto S_1 y n elementos del subconjunto S_2, o 1 elemento del subconjunto S_1 y $n-1$ elementos S_2, o 2 elementos de S_1 y $n-2$ elementos del subconjunto S_2, y así sucesivamente, hasta considerar n elementos seleccionados del subconjunto S_1 y 0 del subconjunto S_2. La suma de los casos anteriores es igual a

$$\sum_{i=0}^{n}\binom{A}{i}\binom{N-A}{n-i}.$$

De este modo, queda justificada la igualdad C.8-

La anterior suma es utilizada en la distribución hipergeométrica, véase la sección 3.4.

Otra suma importante que es utilizada en temas de probabilidad es el desarrollo del binomio $(1-a)$ a una potencia negativa, esto es,

$$(1-a)^{-n}=\sum_{i=0}^{\infty}\binom{n+i-1}{i}a^{i}, \quad \text{donde } |a|<1. \tag{C.9}$$

Este resultado es usado para obtener algunas propiedades de la distribución binomial negativa, véase la sección 3.6.

Ejemplo C.6. Desarrollar los siguientes binomios en sumas:
a) $(x-1)^5$.
b) $(1-x)^{-5}$, donde |x|<1.

a) El desarrollo queda así

$$(x-1)^5 = x^5 - 5x^4 + 10x^3 - 10x^2 + 5x - 1.$$

b) El desarrollo del segundo binomio

$$\begin{aligned}
(1-x)^{-5} &= \sum_{i=0}^{\infty} \binom{5+i-1}{i} x^i \\
&= \binom{4}{0}x^0 + \binom{5}{1}x + \binom{6}{2}x^2 + \binom{7}{3}x^3 + \binom{8}{4}x^4 + \binom{9}{5}x^5 + \cdots \\
&= 1 + 5x + 15x^2 + 35x^3 + 70x^4 + 126x^5 + \cdots
\end{aligned}$$

C.4.5. Otra serie

La siguiente serie es útil en el tema de variables aleatorias discretas

$$\sum_{i=1}^{\infty} \frac{1}{i^2} = 1 + \frac{1}{4} + \frac{1}{9} + \cdots = \frac{\pi}{6}. \tag{C.10}$$

C.5. Series de Taylor

En esta sección se presentará el desarrollo de una función por medio de series de Taylor. En particular, se presentará de manera general el desarrollo de una función por series de Maclaurin que es un caso particular de las series de Taylor. Se darán algunos ejemplos que son usados en diversos temas de probabilidad.

Definición C.3. Si una función $f(x)$ tiene derivadas de todos los órdenes en un intervalo I que contiene a c, entonces, la *serie de Taylor* para $f(x)$ en c se define como

$$\begin{aligned}
f(x) &= \sum_{i=0}^{\infty} \frac{f^{(i)}(c)}{i!}(x-c)^i \\
&= f(c) + f^{(1)}(c)(x-c) + \cdots + \frac{f^{(n)}(c)}{n!}(x-c)^n + \cdots,
\end{aligned}$$

donde $f^{(n)}(c)$ es la n-ésima derivada valuada en c, en el caso particular de $f^{(0)}(c)$, es la función $f(x)$ valuada en c.

Definición C.4. La serie de Taylor de $f(x)$ en $c = 0$ se conoce como la *serie de Maclaurin* de $f(x)$, esto es,

$$f(x) = f(0) + f^{(1)}(0)x + \frac{f^{(2)}(0)}{2!}x^2 + \cdots + \frac{f^{(n)}(0)}{n!}x^n + \cdots \tag{C.11}$$

Definición C.5. Si una función $f(x)$ tiene derivadas hasta de orden n en un intervalo I que contiene a c, entonces, el *n-ésimo polinomio de Taylor* para $f(x)$ en c, se denota como $p_n(x)$ y se define como

$$\begin{aligned} p_n(x) &= \sum_{i=0}^{n} \frac{f^{(i)}(c)}{i!}(x-c)^i \\ &= f(c) + f^{(1)}(c)(x-c) + \frac{f^{(2)}(c)}{2!}(x-c)^2 + \cdots + \frac{f^{(n)}(c)}{n!}(x-c)^n. \end{aligned}$$

El *n-ésimo polinomio de Maclaurin* para $f(x)$, es el n-ésimo polinomio de Taylor en $c = 0$.

Definición C.6. Sea $f(x)$ una función y $p_n(x)$ el n-ésimo polinomio de Taylor para $f(x)$ en c, el *residuo* después de $n+1$ sumandos, se denota como $R_n(x)$ y se define como

$$R_n(x) = f(x) - p_n(x).$$

Entonces,

$$f(x) = p_n(x) + R_n(x).$$

Esto es,

$$f(x) = f(c) + f^{(1)}(c)(x-c) + \cdots + \frac{f^{(n)}(c)}{n!}(x-c)^n + R_n(x).$$

Teorema C.1. *Sea $f(x)$ una función continua en el intervalo $I = [a, b]$ y con derivadas continuas hasta de orden n también en el intervalo I. Supongamos que la derivada $n+1$ existe para toda x en el intervalo (a, b). Entonces, para todo x y c en el intervalo (a, b) se cumple*

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-c)^{n+1},$$

donde ξ es un número que se encuentra entre c y x.

Lo que significa que $f(x)$ se puede expresar de la siguiente manera:

$$\begin{aligned} &f(x) = \\ &f(c) + f^{(1)}(c)(x-c) + \cdots + \frac{f^{(n)}(c)}{n!}(x-c)^n + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-c)^{n+1}, \end{aligned}$$

donde ξ se encuentra entre c y x.

Ahora, considerando $c = 0$, la función $f(x)$ queda expresada de la siguiente manera (suma del n-ésimo polinomio de Maclaurin con su respectivo residuo):

$$f(x) = f(0) + f^{(1)}(0)x + \cdots + \frac{f^{(n)}(0)}{n!}x^n + \frac{f^{(n+1)}(\xi)}{(n+1)!}x^{n+1}, \qquad \text{(C.12)}$$

donde ξ está entre 0 y x.

El siguiente teorema indica cuando una serie de Taylor de una función converge a la función.

Teorema C.2. *Sea $f(x)$ una función que tiene las derivadas de todos los órdenes en un intervalo I que contiene a c, entonces, la serie de Taylor para $f(x)$ en c converge a $f(x)$ para todo x en el intervalo I si y solo si $\lim\limits_{n\to\infty} R_n(x) = 0$, para toda x en el intervalo I.*

Ejemplo C.7. Para la función $f(x) = e^x$.
a) Desarrollar $f(x)$ en serie de Maclaurin.
b) Desarrollar $f(x)$ en polinomio de Maclaurin de grado n más su residuo.

a) Obsérvese

$$f^{(n)}(x) = e^x, \quad \text{para} \quad n = 0, 1, 2, \ldots$$

Por lo tanto,

$$f^{(n)}(0) = e^0 = 1, \quad \text{para} \quad n = 0, 1, 2, \ldots$$

Entonces,

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots = \sum_{i=0}^{\infty} \frac{x^i}{i!}. \qquad \text{(C.13)}$$

b) El polinomio de Maclaurin de grado n es

$$p_n(x) = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} = \sum_{i=0}^{n} \frac{x^i}{i!}.$$

Y el residuo después de $n+1$ sumandos es

$$R_n(x) = \frac{e^{\xi}}{(n+1)!}x^{n+1}, \text{donde}\, 0 < \xi < x.$$

Por lo tanto,

$$e^x = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \frac{x^{n+1}}{(n+1)!}e^{\xi}, \text{ donde } 0 < \xi < x. \qquad \text{(C.14)}$$

Apéndice D

Conceptos básicos de medida

El propósito de este apéndice es proporcionar al lector algunos conceptos básicos de teoría de la medida, los cuales son importantes para la teoría de probabilidad.

Es importante aclarar que los resultados que se presenten en este apéndice no se van a demostrar. Si el lector quiere profundizar se le recomienda literatura sobre el tema, por ejemplo, Capinski y Kopp [5], Galaz F. [9].

D.1. Medida de Lebesgue

A continuación, se presentará la definición de medida exterior de Lebesgue, esta es sumamente importante, ya que la definición de medida de Lebesgue depende de esta.

Definición D.1. Sea A un subconjunto de $\mathbb{R}$. La *medida exterior de Lebesgue* de A se denota como $m^*(A)$ y se define como

$$m^*(A) = inf \left\{ \sum_{i=1}^{\infty} l(I_i), \text{ donde } A \subset \bigcup_{i=1}^{\infty} I_i, \ I_i \text{ son intervalos abiertos} \right\},$$

donde $l(I_i)$ es la longitud del intervalo I_i y el ínfimo se toma sobre todas las colecciones numerables de intervalos abiertos tal que $A \subset \bigcup_{i=1}^{\infty} I_i$ (que cubren a A).

Observación. La definición de medida exterior de Lebesgue se puede hacer indistintamente con colecciones numerables de intervalos abiertos, cerrados o semiabiertos.

A continuación, se presentarán algunas propiedades de la medida exterior de Lebesgue.

Teorema D.1. *La medida exterior de Lebesgue m^* cumple con las siguientes propiedades:*

I. $m^*(\emptyset) = 0$.

II. *Si $A \subset B$, entonces, $m^*(A) \leq m^*(B)$ (monotonía).*
III. *Si $A = I$ es un intervalo, entonces, $m^*(A) = l(I)$, donde $l(I)$ es la longitud de I.*
IV. *Si $A_1, A_2, \ldots$ es una colección numerable de subconjuntos de $\mathbb{R}$, entonces, $m^*\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} m^*(A_i)$ (subaditividad).*
V. *Si $A \subset \mathbb{R}$ es a lo más numerable, entonces, $m^*(A) = 0$. En particular, $m^*(\mathbb{N}) = 0$ y $m^*(\mathbb{Q}) = 0$.*
VI. *Para $A \subset \mathbb{R}$ y $x \in \mathbb{R}$, $m^*(A+x) = m^*(A)$ (propiedad de invarianza por traslaciones).*

Si un conjunto tiene medida exterior de Lebesgue cero (longitud cero), se dice que es un *conjunto nulo*.

Observación. Una propiedad básica que es deseable que una medida cumpla, es la propiedad σ-aditiva, véase la igualdad D.1 de la definición D.10 de este apéndice. La medida exterior de Lebesgue no siempre cumple con esta igualdad, existen subconjuntos de $\mathbb{R}$ que no cumplen esta propiedad, véase apéndice el de Capinski y Kopp [5].

Debido a la situación anterior, solo se van a considerar a los subconjuntos de $\mathbb{R}$, que si cumplen con la propiedad σ-aditiva (igualdad D.1), estos conjuntos son llamados conjuntos *medibles-Lebesgue* o simplemente *medibles*. Una definición típica de un conjunto medible, encontrada en la literatura, debida a Caratheodory, se da a continuación:

Definición D.2. (Caratheodory) un conjunto A se dice *medible-Lebesgue* (o simplemente *medible*) si y solo si para todo $E \subset \mathbb{R}$ se cumple la siguiente igualdad:

$$m^*(E) = m^*(E \cap A) + m^*(E \cap A^c).$$

La familia de subconjuntos de $\mathbb{R}$ que son medibles será denotada como $\mathcal{M}$. El siguiente teorema nos indica las principales propiedades que cumple esta familia.

Teorema D.2. *Para la familia de subconjuntos $\mathcal{M}$, se cumplen las siguientes propiedades:*
I. *Si A es un conjunto nulo, entonces, $A \in \mathcal{M}$.*
II. *Si I es un intervalo, entonces, $I \in \mathcal{M}$.*
III. *$\mathbb{R} \in \mathcal{M}$.*
IV. *Si $A \in \mathcal{M}$, entonces, $A^c \in \mathcal{M}$.*
V. *Si $A_i \in \mathcal{M}$, para toda $i = 1, 2, \ldots$, entonces, $\cup_{i=1}^{\infty} A_i \in \mathcal{M}$.*
VI. *Si $A_i \in \mathcal{M}$, para toda $i = 1, 2, \ldots$, entonces, $\cap_{i=1}^{\infty} A_i \in \mathcal{M}$.*
VII. *Si A es un conjunto abierto o cerrado de $\mathbb{R}$, entonces, $A \in \mathcal{M}$.*

Obsérvese que las condiciones III, IV y V del teorema anterior nos indican que la familia de subconjuntos $\mathcal{M}$ es un σ-álgebra, lo que significa que los conjuntos que pertenecen a la familia $\mathcal{M}$ son cerrados bajo las operaciones de complemento y uniones numerables. También, los conjuntos de esta familia $\mathcal{M}$ es cerrada bajo operaciones de intersecciones numerables.

Definición D.3. La medida exterior de Lebesgue restringida a la colección de conjuntos $\mathcal{M}$ es denotada como m y es llamada *medida de Lebesgue*. Esto es, $m : \mathcal{M} \to [0, \infty]$.

Evidentemente, si A es un conjunto medible, entonces se cumple $m(A) = m^*(A)$.

Los siguientes teoremas tratan con las principales propiedades de la medida de Lebesgue.

