Sonia Salvo Garrido
Antonio Sanhueza
Rodrigo Puchi Ancasay

Probabilidad y estadística

Con aplicaciones para ingeniería y ciencias

COLECCIÓN INGENIERÍA Y CIENCIAS

Ed. UFRO UNIVERSITY PRESS

edulp EDITORIAL DE LA UNLP

Probabilidad y estadística

Con aplicaciones para ingeniería y ciencias

Sonia Salvo Garrido
Antonio Sanhueza
Rodrigo Puchi Ancasay

Probabilidad y estadística

Con aplicaciones para ingeniería y ciencias

Ed. UFRO UNIVERSITY PRESS
edulp EDITORIAL DE LA UNLP

PROBABILIDAD Y ESTADÍSTICA
Con aplicaciones para ingeniería y ciencias

Sonia Salvo Garrido
Antonio Sanhueza
Rodrigo Puchi Ancasay

EDICIONES UNIVERSIDAD DE LA FRONTERA
Colección Ingeniería y Ciencias
Primera edición: octubre de 2020
ISBN: 978-956-236-389 1

UNIVERSIDAD DE LA FRONTERA
Av. Francisco Salazar 01145, casilla 54-D, Temuco

Diseño de portada: Ediciones UFRO
Impreso en Andros Impresores

Las matemáticas, incluyendo las probabilidades, son teorías formales. La estadística llega más lejos. Se apoya en ambas disciplinas y se concreta en la observación de datos y en la resolución de problemas.

Carles M. Cuadras

Índice

Prefacio

Los sucesos naturales, los experimentos y sus resultados que están regidos por las leyes del azar son mucho más abundantes que los debidos a causas deterministas. Se admite que esto es lo que ocurre en la naturaleza, la sociedad, la economía, la técnica y la ciencia. La probabilidad y la estadística proporcionan métodos para el análisis de los datos de origen aleatorio y constituyen un instrumento imprescindible en todas las áreas científico-técnicas. Su ubicuidad se debe al papel preponderante que tienen en el desarrollo de las ciencias experimentales y sociales.

La estadística moderna requiere un buen conocimiento de la probabilidad y sus aplicaciones, a fin de poder modelar y cuantificar los fenómenos aleatorios, que deben ser bien comprendidos para ser tratados correctamente.

Probabilidad y estadística. Con aplicaciones para ingeniería y ciencias, de los profesores Salvo, Sanhueza y Puchi, es un texto docente con un eficaz y riguroso planteamiento de ambos temas, que son tratados mediante desarrollos teóricos y numerosos ejemplos y problemas aplicados.

Se trata de una obra que me satisface prologar y recomendar a los profesores y estudiantes de ingeniería y ciencias que deseen introducirse en la comprensión, descripción y resolución de muchos de los problemas que surgen en los principales ámbitos científico-técnicos.

Dr. Carles M. Cuadras, Universidad de Barcelona,
miembro electo del International Statistical Institute.

Presentación

Este texto corresponde a apuntes de clases de materias de probabilidad y estadística impartidas en carreras de ingeniería y ciencias realizadas por los autores durante su trayectoria académica, complementados por una recopilación de distintas publicaciones. Contar con una versión estructurada y digitalizada de los apuntes ha sido un anhelo permanente, pensando en que sea un instrumento de apoyo a la labor docente y, sobre todo, que sea una guía útil para el trabajo autónomo del estudiante.

Los contenidos están organizados en los siguiente capítulos:

Capítulo 1. Probabilidad: Aborda el concepto de probabilidad desde su definición axiomática, sobre la base del aporte de Kolmogorov. Hace un recorrido por temáticas afines como probabilidad clásica, probabilidad condicional y los teoremas de probabilidad total y de Bayes.

Capítulo 2. Variables aleatorias: Considera para el caso discreto y continuo, conceptos como función de distribución y de probabilidad, valor esperado y varianza.

Capítulo 3. Distribuciones: Estudia las distribuciones de variables aleatorias ampliamente conocidas y usadas para modelar situaciones prácticas, en particular algunas de sus propiedades y formas de aplicarlas al desarrollo de problemas prácticos. Además, presenta algunas transformaciones de variables aleatorias desde y hacia distribuciones usuales.

Capítulo 4. Vectores aleatorios: Es una extensión del capítulo 2 a dos o más dimensiones, desde vectores aleatorios bivariados hasta vectores aleatorios multivariados.

Capítulo 5. Estimación de parámetros: Comienza el estudio de métodos para obtener estimaciones de parámetros poblacionales, medidas de calidad de las estimaciones, estimación por intervalos de confianza y teoremas límites para obtener aproximaciones de estimaciones.

Capítulo 6. Pruebas de hipótesis: Aborda métodos para evaluar la veracidad de hipótesis estadísticas, desde las definiciones de hipótesis estadísticas y la construcción de test estadísticos hasta pruebas de hipótesis de uso habitual, como pruebas de comparación de dos grupos independientes (medias, varianzas y proporciones) y de comparación de medidas correlacionadas.

Capítulo 7. Tópicos complementarios: Considera los tópicos de asociación o causalidad entre dos variables aleatorias: regresión lineal simple, anova de un factor y asociación entre variables cualitativas.

Esperamos que este texto sea un aporte al estudio de la probabilidad y la estadística.

Los autores.

Capítulo 1

Probabilidad

Este primer capítulo tiene por objetivo introducir las nociones básicas de probabilidad, a partir de un experimento aleatorio hasta realizar un desarrollo teórico del concepto de probabilidad por medio de los axiomas de Kolmogorov, incluyendo los teoremas de Bayes y de probabilidad total.

1.1. Experimento aleatorio y espacio muestral

En esta primera sección, se presentan elementos básicos relacionados con un experimento aleatorio, espacio muestral y suceso aleatorio, los cuales nos permitirán modelar diversas situaciones condicionadas por el azar mediante una expresión matemática.

Por ejemplo, la mayoría de las personas en algún momento nos preguntamos: ¿será que hoy va a llover? De alguna manera logramos asignar una chance de llover y luego decidimos qué ropa utilizar o si llevar o no un paraguas. Y así podríamos imaginar otras situaciones en que sentimos incertidumbre y dependemos de que suceda una de las alternativas que pueden ocurrir.

Experimento aleatorio

El primer concepto que aparece en la descripción anterior es el de experimento aleatorio. Este consiste en darse una serie de circunstancias y observar lo que sucede, bajo la condición de que se conozca el conjunto

de resultados posibles, el cual debe ser no vacío, pero sin tener la certeza de cuál ocurrirá al efectuar el experimento.

Este tipo de experimento es opuesto a los experimentos determinísticos, en los cuales se conoce de antemano el resultado que se obtendrá.

En síntesis, se llamará experimento aleatorio a todo experimento que satisfaga las siguientes condiciones:

(a) Se conocen todos los posibles resultados del experimento.

(b) Se desconoce un resultado del experimento antes de ejecutarlo.

(c) Se puede repetir el experimento un gran número de veces bajo las mismas condiciones.

Ejemplo **1.1** Algunos ejemplos de experimentos aleatorios son:

(a) Lanzar una moneda al aire y observar el resultado.

(b) Lanzar n monedas al aire y observar el resultado de cada una ($n \in \mathbb{N}$).

(c) Lanzar un dado sobre una superficie y observar el número que aparece.

(d) Seleccionar aleatoriamente 3 fichas desde una tómbola que contiene n fichas ($n > 3$).

(e) Registrar la cantidad de accidentes automovilísticos que ocurren diariamente en una ciudad.

(f) Seleccionar al azar un punto desde un círculo con centro en el origen y radio 1. □

Una rifa es un juego en el cual un boleto de entre todos los vendidos es premiado (elegido usando un método que depende completamente del azar); la rifa entrega 3 premios a 3 boletos distintos, entonces este tipo de experimento corresponde al planteado en el punto d del ejemplo 1.1. Si alguien que compró (o pretende comprar) un boleto está interesado en conocer las chances de que su boleto sea elegido, es necesario identificar todas las posibilidades que arroja la realización de este experimento, que es lo que se presenta a continuación.

Espacio muestral y eventos aleatorios

El espacio muestral (Ω) asociado a un experimento aleatorio es el conjunto de todos los resultados posibles de dicho experimento. Es decir, se tendrá que $\omega \in \Omega$ si y solo si ω es un resultado posible del experimento.

Los conjuntos que se forman con cada elemento $\omega \in \Omega$ son llamados *eventos* o *sucesos elementales*. Así, el conjunto de todos los eventos elementales constituye el espacio muestral.

Ejemplo **1.2** Del ejemplo 1.1, para cada experimento aleatorio descrito el espacio muestral es, respectivamente:

(a) $\Omega = \{c, s\}$, siendo c: cara y s: sello.

(b) Para $n = 2$, $\Omega = \{(c, c), (c, s), (s, c), (s, s)\}$.

(c) $\Omega = \{1, 2, 3, 4, 5, 6\}$.

(d) Hay dos modalidades para este experimento: una es sin reemplazo, donde la ficha seleccionada no se devuelve a la tómbola, y la otra es con reemplazo, donde la ficha seleccionada se devuelve antes de la siguiente extracción.

Selección sin reemplazo: para la primera selección hay n posibilidades, para la segunda (al no devolver la ficha seleccionada en el paso anterior) hay $n - 1$ posibilidades y, finalmente, para la tercera selección habrá $n - 2$ posibilidades.

Selección con reemplazo: cada una de las 3 realizaciones del experimento tendrá n posibilidades.

(e) $\Omega = \{1, 2, 3, \cdots\}$.

(f) $\Omega = \big\{(a, b) | a^2 + b^2 \leqslant 1\big\}$. □

De los ejemplos anteriores (ejemplo 1.2), se desprende que un espacio muestral puede ser de diferente naturaleza según el tipo de elementos que definen el conjunto. Así, por ejemplo, en los cinco primeros casos, de los cuales los cuatro primeros son finitos y el quinto es infinito, el espacio muestral es un conjunto numerable o discreto, mientras que en

el sexto caso se trata de un espacio muestral continuo. Aun cuando se podría definir un tercer grupo que fuese una combinación entre ambos tipos (discreto y continuo), en este texto el interés está solo en los dos casos descritos.

En ocasiones el interés estará en algunos (no todos) los posibles resultados del experimento aleatorio. Por ejemplo, en un juego de dados, el interés puede estar en que al realizar el experimento aparezca un número par. Entonces, se define con el concepto de evento (o suceso) a cualquier agrupación de posibles resultados del experimento aleatorio.

Así, todo A tal que $A \subseteq \Omega$ se llamará evento. Además, si a todos los elementos de A es posible atribuirles una posibilidad de ocurrencia, entonces A será llamado evento aleatorio.

Definición **1.1** $A = \{\omega_i | \omega_i \in \Omega\} \subset \Omega$ es un evento aleatorio si y solo si es posible atribuirle una posibilidad de ocurrencia.

Ejemplo **1.3** Se lanza una moneda dos veces. Para este experimento determinar el espacio muestral y el evento que se define por *el número de caras es igual al número de sellos.*

Solución: Para este experimento aleatorio el espacio muestral es:

$$\Omega = \{(C, C), (C, S), (S, C), (S, S)\} .$$

Por su parte, si se llama al evento A *el número de caras es igual al número de sellos*, entonces:

$$A = \{(C, S), (S, C)\} .\square$$

Definición **1.2** Algunos eventos que tienen cierta particularidad son:

- $\forall \omega \in \Omega$, $A = \{\omega\}$ es un evento elemental.
- $A = \Omega$ es el evento seguro.
- $A = \phi$ es el evento imposible.
- A^c es el evento complementario (o contrario) al evento A. Es decir, A^c es el evento que ocurre cuando no ocurre el evento A.