Teorema D.3. *La medida de Lebesgue m cumple con las siguientes propiedades:*

I. *Si A es un conjunto nulo, entonces, $m(A) = 0$.*
II. *Si $B \subset A$ y $m(A) = 0$, entonces, $m(B) = 0$.*
III. *Si I es un intervalo, entonces, $m(I) = l(I)$, donde $l(I)$ es la longitud del intervalo I.*
IV. *Si A, $B \in \mathcal{M}$ y $A \subset B$, entonces, $m(A) \leq m(B)$.*
V. *Si $A_i \in \mathcal{M}$, para toda $i = 1, 2, \ldots$, entonces, $m\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} m(A_i)$ (subaditividad).*
VI. *Si $A_i \in \mathcal{M}$, para toda $i = 1, 2, \ldots$, y son excluyentes por pares, entonces, $m\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} m(A_i)$ (σ-aditiva).*
VII. *Si $A \subset \mathbb{R}$ es a lo más numerable, entonces, $m(A) = 0$.*
VIII. *Para $A \subset \mathbb{R}$ y $x \in \mathbb{R}$, $m(A + x) = m(A)$ (propiedad de invarianza por traslaciones).*

Teorema D.4. *Supongamos $A_i \in \mathcal{M}$, para toda $i = 1, 2, \ldots$, entonces,*

I. *Si $A_n \subset A_{n+1}$ para toda n, entonces,*

$$m\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{n\to\infty} m(A_n).$$

II. *Si $A_n \supset A_{n+1}$ para toda n, entonces,*

$$m\left(\bigcap_{i=1}^{\infty} A_i\right) = \lim_{n\to\infty} m(A_n).$$

D.2. Conjuntos de Borel

A continuación, se tratará con la familia de conjuntos de Borel, esta colección también es de gran utilidad en temas de probabilidad.

Si el lector quiere profundizar más en los conjuntos de Borel, se recomienda que consulte bibliografía sugerida al inicio de este apéndice.

Teorema D.5. *La intersección de dos o más σ-álgebras es un σ-álgebra.*

Definición D.4. Sea $\mathcal{A}$ una familia de conjuntos, decimos que $\mathcal{C}$ es el *σ-álgebra generada* por $\mathcal{A}$, si

$$\mathcal{C} = \bigcap \{\mathcal{F} : \mathcal{F} \text{ es un } \sigma\text{-álgebra tal que } \mathcal{F} \supset \mathcal{A}\}.$$

Al intersectar σ-álgebras, tendremos como resultado una familia que también es un σ-álgebra, y es importante aclarar que esta familia está contenida en todas los σ-álgebras que contienen a $\mathcal{A}$. Obviamente $\mathcal{C}$ es la *mínima σ-álgebra* que contiene a la familia $\mathcal{A}$.

Definición D.5. Sea

$$\mathcal{B} = \bigcap \{\mathcal{F} : \mathcal{F} \text{ es un } \sigma\text{-álgebra que contiene a todos los intervalos de } \mathbb{R}\}.$$

$\mathcal{B}$ es el σ-álgebra generada por todos los intervalos de $\mathbb{R}$, y es llamado σ-álgebra de los *conjuntos de Borel*.

Claramente $\mathcal{B}$ es la mínima σ-álgebra que contiene a todos los intervalos de $\mathbb{R}$.

Observación. Los siguientes conjuntos pertenecen a la familia de conjuntos de Borel $\mathcal{B}$:

I. Cualquier conjunto abierto pertenece también a $\mathcal{B}$, ya que un conjunto abierto es la unión numerable de intervalos abiertos.
II. Cualquier conjunto numerable pertenece a $\mathcal{B}$. Este conjunto se puede expresar como la unión numerable de intervalos de la forma $[a, a]$. En particular los conjuntos de números naturales $\mathbb{N}$ y números racionales $\mathbb{Q}$ pertenecen a $\mathcal{B}$.
III. Los números irracionales también pertenecen a $\mathcal{B}$, ya que este conjunto se complemento de los números racionales.
IV. Los conjuntos finitos también pertenecen a $\mathcal{B}$ y sus respectivos conjuntos complementos.

Teorema D.6. *La familia de conjuntos de Borel $\mathcal{B}$ también puede generarse como la mínima σ-álgebra de*

I. *Todos los intervalos abiertos.*
II. *Todos los intervalos cerrados.*
III. *Todos los intervalos de la forma $(-\infty, a)$ (o de la forma $(-\infty, a]$).*
IV. *Todos los intervalos de la forma (a, ∞) (o de la forma $[a, \infty)$).*
V. *Todos los intervalos de la forma $(a, b]$ (o de la forma $[a, b)$.*
VI. *Todos los conjuntos abiertos.*
VII. *Todos los conjuntos cerrados.*

Los conjuntos de Borel tienen una aplicación importante en la probabilidad, cuando el espacio muestral (véase capítulo 1) es finito o infinito numerable, se puede considerar como espacio de eventos, el conjunto potencia. Pero cuando el espacio muestral es el conjunto de los números reales o un subconjunto de los reales como un intervalo, el espacio de eventos no puede considerarse como el conjunto potencia, en este caso, el σ-álgebra de los conjuntos de Borel $\mathcal{B}$ es una buena opción.

Por otro lado, por el teorema D.2, los intervalos pertenecen a la familia $\mathcal{M}$, pero la mínima σ-álgebra que contiene a los intervalos de $\mathbb{R}$, es la familia $\mathcal{B}$, de esta manera podemos afirmar que $\mathcal{B} \subset \mathcal{M}$. Dicho de otra manera, cada conjunto de Borel es medible.

La inclusión anterior es propia, hay conjuntos medibles que no son conjuntos de Borel. Para la construcción de un conjunto que está en $\mathcal{M}$ y no está en $\mathcal{B}$, véase apéndice de Capinski y Kopp [5].

Sin embargo, si $A \in \mathcal{M}$, es posible encontrar un conjunto de Borel B tal que $A \subset B$ (B cubre a A) y donde el conjunto B se puede expresar como $B = \cap_{i=1}^{\infty} O_i$, donde $\{O_n, n = 1, 2, \ldots\}$ es una secuencia de subconjuntos abiertos de $\mathbb{R}$, y además, $m(A) = m(B)$. Dicho de otra manera, $m(B \Delta A) = m(B \cap A^c) = 0$, véase capítulo 2 de Capinski y Kopp [5]. Esto significa que la medida de Lebesgue m no distingue entre la medida del conjunto A y el conjunto de borel B construido por la intersección numerable de conjuntos abiertos O_i.

Entonces, dado cualquier conjunto $A \in \mathcal{M}$, existe un conjunto $B \in \mathcal{B}$ tal que $A \subset B$ y la diferencia simétrica entre A y B es un conjunto nulo.

Observación. Los conjuntos que están en $\mathcal{M}$ y no están en $\mathcal{B}$, son conjuntos nulos, para cada uno de estos conjuntos, existe un conjunto de Borel tal que lo cubre y su medida es cero.

Definición D.6. La medida de Lebesgue restringida a los conjuntos del σ-álgebra de los conjuntos de Borel es llamada *medida de Borel* .

Observación. Las definiciones anteriores se pueden generalizar, en lugar del conjunto de los números reales $\mathbb{R}$, se puede considerar $\mathbb{R}^n$, en este caso, en lugar de intervalos se pueden considerar productos cartesianos de rectángulos en $\mathbb{R}^n$.

D.3. Premedidas y medidas

A continuación, se presenta en términos generales las definiciones de premedida y de medida. En lugar de considerar los números reales, se considera un conjunto en general Ω y un álgebra $\mathcal{A}$ de conjuntos de Ω.

Definición D.7. Sea Ω un conjunto y $\mathcal{A}$ un álgebra de subconjuntos de Ω. Una función denotada como μ con dominio $\mathcal{A}$ y contradominio el intervalo $[0, \infty]$ se llama *premedida* en $\mathcal{A}$, si para A y B que pertenecen a $\mathcal{A}$, los cuales son mutuamente excluyentes, se cumple

$$\mu(A \cup B) = \mu(A) + \mu(B).$$

Definición D.8. Sea Ω un conjunto, $\mathcal{A}$ un álgebra de subconjuntos de Ω y μ una premedida en $\mathcal{A}$, si se cumple $\mu(\Omega) < \infty$, se dice que μ es una *premedida finita.*

Definición D.9. Sea Ω un conjunto, $\mathcal{A}$ un álgebra de subconjuntos de Ω y μ una premedida en $\mathcal{A}$, si se cumple $\mu(\Omega) = 1$, se dice que μ es una *premedida de probabilidad.*

Definición D.10. Sea Ω un conjunto, $\mathcal{A}$ un álgebra de subconjuntos de Ω y μ una premedida en $\mathcal{A}$, si para una colección numerable de subconjuntos mutuamente excluyentes $A_1, A_2, \dots \in \mathcal{A}$ se cumple $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$, y además la igualdad

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i), \tag{D.1}$$

entonces, se dice que μ es una *medida en* $\mathcal{A}$.

La igualdad D.1 es conocida como la propiedad σ-*aditiva*, y es una propiedad fundamental que cumple toda medida.

Observación. En la definición anterior, si se considera que $\mathcal{A}$ es un σ-álgebra, en lugar de álgebra, entonces, no es necesario pedir que se cumpla la condición $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$.

Definición D.11. Sea Ω un conjunto, $\mathcal{A}$ un álgebra de subconjuntos de Ω y μ una medida en $\mathcal{A}$, si se cumple $\mu(\Omega) < \infty$, se dice que μ es una *medida finita.*

Definición D.12. Sea Ω y conjunto y $\mathcal{A}$ un σ-álgebra de subconjuntos de Ω, entonces, el par $(\Omega, \mathcal{A})$ se le conoce como *espacio medible*, y a la terna $(\Omega, \mathcal{A}, \mu)$ se le conoce como *espacio de medida*, donde μ es una medida.

En particular la terna $(\mathbb{R}, \mathcal{M}, m)$, se le conoce como *espacio de medida de Lebesgue.*

D.4. Medidas de probabilidad

A continuación, se darán las definiciones de medida de probabilidad y de espacio de probabilidad.

Definición D.13. Sea Ω un conjunto, $\mathcal{A}$ un σ-álgebra de subconjuntos de Ω y P una medida en $\mathcal{A}$, si se cumple la igualdad $P(\Omega) = 1$, se dice que P es una *medida de probabilidad*, y a la terna $(\Omega, \mathcal{A}, P)$ se le conoce como *espacio de probabilidad.*

A continuación, se presentarán algunos ejemplos de medida, en particular se darán ejemplos de medida de probabilidad y un caso particular de medida de Lebesgue.

Ejemplo D.1. Sean $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$, $\mathcal{A}$ el conjunto potencia de Ω y $A \in \mathcal{A}$. Definimos P de la siguiente manera:

$$P(A) = \frac{\#A}{\#\Omega},$$

donde $\#A$ y $\#\Omega$ son el número de elementos de A y de Ω, respectivamente. Indicar el espacio de probabilidad.

Obsérvese que para cualquier $A \in \mathcal{A}$, se cumple $0 \leq P(A) \leq 1$. También, se puede verificar que P cumple la propiedad σ-aditiva, como también $P(\Omega) = 1$. De esta manera, se puede concluir que $(\Omega, \mathcal{A}, P)$ es un espacio de probabilidad.

El espacio de probabilidad anterior, se aplica al cálculo de probabilidades para un experimento, donde el número de resultados posibles es finito y equiprobable, siendo esta la definición clásica de probabilidad, la cual se presenta en el capítulo 1.

Ejemplo D.2. Sea $\Omega = \{0, 1, \ldots, n\}$ y $\mathcal{A}$ el conjunto potencia de Ω. Definimos la función f, con dominio Ω y contradominio el intervalo $[0, 1]$, de la siguiente manera: $f(i) = \binom{n}{i}(0.5)^n$, donde $i \in \Omega$. Para $A \in \mathcal{A}$, se define $P(A)$ como:

$$P(A) = \sum_{i \in A} f(i).$$

Indicar el espacio de probabilidad.

Obsérvese que para cualquier conjunto A que pertence a $\mathcal{A}$, se cumple $P(A) \geq 0$, también se puede comprobar, aplicando el teorema del binomio (véase igualdad C.7 del apéndice C), que se cumple $P(\Omega) = 1$. Por otro lado, se puede observar que la función P cumple con la propiedad σ-aditiva. De esta manera, podemos concluir que la terna $(\Omega, \mathcal{A}, P)$ es un espacio de probabilidad.

En el anterior ejemplo, la función f corresponde a un caso particular de la distribución binomial, véase definición 3.5 del capítulo 3. Obsérvese, que se puede generalizar la función f, en lugar de considerar esta distribución binomial, se puede definir f como cualquier otra función de densidad discreta con recorrido finito (véase algunas opciones en el capítulo 3). Y en el caso de que el recorrido sea infinito numerable, el σ-álgebra $\mathcal{A}$ puede seguir siendo el conjunto potencia, y la función f puede ser una función de densidad discreta con recorrido infinito numerable, en el capítulo 3, también se exponen varias opciones.

Los siguientes ejemplos, consideran, principalmente, ejemplos de distribuciones de variables aleatorias continuas.