Nótese que usando la definición 1.2 en el ejemplo 1.3 un evento elemental sería que aparecezca un doble sello, un evento seguro sería que salga cara o sello en los lanzamientos, un evento imposible sería que salgan tres caras y, en el caso del evento complementario, sería: $A^c = \{(C, C), (S, S)\}$ (*el número de caras es distinto al número de sellos*).

Luego, si se cuenta con dos eventos A y B de un mismo espacio muestral Ω, se pueden establecer relaciones, entre las cuales están:

- $A \cup B$: ocurre A o B, incluso ambos. Si $\omega \in A \cup B$, entonces $\omega \in A$ o $\omega \in B$.
- $A \cap B$: ocurre A y B. Si $\omega \in A \cap B$, entonces $\omega \in A$ y $\omega \in B$.
- $A \subset B$: si ocurre A entonces también ocurre B. Para todo $\omega \in A$ se tiene que $\omega \in B$, luego la ocurrencia de A es condición inevitable de la ocurrencia de B.
- $A = B \Leftrightarrow A \subset B$ y $B \subset A$.
- $A \cap B = \phi$: A y B son eventos disjuntos o incompatibles.

Observación **1.1** En la interacción entre dos o más eventos de un mismo espacio muestral son válidas todas las leyes o propiedades de la teoría de conjuntos. Entre estas se pueden destacar, siendo A, B y C eventos de un espacio muestral Ω:

(a) $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$.

(b) $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$.

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

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

Ejemplo **1.4** Una tómbola contiene bolitas numeradas del 1 al 15. Aleatoriamente se extrae una bolita y se registra su número. Sea el evento A: *el número de la bolita seleccionada es par* y B: *el número de la bolita seleccionada es múltiplo de 3*, determinar el espacio muestral y los eventos $A \cup B$, $A \cap B$ y A^c.

Solución: El espacio muestral asociado a este experimento es: $\Omega = \{1, 2, \cdots, 15\}$. Para los eventos se tiene:

$$\begin{aligned} A &= \{2, 4, 6, 8, 10, 12, 14\} \\ A^c &= \{1, 3, 5, 7, 9, 11, 13, 15\} \\ B &= \{3, 6, 9, 12, 15\} \\ A \cap B &= \{6, 12\} \\ A \cup B &= \{2, 3, 4, 6, 7, 9, 10, 12, 14, 15\} \,. \square \end{aligned}$$

Además, se presentan las operaciones de unión e intersección enumerable de eventos de un cierto espacio muestral, según:

(a) $\bigcup_{i=1}^{\infty} A_i$: ocurre si al menos un evento A_i, $i = 1, 2, \cdots$ ocurre.

(b) $\bigcap_{i=1}^{\infty} A_i$: ocurre si todos los eventos A_i, $i = 1, 2, \cdots$ ocurren.

Ejercicio propuesto **1.1** Demostrar los resultados de la observación 1.1.

Hasta el momento se cuenta con un espacio muestral Ω y una colección de subconjuntos del espacio muestral de la que es posible saber las posibilidades de ocurrencia. Ahora se definirán elementos que nos permitan determinar el número de posibilidades del espacio muestral y de cualquier evento aleatorio.

Principio de la multiplicación

Esta propiedad resuelve el problema de saber cuál es la cantidad total de resultados posibles que tiene asociado un experimento aleatorio finito. Por ejemplo, una aplicación de este principio es determinar el número de posibilidades del espacio muestral, para el experimento aleatorio de seleccionar 3 fichas una a una, sin reemplazo, desde una tómbola con n fichas. En efecto, para la primera ficha seleccionada hay un total de n posibilidades, por cada una de estas posibilidades para la segunda hay $n-1$ posibilidades (descontando la ficha extraída en la primera selección); hasta aquí se tienen $n \cdot (n-1)$ posibilidades, luego para cada una de estas habrá

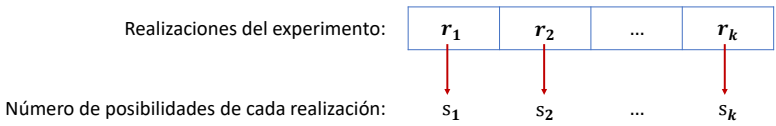
$n-2$ posibilidades (descontando las 2 fichas ya seleccionadas en los pasos anteriores). Por lo tanto, el total de posibilidades será $n \cdot (n-1) \cdot (n-1)$.

En general, utilizando el mismo análisis anterior, si se considera un experimento aleatorio con r_j realizaciones, $1 \leqslant j \leqslant k$, donde cada una tiene una posibilidad de ocurrencia s_j (tal como se representa en la figura 1.1), entonces el principio de la multiplicación señala que la cantidad total de resultados posibles que tiene un experimento aleatorio se calcula por:

$$\prod_{j=1}^{k} s_j.$$

Nótese que este último resultado en teoría de conjuntos se denomina cardinalidad del conjunto y se denota con $\#(A)$

Figura 1.1: Esquema del principio de la multiplicación



Fuente: elaboración propia.

Para determinar la cantidad de posibilidades para el conjunto A se utiliza un elemento matemático llamado *arreglo matemático* o *arreglo*.

Definición **1.3** Sean $x_1, x_2, \cdots, x_k$ las realizaciones de un experimento (que determina el conjunto A) y sean $s_1, s_2, \cdots, s_k$ las posibilidades respectivas de las realizaciones del experimento, lo cual se denota como arreglo matemático:

$$A = \{(x_1, \cdots, x_k) | x_i = s_i \text{ posibilidades}, \, i = 1, 2, \cdots, k\}.$$

Utilizando el principio multiplicativo, se tiene desde el resultado anterior que la cardinalidad del conjunto A (o la cantidad de posibilidades del experimento que determina A) es:

$$\#(A) = s_1 \cdot s_2 \cdots s_k = \prod_{x_i \in A} s_i.$$

Ejemplo **1.5** Del ejemplo 1.2, la cardinalidad del espacio muestral en cada caso es:

(a) $\Omega = \{cara, sello\} = \{(x_1)|x_1 : \text{ 2 pos.}\} \Rightarrow \#(\Omega) = 2$.

(b) Para $n = 2$, $\Omega = \{(x_1, x_2)|x_i = \text{ 2 pos. }\} \Rightarrow \#(\Omega) = 2^2$.

$\forall n \in \mathbb{N}$: $\Omega = \{(x_1, x_2, \cdots, x_n)|x_i = \text{ 2 pos. } i = 1, 2, \cdots, n\}$. Entonces, $\#(\Omega) = 2 \cdot 2 \cdots 2 = 2^n$.

(c) $\Omega = \{1, 2, 3, 4, 5, 6\} = \{(x_1)|x_1 : \text{ 6 pos.}\} \Rightarrow \#(\Omega) = 6$.

(d) Con reemplazo:
$\Omega = \{(x_1, x_2, x_3)|x_i = n \text{ pos.}; i = 1, 2, 3\} \Rightarrow \#(\Omega) = n \cdot n \cdot n = n^3$.

Sin reemplazo:
$\Omega = \{(x_1, x_2, x_3)|x_1 = n \text{ pos. }; x_2 = n - 1 \text{ pos. }; x_3 = n - 2 \text{ pos. }\}$
$\Rightarrow \#(\Omega) = n \cdot (n - 1) \cdot (n - 2)$. □

Ejercicio propuesto **1.2** Cuántos números de 4 dígitos pueden formarse con las cifras $0, 1, \cdots, 9$:

(a) Si se permiten repeticiones.

(b) Si no se permiten repeticiones.

(c) Si el último dígito ha de ser par y no se permiten repeticiones.

Ejercicio propuesto **1.3** Sea el conjunto $\{a, b, c, d\}$, determinar la cantidad de distintas combinaciones que pueden obtenerse usando 3 de las letras del conjunto dado, considerando que (1) no se permiten repeticiones y (2) se permiten repeticiones.

Ejercicio propuesto **1.4** Suponer que la fecha de cumpleaños de una persona está igualmente distribuida entre los 365 días del año. Si en una sala hay n personas, determinar el número de posibilidades del espacio muestral asociado y del evento *todas las personas han nacido en días diferentes.*

Ejercicio propuesto **1.5** Se lanza un icosaedro regular 4 veces (las caras del poliedro están numeradas del 1 al 20). Determinar el número de posibilidades que tiene el espacio muestral y el evento *no aparece ningún número repetido en los lanzamientos.*

Ejemplo **1.6** Se tienen 5 bolitas en una tómbola numeradas del 1 al 5, donde las bolitas 1, 2 y 3 son azules (A) y las restantes son blancas (B). El experimento aleatorio consiste en extraer 2 bolitas desde la tómbola sin reposición. Obtener el número de posibilidades:

(a) Totales para las 2 bolitas seleccionadas.

(b) Que se generan con las 2 bolitas seleccionadas.

(c) Si solo se considera el color de la bolita (y no el número).

Solución: Nótese que para (a) y (b) las 5 bolitas se consideran como distintas y, por ende, una selección de la bolita 1 y la bolita 2 es distinta a la bolita 2 y la bolita 1. El número de posibilidades es respectivamente:
(a) $5 \cdot 4 = 20$ posibilidades.
(b) $2 \cdot 1 = 2! = 2$ posibilidades.
(c) Posibilidades de (A, B): 6, (A, A): 3 y (B, B): 1; total: 10. □

Observación **1.2** Nótese que para el ejemplo 1.6 (usando la letra correspondiente al listado) se cumple que $a = b \cdot c$. Se puede mostrar que esta regla se cumple también en general, es decir, para N objetos de los cuales se extraen n y el resultado será:

$$N \cdot (N-1) \cdots (N-(n-1)) = \frac{N!}{(N-n)!} = n! \cdot c \Rightarrow c = \frac{N!}{n!(N-n)!} = \binom{N}{n}.$$

Esta última expresión se llama *coeficiente binomial* (o símbolo binomial) y se utiliza para calcular combinaciones de n entre N objetos.

Hasta aquí se cuenta con un espacio muestral Ω, eventos aleatorios y técnicas para el conteo del número de posibilidades de un evento. Se avanzará a determinar la probabilidad de ocurrencia de un evento aleatorio, para lo cual se definirán conceptos y leyes que permitan alcanzar este objetivo.

Sigma-álgebra

Sea Ω un conjunto no vacío. Una colección de eventos aleatorios $\mathcal{F}$ de subconjuntos de Ω se llama σ-álgebra si y solo si se cumple:

(a) $\Omega \in \mathcal{F}$.

(b) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$.

(c) $A_1, A_2, \cdots$ es una sucesión de eventos: $A_i \in \mathcal{F}$, $\Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Observación **1.3** $A \in \mathcal{F} \Leftrightarrow A \subseteq \Omega$.

Ejemplo **1.7** Considerar el espacio muestral dado por: $\Omega = \{1, 2, 3, 4, 5, 6\}$ y la colección de subconjuntos de Ω, correspondiente a su conjunto potencia $\wp(\Omega)$, luego se tiene que la σ-álgebra $\mathcal{F}$ es:

$$\mathcal{F} = \wp(\Omega) = \{\Omega, \phi, \{1\}, \{2\}, \cdots, \{6\}, \cdots, \{1, 2, 3, 4, 5, 6\}\}.$$

Se afirma que $\mathcal{F}$ es una σ-álgebra. En efecto, (a) y (b) son inmediatas, mientras que (c) se puede deducir como sigue:

Sea $A_i = \{i\}, i = 1, 2, \cdots, 6$ y $A_i = \{\phi\}$ si $i = 7, 8, \cdots$. Luego, $\bigcup_{i=1}^{\infty} A_i = \Omega \in \mathcal{F}$. □

Ejercicio propuesto **1.6** Se define $\mathcal{F} = \{\phi, \Omega\}$, llamada la menor σ-álgebra asociada a Ω. Demostrar que es efectivamente una σ-álgebra.

Ejercicio propuesto **1.7** Sea $\mathcal{F}$ una σ-álgebra, mostrar que:

(1) $\phi \in \mathcal{F}$.

(2) Si $A_1, \cdots, A_n \in \mathcal{F}$, se tiene que $\bigcup_{i=1}^{n} A_i \in \mathcal{F}$ y $\bigcap_{i=1}^{n} A_i \in \mathcal{F}$.

(3) Si $A_1, A_2, \cdots \in \mathcal{F}$ se tiene que $\bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$.

(4) Sea $A, B \in \mathcal{F}$, entonces $A - B = A \cap B^c \in \mathcal{F}$.

Ejercicio propuesto **1.8** Considerar el experimento aleatorio de lanzar un dado (equilibrado) 2 veces, el espacio muestral está dado por:

$$\Omega = \{(x_1, x_2) | x_1 = 6 \text{ pos.}, x_2 = 6 \text{ pos.}\},$$

y sea $\mathcal{F} = \wp(\Omega)$ (conjunto potencia de Ω). Demostrar que $\mathcal{F}$ es una σ-álgebra.

Observación **1.4** $(\mathcal{F}, \Omega)$ constituyen un espacio medible. Finalmente, se define la medida de probabilidad para construir el espacio de probabilidad $(\mathcal{F}, \Omega, \mathbb{P})$.

Ejemplo **1.8** Se define el experimento con 2 resultados posibles: éxito (E) y fracaso (F) y se repite sucesivamente. Si el experimento se detiene cuando ocurre el primer éxito y es de interés observar el número de fracasos obtenidos, entonces el espacio muestral está dado por:

$$\Omega = \{\omega_1, \omega_2, \cdots\},$$

donde ω_i representa el resultado experimental cuyas $i-1$ primeras componentes son F:

$$\omega_i = (\underbrace{F,\ F,\ F,\ \cdots,\ F}_{i-1 \text{ componentes}}, E\,).$$

Ahora se desea asignar una probabilidad (medida de incertidumbre) al evento A: *el primer éxito ocurre después de un número par de intentos*:

$$A = \{w_2\} + \{w_4\} + \cdots.$$

Como A es una unión numerable de eventos elementales (y por tanto disjuntos), se requiere que $A \in \mathcal{F}$ para que sea posible atribuirle una probabilidad. □

1.2. Axiomas de probabilidad

Sea el espacio muestral Ω y $\mathcal{F}$ la σ-álgebra asociada a Ω. Se define la función $\mathbb{P}$, llamada función de probabilidad sobre $(\mathcal{F}, \Omega)$, que va desde la σ-álgebra de subconjuntos de Ω a $[0,1] \subset \mathbb{R}$.

$$\begin{array}{rccc} \mathbb{P}: & \mathcal{F} & \longrightarrow & \mathbb{R} \\ & A \in \mathcal{F} & \longmapsto & \mathbb{P}(A) \in \mathbb{R}. \end{array}$$

Es decir, la función asigna a cada $A \in \mathcal{F}$ un número $\mathbb{P}(A) \in [0,1]$ que se interpreta como *la probabilidad de que ocurra el evento* A una vez realizado el experimento y que satisface los axiomas (de Kolmogorov) siguientes:

Axioma 1. $\mathbb{P}(A) \geqslant 0$, $\forall A \in \mathcal{F}$.

Axioma 2. $\mathbb{P}(\Omega) = 1$.

Axioma 3. Dada una sucesión numerable de eventos disjuntos, es decir, $A_i \cap A_j = \phi$ si $i \neq j$, se verifica que:

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}\left(A_i\right).$$

Observación **1.5** Los elementos $(\Omega, \mathcal{F}, \mathbb{P})$ definen completamente una probabilidad. Unen lo que se definió como experimento aleatorio (a través del espacio muestral) y aquellos elementos que permiten asignar un valor a la ocurrencia de un evento.

Ejemplo **1.9** Considerar el espacio muestral asociado a un experimento aleatorio dado por $\Omega = \{0, 1, 2, \cdots, n\}$ y $\mathcal{F}$ una σ-álgebra asociada a Ω. Se mostrará que la siguiente función es una función de probabilidad:

$$\mathbb{P}(A) = \frac{1}{2^n} \sum_{x \in A} \binom{n}{x}, \ A \in \mathcal{F}.$$

Solución: Para probar que $\mathbb{P}(A)$ es una función de probabilidad se verifican los axiomas de probabilidad:

(a) Dado que $\frac{1}{2^n} > 0$ y $\binom{n}{x} = \frac{n!}{x!(n-x)!} > 0$, entonces $\mathbb{P}(A) > 0$.

(b) $\mathbb{P}(\Omega) = \dfrac{1}{2^n} \displaystyle\sum_{x \in \Omega} \binom{n}{x} = \frac{1}{2^n} \sum_{x=0}^{n} \binom{n}{x} = 1.$

(c) Sea $A_1 = \{1\}, A_1 = \{2\}, \cdots, A_n = \{n\}$ y sea $A_i = \phi$, para $i = n+1, \cdots$.

$$\begin{aligned}
\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) =& \frac{1}{2^n} \sum_{x \in \bigcup_{i=1}^{\infty} A_i} \binom{n}{x} \\
=& \frac{1}{2^n} \sum_{x \in \{1,2,\cdots,n\}} \binom{n}{x} \\
=& \frac{1}{2^n} \left(\sum_{x \in 1} \binom{n}{x} + \cdots + \sum_{x \in n} \binom{n}{x} + \sum_{x \in n+1} \binom{n}{x} \cdots \right) \\
=& \mathbb{P}(\{1\}) + \mathbb{P}(\{2\}) + \cdots + \mathbb{P}(\{n\}) \\
=& \mathbb{P}(A_1) + \mathbb{P}(A_2) + \cdots + \mathbb{P}(A_n) + \mathbb{P}(A_{n+1}) + \cdots
\end{aligned}$$

Como $A_{n+1} = \phi \Rightarrow \mathbb{P}(A_{n+1}) = 0$ y para los eventos sucesivos:

$$= \sum_{i=1}^{\infty} \mathbb{P}(A_i).\square$$

Ejercicio propuesto **1.9** Considerar el experimento aleatorio de lanzar un dado (equilibrado) 2 veces, el espacio muestral está dado por:

$$\Omega = \{(x_1, x_2) | x_1 = 6 \text{ pos.}, x_2 = 6 \text{ pos.}\},$$

y sea $\mathcal{F} = \wp(\Omega)$ (conjunto potencia de Ω). Mostrar que $\mathbb{P}(A) = \frac{\#(A)}{36}$ es una función de probabilidad, $\forall A \subseteq \Omega$.

Propiedades de la función de probabilidad

La función de probabilidad $\mathbb{P}(\cdot)$ o simplemente $\mathbb{P}$ tiene asociadas ciertas propiedades que serán de utilidad para obtener el cálculo de probabilidad de ocurrencia de eventos y posteriormente la resolución de problemas de aplicación. Estas propiedades se refieren a los límites de la función $\mathbb{P}$ dado un cierto evento o suceso y, por ende, tienen relación con la teoría de conjuntos. A continuación se estudian algunas propiedades.

Propiedad **1.1** Probabilidad de ocurrencia del evento imposible: si ϕ es el evento imposible, entonces $\mathbb{P}(\phi) = 0$.

Demostración: Utilizando el tercer axioma de probabilidad se puede demostrar la propiedad. En efecto, sea:

$$A_1 = \phi, A_2 = \phi, \cdots, A_n = \phi, A_{n+1} = \phi, \cdots$$

Luego, $\bigcup_{i=1}^{\infty} A_i = \phi$. Desde el tercer axioma de probabilidad:

$$\mathbb{P}(\phi) = \mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(\phi) \Rightarrow \mathbb{P}(\phi) = 0.$$

Propiedad **1.2** Sea $A_1, A_2, \cdots, A_n$ una sucesión finita de eventos disjuntos. Entonces:

$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}(A_i).$$

Demostración: Considérese la sucesión finita a la cual se agregan sucesivos eventos imposibles hasta el infinito, es decir, $A_{n+1} = \phi, A_{n+2} = \phi, \cdots$. Luego, por el axioma 3, se tiene que:

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i).$$

Por la condición inicial sobre los eventos $A_{n+1}, \cdots$, $\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{n} A_i$:

$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}(A_i) + \sum_{i=n+1}^{\infty} \mathbb{P}(A_i),$$

donde $\mathbb{P}(A_i) = 0$ para $i = n+1, n+2, \cdots$, con lo cual se demuestra el resultado.

Propiedad **1.3** Sean A y B dos eventos de un mismo espacio muestral, entonces se cumple que:

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

Demostración: Un primer resultado se obtiene al expresar $A \cup B$ en términos de eventos disjuntos. Así, se definen los subconjuntos disjuntos S_1, S_2 y S_3 como $S_1 = A \cap B^c$, $S_2 = A \cap B$ y $S_3 = A^c \cap B$ (tal como en el diagrama de la figura 1.2). Para comprobar que son disjuntos basta con verificar que la intersección dos a dos es vacía. Luego se puede escribir:

$$A \cup B = S_1 \cup S_2 \cup S_3 = (A \cap B^c) \cup (A \cap B) \cup (A^c \cap B).$$

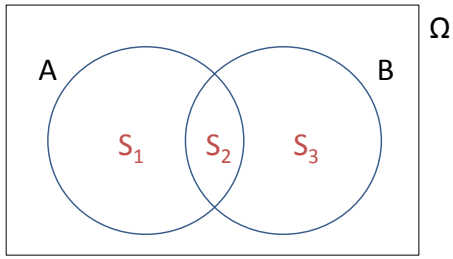
Aplicando la función de probabilidad y la propiedad 1.2 se obtendrá el resultado, el cual se usará en la otra parte del desarrollo:

$$\mathbb{P}(A \cup B) = \mathbb{P}(A \cap B^c) + \mathbb{P}(A \cap B) + \mathbb{P}(A^c \cap B).$$

Por otro lado, $A = S_1 \cup S_2 = (A \cap B^c) \cup (A \cap B)$ y $B = S_2 \cup S_3 = (A \cap B) \cup (A^c \cap B)$. Aplicando la función de probabilidad:

$$\begin{aligned} \mathbb{P}(A) =& \mathbb{P}(A \cap B^c) + \mathbb{P}(A \cap B). \\ \mathbb{P}(B) =& \mathbb{P}(A^c \cap B) + \mathbb{P}(A \cap B). \end{aligned}$$

Figura 1.2: Diagrama con las relaciones entre los eventos A y B



Fuente: elaboración propia.

Resolviendo el sistema de ecuaciones anterior:

$$\begin{aligned}\mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) =& \mathbb{P}(A \cap B^c) + \mathbb{P}(A^c \cap B) + \mathbb{P}(A \cap B) \\ =& \mathbb{P}(A \cup B).\end{aligned}$$

Propiedad **1.4** Si A es un evento cualquiera y A^c es su complemento, entonces $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.

Demostración: Por una parte se sabe que: $\mathbb{P}(A \cup A^c) = \mathbb{P}(\Omega) = 1$. Además, por la propiedad 1.3:

$$\mathbb{P}(A \cup A^c) = \mathbb{P}(A) + \mathbb{P}(A^c) - \mathbb{P}(A \cap A^c).$$

Finalmente, por la propiedad 1.1 $\mathbb{P}(A \cap A^c) = \mathbb{P}(\phi)$. Entonces:

$$\mathbb{P}(A) + \mathbb{P}(A^c) = 1.$$

A partir de último resultado es inmediato que $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.