Ejemplo D.3. Comprobar que la medida de Lebesgue, restringida al intervalo $[0, 1]$ es una medida de probabilidad. Indicar cuál es el espacio de probabilidad.

En este caso, $\Omega = [0,1]$. Se sabe que $m(\Omega) = m([0,1]) = l([0,1]) = 1$. Entonces, m es una medida de probabilidad, denotemos $m = P$.

El σ-álgebra es la familia de conjuntos $\mathcal{M}$ restringida al intervalo $[0,1]$, denotemos a esta familia como $\mathcal{M}_0$. Entonces, el espacio de probabilidad es $([0,1], \mathcal{M}_0, P)$.

En la práctica, este último ejercicio corresponde a seleccionar un número al azar del intervalo $[0,1]$, y la medida de probabilidad corresponde al cálculo de probabilidades de una distribución uniforme continua en el intervalo $[0,1]$, véase capítulo 4, por ejemplo, sean a y b tal que $0 < a < b < 1$, entonces, $P[a,b] = l[a,b] = b - a$. Por medio, de la distribución uniforme,

$$P[a,b] = \int_a^b 1dx = b - a.$$

Se puede generalizar el ejemplo anterior, consideremos el intervalo $[c,d]$ y sean a y b tales que $c < a < b < d$. Definamos la función P de la siguiente manera:

$$P([a,b]) = m[a,b]\left(\frac{1}{m[c,d]}\right) = \frac{b-a}{d-c}.$$

Nótese que se puede comprobar con facilidad que P es una medida de probabilidad, la cual corresponde a la distribución uniforme continua sobre el intervalo $[c,d]$.

Esta última probabilidad se puede calcular, usando la función de densidad $f_X(x)$ o la función de distribución $F_X(x)$ de la distribución uniforme continua sobre el intervalo $[c,d]$ (véase sección 4.1), esto es,

$$P[a,b] = \int_a^b f_X(x)dx = \int_a^b \frac{1}{d-c}dx = \frac{b-a}{d-c} = F_X(b) - F_X(a).$$

Por otro lado, en términos generales, se puede comprobar que la función de distribución evaluada en un valor x (véase definición en el capítulo 2), $F_X(x) = m_0((\infty, x])$, es una medida de probabilidad, la cual está bien definida para cualquier conjunto de Borel $B \in \mathcal{B}$. En particular, si $a, b \in \mathbb{R}$ y además, $a < b$, $m_0((a,b]) = F_X(b) - F_X(a)$. En este caso, en lugar de proporcionar la longitud del intervalo (medida de Lebesgue), la medida proporciona la probabilidad de que la variable aleatoria tome un valor en el intervalo $(a,b]$. Ahora, como se vio en el ejemplo D.3, si la función de distribución es la de una uniforme sobre el intervalo $[0,1]$, esto es, $F_X(x) = xI_{[0,1]}(x) + I_{(1,\infty)}$, la medida de probabilidad coincide con la medida de Lebesgue en el intervalo $[0,1]$.

D.5. Funciones medibles

En esta sección se presentará la definición de función medible y algunas de sus propiedades más importantes. Este concepto es de mucha trascendencia para la teoría de probabilidad, de hecho, la definición d e f unción medible coincide con la definición d e v ariable a leatoria, v éase e n e l c apítulo 2 , la definición 2.1. En seguida, se mostrará la definición de función medible.

Definición D.14. Dados dos espacios medibles $(\Omega_1, \mathcal{A})$ y $(\Omega_2, \mathcal{B})$, decimos que $f\colon \Omega_2 \to \Omega_2$ es una función medible, si para todo $B \in \mathcal{B}$ tenemos que $f^{-1}(B) \in \mathcal{A}$.

En términos coloquiales, f es medible, si la preimagen de conjuntos medibles, también es medible.

Podemos observar que un caso relevante es considerar una función con los espacios medibles $(\mathbb{R}, \mathcal{M})$ y $(\mathbb{R}, \mathcal{B})$, con $\mathcal{M}$ los conjuntos Lebesgue medibles y $\mathcal{B}$ el álgebra de Borel. En este caso, una función medible $f\colon \mathbb{R} \to \mathbb{R}$ se llama *función Lebesgue medible.* Recordando que el álgebra de Borel es el álgebra generada por los intervalos, tenemos a continuación, la siguiente definición equivalente:

Definición D .15. Sea $\Omega \subset \mathbb{R}$ un conjunto medible y $f\colon \Omega \to \mathbb{R}$ una función, decimos que f es una *función Lebesgue medible*, si para cada $r \in \mathbb{R}$, el conjunto $A_r = \{x \in \Omega | f(x) \leq r\} = f^{-1}((-\infty, r]) \in \mathcal{M}$, esto es, A_r es un conjunto medible.

Teorema D.7. *Las siguientes afirmaciones s on equivalentes:*

I. *f es Lebesgue medible.*
II. *Para todo r, $f^{-1}((-\infty, r])$ es Lebesgue medible.*
III. *Para todo r, $f^{-1}((-\infty, r))$ es Lebesgue medible.*
IV. *Para todo r, $f^{-1}([r, \infty))$ es Lebesgue medible.*
V. *Para todo r, $f^{-1}([r, \infty))$ es Lebesgue medible.*

A continuación, se darán las principales propiedades básicas de funciones Lebesgue medibles.

Teorema D.8. *Sean f y g funciones reales definidas en $\Omega \subset \mathbb{R}$ medibles, tales que son funciones Lebesgue medibles, entonces,*

I. *$f + g$ es Lebesgue medible.*
II. *fg es Lebesgue medible.*
III. *cf es Lebesgue medible, donde c es una constante.*
IV. *$|f|$ es Lebesgue medible.*

Otro caso sumamente importante está dado por considerar un espacio de probabilidad $(\Omega, \mathcal{A}, \mathcal{P})$ y $(\mathbb{R}, \mathcal{B})$, con $\mathcal{B}$ el álgebra de Borel. Entonces, una función $X\colon \Omega \to \mathbb{R}$ es medible si y solo si es una variable aleatoria, véase definición 2.1.

Por último, cabe mencionar que en probabilidad la integración de funciones es de suma importancia, es necesaria en varios tópicos de probabilidad, se utiliza para calcular probabilidades, valores esperados, funciones generadoras de momentos de una variable aleatoria, entre otros, véase capítulo 2. La integral de Riemann no es suficiente para algunas funciones de densidad u otro tipo de funciones usadas en probabilidad. En cambio, la integral de Lebesgue sí cumple los requerimientos que se necesitan para las diferentes funciones usadas en probabilidad. Para la integral de Lebesgue, es muy importante considerar que las funciones sean medibles sobre conjuntos medibles, véase capítulo 4 del libro de Capinski y Kopp [5]. En conclusión, la teoría de la medida es fundamental para la integración de funciones y por lo tanto para la probabilidad.

Una recomendación para el lector, si le interesa estudiar diferentes temas de probabilidad a un nivel más avanzado, es estudiar con más profundidad diferentes temas de la teoría de la medida. En este apéndice se presentaron algunos de los temas de más importancia.

Apéndice E

Actuaría: seguros y finanzas

Es muy amplia la diversidad de aplicaciones que existen de la probabilidad en las diferentes áreas de la actuaría. En esta sección se presentarán, primero algunos conceptos básicos sobre dos de las principales áreas de la actuaría: seguros y finanzas. En segunda instancia, se tratarán algunos ejemplos de seguros y finanzas en forma muy general, con la idea de que el lector se empiece a familiarizar con las aplicaciones de la probabilidad en estas áreas.

Para profundizar más en las aplicaciones de la probabilidad en estas áreas de la actuaría, se le recomienda al lector consultar literatura especializada, por ejemplo, Yiu-Kuen Tse [29].

E.1. Seguros

Supongamos que una persona tiene el riesgo de tener una pérdida financiera, debido a la ocurrencia de un siniestro. Hoy en día, las personas pueden reducir este impacto financiero, contratando una protección, la cual es llamada *seguro*. La persona que adquiere el seguro para protegerse es llamada *asegurado*, y la parte que otorga la protección es llamada *asegurador* o *compañía de seguros*. El contrato entre asegurado y asegurador, donde se especifican las características y condiciones de la protección es llamado *póliza de seguro* o simplemente *póliza*, En este contrato se especifica el costo para el asegurado por esta protección, el cual se llama *prima*, además, también se especifican las condiciones de pago, en caso de ocurrir un siniestro.

Generalmente, se puede acordar que la pérdida por un siniestro sea pagada en forma compartida por ambas partes: por el asegurador y por el asegurado. Esto es, se estipula que un porcentaje de la pérdida sea pagada por el asegurado, esta cantidad es llamada *deducible*. Una modalidad muy frecuente es que si la cantidad reclamada es menor o igual al deducible, el asegurador no paga cantidad alguna al asegurado, si la cantidad reclamada es mayor que el deducible, entonces, la cantidad pagada por el asegurador es la cantidad reclamada menos el deducible acordado.

También, se puede acordar una cantidad máxima de pago en caso de una reclamación, esta cantidad se llama *suma asegurada* y cuando un reclamo sea superior a esta cantidad, la compañía solo pagará la cantidad máxima estipulada en la póliza (menos el deducible, si este se estipuló en el contrato).

Ahora, consideremos las siguientes variables aleatorias (véase capítulo 2): $X =$ Pérdida o cantidad reclamada por el asegurado y $Y =$ Cantidad pagada por la compañía al asegurado. Si no se acuerda un deducible y tampoco una suma asegurada, ambas variables son iguales. Pero en caso de que se estipule un deducible y/o suma asegurada, ambas variables son diferentes y ambas tendrán que ser consideradas.

Es importante mencionar que existen varios posibles modelos, tanto discretos como continuos, para variables aleatorias como X y Y, véanse los capítulos 2, 3 y 4. Un modelo más adecuado para variables como Y es una mezcla de dos distribuciones, una discreta y una continua. También se puede considerar, dependiendo de las condiciones de pago, la mezcla de dos distribuciones discretas, véase capítulo 2, sección 2.14.

En ocasiones, se pueden presentar reclamos extremos, debido a algún evento catastrófico, como un huracán, terremoto, entre otros. Por lo mismo, se tendrá que considerar para estos casos, alguna distribución con cola pesada, véase capítulo 2, sección 2.15, también Reiss y Thomas [24].

Finalmente, consideremos un conjunto de pólizas de seguro de no vida con cobertura en un periodo de tiempo fijo. Para el análisis de este grupo de pólizas, será de interés estudiar las siguientes variables: el número de reclamos, conocido como *frecuencia de reclamos*; la cantidad monetaria reclamada por cada póliza, conocida como *severidad o tamaño de reclamo*, y la suma de las pérdidas monetarias por todos los reclamos, conocida como *reclamaciones acumuladas*, véase Yiu-Kuen Tse [29].

Cabe señalar que las distribuciones de variables aleatorias discretas con valores enteros no negativos son posibles modelos para la frecuencia de reclamos, en particular, las distribuciones binomial, geométrica, binomial negativa y Poisson, véase capítulo 3 y capítulo 1 de Yiu-Kuen Tse [29].

Por otro lado, la severidad de reclamo puede ser modelada por una distribución continua con recorrido, los números reales no negativos, en particular, las distribuciones exponencial, gamma, de Weibull y de Pareto pueden ser una opción, véase el capítulo 4 y el capítulo 2 de Yiu-Kuen Tse [29]. Si en la póliza se considera un deducible y/o suma asegurada limitada, una mejor forma de modelar el tamaño de un reclamo es por medio de una mezcla de dos distribuciones, una discreta y una continua, véase sección 2.14.

En el capítulo 3 de Yiu-Kuen Tse [29], se tratan diferentes formas de modelar la variable: reclamaciones acumuladas. Por ejemplo, suponiendo que el tamaño de reclamos es fijo, la distribución de esta variable se logra encontrando la distribución de una suma de variables aleatorias independientes e idénticamente distribuidas, véase capítulos 3 y 6 de Yiu-Kuen Tse [29].

E.2. Finanzas

En el tema de finanzas, es importante dejar claro varios conceptos, ya que estos son parte del lenguaje financiero y nos ayudarán a comprender de mejor manera algunos ejemplos y ejercicios que se presentarán en algunos temas de aplicaciones de la probabilidad a las finanzas.

Un *préstamo* es una operación financiera, por el cual una persona (empresa, banco, etc.), llamada *prestamista* o *comprador*, otorga una cantidad de dinero (bien o servicio) a otra persona (empresa, entidad de gobierno, persona física), llamada *prestatario*, *vendedor* o *emisor*, a cambio de un rédito, interés o beneficio (costo financiero del préstamo). Por lo general, esta operación se realiza por medio de un contrato, llamado *instrumento financiero*, dando lugar un derecho al comprador llamado *activo financiero* y una obligación al vendedor, llamado *pasivo financiero*.

En otras palabras, un *activo* es algo que tiene valor o puede generar ingresos. Existen los *activos tangibles*, como un auto, una máquina o un inmueble, mientras que los activos financieros son documentos o títulos que tienen un valor, y además, pueden producir rendimientos. El activo financiero es emitido por una empresa, banco o institución pública (emisor) y donde el comprador espera un rendimiento. De esta manera, el comprador adquiere un derecho de cobrar alguna suma, el emisor adquiere una obligación de pagar al comprador un retorno por la inversión adquirida.