Propiedad **1.5** Si $A \subset B$, entonces $\mathbb{P}(A) \leqslant \mathbb{P}(B)$.

Demostración: Si se expresa el evento B como la unión de dos eventos disjuntos (tal como en la demostración de la propiedad 1.3), $A \cap B$ y $A^c \cap B$. Es decir:

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

Como $A \subset B \Rightarrow A \cap B = A$, uniendo estos dos últimos resultados:

$$\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(A^c \cap B).$$

Y como $\mathbb{P}(A^c \cap B) \geqslant 0$, se obtendrá que $\mathbb{P}(B) \geqslant \mathbb{P}(A)$.

Con las propiedades anteriormente definidas se cuenta con un conjunto de resultados que permite aplicarlos a la resolución de situaciones problemáticas. Estas situaciones pueden estar dadas por demostraciones de algún resultado (que se construyen a partir de los ya vistos y las propiedades de conjuntos) y por aplicaciones que impliquen situaciones en que está presente la incertidumbre ante un experimento aleatorio.

Ejemplo **1.10** La probabilidad de que un estudiante apruebe el curso de Probabilidades es de 0,55, la probabilidad de que apruebe el curso de Cálculo es de 0,40 y la probabilidad de que apruebe ambos cursos es de 0,25. Calcular la probabilidad de que:

(a) Apruebe por lo menos uno de los dos cursos.

(b) Apruebe solo el curso de Probabilidades.

(c) Apruebe solo uno de los dos cursos.

(d) No apruebe ninguno de los dos cursos.

Solución: En este ejercicio se utilizarán las propiedades definidas. Desde el enunciado se pueden obtener algunos resultados inmediatos, no sin antes definir los eventos de interés:
A: *el estudiante aprueba el curso de Probabilidades.*
B: *el estudiante aprueba el curso de Cálculo.*

Luego, se pueden deducir los siguientes resultados (obtenidos de la propiedad 1.4):

$$\begin{aligned} \mathbb{P}(A) = 0{,}55 &\Rightarrow \mathbb{P}(A^c) = 0{,}45. \\ \mathbb{P}(B) = 0{,}40 &\Rightarrow \mathbb{P}(B^c) = 0{,}60. \\ \mathbb{P}(A \cap B) &= 0{,}25. \end{aligned}$$

Luego, teniendo en consideración que *aprobar por lo menos un curso* es el evento complementario a *no aprobar ningún curso*:

(a) $1 - \mathbb{P}(A^c \cap B^c) = \mathbb{P}(A \cup B) = 0{,}55 + 0{,}40 - 0{,}25 = 0{,}70.$

(b) $\mathbb{P}(A \cap B^c) = \mathbb{P}(A) - \mathbb{P}(A \cap B) = 0{,}55 - 0{,}25 = 0{,}30.$

(c) $\mathbb{P}\left((A \cap B^c) \cup (A^c \cap B)\right) = \mathbb{P}(A \cap B^c) + \mathbb{P}(A^c \cap B) =$
$= 0{,}30 + \mathbb{P}(B) - \mathbb{P}(A \cap B) = 0{,}30 + 0{,}40 - 0{,}25 = 0{,}45.$

(d) $\mathbb{P}(A^c \cap B^c) = 1 - \mathbb{P}(A \cup B) = 1 - 0{,}70 = 0{,}30.$

Este ejemplo constituye una muestra de cómo las propiedades enunciadas anteriormente pueden utilizarse para resolver problemas. □

Propiedad **1.6** Para todo evento $A \in \mathcal{F}$, donde $\mathcal{F}$ es una σ-álgebra, se tiene que $\mathbb{P}(A) \leqslant 1$.

Demostración: Como $A \in \mathcal{F} \Rightarrow A \subseteq \Omega$, utilizando la propiedad 1.5, resulta inmediato que $\mathbb{P}(A) \leqslant \mathbb{P}(\Omega) = 1$.

Observación **1.6** A partir de las propiedades anteriores se puede obtener que $0 \leqslant \mathbb{P}(A) \leqslant 1$.

Esta última expresión es importante en probabilidades pues restringe el recorrido de la función de probabilidad a un conjunto de infinitos valores en el intervalo $[0, 1]$, cualquiera sea el evento que se defina. Si bien este resultado se había empleado anteriormente, ahora se encuentra explicitado y, además, permite hacer la relación de equivalencia entre probabilidad y porcentaje. Ahora la probabilidad de ocurrencia puede medirse como una proporción o como un porcentaje.

Propiedad **1.7** Sean A_1, A_2, A_3 eventos cualesquiera, entonces $\mathbb{P}(A_1 \cup A_2 \cup A_3) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + \mathbb{P}(A_3) - \mathbb{P}(A_1 \cap A_2) - \mathbb{P}(A_1 \cap A_3) - \mathbb{P}(A_2 \cap A_3) + \mathbb{P}(A_1 \cap A_2 \cap A_3)$.

La demostración se puede obtener considerando un evento $B = A_2 \cup A_3$ y aplicando la propiedad 1.3 (queda como ejercicio para el lector). Se puede obtener una expresión para la probabilidad de la unión de n eventos a partir de la propiedad 1.7.

1.3. Probabilidad en espacios muestrales finitos

Si se considera un experimento aleatorio que tiene n resultados posibles, entonces el espacio muestral Ω se expresa por:

$$\Omega = \{\omega_1, \omega_2, \cdots, \omega_n\}, \text{ con } \#(\Omega) = n$$

La σ-álgebra asociada en este caso es el conjunto potencia de Ω, $\wp(\Omega)$, el cual tiene una cardinalidad de 2^n.

Ahora se asume que cada resultado posible del experimento aleatorio tiene una probabilidad conocida, es decir:

$$\mathbb{P}(\{\omega_j\}) = p_j, \ j = 1, 2, \cdots, n \text{tal que } \sum_{j=1}^{n} p_j = 1. \tag{1.1}$$

Sea $A = \{\omega_{i1}, \omega_{i2}, \cdots, \omega_{ir}\}$, un evento de interés, luego:

$$\begin{aligned}
\mathbb{P}(A) =& \mathbb{P}\left(\{\omega_{i1}, \omega_{i2}, \cdots, \omega_{ir}\}\right) \\
=& \mathbb{P}\left(\{\omega_{i1}\} \cup \{\omega_{i2}\} \cup \cdots \cup \{\omega_{ir}\}\right) \\
=& \mathbb{P}\left(\{w_{i1}\}\right) + \mathbb{P}\left(\{w_{i2}\}\right) + \cdots + \mathbb{P}\left(\{w_{ir}\}\right) \\
=& p_{i1} + p_{i2} + \cdots + p_{ir}.
\end{aligned}$$

Observación **1.7** El supuesto de la expresión 1.1, reemplazado por:

$$\mathbb{P}(\{\omega_j\}) = p; \ j = 1, 2, \cdots, n; \text{ tal que } \sum_{j=1}^{n} p = 1,$$

desde donde se tiene que $p = \frac{1}{n}$, que se denomina *supuesto de equiprobabilidad*. Este supuesto implica que:

$$\mathbb{P}(A) = p_{i1} + p_{i2} + \cdots + p_{ir} = \frac{1}{n} + \frac{1}{n} + \cdots + \frac{1}{n} = \frac{\#(A)}{\#(\Omega)}.$$

Esta última expresión $\mathbb{P}(A) = \frac{\#(A)}{\#(\Omega)}$, con $A \subseteq \Omega$, es conocida como *regla de Laplace* o probabilidad clásica.

Ejemplo **1.11** Se realiza un experimento aleatorio de lanzar un dado. Se desea calcular la probabilidad de que el número que aparece en la cara superior del dado sea un número impar.

Solución: Espacio muestral $\Omega = \{1, 2, 3, 4, 5, 6\} \Rightarrow \#(\Omega) = 6$. Es conveniente notar que el ejercicio está en un contexto de equiprobabilidad de ocurrencia de eventos elementales.

Evento de interés: $A = \{1, 3, 5\} \Rightarrow \#(A) = 3$, con lo que se obtiene que $\mathbb{P}(A) = \frac{3}{6} = 0{,}5$. □

Ejemplo **1.12** A partir del ejemplo 1.4, calcular la probabilidad de ocurrencia de los eventos A, B, $A \cup B$, $A \cap B$ y $(A \cap B)^c$.

Solución: Como la tómbola contiene bolitas numeradas del 1 al 15, la probabilidad de ocurrencia de una bola cualquiera es $\frac{1}{15}$. Así, a partir de los resultados obtenidos de la solución del ejemplo 1.4 y usando la regla definida anteriormente, se tienen los siguientes resultados:

$$\mathbb{P}(A) = \frac{7}{15},\ \mathbb{P}(B) = \frac{5}{15} = \frac{1}{3} \text{ y } \mathbb{P}(A \cap B) = \frac{2}{15}.$$

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) = \frac{7}{15} + \frac{5}{15} - \frac{2}{15} = \frac{10}{15} = \frac{2}{3}.$$

$$\mathbb{P}((A \cup B)^c) = 1 - \mathbb{P}(A \cup B) = 1 - \frac{2}{3} = \frac{1}{3}.\square$$

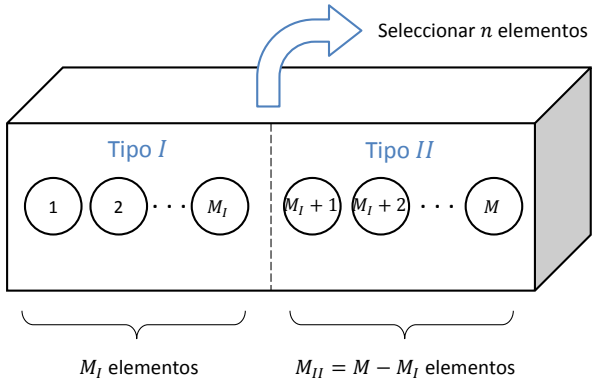
Ejercicio propuesto **1.10** Sea el experimento aleatorio de lanzar un dado (equilibrado) 2 veces, se definen los eventos A: *sale 3 en la primera tirada* y B: *la suma de ambos lanzamientos es superior a 6*.

Calcular: (a) $\mathbb{P}(A)$, (b) $\mathbb{P}(B)$ y (c) $\mathbb{P}(A \cap B)$.

Principio de la urna

Este principio agrupa a todos aquellos experimentos que se basan en observar el resultado de la extracción de un elemento, dado que se sabe que el conjunto completo está dividido por elementos de dos tipos (bolas rojas y azules, hombres y mujeres, personas a las que les gusta algo y a las que no les gusta, etc.). Los cuestionamientos de interés para este experimento tienen que ver con la cantidad de elementos observados del primer tipo (se sabe que solo hay de dos tipos por lo que los restantes deben ser del otro tipo) y el orden en que aparecen en la selección. En general, se puede entender que se tiene una urna con M elementos donde M_I son del *tipo I* y los restantes $M - M_I$ elementos son del *tipo II*, desde la que se extraen n elementos, de uno en uno, bajo la condición excluyente de reemplazo o sin reemplazo de la selección (véase esquema de la figura 1.3).

Figura 1.3: Esquema del principio de la urna



Fuente: elaboración propia.

Propiedad **1.8** Considerar una urna que contiene M elementos de la misma forma y numerados del 1 al M, donde los M_I primeros son del tipo I y los restantes $M_{II} = M - M_I$, del tipo II. Se seleccionan aleatoriamente n elementos de uno en uno y se definen los eventos U_k y V_k, tal que:

U_k: los k ($k \leqslant n$) primeros elementos seleccionados son del tipo I y los restantes $n - k$ son del tipo II.

V_k: hay exactamente k ($k \leqslant n$) elementos del tipo I entre los n seleccionados (y los restantes $n - k$ del tipo II).

Si la selección es con reemplazo, entonces la probabilidad de U_k y V_k es:

$$\mathbb{P}(U_k) = \left(\frac{M_I}{M}\right)^k \cdot \left(1 - \frac{M_I}{M}\right)^{n-k} \tag{1.2}$$

$$\mathbb{P}(V_k) = \binom{n}{k} \cdot \left(\frac{M_I}{M}\right)^k \cdot \left(1 - \frac{M_I}{M}\right)^{n-k}. \tag{1.3}$$

Si la selección se realiza sin reemplazo, entonces:

$$\mathbb{P}(U_k) = \frac{M_I!(M - M_I)!(M - n)!}{(M_I - k)!((M - M_I) - (n - k))!M!} \tag{1.4}$$

$$\mathbb{P}(V_k) = \frac{\binom{M_I}{k}\binom{M-M_I}{n-k}}{\binom{M}{n}}. \tag{1.5}$$

Demostración: Se pueden obtener las expresiones anteriores aplicando arreglos matemáticos y el principio de la multiplicación. Por simplificación solo se construirá la demostración de la probabilidad de U_k con reemplazo (expresión 1.2). Así, el arreglo para el espacio muestral y para U_k son:

$$\begin{aligned}\Omega &= \{(x_1, x_2, \cdots, x_n) \,|\text{pos. } x_1 = M, x_2 = M, \cdots, x_n = M\} \\ U_k &= \{(x_1, x_2, \cdots, x_n) \,|\text{pos. } x_1 = M_I, x_2 = M_I, \cdots, \\ &\qquad x_k = M_I, x_{k+1} = M_{II}, x_{k+2} = M_{II}, \cdots, x_n = M_{II}\}.\end{aligned}$$

Luego, $\#(\Omega) = M^n = M^k \cdot M^{n-k}$ y $\#(U_k) = M_I^k \left(M - M_I\right)^{n-k}$. Con esto, la probabilidad de U_k es:

$$\mathbb{P}\left(U_k\right) = \frac{\#(U_k)}{\#(\Omega)} = \frac{M_I^k \cdot \left(M - M_I\right)^{n-k}}{M^k \cdot M^{n-k}} = \left(\frac{M_I}{M}\right)^k \cdot \left(1 - \frac{M_I}{M}\right)^{n-k}.$$

Las restantes expresiones pueden obtenerse mediante el mismo esquema de resolución.

Ejemplo **1.13** En una sala hay 20 estudiantes, de los cuales 9 dicen que les gustan las matemáticas y los 11 restantes, que no les gustan. Se extrae una muestra de 5 estudiantes (de uno en uno y sin repetición). Determinar la probabilidad de que de los estudiantes seleccionados:

(a) A los 2 primeros les gusten las matemáticas.

(b) Haya exactamente 2 a los que les gustan las matemáticas.

Solución: El enunciado es un ejemplo de una aplicación de la propiedad 1.8. Llevando los datos a la nomenclatura usada por la propiedad, se tiene que: $M = 20$, $M_I = 9$ y $M_{II} = 11$. Luego: $\mathbb{P}(U_2) = \frac{9!11!15!}{7!8!20!} = 0{,}038$ $\mathbb{P}(V_2) = \frac{\binom{9}{2}\binom{11}{3}}{\binom{20}{5}} = 0{,}383$.

Con estos resultados, (a) la probabilidad de que a los 2 primeros estudiantes les gusten las matemáticas es de 4 % y (b) la probabilidad de que a 2 estudiantes les gusten las matemáticas es de 38 %. □

1.4. Probabilidad condicional

La situación es la siguiente (como ejemplo): se lanzan 2 dados distintos y se desea calcular la probabilidad de que aparezcan los números 4 y 5.

Entonces, la cardinalidad del espacio muestral y del evento de interés es:

$$\Omega = \{(x_1, x_2) | x_i = 6 \text{ pos.}, i = 1, 2\} \Rightarrow \#(\Omega) = 6^2 = 36$$

$$A = \{(4,5), (5,4)\} \Rightarrow \#(A) = 2$$
$$\therefore \mathbb{P}(A) = \frac{2}{36} = 0{,}0556.$$

Ahora, si se sabe que el resultado en uno de los dados es el número 5, entonces se modifica el espacio muestral y, por ende, la probabilidad de ocurrencia del evento. En efecto, si Ω^* es el nuevo espacio muestral:

$$\Omega^* = \{(1,5), (2,5), (3,5), (4,5), (5,5), (6,5), (5,1), (5,2), (5,3), (5,4), (5,6)\}$$
$$\Rightarrow \#(\Omega^*) = 11$$
$$\Rightarrow \mathbb{P}(A) = \frac{\#(A)}{\#(\Omega^*)} = \frac{2}{11} = 0{,}182.$$

Definición **1.4** Sean los eventos A y B tal que $\mathbb{P}(B) > 0$. La probabilidad condicional del evento A dado el evento B es definida por la expresión:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Ejemplo **1.14** Para la situación introductoria a la sección considerando la definición 1.4, se tendrá que A: *aparecen los números 4 y 5 al efectuar el experimento de lanzar 2 dados* y B: *en cualquiera de los 2 dados ha aparecido el número 5.* □

Ejemplo **1.15** Considerar el experimento aleatorio de lanzar un dado (equilibrado) 2 veces. El espacio muestral Ω verifica que $\#(\Omega) = 6 \cdot 6 = 36$. Si los eventos A y B se definen como: A: *sale 3 en la primera tirada* y B: *los números de ambas tiradas suman más que 6.* Hallar $\mathbb{P}(A \cap B)$.

Solución: Aquí se tiene que $\mathbb{P}(A) = \frac{6}{36}$, $\mathbb{P}(B) = \frac{21}{36}$ y luego $\mathbb{P}(A \cap B) = \frac{3}{36}$. Con estos resultados se tiene que $\mathbb{P}(A|B) = \frac{3}{21}$. □

Observación **1.8** Se puede probar que $\mathbb{P}(A|B)$ satisface las condiciones para ser una función de probabilidad. En efecto:

(a) Como $\mathbb{P}(A \cap B) \geqslant 0$ y $\mathbb{P}(B) > 0$, entonces $\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \geqslant 0$.

(b) $\mathbb{P}(\Omega|B) = \frac{\mathbb{P}(\Omega \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B)}{\mathbb{P}(B)} = 1$.

(c) Sea $A_1, A_2, \cdots$ una sucesión infinita de eventos disjuntos, entonces:

$$\mathbb{P}\left(\bigcup_{j=1}^{\infty} A_j|B\right) = \frac{\mathbb{P}\left(\left(\bigcup_{j=1}^{\infty} A_j\right) \cap B\right)}{\mathbb{P}(B)}$$
$$= \frac{\mathbb{P}\left((A_1 \cup A_2 \cup \cdots) \cap B\right)}{\mathbb{P}(B)}$$

Como $(A_1 \cup A_2 \cup \cdots) \cap B = (A_1 \cap B) \cup (A_2 \cap B) \cup \cdots$:

$$= \frac{\mathbb{P}\left(\bigcup_{j=1}^{\infty}(A_j \cap B)\right)}{\mathbb{P}(B)}$$

Los eventos $A_j \cap B$ son disjuntos para $j = 1, 2, \cdots$, entonces por la propiedad 1.2:

$$= \frac{\sum_{j=1}^{\infty} \mathbb{P}\left(A_j \cap B\right)}{\mathbb{P}(B)} = \sum_{j=1}^{\infty} \frac{\mathbb{P}\left(A_j \cap B\right)}{\mathbb{P}(B)}$$
$$= \sum_{j=1}^{\infty} \mathbb{P}\left(A_j|B\right).$$

Propiedad **1.9** Sean dos eventos A y B, entonces:

$$\mathbb{P}(A \cap B) = \mathbb{P}(B) \cdot \mathbb{P}(A|B).$$

Demostración: se obtiene directamente desde la definición 1.4.

Ejemplo **1.16** Se seleccionan 2 estudiantes (sin reemplazo) y se desea calcular la probabilidad de que ambos sean fumadores. De la muestra obtenida se tiene que el conjunto de estudiantes es de 32, de los cuales 4 fuman y los 28 restantes no fuman.

Solución: Sean los eventos A: *primer estudiante seleccionado es fumador* y B: *segundo estudiante seleccionado es fumador*.

Luego, la probabilidad de que ambos sean fumadores es $\mathbb{P}(A \cap B)$ y se puede calcular mediante la propiedad 1.9:
$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B|A) = \frac{4}{32} \cdot \frac{3}{31} = 0{,}0121.$ □

Propiedad **1.10** Se puede obtener una generalización de la propiedad 1.9, para esto sean $A_1, A_2, \cdots, A_n$ eventos cualesquiera, luego se define el evento $B = A_1 \cap A_2 \cap \cdots \cap A_n$, se cumplirá que:

$$\mathbb{P}(B) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2|A_1) \cdot \mathbb{P}(A_3|(A_1 \cap A_2)) \cdot \cdots \cdot \mathbb{P}\left(A_n \left| \bigcap_{j=1}^{n-1} A_j \right.\right).$$

Demostración: Puede realizarse mediante inducción matemática.

Ejemplo **1.17** Se seleccionan 4 estudiantes (sin reemplazo) y se desea calcular la probabilidad de que solo el primer y el tercer seleccionado sean fumadores. De la muestra obtenida se tiene que el conjunto de estudiantes es de 32, de los cuales 4 fuman y los 28 restantes no fuman.

Solución: Esta es una extensión del ejemplo 1.16, convenientemente se definen los siguientes eventos:
A_1: primer estudiante seleccionado es fumador.
A_2: segundo estudiante seleccionado es no fumador.
A_3: tercer estudiante seleccionado es fumador.
A_4: cuarto estudiante seleccionado es no fumador.

Siendo $B = A_1 \cap A_2 \cap A_3 \cap A_4$, la solución al problema corresponde a calcular la probabilidad de ocurrencia del evento B:

$$\begin{aligned}\mathbb{P}(B) =& \mathbb{P}(A_1) \cdot \mathbb{P}(A_2|A_1) \cdot \mathbb{P}(A_3|(A_1 \cap A_2)) \cdot \mathbb{P}(A_4|(A_1 \cap A_2 \cap A_3)) \\ =& \frac{4}{32} \cdot \frac{28}{31} \cdot \frac{3}{30} \cdot \frac{27}{29} = 0{,}0105.\end{aligned}$$

Con este resultado, la probabilidad de que solo el primer y el tercer seleccionado sean fumadores es de 0,0105. □

Ejercicio propuesto **1.11** Considerar una urna compuesta por a fichas azules y b fichas blancas. Se realiza el experimento (secuencial y repetitivo) con las siguientes condiciones:

(1) Se extrae una ficha y se observa su color.

(2) La ficha se devuelve con c fichas del mismo color.

(3) Luego se repite (1) y (2).

Para 3 repeticiones del experimento, definir los eventos A_i: *la i-ésima ficha es azul* y calcular la probabilidad de que en las 3 ocasiones la ficha seleccionada sea de color azul.

Independencia de eventos

La probabilidad condicional surge como respuesta a la actualización de la información acerca de la ocurrencia de un evento A mediante lo que se sabe de otro evento B. Es decir, $\mathbb{P}(A|B)$ modifica la probabilidad de ocurrencia del evento A, $\mathbb{P}(A)$. Sin embargo, en algunas situaciones ocurre que $\mathbb{P}(A|B) = \mathbb{P}(A)$, es decir, la información que se tiene de A a través del evento B no cambia la probabilidad de A. Este concepto se denomina independencia de eventos.

Definición **1.5** Sean A y B eventos del espacio muestral Ω. Luego, A y B serán eventos independientes si:

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B).$$

Ejemplo **1.18** Considerar el experimento aleatorio de lanzar un dado dos veces y los eventos A: *el número obtenido en ambas tiradas es impar* y B: *la suma en ambas tiradas es impar*. Probar la independencia de los eventos A y B.