Cabe mencionar que los préstamos pueden ser pagados en una sola exhibición con su interés correspondiente o mediante pagos periódicos, amortizando la deuda; cada uno de estos pagos, con su interés correspondiente.

Los instrumentos financieros (activos financieros) tienen los siguientes elementos que los caracterizan:

-*Liquidez*. La capacidad del instrumento en convertirse lo más pronto posible en dinero y sin pérdidas.

-*Riesgo*. De que el prestatario cumpla con las obligaciones en el plazo establecido.

-*Rentabilidad*. La producción de rendimientos del instrumento.

-*Vencimiento*. Tiempo estipulado de la duración del contrato financiero.

Cabe mencionar que, por lo general, entre más rendimientos tenga un activo financiero, mayor riesgo tendrá de no cumplir las obligaciones. El comprador deberá hacer un balance, donde logre rendimientos interesantes sin tanto riesgo.

Algunos ejemplos de activos financieros son:

-*Dinero en curso legal*. Son las monedas y billetes, es lo más líquido que existe, ya que se puede cambiar en cualquier momento por bienes.

-*Dinero en los bancos*. Depósitos a la vista, depósitos de ahorro y depósitos a plazo fijo. Por lo general, el cliente puede disponer de su dinero en su totalidad o parcialmente, sin mucho trámite.

-*Letras del Tesoro*. Son títulos de deuda pública a corto plazo y a renta fija, emitidos por el Estado, como un modo de financiamiento.

-*Bonos y obligaciones del Tesoro*. También son títulos de deuda pública a mediano y largo plazo, respectivamente. Se usan para el financiamiento del Estado.

-*Activos o pagarés emitidos por empresas*. Son emitidos por las empresas, donde la renta es fija.

-*Acciones y derivados*. Las acciones son un instrumento de renta variable, donde no hay una retribución segura al accionista. Dependerá de la situación de la empresa (emisora) de la repartición de dividendos. Un derivado es un activo financiero del cual su valor depende de otro activo financiero.

E.3. Retorno de la inversión

Un indicador financiero muy usado que mide la rentabilidad de una inversión en un lapso de tiempo es el *retorno o rendimiento de la inversión*, el cual en términos generales se denota como ROI y se define de la siguiente manera:

$$ROI = \frac{V_f - V_i}{V_i},$$

donde V_i es el valor inicial de la inversión y V_f es el valor final de la inversión, considerando los intereses y dividendos.

Obsérvese que el ROI puede ser positivo, logrando en este caso una ganancia, también puede ser negativo, siendo en este caso una pérdida, o puede ser cero, no logrando ni pérdida ni ganancia.

Supongamos que para los precios de una acción, divisas extranjeras o alguna inversión, se evalúan en los siguientes momentos $t = 0, 1, 2, \ldots$, donde estos pueden ser días, semanas, meses, etc. Sea S_t el precio de la inversión, para $t = 0, 1, 2, \ldots$, siendo S_0 la inversión inicial. Entonces, para cada momento se puede calcular el ROI, esto es,

$$R_t = \frac{S_t - S_{t-1}}{S_{t-1}},\ t = 1, 2, \ldots,$$

donde R_t es el ROI en el momento t.

Obsérvese que para un periodo t, la inversión inicial es la inversión final del anterior periodo.

En varios análisis financieros, es de interés llevar un registro de la serie de tiempo: $R_0, R_1, R_2, \ldots$ Con la idea de encontrar un modelo estadístico para esta serie y poder hacer pronósticos para valores futuros de la serie, véase Guerrero V. [10].

Puede ser de interés para una inversión evaluar las pérdidas, es importante tener en cuenta los modelos de cola pesada para este tipo de variables, por la ocurrencia de algún evento, la pérdida puede tomar un valor extremo, véase capítulo 2, sección 2.15, también Reiss y Thomas [24].

Otro indicador financiero que es usado cuando la capitalización es continua son los *retornos o rendimientos logarítmicos*. Para un determinado t, el retorno logarítmico se denota como RG_t y se define de la siguiente manera:

$$RG_t = ln\left(\frac{S_t}{S_{t-1}}\right) = ln(S_t) - ln(S_{t-1}).$$

Es común ver en la literatura financiera que los retornos R_t, son llamados *retornos aritméticos* y los retornos logarítmicos RG_t son llamados *retornos geométricos*

E.4. Administración de riesgos financieros

Los actuarios históricamente han sido expertos en la administración de los riesgos a las personas con los seguros de vida, los seguros médicos y planes de jubilación. También a los patrimonios de las personas con los seguros de automóviles, los seguros de casa y de algún otro bien. En este tipo de riesgos, se consideran también los bienes de las empresas ante un evento catastrófico, como un terremoto, un incendio, explosión, etc.

Otro campo actuarial, hoy en día, es la *administración de los riesgos financieros*. El riesgo financiero hace referencia a la incertidumbre producida en el rendimiento de una inversión, cartera de créditos o portafolio de productos financieros derivados. Esta incertidumbre es debida a los cambios producidos en el sector en el que se opera.

Los riesgos financieros se clasifican en cuatro: riesgo de mercado, riesgo de contraparte, riesgo de liquidez y riesgo operativo.

El *riesgo de mercado* se origina por las pérdidas potenciales ante movimientos adversos en los precios de mercado. Por ejemplo, consideremos un país que es exportador de petróleo, supongamos una caída en los precios del petróleo en el mercado internacional, entonces, se tendrán menores ingresos por la misma cantidad vendida de petróleo.

El *riesgo de contraparte* es aquel que se origina de un contrato, y ocurre en caso de que algunas de las partes no cumplan con el trato pactado. Por ejemplo, en un contrato de crédito (de hecho, se llama *riesgo de crédito*, pero es un subconjunto del riesgo de contraparte), el prestamista o prestatario no cumple con lo acordado. Otro caso del riesgo de contraparte es un contrato financiero derivado que se negocia entre dos contrapartes fuera del mercado centralizado, por lo que puede no cumplirse al vencimiento (situación que ha ocurrido).

El *riesgo de liquidez* se origina cuando una empresa se queda corta de dinero y tiene compromisos que pagar al corto plazo, por lo que podría enfrentar pérdidas al verse en la necesidad de vender, de manera anticipada, sus activos a menor valor.

El *riesgo operativo* es aquel que impide a una empresa operar en forma normal, ya que es afectada por un acontecimiento, como un ciberataque, un fraude, una huelga, un evento de terrorismo, entre otros.

Entenderemos por una *cartera* o un *portafolio* de inversiones al conjunto de activos financieros que posee un inversionista (persona o institución). La selección adecuada de los elementos que componen un portafolio conlleva a buscar un equilibrio entre los rendimientos y el riesgo que caracterizan a los componentes del portafolio. Las acciones y bonos son los activos más utilizados para formar una cartera de inversiones, pero pueden usarse también el oro, petróleo, bienes raíces, etc.

Para administrar el riesgo de mercado, se utiliza el indicador: *Valor en Riesgo* (*Value at Risk*, VaR), el cual cuantifica la exposición de pérdidas ante una posición de inversión, para lo cual se utilizan técnicas estadísticas. Este método fue desarrollado por el banco de inversión JP Morgan a principios de los años 90, y fue adoptado rápidamente por todas las firmas financieras gracias a su simplicidad de concepto; por lo que este indicador se encuentra muy desarrollado en el mundo financiero de los mercados capitales, siendo una medida de riesgo que hasta la fecha se sigue utilizando.

Siendo más precisos, el VaR es el cálculo de la máxima pérdida esperada con una determinada confianza y en un tiempo determinado. Para su explicación, consideremos la siguiente variable $X =$ Ganancia o pérdida de un portafolio o activo financiero. La distribución de X es llamada *distribución de pérdidas*. Entonces, el VaR a una confianza del $(1-\alpha)\,\%$, es el cuantil (véase definición 2.20) $\zeta_{(1-\alpha)}$ de esta distribución.

Cabe mencionar, que durante muchos años, una de las técnicas para determinar el VaR ha sido asumir que la distribución de pérdidas sigue una distribución normal (véase sección 4.6), sin embargo, esto no es siempre adecuado, sobre todo, ante la posibilidad de observar pérdidas extremas, en este caso, una distribución más adecuada pudiera ser una distribución de cola pesada (véase sección 2.15). Para abundar más en el tema de valor en riesgo, se recomienda consultar el capítulo 4 del libro de Yiu-Kuen Tse [29] o también el capítulo 22 de Hull J. C. [15].

Para la administración del riesgo de contraparte, se considera una cartera de n contratos, donde p es la probabilidad de incumplimiento por contrato. En este caso, es de interés determinar el número de contratos de la cartera con un estatus de incumplimiento. Entonces, es común que la distribución binomial (véase sección 3.3) sea usada para estimar las pérdidas esperadas por el potencial de incumplimiento de la cartera.

El riesgo de liquidez se administra de manera diferente, generalmente con base en criterios prudenciales.

El riesgo operativo se administra de la misma manera en que las compañías de seguros administran sus carteras de pólizas con base en las frecuencias y las severidades de los eventos adversos se construye una distribución de pérdidas.

Apéndice F

Soluciones

Capítulo 1

1. *a*) A^c.
   b) $(A \cup B \cup C)^c = A^c \cap B^c \cap C^c$.
   c) $A \cap (B \cup C)^c = A \cap B^c \cap C^c$.
   d) $A \cup B \cup C$.
   e) $[A \cap (B \cup C)^c] \cup [B \cap (A \cup C)^c] \cup [C \cap (A \cup B)^c]$.
   f) $[A \cap B \cap (A \cap B \cap C)^c] \cup [A \cap C \cap (A \cap B \cap C)^c] \cup [B \cap C \cap (A \cap B \cap C)^c]$.
   g) $(A \cap B \cap C)^c = A^c \cup B^c \cup C^c$.
3. 766.
6. *a*) $\frac{11}{12}$.
   b) 1.
9. 0.5.
10. $\frac{7}{15}$.
11. *a*) 0.3.
    b) 0.5.
12. 0.49856.
13. *a*) 0.06818.
    b) 0.04545.
14. *a*) 0.83849.
    b) 0.32301.
15. 0.33333.
16. 0.66667.
17. *a*) 36 puntos muestrales.
    b) $\frac{5}{18}$.
    c) Si, $\frac{1}{9} > \frac{1}{12}$.
18. *a*) 0.22222.
    b) 0.29091.
19. Es cierto.
20. No es cierto.

21. No es cierto.
22. Es cierto.
23. 0.57463.
24. Es cierto.
25. Es cierto.
26. No es cierto.
29. Es cierto.
30. *a*) 294.
    b) 168.
    c) 118.
31. *a*) 0.57143.
    b) 0.23761.
32. No es cierto.
33. Es cierto.
35. 0.33333.
36. 0.00952.
37. *a*) 0.2445.
    b) 0.66138.
38. 0.33333.
39. *a*) 0.00282.
    b) 0.00056.
    c) 0.09034.
    d) 0.38114.
40. *a*) 0.775.
    b) 0.26625.
41. Sí son independientes.
42. 0.68.
43. 0.39130.
44. 0.2963.
45. 0.16416.
46. 0.49375.
48. *a*) 0.775.
    b) 0.90323.
49. *a*) Es cierto.
    b) En general, no es cierto.
    c) Es cierto.
50. 0.00038.
51. 0.54925.
52. $\frac{29}{30}$.
53. 0.25479.
55. 0.15.
56. $\frac{5}{6}$.
57. *a*) $\frac{1}{3}$.
    b) 0.25.

59. No es cierto.
60. 0.01633.
61. *a*) 0.11836.
    b) 0.04073.
62. 0.29032.
63. 0.0953.
64. $\frac{1}{3}$.
65. 0.11111.
66. 0.35.
67. *a*) 0.00944.
    b) 0.22223.

Capítulo 2

1. $f_X(x) = 0.1I_{\{0\}}(x) + 0.75I_{\{1\}}(x) + 0.15I_{\{2\}}(x)$.
2. $f_X(x) = 0.1I_{\{3\}}(x) + 0.3I_{\{4\}}(x) + 0.6I_{\{5\}}(x)$.
3. $f_X(x) = \dfrac{\binom{5}{x}\binom{6}{7-x}}{\binom{11}{7}} I_{\{1,2,3,4,5\}}(x)$.
4. *a*) $f_X(x) = \dfrac{\binom{x-1}{1}}{\binom{6}{2}} I_{\{2,3,4,5,6\}}(x)$.
   b) $f_X(x) = \frac{1}{15}I_{\{3,4,10,11\}}(x) + \frac{2}{15}I_{\{5,6,8,9\}}(x) + \frac{3}{15}I_{\{7\}}(x)$.
5. $f_X(x) = (0.6)^x(0.4)I_{\{0,1,2,\ldots\}}(x)$.
6. $F_X(x) = 1 - (0.6)^{x+1}$, para $x = 0, 1, 2, \ldots$
7. $f_X(x) = \binom{20}{x}(\frac{1}{4})^x(\frac{3}{4})^{20-x}I_{\{0,1,2,\ldots,20\}}(x)$.
8. $E[X] = 3.18182$; $V[X] = 0.69421$.
9. -4.40925.
11. $P[\mu - 2\sigma < X < \mu + 2\sigma] = 0.93937 \geq 0.75$.
12. *a*) Sí.
    b) No.
    c) No
    d) Sí.
    e) Sí.
    f) Sí.
    g) No.
    h) No.
    i) Sí.
13. 1.
14. a) En general, no es.
    a) Sí es.
15. 1.10668.
16. $E[X] = 0$;
    Mediana= 0;

Modas$= -\sqrt[4]{\frac{3}{2}}, \sqrt[4]{\frac{3}{2}}$.