Solución: En este caso A y B son eventos disjuntos, es decir, $\mathbb{P}(A \cap B) = \mathbb{P}(\phi) = 0$.

Además, se tiene que $\mathbb{P}(A) = \frac{1}{4}$ y $\mathbb{P}(B) = \frac{1}{2}$, entonces $\mathbb{P}(A) \cdot \mathbb{P}(B) = \frac{1}{8}$. Luego, los eventos A y B no son independientes. □

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

Demostración: Como A y B son independientes, entonces $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$. Luego:

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

Finalmente, de acuerdo con la definición 1.5, A y B^c son eventos independientes. Nótese que lo mismo ocurre entre los eventos A^c y B, y entre A^c y B^c.

Propiedad **1.12** Sean A y B eventos cualesquiera, entonces:

(1) $\mathbb{P}(B) > 0 \Rightarrow A$ y B son independientes $\Leftrightarrow \mathbb{P}(A|B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$.

(2) $\mathbb{P}(B) = 0 \Rightarrow A$ y B son independientes.

Se puede generalizar esta propiedad para un número finito de eventos.

Definición **1.6** $A_1, A_2, \cdots, A_n$ son eventos independientes si y solo si:

$$\mathbb{P}\left(\bigcap_{j=1}^{h} A_j\right) = \prod_{j=1}^{h} \mathbb{P}(A_j), \text{ para cualquier } h \leqslant n.$$

Luego, se puede enunciar una propiedad de independencia a través de la probabilidad condicional.

Propiedad **1.13** Los eventos $A_1, A_2, \cdots, A_n$ son independientes si y solo si:

$$\mathbb{P}\left(A_i \middle| \bigcap_{j=1}^{h} A_j\right) = \mathbb{P}(A_i), \quad i \neq j \text{ y para cualquier } h \leqslant n,$$

con $\mathbb{P}(\bigcap_{j=1}^{h} A_j) \neq 0$.

Propiedad **1.14** Los eventos $A_1, A_2, \cdots, A_n$ son dos a dos independientes si y solo si:

$$\mathbb{P}(A_i|A_j) = \mathbb{P}(A_i), \quad \forall i \neq j, \text{ siendo } \mathbb{P}(A_j) \neq 0, \ \forall j \leqslant n.$$

Particiones del espacio muestral

Definición **1.7** Una colección de eventos $\{A_i | i \in \mathbb{N}\}$ es una partición de Ω si y solo si:

(1) $A_i \cap A_j = \phi$, $\forall i \neq j$ (los elementos A_i son excluyentes).

(2) $\bigcup_{i \in \mathbb{N}} A_i = \Omega$ (los elementos A_i son exhaustivos).

Ejemplo **1.19** La colección de eventos $\{A, A^c\}$, $\forall A \in \mathcal{F}$ constituyen una partición del espacio muestral Ω donde $\mathcal{F}$ es una σ-álgebra de Ω. □

Probabilidad total y teorema de Bayes

Para la propiedad 1.9 si se considera el evento A como una partición $A_1, A_2, \cdots, A_n$ del espacio muestral, entonces la probabilidad de B puede

ser entendida como la probabilidad de ocurrencia del evento B en cada uno de los eventos que constituyen la partición de A definida anteriormente, y el cálculo de $\mathbb{P}(B)$ es la suma de cada probabilidad obtenida en cada partición de acuerdo con la propiedad 1.9. Este resultado se conoce como *teorema de la probabilidad total*.

Otro resultado de interés es la probabilidad de ocurrencia de alguno de los eventos de la partición condicionado en la ocurrencia de B (por ejemplo, la partición puede ser *tres métodos de enseñanza* y el evento B, *alcanzar un porcentaje aceptable de los resultados de aprendizaje por parte de los estudiantes*. Entonces la pregunta de interés es *¿cuál es la probabilidad de que al aplicar el método 1 los estudiantes alcancen un resultado de aprendizaje adecuado?*). Esta situación se conoce como *teorema de Bayes*.

Propiedad **1.15** $A_1, A_2, \cdots, A_n$ definen una partición del espacio muestral Ω. Sea B un evento cualquiera, entonces:

(a) $\mathbb{P}(B) = \sum_{i=1}^{n} \mathbb{P}(B|A_i) \cdot \mathbb{P}(A_i)$ (teorema de probabilidad total).

(b) $\mathbb{P}(A_i|B) = \dfrac{\mathbb{P}(B|A_i) \cdot \mathbb{P}(A_i)}{\mathbb{P}(B)}$, $i = 1, 2, \cdots, n$ (teorema de Bayes).

Demostración: (a) como los eventos A_i constituyen una partición del espacio muestral Ω y $B \subset \Omega$, entonces el evento B se puede expresar como la unión de este con cada uno de los eventos de la partición:

$$B = \bigcup_{i=1}^{n} (A_i \cap B) \Rightarrow \mathbb{P}(B) = \mathbb{P}\left(\bigcup_{i=1}^{n} (A_i \cap B)\right).$$

Como $(A_i \cap B) \cap (A_j \cap B) = \phi$ para $i \neq j$ y $\mathbb{P}(A_i \cap B) = \mathbb{P}(A_i) \cdot \mathbb{P}(A_i|B)$, entonces:

$$\mathbb{P}(B) = \sum_{i=1}^{n} \mathbb{P}(A_i \cap B) = \sum_{i=1}^{n} \mathbb{P}(A_i)\mathbb{P}(B|A_i).$$

(b) Aquí se utiliza la propiedad 1.9 y la propiedad 1.15(a). En efecto:

$$\mathbb{P}(A_i|B) = \frac{\mathbb{P}(A_i \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A_i) \cdot \mathbb{P}(B|A_i)}{\sum_{j=1}^{n} \mathbb{P}(A_j) \cdot \mathbb{P}(B|A_j)}.$$

Ejemplo **1.20** En una fábrica existen 3 máquinas que producen el 100 % de los artículos que se elaboran. La máquina M_1 produce el 20 %; la máquina M_2, el 30 %, y la máquina M_3, el 50 % restante. Por otro lado, se sabe que el 1 %, el 2 % y el 3 % de los artículos fabricados por las máquinas M_1, M_2 y M_3, respectivamente, son defectuosos. Si se selecciona un artículo de la producción total, determinar la probabilidad de que el artículo seleccionado provenga de la máquina M_1 si se sabe que este es defectuoso.

Solución: Se definen algunos eventos de interés para el problema. A_i: *el artículo seleccionado proviene de la máquina* M_i, $i = 1, 2, 3$, y B: *el artículo seleccionado es defectuoso.*

$$\begin{aligned}\mathbb{P}(B) = &\sum_{j=1}^{3} \mathbb{P}(B|A_j) \cdot \mathbb{P}(A_j) \\ =&\mathbb{P}(B|A_1) \cdot \mathbb{P}(A_1) + \mathbb{P}(B|A_2) \cdot \mathbb{P}(A_2) + \mathbb{P}(B|A_3) \cdot \mathbb{P}(A_3) \\ =&0{,}01 \cdot 0{,}20 + 0{,}02 \cdot 0{,}30 + 0{,}03 \cdot 0{,}50 = 0{,}023.\end{aligned}$$

Nótese que el evento de interés, para responder a lo que se pide determinar, es $A_1|B$. Luego:

$$\mathbb{P}(A_1|B) = \frac{\mathbb{P}(B|A_1) \cdot \mathbb{P}(A_1)}{\sum_{j=1}^{3} \mathbb{P}(B|A_j) \cdot \mathbb{P}(A_j)} = \frac{0{,}01 \cdot 0{,}20}{0{,}023} = 0{,}087.$$

Entonces, al observar un artículo defectuoso, la probabilidad de que provenga de la máquina M_1 es de 8,7 %. □

Observación **1.9** En el teorema de Bayes, $\mathbb{P}(A_i)$ es la probabilidad *a priori* de ocurrencia del evento A_i, y $\mathbb{P}(A_i|B)$ es la probabilidad *a posteriori* de ocurrencia del evento A_i con la información que aporta la ocurrencia del evento B.

1.5. Ejercicios resueltos

1. Sean A, B y C tres eventos en un espacio de probabilidad. Expresar, utilizando la notación de conjuntos, los siguientes eventos en términos de A, B y C.

   (a) Solo ocurre A.

(b) Ocurren A y B, pero no ocurre C.

(c) Los tres eventos ocurren.

(d) Exactamente uno de los tres eventos ocurre.

(e) Ninguno de los tres eventos ocurre.

(f) Por lo menos dos de los tres eventos ocurren.

Solución:

(a) Si solo ocurre A, entonces no ocurre ni B ni C, es decir, $B^C \cap C^C$. De esta manera, el evento *solo ocurre* A se denota con: $\left(A \cap B^C \cap C^C\right)$ y se lee como *ocurre* A *y no ocurre* B *y no ocurre* C.

(b) Que no ocurra B equivale a su complemento: B^C. *Ocurren* A *y* B, *pero no ocurre* C se denota como $\left(A \cap B \cap C^C\right)$.

(c) Este evento se puede reescribir como *ocurre* A *y ocurre* B *y también ocurre* C, lo que en notación conjuntista se denota con $(A \cap B \cap C)$.

(d) Exactamente uno de los tres eventos ocurre se interpreta como *ocurre* A*, pero no ocurre ni* B *ni* C*; u ocurre* B*, pero no ocurre ni* A *ni* C*; u ocurre* C*, pero no ocurre* A *ni* B. Se denota con:

$$\left(A \cap B^C \cap C^C\right) \cup \left(B \cap A^C \cap B^C\right) \cup \left(C \cap A^C \cap B^C\right).$$

(e) Ninguno de los tres eventos ocurre se interpreta como que *no ocurre ni* A *ni* B *y tampoco ocurre* C, esto es, $\left(A^C \cap B^C \cap C^C\right)$.

(f) Por lo menos dos de los tres eventos ocurren implica una disyunción entre *ocurren exactamente dos eventos* y *ocurren los tres a la vez*. Ocurren exactamente dos de los tres eventos es:

$$\left(A \cap B \cap C^C\right) \cup \left(A \cap C \cap B^C\right) \cup \left(B \cap C \cap A^C\right).$$

Mientras que el evento *ocurren los tres a la vez* se denota con $A \cap B \cap C$. Por lo tanto, el evento de interés es la unión de estos dos eventos.

$$\left(A \cap B \cap C^C\right) \cup \left(A \cap C \cap B^C\right) \cup \left(B \cap C \cap A^C\right) \cup (A \cap B \cap C).$$

2. Se extraen 4 cartas de una baraja con 52 cartas. ¿Cuál es la probabilidad de que 2 sean negras y 2 sean rojas?

   Solución: Teniendo en cuenta que el orden en que van apareciendo los colores es irrelevante y, además, que de las 52 cartas hay 26 rojas y 26 negras, el espacio muestral es:

$$\begin{aligned}\Omega = &\{(x_1, x_2, x_3, x_4) | x_i = 52 - (i-1) \text{ pos.}, i = 1, 2, 3, 4\} \\ \Rightarrow \#(\Omega) = &52 \cdot 51 \cdot 50 \cdot 49.\end{aligned}$$

   Debido a que el orden es irrelevante, es posible definir el evento de interés A: *las primeras 2 cartas son rojas y las siguientes 2 cartas son negras*. Entonces:

$$\begin{aligned}A = &\{(x_1, x_2, x_3, x_4) | x_1 = 26 \text{ pos. }, x_2 = 25 \text{ pos.} \\ &x_3 = 26 \text{ pos.}, x_4 = 25 \text{ pos.}\} \\ \Rightarrow &\#(A) = 26 \cdot 25 \cdot 26 \cdot 25.\end{aligned}$$

   Luego, la probabilidad de A es:

$$\mathbb{P}(A) = \frac{\#(A)}{\#(\Omega)} = \frac{26 \cdot 25 \cdot 26 \cdot 25}{52 \cdot 51 \cdot 50 \cdot 49} = \frac{325}{833}.$$

3. ¿Cuál es la probabilidad de que los cumpleaños de 12 personas sean en meses diferentes?

   Solución: Sea i el número de mes, siendo enero el primero, febrero el segundo, etc., el espacio muestral es:

$$\Omega = \{x_1, \cdots, x_{12} | x_i = 12 \text{ pos.}\} \Rightarrow \#(\Omega) = 12 \cdots 12 = 12^{12}.$$

   El evento de interés puede definirse como A: *cada una de las 12 personas tiene cumpleaños en un mes diferente al resto*:

$$\begin{aligned}A = &\{(x_1, \cdots, x_{12} | x_i = 12 - (i-1) \text{ pos., } i = 1, \cdots, 12)\} \\ \Rightarrow &\#(A) = 12 \cdot 11 \cdots 1 = 12!\end{aligned}$$

   Por lo tanto, la probabilidad de que ocurra A es:

$$\mathbb{P}(A) = \frac{12!}{12^{12}} = \frac{1925}{35\,831\,808} = 5{,}4 \cdot 10^{-5}.$$

4. Una caja contiene 40 tornillos en buen estado y 10 tornillos defectuosos. Se selecciona una muestra de 5 tornillos. Calcular las probabilidades de ocurrencia de los siguientes eventos.

   (a) Ningún tornillo de la muestra es defectuoso.

   (b) Ninguno, uno o dos tornillos de la muestra son defectuosos.

   (c) La muestra contiene, por lo menos, un tornillo en buen estado.

   Solución: Este problema corresponde al principio de la urna con selección sin reemplazo. Luego, el espacio muestral es $N = 50$ que corresponde a la cantidad de tornillos, en el cual se pueden identificar $M_I = 40$ elementos del tipo I (tornillos en buen estado) y $M_{II} = 10$ del tipo II (defectuosos). Se extrae una muestra de $n = 5$ elementos (tornillos). Entonces, si se define el evento V_k: *existen exactamente k tornillos defectuosos*, por lo que se tiene el evento V_k^c: *existen exactamente k tornillos en buen estado.*

   (a) El evento de interés *ningún tornillo es defectuoso* es equivalente a V_0: *hay exactamente 0 tornillos defectuosos.* Luego, la probabilidad de V_0 queda determinada por:

   $$\mathbb{P}(V_0) = \frac{\binom{40}{5}\binom{10}{0}}{\binom{50}{5}} = \frac{82\,251}{264\,845} \approx 0{,}310$$

   Entonces, la probabilidad de no encontrar tornillos defectuosos es de 31 %.

   (b) Según el problema se pueden identificar tres eventos: V_0, V_1 y V_2. Ya que V_0 fue obtenido en el punto (a), de manera análoga se puede obtener la probabilidad de los eventos V_1 y V_2. En efecto:

   $$\mathbb{P}(V_1) = \frac{\binom{40}{4}\binom{10}{1}}{\binom{50}{5}}$$

   $$\mathbb{P}(V_2) = \frac{\binom{40}{3}\binom{10}{2}}{\binom{50}{5}}.$$

Como los 3 eventos son independientes, la probabilidad de queda determinada por la suma directa. Así:

$$\mathbb{P}(V_0 \cup V_1 \cup V_2) = \frac{\binom{40}{5}\binom{10}{0}}{\binom{50}{5}} + \frac{\binom{40}{4}\binom{10}{1}}{\binom{50}{5}} + \frac{\binom{40}{3}\binom{10}{2}}{\binom{50}{5}}$$

$$\mathbb{P}(V_0 \cup V_1 \cup V_2) = \frac{504\,127}{529\,690} = 0{,}952.$$

Entonces, la probabilidad de tener como máximo 2 tornillos defectuosos es de 95 %.

(c) El problema se puede resolver de dos formas diferentes:

Forma 1: Sea $B = V_1^c \cup V_2^c \cup V_3^c \cup V_4^c \cup V_5^c$ el evento de interés, donde V_i, $i = 1, 2, 3, 4, 5$ son eventos excluyentes, por lo que la probabilidad de ocurrencia de B se obtiene:

$$\begin{aligned}\mathbb{P}(B) =& \frac{\binom{40}{1}\binom{10}{4}}{\binom{50}{5}} + \frac{\binom{40}{2}\binom{10}{3}}{\binom{50}{5}} + \frac{\binom{40}{3}\binom{10}{4}}{\binom{50}{5}} + \frac{\binom{40}{4}\binom{10}{1}}{\binom{50}{5}} + \frac{\binom{40}{5}\binom{10}{0}}{\binom{50}{5}} \\ =& \frac{75661}{75660} = 0{,}999.\end{aligned}$$

Forma 2: Esta forma es más directa; se usa el complemento del evento de interés. El complemento del evento *por lo menos un tornillo está en buen estado* es *ninguno de los tornillos está en buen estado* y este evento es equivalente con *los 5 tornillos son defectuosos.*

$$\begin{aligned}\mathbb{P}(V_1^c \cup V_2^c \cup V_3^c \cup V_4^c \cup V_5^c) &= 1 - \mathbb{P}(V_5) \\ &= \frac{75\,661}{75\,660} = 0{,}999.\end{aligned}$$

5. Se distribuyen, aleatoriamente, 12 monedas de $\$100$ entre 5 niños llamados Abel, Bolzano, Cauchy, Dirichlet y Euclides. Calcular la probabilidad de que Abel reciba exactamente 3 monedas.

Solución: El espacio muestral es el conjunto formado por todas las diferentes maneras de repartir las 12 monedas entre los 5 niños, considerando que puede haber al menos 1 niño que no reciba monedas.

Se define el suceso A_i como el número de monedas que recibe el i-ésimo niño y cuyos valores pueden ser un entero entre 0 y 12, es decir, $A_i = \{0, 1, 2, 3, 4, ..., 12\}$, y considerando a Abel como el primer niño, a Bolzano el segundo, a Cauchy el tercero, a Dirichlet el cuarto y a Euclides el quinto (ejemplo: si $A_2 = 4$, quiere decir que el segundo niño, Bolzano, recibe 4 monedas).

El espacio muestral se denota con:

$$\Omega = \{(x_1, x_2, x_3, x_4, x_5) | x_1 + x_2 + x_3 + x_4 + x_5 = 12\}.$$

Así, se tiene que la cardinalidad del espacio muestral es:

$$\#(\Omega) = \binom{12 + 5 - 1}{5 - 1} = 1820.$$

El cálculo de los casos favorables, teniendo en cuenta que Abel debe recibir exactamente 3 monedas, por lo que se requiere repartir las restantes 9 monedas entre los otros 4 amigos, se puede hacer de $\binom{9 + 4 - 1}{4 - 1} = 220$ maneras diferentes.
Por lo tanto, la probabilidad buscada es:

$$\mathbb{P}(A_1 = 3) = \frac{220}{1820} = 0{,}121.$$

6. Un contador tiene sobre su mesa 2 grupos de 20 planillas cada uno. En el primer grupo existen 2 planillas con errores de cálculo y en el segundo hay 3. Una corriente de aire hace que las planillas caigan de la mesa y, al recogerlas, una del primer grupo se mezcla con las del segundo grupo. ¿Cuál es la probabilidad de que, al revisar una planilla del segundo grupo, el contador encuentre un error?

   Solución: Este es un de problema de cálculo de probabilidad que involucra condicionalidad. Primero, se calcula la probabilidad de que la planilla que pasó del grupo 1 al grupo 2 contenga un error de cálculo. Para esto, se definen los siguientes eventos:

   E_1: la planilla que pasa del primer grupo al segundo contiene un error de cálculo. De esta manera, se tiene que:

$$\mathbb{P}(E_1) = \frac{2}{20} = \frac{1}{10} \Rightarrow \mathbb{P}(E_1^c) = \frac{18}{20} = \frac{9}{10}.$$

Se define el evento E: *la planilla seleccionada por el contador, del grupo 2, contiene un error de cálculo*. Luego, se analizan las probabilidades en el segundo grupo, sabiendo que hay una planilla más que al inicio, por lo que el total de planillas en el grupo 2 es 21. Si la planilla que viene del grupo 1 contiene errores de cálculo, el grupo 2 ahora tiene 4 planillas con errores y 17 sin errores; si la planilla que viene del grupo 1 no contiene errores, entonces en el grupo 2 hay 3 planillas con errores y 18 sin errores. Esto último, permite calcular las siguientes probabilidades condicionales:

$$\mathbb{P}(E|E_1) = \frac{4}{21} \quad \text{y} \quad \mathbb{P}(E|E_1^c) = \frac{18}{20}.$$

Por último, para determinar la probabilidad pedida, se usa el teorema de probabilidad total. Así:

$$\begin{aligned}\mathbb{P}(E) =& \mathbb{P}(E|E_1) \cdot \mathbb{P}(E_1) + \mathbb{P}(E|E_1^c) \cdot \mathbb{P}(E_1^c) \\ =& \frac{4}{21} \cdot \frac{1}{10} + \frac{3}{21} \cdot \frac{9}{10} = \frac{31}{210}.\end{aligned}$$

7. En un armario hay 10 pares diferentes de zapatos. Se escogen al azar y sin reemplazo 4 zapatos. Hallar la probabilidad de que al menos aparezca un par correcto entre los 4 escogidos.

   Solución: Sea el evento A: *al menos aparece un par entre los 4 escogidos*. Para determinar $\mathbb{P}(A)$ es posible usar el complemento del evento, A^c: *entre los 4 zapatos escogidos no aparece ningún par*.

   Para el espacio muestral Ω: *número de maneras diferentes de escoger 4 zapatos entre los 20*, sin reposición, se tiene que su cardinalidad es:

$$\#(\Omega) = 20 \cdot 19 \cdot 18 \cdot 17.$$

   Para el cálculo de $\#(A)$, reflexionar que entre los 10 pares hay que elegir 4 distintos representados por un zapato y este puede ser izquierdo o derecho. Así, se tiene que:

$$\mathbb{P}(A^c) = \frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot 2 \cdot 2 \cdot 2 \cdot 2}{20 \cdot 19 \cdot 18 \cdot 17} = \frac{20 \cdot 18 \cdot 16 \cdot 14}{20 \cdot 19 \cdot 18 \cdot 17} = \frac{224}{323}.$$

   Finalmente, la probabilidad que se buscaba es:

$$\mathbb{P}(A) = 1 - \mathbb{P}(A^c) = 1 - \frac{224}{323} = \frac{99}{323} = 0{,}3065.$$

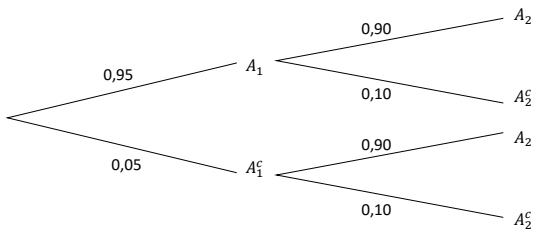
8. Una empresa produce piezas de dos máquinas I y II, que pueden presentar desajustes con probabilidad de ocurrencia 0,05 y 0,10 respectivamente. En el inicio del día de operar una prueba es realizada y, en caso que la máquina estuviera desajustada, ella quedará sin operar ese día pasando por revisión técnica. Para cumplir el nivel mínimo de producción, por lo menos una de las máquinas debe operar. Determinar la probabilidad de que la empresa cumpla con sus metas de producción.