18. *a*) La función de distribución para valores del recorrido es

x	$F_X(x)$
0	0.53978
1	0.89963
2	0.98959
3	0.99958
4	1

b) $f_X(x) = \binom{4}{x}(\frac{1}{7})^x(\frac{6}{7})^{4-x} I_{\{0,1,2,3,4\}}(x)$.
c) $E[X] = 0.57143$; $V[X] = 0.4898$.
d) 0.
e) 0.

19. $M_X(t) = \frac{1}{8}(t+1)^3$; $E[X] = 1.5$; $V[X] = 0.75$.

20. *a*) $f_X(x) = (\frac{2}{3})^{x-1}(\frac{1}{3}) I_{\{1,2,\ldots,\}}(x)$.
b) $m_X(t) = e^t(3-2e^t)^{-1}$, si $t < ln(\frac{3}{2})$.
c) $M_X(t) = t(3-2t)^{-1}$, si $t < \frac{3}{2}$.
d) $E[X] = 3$; $V[X] = 6$.
e) 1.

21. *a*) 2.
b) $F_X(x) = (x - \frac{x^2}{4}) I_{[0,2]}(x) + I_{(2,\infty)}(x)$.
c) $2 - \sqrt{(2)}$.
d) 0.

22. b) $F_X(x) = \frac{1}{2}(\frac{x}{2} + 1 - \frac{a}{2})^2 I_{(a-2,a)}(x) + (1 - \frac{1}{2}(1 + \frac{a}{2} - \frac{x}{2})^2 I_{[a,a+2)}(x) + I_{[a+2,\infty)}(x)$.
c) a.
d) a.
e) $a + 2 - \sqrt{1.6}$.

23. *a*) $c = \frac{3}{4}$, esto es,$f_X(x) = \frac{3}{4}(1 - (x-a)^2) I_{(a-1,a+1)}(x)$.
b) a.
c) a.

24. a) I) $\frac{1}{2}$.
II) $F_X(x) = 0.25x I_{(0,1)}(x) + (0.5x - 0.25) I_{[1,2]}(x) + (0.25x + 0.25) I_{(2,3)}(x) + I_{(3,\infty)}(x)$.
III) 1.5.
b) I) $\frac{1}{2}$.
II) $F_X(x) = \frac{\lambda}{2} x I_{(0,1)}(x) + \frac{1}{2}(\lambda - 1 + x) I_{[1,2]}(x) + \frac{1}{2}(3\lambda - \lambda x + x - 1) I_{(2,3)}(x) + I_{[3,\infty)}(x)$.
III) $2 - \lambda$.

25. *a*) 0.9606.
b) 0.00018.

27. 0.88889.

29. 1.25.
30. $\frac{1}{2}$.
31. *a*) $|c| \leq \frac{1}{8}$.
b) $\frac{16}{3}c$.
c) $(\frac{c}{2}x^2 + \frac{1}{4}x - 2c + \frac{1}{2})I_{[-2,2]}(x) + I_{(2,\infty)}(x)$.
d) Si $c > 0$, la moda es 2; si $c < 0$, la moda es -2; si $c = 0$, no hay modas.
33. *a*) $f_X(x) = I_{\{c\}}(x)$.
b) $F_X(x) = I_{[c,\infty)}(x)$.
c) $E[X] = c$; $V[X] = 0$.
d) $m_X(t) = e^{ct}$.
35. 2.
36. $f_X(x) = ((\frac{1}{6})(\frac{x-1}{12}) + \frac{1}{144})I_{\{1,2,\ldots,12\}}(x)$.
37. $f_X(x) = \dfrac{\binom{x-1}{1}}{\binom{12}{2}} I_{\{1,2,\ldots,12\}}(x)$.
38. $P[X \leq x] = \dfrac{\binom{x}{n}}{\binom{r}{n}}$, para $x = n, n+1, \ldots, r$.
39. $f_X(k) = \dfrac{c^{\frac{k}{2}}e^{-c}}{k!}$, para $k = 0, 2, 4, \ldots$; 0 en otro caso.
40. $\frac{1}{3}$.
41. *a*) $\frac{2}{n(n+1)}$.
b) $\frac{2n+1}{3}$.
c) n.
42. *a*) $\sqrt{ln(2)}$.
b) $\frac{1}{\sqrt{ln(2)}}$.
c) $\sqrt{ln(2)}$.
d) 0.
43. *a*) $\frac{3}{8}$.
b) 2.
44. *a*) 1.
b) $F_X(x) = (1 - e^{-x})I_{[0,\infty)}$.
c) $\frac{6}{7}$.
45. 2.
46. $E[X^n] = \beta^n(k)(k+1)\cdots(k+n-1)$.
47. $m_X(t) = \frac{1}{4}(3e^{2t} + e^{3t})$.
48. $\sqrt{8}$.
49. 2.
50. b) $\frac{1}{3}$.
c) $\frac{1}{16}$.
d) $\frac{1}{7}$.
51. *a*) $M_X(t) = \frac{t}{2-t}$, si $|t| < 2$.
b) 2.

c) 3.
d) $F_X(x) = 1 - (1/2)^x$, para $x = 1, 2, \ldots$

52. *a*) $\frac{1}{2}$.
b) $F_X(x) = \frac{x^2}{4} I_{[0,1]}(x) + (\frac{x}{2} - \frac{1}{4}) I_{(1,2)}(x)$
$+ (1 - \frac{1}{4}(3-x)^2) I_{(2,3]}(x) + I_{(3,\infty)}(x)$.
c) 1.5.
d) 1) 0.375.
2) 0.875.

53. *a*) p.
b) 1.
c) $F_X(x) = 1 - (1-p)^x$, para $x = 1, 2, \ldots$

54. *a*) $3ln(4)$.
b) 0.
c) $\frac{ln(2)}{3ln(4)}$.

55. 30.59592.

56. 11,530.65.

57. *a*) 6.
b) $\frac{1}{2}$.
c) $F_X(x) = (3x^2 - 2x^3) I_{(0,1)}(x) + I_{[1,\infty)}(x)$.
d) 0.45744.
e) $m = 0.5$.

60. Media=Mediana= 0.5.

61. *a*) $\frac{2}{7}$.
b) $\frac{4}{21}$.
c) $F_X(x) = \frac{1}{7}(3 + 4x + x^2) I_{(-1,0)}(x)$
$+ \frac{1}{7}(3 + 4x - x^2) I_{[0,2)}(x) + I_{[2,\infty)}(x)$.
d) $2 - \frac{\sqrt{14}}{2}$.
e) 0.

62. *a*) 0.3.
b) 0.
c) 0.675.
d) 0.25.

63. $(1 - \frac{x}{2}) I_{[0,2]}(x)$.

64. $(\frac{1}{2})^{x-3} I_{\{4,5,6,\ldots\}}(x)$.

65. 6.30887.

67. *a*) 46,538.
b) 740.11283.

68. *a*) $2/\sqrt{11}$.
b) $f_X(x) = (1/\sqrt{11} + x/2) I_{(0,2/\sqrt{11})}(x)$
$+ (4/\sqrt{11}) I_{[2/\sqrt{11},4/\sqrt{11})}(x)$.
c) 0.62322.
d) 0.74921.

Capítulo 3

1. *a*) 1.555×10^{-9}.
   b) 3.813×10^{-6}.
   c) 3.
   d) 5.
2. 0.999895.
3. 0.15120.
4. *a*) 0.999596.
   b) 0.002517.
   c) 10.
5. 0.64464.
6. *a*) 0.0189.
   b) 5.9049×10^{-6}.
   c) 0.05292.
   d) 0.16308.
   e) 0.9919.
7. *a*) 0.01099.
   b) 0.90485.
   c) 5.46; 0.4914.

8.- a) $1-(1-p)^x, x = 1, 2, \ldots$
   d) $pe^t(1-qe^t)^{-1}, t < -ln(q)$.
   e) $\frac{1}{p}$; $\frac{q}{p^2}$.

11. *a*) 0.06294.
    b) 0.97702.
    c) 1.78571; 0.89148.
    d) 2.
12. *a*) 0.08192.
    b) 0.32768.
    c) 0.03072.
    d) 0.5904.
13. $p(1-qt)^{-1}$, donde $t < \frac{1}{q}$.
14. *a*) 0.42319.
    b) 0.04130.
    c) 0.95021.
    d) 0.002205.
15. *a*) 0.014599.
    b) 0.93194.
    c) 0.003304.
    d) 0.33848.
16. $p^r(1-qt)^{-r}, t < \frac{1}{q}$.
17. *a*) 0.090909.

b) 0.125.
18. *a*) 0.1024.
b) 0.96666.
c) 0.32768.
d) 0.08808.
19. $M_X(t) = (1 - p + pt)^n$.
20. *a*) 0.133602.
b) 0.12462.
21. 0.00141.
22. 0.01402.
23. *a*) 0.56653.
b) 0.59049.
c) 0.99498.
24. *a*) 1.
b) 20; 0.5.
c) 0.40547.
d) $\frac{1}{2}$.
e) 1.
f) 10; 5.
25. *a*) 2.
b) 2.
c) 2.
d) 1.
e) 1.
26. *a*) 2.
b) 2; 3.
c) 1; 2.
d) 0.
e) 1; 2.
27. b) $(pe^t)^r(1 - qe^t)^{-r}, t < -ln(q)$.
c) $\frac{rq}{p}$; $\frac{q}{p^2}$.
28. *a*) 117.
b) 0.37006.
c) 0.45789.
29. *a*) 0.37033.
b) 0.3125.
30. 0.08392.
31. *a*) 3.
b) 1.
c) 3.
32. 0.8008517.
34. *a*) 0.03004.
b) 0.69929.

c) 0.04979.
35. a) 0.38404.
b) 0.41692.
c) 0.999498.
d) 0.32927.
36. a) 0.55951.
b) 0.99678.
c) 0.77107.
37. 0.095256.
38. 4.
39. 0.362137.
40. Poisson con $\lambda = 7$.
41. $\frac{1}{2}$.
42. 0.00461.
43. a) 6.66667; 1.11111.
b) 7.
c) 7.
d) 6.
44. 0.34375.
45. 6.
46. a) 0.21208.
b) 1.
47. 0.99364.
48. $n = 6$; $p = 0.4$.
49. $\frac{3}{16}$.
50. a) 0.32768; 0.99968.
b) 6.
51. 0.00101.
52. $1 - (\frac{1}{6})^3[1 + \binom{3}{1}(\frac{5}{6}) + \binom{4}{2}(\frac{5}{6})^2 + \cdots + \binom{n-1}{n-3}(\frac{5}{6})^{n-3}]$.
53. a) 15; 90.
b) 10.
c) 5.
54. 0.5121.
55. $\frac{10}{11}$; $\frac{100}{121}$.
56. 17.88819 o 0.11181.
57. $e^{\lambda(t-1)}$.
60. $r = 6$; $p = 0.4$.
61. a) 0.00015.
b) 0.19876.
64. $[\binom{10}{4} - \binom{6}{4}]^{-1}[\binom{4}{x}\binom{6}{4-x}]I_{\{1,2,3,4\}}(x)$.
65. 1.72308; 0.47716.
66. $[1 + rq + \binom{r+1}{2}q^2 + \binom{r+2}{3}q^3 + \binom{r+3}{4}q^4]^{-1}[\binom{x+r-1}{x}q^x]I_{\{0,1,2,3,4\}}(x)$.
67. 1.47368; 1.61773.
68. a) 0.001234.

- *b*) 0.91136.
- *c*) 2.

69. *a*) 0.37199.
    b) 0.45455; 0.41605; (0.43388 binomial).
    c) 0.00712.
70. 0.35358.
71. 0.75449.

Capítulo 4

1. *a*) 0.5.
   b) 0.20.
2. *a*) 0.0465.
   b) 0.2877.
   c) 0.42436.
   d) 14.55279.
3. *a*) 0.0091.
   b) 0.95532.
   c) 0.00355.
4. *a*) 0.30119.
   b) 1.73287.
   c) 8.04719.
5. r/λ; r/λ^2.
6. *a*) 1/10.
   b) $F_X(x) = (x+5)/10I_{(-5,5)}(x) + I_{[5,\infty)}(x)$.
   c) $E[X] = 0$; $V[X] = 25/3$.
   d) 2/9.
   e) 3.
7. *a*) 243.01.
   b) 0.5.
8. μ; σ^2.
9. *a*) 0.33333; 0.03175.
   b) 0.25.
   c) 0.16861.
   d) 0.33696.
   e) 0.26411.
10. 0.9544.
11. $E[X] = ab/(a-1)$, si $a > 1$; $V[X] = ab^2/((a-1)^2(a-2))$, si $a > 2$.
12. *a*) 0.67668.
    b) 0.23533.
13. *a*) 0.47799.
    b) 9,740.03.
14. *a*) 0.77868.
    b) 0.02313.

c) 0.7555.
15. 0.75.
16. 15.625.
17. *a*) 0.98168.
b) 0.01034.
18. No es posible.
19. *a*) 8.13917.
b) 0.13863.
c) 100; 29.85075.
20. 0.0668.
22. *a*) 90.01713; 117.9974.
b) 0.8438.
c) 0.2066.
d) 0.61972.
23. *a*) 0.13534.
b) 52.59837.
c) 0.14127.
24. $\sqrt{ln(2)/\pi}$.
25. 12.52632; 5.26316.
27. *a*) 0.32439.
b) 0.85248.
28. *a*) 0.0195.
b) $1,648,835.78$.
29. $a = -2$; $b = 3$.
30. *a*) 0.25; 0.035.
b) 0.125.
c) 0.46867.
d) 22.1; 1.67886.
31. *a*) 0.91792.
b) 0.22359.
32. $6,001.25$.
33. 0.7372.
34. *a*) 0.12457.
b) 46.3; 3,905.64
35. $(k+1)/(2k+1)$, si $k > 0$.
37. *a*) 0.13534.
b) 0.01485.
c) 0.28632.
d) 300; 1,307.67.
38. 141.42136.
39. $\{1/[(b-a)(k+1)]\}(b^{k+1} - a^{k+1})$.
40. 20.36364; 1,162.19014.
42. 132; 0.06923.
44. *a*) λ.

  b) $(b/a^b)x^{b-1}$.

45. *a*) 2.
  b) 6; 12.
  c) 0.79629.
  d) 10.64464.

46. 0.528.
47. 1,222.02.
48. $e^{t^2/2}$; $(1/\sqrt{2\pi})e^{-x^2/2}$.
49. $(-0.1, 0.2)$; $(0.4, 0.7)$.
50. *a*) 9.985.
  b) 14.41428.
  c) 0.0749.

51. a/x.
52. 0.99227.
53. 14.75707.
54. Garantía 3.
55. $\sqrt{2/\pi}e^{-x^2/2}I_{(0,\infty)}$.
56. $\lambda e^{-\lambda(x-\theta)}I_{(\theta,\infty)}(x)$.
57. $[1/(b-a-(d-c))]I_{(a,c)\cup(d,b)}(x)$.
58. $[1/(\sqrt{2\pi}\sigma A)]e^{-(x-\mu)^2/2\sigma^2}I_{(0,1)}(x)$, donde
  $A = P[-\mu/\sigma < Z < (1-\mu)/\sigma]$.
59. $(1/A)(b/a^b)x^{b-1}e^{-(x/a)^b}I_{(c,d)}(x)$, donde $A = e^{-(c/a)^b} - e^{-(d/a)^b}$.
61. *a*) 0.0004.
  b) ≈ 1.
  c) 0.01523.

62. *a*) 788.2405.
  b) 0.19452.
  c) 0.00018.

Capítulo 5

2. $f_{\underline{X}}(x_1, x_2) = \dfrac{\binom{5}{x_1}\binom{2}{x_2}\binom{3}{4-x_1-x_2}}{\binom{10}{4}} I_{\{0,1,\ldots 4-x_2\}}(x_1)I_{\{0,1,\ldots,4\}}(x_2)$.

3. *a*) $f_{\underline{X}}(x_1, x_2) = kI_{[0,2-x_2]}(x_1)I_{[0,2]}(x_2)$.
  b) $\frac{1}{2}$.
  c) 0.875.
  d) 0.5.
  e) $\frac{2}{3}$.

4. *a*) $f_{\underline{X}}(x_1, x_2) = ke^{-x_1}I_{[0,x_1]}(x_2)I_{[0,\infty)}(x_1)$.
  b) $k = 1$.
  c) 0.58233.

d) 0.86466.
e) 0.84518.

5. *a*) $f_{\underline{X}}(x_1, x_2) = \dfrac{1}{\pi r^2} I_{[-\sqrt{r^2-x_2^2}, \sqrt{r^2-x_2^2}]}(x_1) I_{[-r,r]}(x_2).$
b) 0.25.
c) 0.5.
d) 0.75.

6. *a*) $f_{\underline{X}}(x_1, x_2) = k x_1 x_2 e^{-\frac{x_1+x_2}{2}} I_{(0,\infty)}(x_1) I_{(0,\infty)}(x_2).$
b) $k = \frac{1}{16}$.
c) 0.08207.
d) 0.55783.

7. *a*) $f_{X_1}(x_1) = \dfrac{\dbinom{5}{x_1}\dbinom{5}{4-x_1}}{\dbinom{10}{4}} I_{\{0,1,\ldots,4\}}(x_1).$
b) $f_{X_2|X_1}(x_2|x_1) = \dfrac{\dbinom{2}{x_2}\dbinom{3}{4-x_1-x_2}}{\dbinom{5}{4-x_1}} I_{\{0,1,\ldots,4-x_1\}}(x_1).$
c) 0.6.
d) 0.53571.

8. *a*) $k = 12$.
b) $f_{X_1}(x_1) = 3x_1^2 I_{(0,1)}(x_1)$; $f_{X_2}(x_2) = 4x_2^3 I_{(0,1)}(x_2)$.
c) 0.064.
d) $f_{X_1|X_2}(x_1|x_2) = 3x_1^2 I_{(0,1)}(x_1)$; $f_{X_2|X_1}(x_2|x_1) = 4x_2^3 I_{(0,1)}(x_2)$.
e) 0.343.

9. *a*) 0.42857.
b) $\frac{1}{3}$.

10. *a*) $f_{X_1}(x_1) = x_1 e^{-x_1} I_{(0,\infty)}(x_1)$; $f_{X_2}(x_2) = e^{-x_2} I_{(0,\infty)}(x_2)$.
b) $f_{X_1|X_2}(x_1|x_2) = e^{-x_1} e^{x_2} I_{[x_2,\infty)}(x_1)$, donde $0 \leq x_2 < \infty$;
$f_{X_2|X_1}(x_2|x_1) = \frac{1}{x_1} I_{[0,X_1]}(x_2)$, donde $0 < x_1 < \infty$.
c) 0.36788.
d) $E[X_1] = 2$; $E[X_2] = 1$.

11. *a*) $f_X(x) = \frac{1}{11} I_{\{0,1,\ldots,10\}}(x)$.
b) 0.27273.
c) 5.
d) 5.

12. *a*) $f_{\underline{X}}(x_1, x_2) = \frac{10!}{x_1! x_2! (10-x_1-x_2)!} 0.3^{x_1} 0.5^{x_2} 0.2^{10-x_1-x_2} \times$
$I_{\{0,1,\ldots,10-x_2\}}(x_1) I_{\{0,1,\ldots,10\}}(x_2)$.
b) 0.00159.
c) 0.00041.

13. *a*) $f_{\underline{X}}(x_1, x_2) = \frac{1}{100} e^{-\frac{x_1+x_2}{10}} I_{(0,\infty)}(x_1) I_{(0,\infty)}(x_2)$.
b) 0.26424.

14. 0.

15. *a*) X_1 y X_2 son variables aleatorias incorrelacionadas.

    b) X_1 y X_2 son variables aleatorias dependientes.

16. *a*) $f_{\underline{X}}(x_1,x_2) = \frac{5!}{x_1!x_2!(5-x_1-x_2)!}(0.5)^{x_1}(0.3665)^{x_2}(0.1335)^{5-x_1-x_2} \times I_{\{0,1,\ldots,5-x_2\}}(x_1)I_{\{0,1,\ldots,5\}}(x_2)$.

    b) 0.065721.

    c) 0.37629.

    d) 0.00142.

18. *a*) $f_{\underline{X}}(1,1) = \frac{4}{9}$; $f_{\underline{X}}(1,2) = \frac{2}{9}$; $f_{\underline{X}}(2,2) = \frac{2}{9}$; $f_{\underline{X}}(2,4) = \frac{1}{9}$.

    b) $Cov(X_1, X_3) = 0.296296$.

    c) X_1 y X_3 son variables aleatorias dependientes.

19. *a*) $f_{\underline{X}}(-2,-1) = \frac{1}{12}$; $f_{\underline{X}}(-2,1) = \frac{1}{12}$; $f_{\underline{X}}(-2,2) = \frac{1}{12}$; $f_{\underline{X}}(-1,-1) = \frac{1}{12}$; $f_{\underline{X}}(-1,1) = \frac{1}{12}$; $f_{\underline{X}}(-1,2) = \frac{1}{12}$; $f_{\underline{X}}(1,1) = \frac{1}{6}$; $f_{\underline{X}}(1,2) = \frac{1}{12}$; $f_{\underline{X}}(2,2) = \frac{1}{4}$.

    b) $Cov(X_1, X_3) = \frac{5}{6}$.

    c) X_1 y X_3 son variables aleatorias dependientes.

20.
$$f_{X_2|X_1}(x_2|x_1) = \begin{cases} \frac{1}{4}I_{\{1,2,3,4\}}(x_2), & \text{si } x_1 = 1, \\ \\ \dfrac{1}{2}I_{\{2\}}(x_2) + \dfrac{1}{4}I_{\{3,4\}}(x_2), & \text{si } x_1 = 2, \\ \\ \dfrac{3}{4}I_{\{3\}}(x_2) + \dfrac{1}{4}I_{\{4\}}(x_2), & \text{si } x_1 = 3, \\ \\ I_{\{4\}}(x_2), & \text{si } x_1 = 4. \end{cases}$$

21. $E[X_2] = 1.66667$.

22. *a*) $f_{\underline{X}}(0,0) = 0.48225$; $f_{\underline{X}}(0,1) = 0.09645$; $f_{\underline{X}}(1,0) = 0.09645$; $f_{\underline{X}}(1,1) = 0.21219$; $f_{\underline{X}}(1,2) = 0.0385$; $f_{\underline{X}}(2,1) = 0.0385$; $f_{\underline{X}}(2,2) = 0.02701$; $f_{\underline{X}}(2,3) = 0.00386$; $f_{\underline{X}}(3,2) = 0.00386$; $f_{\underline{X}}(3,3) = 0.00077$.

    b) $f_{X_i}(0) = 0.5787$; $f_{X_i}(1) = 0.34722$; $f_{X_i}(2) = 0.06945$; $f_{X_i}(3) = 0.00463$, $i = 1, 2$.

    c) $f_{X_2|X_1}(x_2|x_1) =$

$$\begin{cases} 0.83333I_{\{0\}}(x_2) + 0.16667I_{\{1\}}(x_2), & \text{si } x_1 = 0, \\ \\ 0.27778I_{\{0\}}(x_2) + 0.61111I_{\{1\}}(x_2) + 0.11111I_{\{2\}}(x_2), & \text{si } x_1 = 1, \\ \\ 0.55615I_{\{1\}}(x_2) + 0.38936I_{\{2\}}(x_2) + 0.05564I_{\{3\}}(x_2), & \text{si } x_1 = 2, \\ \\ 0.83369I_{\{2\}}(x_2) + 0.16631I_{\{3\}}(x_2), & \text{si } x_1 = 3. \end{cases}$$

d) X_1 y X_2 son variables aleatorias dependientes.

23. *a*) $\frac{n}{2}$.
b) $f_{X_2}(x_2) = \frac{1}{n+1} I_{\{0,1,\ldots,n\}}(x_2)$.
c) $\frac{n}{2}$.

24. b) $f_{X_i}(x_i) = I_{[0,1]}(x_i)$, $i = 1, 2$.
c) $Cov(X_1, X_2) = -\frac{\alpha}{36}$.

25. *a*) $f_{\underline{X}}(1,1) = \frac{1}{15}$; $f_{\underline{X}}(1,2) = \frac{4}{15}$; $f_{\underline{X}}(1,3) = \frac{4}{15}$;
$f_{\underline{X}}(2,2) = \frac{1}{15}$; $f_{\underline{X}}(2,3) = \frac{4}{15}$; $f_{\underline{X}}(3,3) = \frac{1}{15}$.
b) $f_{X_1}(x_1) = \frac{3}{5} I_{\{1\}}(x_1) + \frac{1}{3} I_{\{2\}}(x_1) + \frac{1}{15} I_{\{3\}}(x_1)$,
$f_{X_2}(x_2) = \frac{1}{15} I_{\{1\}}(x_2) + \frac{1}{3} I_{\{2\}}(x_2) + \frac{3}{5} I_{\{3\}}(x_2)$.
c) $Cov(X_1, X_3) = 0.15111$.

26. *a*) $E[X_1] = \frac{7}{3}$; $V[X_1] = \frac{5}{9}$.
b) $E[X_2] = 0.23333$.
c) $f_{\underline{X}}(x_1, x_2) = \binom{x_1}{x_2} \frac{x_1}{6} (0.1)^{x_2} (0.9)^{x_1 - x_2} I_{\{0,1,\ldots,x_1\}}(x_2) I_{\{1,2,3\}}(x_1)$.

27. b) $E[X_1] = np_1$; $V[X_1] = np_1(1 - p_1)$.
c) $E[X_2] = np_2$; $V[X_2] = np_2(1 - p_2)$.
d) $Cov(X_1, X_2) - np_1p_2$.
e) $B(n, p_2)$.
f) $B(n - x_1, \frac{p_2}{1-p_1})$.
g) $E[X_2|x_1] = (n - x_1)\frac{p_2}{1-p_1}$.

28. *a*) $NB(0, 0, 2, 2, 0)$.
b) $n(0, 2)$.
c) X_1 y X_2 son variables incorrelacionadas, ya que $\rho = 0$.
d) X_1 y X_2 son variables aleatorias independientes, ya que (X_1, X_2) se distribuye normal bivariada y $\rho = 0$.

29. *a*) $c_1 = 3$; $c_2 = 4$.
b) $f_{\underline{X}}(x_1, x_2) = 12x_1^2 I_{(0,x_2)}(x_1) I_{(0,1)}(x_2)$.
c) 0.448.
d) 0.26172.

30. -0.71811.

31. *a*) $\frac{1}{30}$.
b) $\frac{2}{15}$.
c) $\frac{1}{3}$.

32. 0.25.

33. $$E[X_2|x_1] = \begin{cases} 2.19048, & \text{si } x_1 = 1, \\ \\ 2.13333, & \text{si } x_1 = 2. \end{cases}$$
$$V[X_2|x_1] = \begin{cases} 0.63039, & \text{si } x_1 = 1, \\ \\ 0.64889, & \text{si } x_1 = 2. \end{cases}$$

34. *a*) $f_{X_i}(x_i) = \frac{10}{21} I_{\{0\}}(x_i) + \frac{7}{42} I_{\{1,2\}}(x_i) + \frac{4}{21} I_{\{3\}}(x_i)$, donde $i = 1, 2$.

b) $f_{X_1|X_2}(x_1|x_2) =$
$$\begin{cases} \frac{1}{10}I_{\{0\}}(x_1) + \frac{2}{10}I_{\{1\}}(x_1) + \frac{3}{10}I_{\{2\}}(x_1) + \frac{4}{10}I_{\{3\}}(x_1), & \text{si } x_2 = 0, \\ \frac{4}{7}I_{\{0\}}(x_1) + \frac{2}{7}I_{\{1\}}(x_1) + \frac{1}{7}I_{\{2\}}(x_1), & \text{si } x_2 = 1, \\ \frac{6}{7}I_{\{0\}}(x_1) + \frac{1}{7}I_{\{1\}}(x_1), & \text{si } x_2 = 2, \\ I_{\{0\}}(x_1), & \text{si } x_2 = 3. \end{cases}$$

$f_{X_2|X_1}(x_2|x_1) =$
$$\begin{cases} \frac{1}{10}I_{\{0\}}(x_2) + \frac{2}{10}I_{\{1\}}(x_2) + \frac{3}{10}I_{\{2\}}(x_2) + \frac{4}{10}I_{\{3\}}(x_2), & \text{si } x_1 = 0, \\ \frac{4}{7}I_{\{0\}}(x_2) + \frac{2}{7}I_{\{1\}}(x_2) + \frac{1}{7}I_{\{2\}}(x_2), & \text{si } x_1 = 1, \\ \frac{6}{7}I_{\{0\}}(x_2) + \frac{1}{7}I_{\{1\}}(x_2), & \text{si } x_1 = 2, \\ I_{\{0\}}(x_2), & \text{si } x_1 = 3. \end{cases}$$

c) $E[X_1|x_2] = \begin{cases} 2, & \text{si } x_2 = 0, \\ \frac{4}{7}, & \text{si } x_2 = 1, \\ \frac{1}{7}, & \text{si } x_2 = 2, \\ 0, & \text{si } x_2 = 3. \end{cases}$

$E[X_2|x_1] = \begin{cases} 2, & \text{si } x_1 = 0, \\ \frac{4}{7}, & \text{si } x_1 = 1, \\ \frac{1}{7}, & \text{si } x_1 = 2, \\ 0, & \text{si } x_1 = 3. \end{cases}$

d) $E[X_2] = 1.07143$.

37. b) $f_{X_2}(0) = 0.1625$; $f_{X_2}(1) = 0.3625$; $f_{X_2}(2) = 0.3125$; $f_{X_2}(3) = 0.1375$; $f_{X_2}(4) = 0.025$.
    c) $E[X_2] = 1.5$.
    d) $E[E[X_2|X_1]] = 1.5$.
38. b) $m_{\underline{X}}(t_1, t_2) = e^{e^{t_2}(1+e^{t_1})-2}$.
    c) $E[X_1] = 1$; $E[X_2] = 2$; $V[X_1] = 1$; $V[X_2] = 2$.

d) 0.70711.
e) $E[X_1|x_2] = \frac{x_2}{2}, \quad x_2 = 0, 1, \ldots$
f) $E[E[X_1|X_2]] = 1$.

39. $\rho_{X_1,X_2} = \pm 0.8$.
40. X_i tiene distribución $N(0,1)$, para $i = 1, 2$.
41. *a*) $f_{\underline{X}}(x_1, x_2) = \frac{1}{2} I_{(0,x_1)}(x_2) I_{(0,2)}(x_1)$.
b) 0.5.
c) $f_{X_1}(x_1) = 0.5x_1 I_{(0,2)}(x_1)$; $f_{X_2}(x_2) = (1 - 0.5x_2) I_{(0,2)}(x_2)$.
d) $f_{X_1|X_2}(x_1|x_2) = \frac{1}{2-x_2} I_{(x_2,2)}(x_1)$, donde $0 \le x_2 < 2$.
$f_{X_2|X_1}(x_2|x_1) = \frac{1}{x_1} I_{(0,x_1)}(x_2)$, donde $0 < x_1 \le 2$.
42. *a*) 2/3.
b) 1/3.
c) 1/3.
43. ± 0.57735.
44. *a*) 0.5.
b) 0.25.
45. 265.
46. X_1 y X_2 son variables aleatorias dependientes.
47. X_1 y X_2 son variables aleatorias dependientes.
48. *a*) 18.
b) $\frac{3}{5} x_2^2$, donde $0 < x_2 \le 1$.
49. 0.99723.
50. $Beta(2, 2)$.
51. *a*) 3/2.
b) $f_{X_1}(x_1) = \frac{3}{2} \sqrt{x_1} I_{(0,1)}(x_1)$.
c) $f_{X_2|X_1}(x_2|x_1) = \frac{1}{\sqrt{x_1}} I_{(0,\sqrt{x_1})}(x_2)$, donde $0 < x_1 \le 1$.
52. *a*) 2.
b) $Exp(2)$.
c) $f_{X_2|X_1}(x_2|x_1) = e^{-x_2+x_1} I_{(x_1,\infty)}(x_2)$.
d) $E[X_2|x_1] = 1 + x_1$.
53. *a*) $f_{\underline{X}}(x_1, x_2) = \dfrac{1}{\pi} I_{[1-\sqrt{1-(1-x_1)^2}, 1+\sqrt{1-(1-x_1)^2}]}(x_2) I_{(0,2)}(x_1)$.
b) $f_{X_1}(x_1) = \frac{2}{\pi} \sqrt{1 - (1 - x_1)^2} I_{(0,2)}(x_1)$.
c) $f_{X_2|X_1}(x_2|x_1) = \frac{1}{2\sqrt{1-(1-x_1)^2}} I_{[1-\sqrt{1-(1-x_1)^2}, 1+\sqrt{1-(1-x_1)^2}]}(x_2)$, si $0 < x_1 < 2$.
d) 1.
54. $E[X_2|x_1] = \frac{2}{3} x_1$, donde $0 < x_1 \le 1$.
55. $a = 2$.
57. $E[X_2] = 1/2$.
58. $E[X_2] = 25/8$.
59. $V[X_2] = 945$.
61. $E[X_1] = \mu_1$; $E[X_2] = \mu_2$; $V[X_1] = \sigma_1^2$; $V[X_2] = \sigma_2^2$; $\rho_{X_1,X_2} = \rho$.
63. *a*) 0.31166.

b) 0.31166.

64. *a*) $f_{\underline{X}}(x_1, x_2) = \frac{1}{4} I_{[0,-x_1]}(x_2) I_{[-2,0]}(x_1) + \frac{1}{4} I_{[0,x_1]}(x_2) I_{(0,2]}(x_1).$
$= \frac{1}{4} I_{[-2,-x_2]\cup[x_2,2]}(x_1) I_{[0,2]}(x_2).$
b) 3/4.
c) $f_{X_1}(x_1) = -\frac{x_1}{4} I_{[-2,0]}(x_1) + \frac{x_1}{4} I_{(0,2]}(x_1);$
$f_{X_2}(x_2) = \frac{1}{2}(2 - x_2) I_{[0,2]}(x_2).$
d) $X_2|x_1 \sim U(0, -x_1)$, si $-2 \leq x_1 < 0$.
$X_2|x_1 \sim U(0, x_1)$, si $0 < x_1 \leq 2$.

65. *a*) $f_{X_1}(x_1) = \frac{4}{5}(ln(5) - ln(4x_1)) I_{(0,5/4)}(x_1).$
b) 5/16.

Capítulo 6

1. *a*) $f_U(u) = \frac{1}{25}|u| I_{[-5,5]}(u).$
b) $f_U(u) = (2 - u) I_{[1,2]}(u) + (u - 2) I_{(2,3]}(u).$
c) $f_U(u) = \frac{1}{u}|ln(u)| I_{[e^{-1},e]}(u).$
d) $f_U(u) = I_{[0,1]}(u).$
2. *a*) $f_U(u) = 0.1 I_{[-2,3]}(u) + 0.20 I_{[3,5.5]}(u).$
b) 2.375.
3. *a*) $f_U(u) = \frac{u}{2} I_{[0,2]}(u).$
b) $f_U(u) = \frac{1}{2}(\frac{u}{2} + 1) I_{[-2,0]}(u) + \frac{1}{2}(1 - \frac{u}{2}) I_{(0,2]}(u).$
4. 0.5.
5. $f_U(u) = p(1 - p)^{ln(u)} I_{\{1,e,e^2,\ldots\}}(u).$
6. *a*) $F_U(u) = \frac{u^2}{8} I_{[0,2]}(u) + \{1 - \frac{(4-u)^2}{8}\} I_{(2,4]}(u) + I_{(4,\infty)}(u).$
b) $f_U(u) = \frac{u}{4} I_{[0,2]}(u) + (1 - \frac{u}{4}) I_{(2,4)}(u).$
7. *a*) $f_U(u) = (u + 1) I_{[-1,0]}(u) + (1 - u) I_{(0,1]}(u).$
b) $f_U(u) = \frac{1}{2} I_{[0,1]}(u) + \frac{1}{2u^2} I_{(1,\infty)}(u).$
8. *a*) $f_U(u) = (u + 1) I_{[-1,0]}(u) + (1 - u) I_{(0,1]}(u).$
b) $f_U(u) = \frac{1}{2} I_{[0,1]}(u) + \frac{1}{2u^2} I_{(1,\infty)}(u).$
9. $f_{\underline{U}}(u_1, u_2) = (\frac{1}{5})^{u_2-2}(\frac{4}{5})^{4-u_2}$, donde $\underline{u} = (-1, 3), (0, 2), (0, 4), (1, 3).$
10. $f_U(u) = \frac{2}{9} I_{\{-1,1\}}(u) + \frac{5}{9} I_{\{0\}}(u).$
11. $U \sim U(0, 27).$
12. $U \sim Exp(\lambda = \frac{1}{4}).$
13. *a*) $f_{\underline{U}}(u_1, u_2) = \frac{1}{2\pi}|u_2| e^{-\frac{1}{2}(u_1^2 u_2^2 + u_2^2)}.$
16. *a*) $NB(2\mu, 0, 2\sigma^2, 2\sigma^2, 0).$
b) Sí son independientes, debido a que $\rho = 0$.
c) $U_1 \sim N(2\mu, 2\sigma^2)$; $U_2 \sim N(0, 2\sigma^2)$.
17. *a*) $f_U(u) = \frac{3}{28} I_{[0,1]}(u) + \frac{11}{28} I_{(1,2]}(u) + \frac{7}{28} I_{(2,4]}(u).$
b) $f_U(u) =$
$\left[\binom{7}{7-\frac{3}{4}u} I_{\{0,1.33,2.67,4\}}(u) + \binom{7}{u} I_{\{3\}}(u) + \binom{7}{2u-2} I_{\{1,1.5,2\}}(u)\right] 0.7^5.$
18. *a*) $\frac{1}{2\pi\sigma^2} exp\left(-\left(5u_1^2 + 2u_2^2 - 6u_1u_2 - 2u_1\mu + 2\mu^2\right)\left(\frac{1}{2\sigma^2}\right)\right).$
b) $NB(2\mu, 3\mu, 2\sigma^2, 5\sigma^2, \frac{3}{\sqrt{10}}).$

19. $f_U(u) = \frac{1}{2\sqrt{u}} I_{(0,1)}(u)$.
20. $f_U(u) = \frac{1}{4\sqrt{u}} I_{[0,1]}(u) + \frac{1}{8\sqrt{u}} I_{(1,3]}(u)$.
21. *a*) $f_{\underline{U}}(u_1, u_2) = \frac{1}{2\pi} e^{-\frac{u_1^2+u_2^2}{2}}$.
    b) $NB(0, 0, 1, 1, 0)$.
    c) Sí son independientes, debido a que $\rho = 0$.
    d) $U_i \sim N(0, 1)$, donde $i = 1, 2$.
22. $f_U(u) = u^{-2} I_{[1,\infty)}(u)$.
23. $U \sim Beta(b, a)$.
24. $f_U(u) = (1-u)^{-2} e^{-\frac{u}{1-u}} I_{(0,1)}(u)$.
25. $f_U(u) = \frac{1}{\pi(1+u^2)}$.
26. *a*) $f_U(u) = \frac{1}{2\theta\sqrt{u}} I_{(0,\theta^2)}(u)$.
    b) $f_U(u) = \frac{1}{2\theta\sqrt{u}} I_{(0,\theta^2)}(u)$.
    c) $f_U(u) = \frac{1}{3\theta\sqrt{u}} I_{(0,\theta^2)}(u) + \frac{1}{6\theta\sqrt{u}} I_{(\theta^2,4\theta^2)}(u)$.
27. $f_U(u) = (2u_1^2 - \frac{2}{3}u_1^3) I_{[0,1]}(u_1) + (\frac{8}{3} - 2u_1^2 + \frac{2}{3}u_1^3) I_{(1,2]}(u)$.
28. *a*) $f_{\underline{U}}(u_1, u_2) = \lambda^2 u_2 e^{-\lambda u_2}(1+u_1)^{-2} I_{[0,\infty)}(u_1) I_{[0,\infty)}(u_2)$.
    b) $U_1 \sim F(2, 2)$.
    c) $U_2 \sim Gamma(r = 2, \lambda)$.
29. *a*) $f_{\underline{U}}(u_1, u_2) = 2e^{-u_2} I_{[0,\frac{u_2}{2}]}(u_1) I_{[0,\infty)}(u_2)$.
    b) $f(u_2) = u_2 e^{-u_2} I_{[0,\infty)}(u_2)$.
30. *a*) $f_{\underline{U}}(u_1, u_2) = \frac{3}{2}u_1 I_{[-u2,1]}(u_1) I_{[-1,0)}(u_2) + \frac{3}{2}u_1 I_{[u_2,1]}(u_1) I_{[0,1]}(u_2)$.
    b) $f_{U_1}(u_1) = 3u_1^2 I_{(0,1)}(u_1)$; $f_{U_2}(u_2) = \frac{3}{4}(1-u_2^2) I_{[-1,1]}(u_2)$.
31. $f_U(u) = 2u^{-3} I_{[1,\infty)}(u)$.
32. *a*) $f_{\underline{U}}(u_1, u_2) = 8u_1u_2^3 I_{(0,1)}(u_1) I_{(0,1)}(u_2)$.
    b) $U_1 \sim Beta(2, 1)$; $U_2 \sim Beta(4, 1)$.
33. *a*) $f_{\underline{U}}(u_1, u_2) = \frac{1}{5}\lambda^2 e^{-\frac{3}{5}\lambda u_1 - \frac{1}{5}\lambda u_2} I_{(0,\frac{u_2}{2})}(u_1) I_{(0,\infty)}(u_2)$.
    b) $f_{U_1}(u_1) = \lambda e^{-\lambda u_1} I_{(0,\infty)}(u_1)$;
    $f_{U_2}(u_2) = \frac{1}{3}\lambda^2 e^{-\frac{1}{5}\lambda u_2}(1 - e^{-\frac{3}{10}u_2}) I_{(0,\infty)}(u_2)$.
34. $f_U(u) = \frac{4}{u^3} I_{[1,\sqrt{2}]}(u)$.
35. $f_U(u_1, u_2) = \frac{\lambda^2}{5} e^{-\frac{\lambda}{5}(u_1+u_2)} I_{[\frac{2}{3}u_2,\frac{3}{2}u_2]}(u_1) I_{(0,\infty)}(u_2)$.
36. *a*) $f_U(u) = uI_{[0,1]}(u) + (2-u) I_{(1,2]}(u)$.
    b) $f_U(u) = uI_{[0,1]}(u) + (2-u) I_{(1,2]}(u)$.
37. $BN(r = 2, p)$.
40. *a*) 53.
    b) $N(53, 16.36)$.
    c) 63,415.23.
41. La f.g.m. de $\underline{U}$ es $m_{\underline{U}}(t_1, t_2) = e^{2\mu t_1 + \sigma^2(t_1^2+t_2^2)}$.
42. *a*) La f.g.m. es $(p + (1-p)e^t)^{n_2}$.
    b) La f.g.m. es $(0.5e^t + 0.5)^{n_1+n_2}$.
43. La f.g.m de $\underline{U}$ es $m_{\underline{U}}(t_1, t_2) = e^{\frac{1}{2}(t_1^2+t_2^2)}$.

44. *a*) Normal bivariada con parámetros $\mu_1 = \mu \sum_{i=1}^{n} a_i, \mu_2 = \mu \sum_{i=1}^{n} b_i,$
$\sigma_1^2 = \sigma^2 \sum_{i=1}^{n} a_i^2, \sigma_2^2 = \sigma^2 \sum_{i=1}^{n} b_i^2, \rho = \frac{\sum_{i=1}^{n} a_i b_i}{\sqrt{\sum_{i=1}^{n} a_i^2 \sum_{i=1}^{n} b_i^2}}.$
b) Son independientes si y solo si $\sum_{i=1}^{n} a_i b_i = 0$.
45. 0.125.
47. La función de densidad conjunta de U_1 y U_2 es igual a
$f_{\underline{U}}(u_1, u_2) = 120(1 - e^{-u_1})e^{-4u_1} I_{(0,\infty)}(u_1)(1 - e^{-u_2})e^{-2u_2} I_{(0,\infty)}(u_2).$
48. $Weibull(a, nb)$.
49. *a*) La función de densidad conjunta de U_1, U_2 y U_3 es
$f_{\underline{U}}(u_1, u_2, u_3) = 48u_1 I_{(0,1)}(u_1)u_2^3 I_{(0,1)}(u_2)u_3^5 I_{(0,1)}(u_3).$
b) $U_1 \sim Beta(2, 1)$; $U_2 \sim Beta(4, 1)$; $U_3 \sim Beta(6, 1)$.
50. 0.40741.
51. *a*) $f_{u_1}(u_1) = \frac{3}{2}u_1^2 I_{(0,1]}(u_1) + (6 - 6u_1 + \frac{3}{2}u_1^2)I_{(1,2)}(u_1).$
b) $f_{u_1}(u_1) = 12u_1^2 I_{(0,0.5]}(u_1) + 12(1 - u_1)^2 I_{(0.5,1)}(u_1).$
52. $f_{\underline{U}}(u_1, u_2) = 9(u_1^2 + u_2^2 - 2u_1^2u_2 - 2u_1u_2^2 + 2u_1^2u_2^2)I_{[0,u_2)}(u_1)I_{[0,1]}(u_2).$
53. *a*) Las función de densidad conjunta de U_1 y U_2 es
$f_{\underline{U}}(u_1, u_2) = e^{-u_1} I_{(0,\infty)}(u_1)e^{-u2} I_{(0,\infty)}(u_2).$
b) La función de densidad conjunta de $U_1, U_2, \ldots, U_n$ es
$f_{\underline{U}}(u_1, \ldots, u_n) = e^{-u_1} I_{(0,\infty)}(u_1)e^{-u_2} I_{(0,\infty)}(u_2) \cdots e^{-u_n} I_{(0,\infty)}(u_n).$
c) $U_i \sim Exp(1)$.
54. $f_R(r) = \frac{1}{\sqrt{\pi}}e^{-\frac{r^2}{4}} I_{(0,\infty)}(r).$
55. *a*) $U \sim N(c + a\mu_1 + b\mu_2, \sigma_1^2 a^2 + \sigma_2^2 b^2 + 2\rho\sigma_1\sigma_2 ab).$
b) $U \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2).$
c) 0.3345.
d) 0.0571.
56. 0.1587.
57. $E[S] = \frac{\sqrt{2}\Gamma(\frac{n}{2})}{\sqrt{n-1}\Gamma(\frac{n-1}{2})}\sigma;\ V[S] = \sigma^2 \left(1 - \frac{2\Gamma^2(\frac{n}{2})}{(n-1)\Gamma^2(\frac{n-1}{2})}\right).$
58. *a*) $N(0, 1)$.
b) $N(0, 1)$.
c) $F(1, 1)$.
d) Ji cuadrada con 1 grado de libertad.
e) t de student con 1 grado de libertad.
f) $F(1, 1)$.
59. *a*) $N(0, \frac{1}{k} + \frac{1}{n-k})$.
b) Ji cuadrada con 2 grados de libertad.
c) $F(1, 1)$.
d) t de student con 1 grado de libertad.
e) Ji cuadrada con $n - 2$ grados de libertad.
f) $F(k - 1, n - k - 1)$.
g) t de student con $n - 1$ grados de libertad.
60. *a*) 1.345
b) 6.388.
c) 3.868.

d) $21.0641\sigma^2$.
e) -0.45469.

61. *a*) 0.75.
b) $1 - 0.5^n$.

62. $\frac{2}{3}$.

63. *a*) $Beta(10, 1)$.
b) 0.90909; 0.00689.

64. *a*) $34(1-2y)^{16} I_{(0,0.5)}(y)$.
b) $Beta(1, 17)$.

65. $Beta(3, 2)$.

66. $Beta(n-1, 2)$.

67. $Exp(n\lambda)$.

68. *a*) 0.97079.
b) 0.32529.

69. $B(n, p_1 + p_2)$.

70. $BN(2, p)$.

71. *a*) $e^{(0.9\mu - 30000)t + \frac{0.81\sigma^2 t^2}{2}}$.
b) $N(0.9\mu - 30000, 0.81\sigma^2)$.

72. *a*) 3.9403.
b) $8.5468\sigma^2$.
c) 5.20916.
d) 0.19748.

73. Dos casos:
$f_y(y) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \frac{1}{d^{a+b-1}} (y-c)^{a-1}(d-(y-c))^{b-1} I_{(c,c+d)}(x)$, si $d \geq 0$.
$f_y(y) = -\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \frac{1}{d^{a+b-1}} (y-c)^{a-1}(d-(y-c))^{b-1} I_{(c+d,c)}(x)$, si $d < 0$.

Capítulo 7

2. $Exp(1)$.
3. μ.
7. $Exp(1)$.
8. $Exp(1)$.
9. $Gamma(r = 2, \lambda = 1)$.
10. $Gamma(r = 2, \lambda = 1)$.
12. 0.
13. 0.
15. No tiene distribución límite.
16. No tiene distribución límite.
20. 0.
24. $1/\lambda$.
25. 0.
26. 1.
27. 0.6826.

Acerca del autor

Antonio González Fragoso estudió la licenciatura en Actuaría y la maestría en Estadística en la Universidad Nacional Autónoma de México (UNAM). Realizó un doctorado en Matemáticas, con tesis doctoral en probabilidad y estadística en la Benemérita Universidad Autónoma de Puebla. Es uno de los fundadores de la licenciatura en Actuaría de la UDLAP y el primer coordinador del programa, donde ha impartido cursos de probabilidad, estadística y simulación, entre otros. Fue director académico del Departamento de Actuaría, Física y Matemáticas. Ha realizado investigación en inferencia estadística para procesos estocásticos aplicados a la actuaría. Además, ha trabajado como consultor e instructor para diversas empresas en estadística aplicada, esencialmente, en temas de mejora continua. Actualmente, sigue colaborando en la UDLAP con la impartición de cursos de probabilidad y estadística.

Probabilidad: un enfoque actuarial
fue preparado por el Departamento de Publicaciones de la Universidad de las Américas Puebla, Ex hacienda Santa Catarina Mártir s/n, San Andrés Cholula, Puebla, 72810, para su publicación en línea en agosto de 2025.

En este libro se presentan los principales temas de probabilidad que un alumno de licenciatura en disciplinas como Actuaría, Matemáticas o Ciencia de Datos deberá estudiar para lograr una base sólida que le permita abordar con facilidad diferentes áreas.

La teoría de la probabilidad es, muchas veces, indispensable para la aplicación de las matemáticas. Su uso abarca, entre otras áreas, la matemática actuarial que se emplea en diversos modelos de seguros; las metodologías estadísticas aplicadas a problemas reales, o los modelos de probabilidad en los que se sustenta, en buena medida, la matemática financiera.

En cada capítulo se explican conceptos y resultados de manera simple, clara y bien estructurada, sin perder la formalidad, con el apoyo de gráficas, simulaciones y ejemplos que ayudan a comprender los temas tratados.

Pensando en los futuros actuarios, se exponen con detalle algunos modelos o distribuciones que tienen aplicaciones en el área de seguros y finanzas. Se consideraron también ejemplos y ejercicios que serán de gran ayuda para los estudiantes interesados en presentar los exámenes de The Society of Actuaries (SOA).

UDLAP®