   Solución: Sea A_i el evento: *la máquina i está funcionando*, con $i = 1, 2$. Luego, para obtener el mínimo de producción diaria, se necesita tener por lo menos una máquina operando. Esto corresponde a la ocurrencia del evento: $B = (A_1 \cap A_2) \cup (A_1 \cap A_2^c) \cup (A_1^c \cap A_2)$. Notar que estos tres eventos son disjuntos, con esto la probabilidad corresponde a:

$$\mathbb{P}(B) = \mathbb{P}(A_1 \cap A_2) + \mathbb{P}(A_1 \cap A_2^c) + \mathbb{P}(A_1^c \cap A_2). \tag{1.6}$$

   Así, quedan por determinar las probabilidades de las intersecciones. Una forma de abordar este tipo de problemas es apoyarse en una representación gráfica, el llamado *árbol de probabilidades*. Cada uno de los caminos del árbol indica una posible ocurrencia, la cual es posible completar considerando la información disponible, esto es, $P(A_1) = 0{,}95$ y $P(A_2) = 0{,}90$ (véase figura 1.4).

Figura 1.4: Árbol de probabilidades



Fuente: elaboración propia.

Al completar el árbol, obsérvese que se asume independencia entre los eventos A_1 y A_2, pues se cree que la eventual falta de ajuste en una máquina no interfiere en el comportamiento de la otra. Nótese también que en caso de independencia, la segunda rama del árbol no es afectada por la ocurrencia de los eventos que aparecen en la primera rama. Por lo tanto, por la definición de independencia, se tiene que $\mathbb{P}(A_2|A_1) = \mathbb{P}(A_2) = 0{,}90$.

Luego, la probabilidad de ocurrencia de A_1 y A_2 está dada por el producto de las dos ramas que genera este evento, es decir: $\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_2|A_1) \cdot \mathbb{P}(A_2) = 0{,}95 \cdot 0{,}90 = 0{,}855$. Usando este análisis es posible obtener las probabilidades de los otros dos eventos de la expresión 1.6, con lo cual se tiene que la probabilidad de mantener el nivel mínimo diario de producción es de 0,995.

9. En un cuarto medio, $\frac{1}{3}$ son hombres. El 20 % de los hombres está en el electivo de Termodinámica, mientras que de las mujeres solo el 10 % está en ese electivo. Si se selecciona un estudiante al azar, calcular la probabilidad de que:

   (a) Estudie Termodinámica.

   (b) Sabiendo que está en el curso Termodinámica, sea mujer.

   Solución: Se definen los eventos M: *el estudiante seleccionado al azar es mujer* y T: *el estudiante seleccionado al azar estudia Termodinámica.*

   (a) Para la primera problemática, se puede usar el teorema de probabilidad total, pues la persona que estudia Termodinámica puede ser hombre o mujer.

$$\begin{aligned}\mathbb{P}(T) =& \mathbb{P}(T|M) \cdot \mathbb{P}(M) + \mathbb{P}(T|M^c) \cdot \mathbb{P}(M^c) \\ =& \frac{1}{10} \cdot \frac{2}{3} + \frac{2}{10} \cdot \frac{1}{3} = \frac{2}{15}.\end{aligned}$$

   (b) Para la segunda parte, la pregunta hace referencia a una condi-

cionalidad, lo que implica el teorema de Bayes.

$$\begin{aligned}\mathbb{P}(M|T) =& \frac{\mathbb{P}(T|M)\cdot\mathbb{P}(M)}{\mathbb{P}(T|M)\cdot\mathbb{P}(M)+\mathbb{P}(T|M^c)\cdot\mathbb{P}(M^c)}\\ =& \frac{\frac{1}{10}\cdot\frac{2}{3}}{\frac{1}{10}\cdot\frac{2}{3}+\frac{2}{10}\cdot\frac{1}{3}} = \frac{1}{2}.\end{aligned}$$

10. Sean 2 monedas, una honesta y otra que resulta cara en cada lanzamiento con probabilidad 0,6. Se escoje una moneda al azar y, después de lanzada 4 veces, se observa cara 3 veces. ¿Cuál es la probabilidad de que la moneda escogida haya sido la honesta?

    Solución: Según el problema se pide calcular la probabilidad de que la moneda elegida haya sido la honesta *sabiendo* que salieron 3 caras en los 4 lanzamientos, esto es, teorema de Bayes. Se define el suceso C: *en cuatro lanzamientos aparecen 3 caras* y H: *la moneda escogida es honesta*. Usando el triángulo de Pascal y el principio multiplicativo, se obtiene que si se lanza la moneda honesta se tendrá que:

    $$\mathbb{P}(C|H) = \binom{4}{3}\cdot(0{,}5)^3\cdot(0{,}5)\,,$$

    pues la probabilidad de salir cara o sello en una moneda honesta es 0,5. En cambio, si lanzamos la moneda cargada:

    $$\mathbb{P}(C|H^c) = \binom{4}{3}\cdot(0{,}6)^3\cdot(0{,}4)\,.$$

    Con estas probabilidades claras, más el hecho de que la probabilidad de elegir la moneda honesta es igual a la de elegir la moneda cargada (0,5), la respuesta al problema se basa en el teorema de Bayes:

    $$\begin{aligned}\mathbb{P}(H|C) =& \frac{\mathbb{P}(C|H)\cdot\mathbb{P}(H)}{\mathbb{P}(C|H)\cdot\mathbb{P}(H)+\mathbb{P}(C|H^c)\cdot\mathbb{P}(H^c)}\\ =& \frac{\binom{4}{3}\cdot(0{,}5)^4}{\binom{4}{3}\cdot(0{,}5)^4+\binom{4}{3}\cdot(0{,}6)^3\cdot(0{,}4)} = 0{,}4197.\end{aligned}$$

11. Se tienen 3 cajas. La caja I contiene 4 bolas blancas y 2 negras, la caja II contiene 3 blancas y 1 negra, y la caja III contiene 1 bola blanca y 2 negras.

(a) Se extrae una bola de cada caja. Determinar la probabilidad de que todas las bolas sean blancas.

(b) Se selecciona una caja y de ella una bola (en ambos casos al azar). Determinar la probabilidad de que la bola extraída sea blanca.

(c) Calcular en (b) la probabilidad de que la primera caja haya sido escogida, dado que la bola extraída es blanca.

Solución:

(a) Extraer una bola de una caja representa ensayos independientes, por lo que se puede usar el principio multiplicativo. Sea el suceso B_i *la bola de la i-ésima caja es blanca*, luego la probabilidad de ocurrencia de los 3 eventos es:

$$\begin{aligned}\mathbb{P}\left(B_1 \cap B_2 \cap B_3\right) =& \mathbb{P}\left(B_1\right) \cdot \mathbb{P}\left(B_2\right) \cdot \mathbb{P}\left(B_3\right) \\ =& \frac{4}{6} \cdot \frac{3}{4} \cdot \frac{1}{3} = \frac{1}{6}.\end{aligned}$$

(b) Que la bola sea blanca depende de qué caja fue seleccionada, esto es, teorema de probabilidad total. Se definen los sucesos B: *la bola seleccionada es blanca* y C_i: *la caja seleccionada es la i*, donde $i = \{1, 2, 3\}$. Además, para cada i, $\mathbb{P}(C_i) = \dfrac{1}{3}$.
Con el resultado anterior, la probabilidad de B es:

$$\begin{aligned}\mathbb{P}(B) =& \mathbb{P}(B|C_1) \cdot \mathbb{P}(C_1) + \mathbb{P}(B|C_2) \cdot \mathbb{P}(C_2) + \mathbb{P}(B|C_3) \cdot \mathbb{P}(C_3) \\ =& \frac{4}{6} \cdot \frac{1}{3} + \frac{3}{4} \cdot \frac{1}{3} + \frac{1}{3} \cdot \frac{1}{3} = \frac{7}{12}.\end{aligned}$$

(c) Utilizando lo realizado en (b), la solución es:

$$\mathbb{P}(C_1|B) = \frac{\mathbb{P}(B|C_1) \cdot \mathbb{P}(C_1)}{\sum_{i=1}^{3} \mathbb{P}(B|C_i) \cdot \mathbb{P}(C_i)} = \frac{\frac{4}{6} \cdot \frac{1}{3}}{\frac{7}{12}} = \frac{8}{21}.$$

12. En un restaurante, 3 cocineros A, B y C preparan un tipo especial de torta, y con probabilidades respectivas de 2 %, 3 % y 5 % su masa de torta no crece (no completa satisfactoriamente el proceso de fermentación de la masa). Se sabe que A prepara el 50 % de las tortas,

B el 30 % y C el 20 % restante. Si se observa una masa de torta que no creció, ¿cuál es la probabilidad de que haya sido preparada por el cocinero A?

Solución: Se quiere calcular la probabilidad de que el cocinero A haya preparado la masa, sabiendo que la masa no creció. Esto indica que se debe trabajar con el teorema de Bayes. Definiendo los sucesos:

C_1: La masa fue preparada por el cocinero A.
C_2: La masa fue preparada por el cocinero B.
C_3: La masa fue preparada por el cocinero C.
N: La masa no creció.

Usando el teorema de Bayes, se tiene que:

$$\begin{aligned}\mathbb{P}(C_1|N) &= \frac{\mathbb{P}(N|C_1)\cdot\mathbb{P}(C_1)}{\sum_{i=1}^{3}\mathbb{P}(N|C_i)\cdot\mathbb{P}(C_i)} \\ &= \frac{\frac{2}{100}\cdot\frac{1}{2}}{\frac{2}{100}\cdot\frac{1}{2}+\frac{3}{100}\cdot\frac{3}{10}+\frac{5}{100}\cdot\frac{1}{5}} = \frac{10}{29}.\end{aligned}$$

1.6. Ejercicios propuestos

1. ¿De cuántas maneras pueden sentarse 15 personas en una fila de 5 sillas disponibles?

2. ¿Cuántos números se pueden formar al arreglar los dígitos del número 4 130 131 (excluyendo los que comienzan con 0)?

3. En un grupo de 10 amigos, ¿cuántas distribuciones de sus fechas de cumpleaños pueden darse al año?

4. Un estudiante debe elegir 7 de las 10 preguntas de un examen. ¿De cuántas maneras puede elegir si (a) no hay repeticiones y (b) las 4 primeras son obligatorias?

5. Entre 5 matemáticos y 7 físicos hay que constituir una comisión de 2 matemáticos y 3 físicos. ¿De cuántas maneras podría formarse si:

   (a) Todos son elegibles.

   (b) 1 físico en particular ha de estar en la comisión.

(c) 2 matemáticos concretos no pueden estar juntos.

6. Hay que poner a 5 hombres y 4 mujeres en una fila de modo que las mujeres ocupen los lugares pares. ¿De cuántas maneras puede hacerse?

7. ¿Cuántos tipos de collares (circulares) pueden formarse con 4 perlas azules y 5 perlas rojas usando todas las perlas disponibles (es decir, modelos de combinaciones de colores que no se repiten)?

8. Sean los eventos A, B y C de un mismo espacio muestral. Probar:

   (a) Sea $C = (A \cap B^c)^c \cap (A^c \cap B)^c$:
   $\mathbb{P}(C) = 1 - [\mathbb{P}(A) - \mathbb{P}(A \cap B)] - [\mathbb{P}(B) - \mathbb{P}(A \cap B)]$.

   (b) $\mathbb{P}[(A \cap B^c) \cup (A^c \cap B)] = \mathbb{P}(A) + \mathbb{P}(B) - 2\,\mathbb{P}(A \cap B)$.

   (c) $1 - \mathbb{P}(A^c) - \mathbb{P}(B^c) \leqslant \mathbb{P}(A \cap B)$ (desigualdad de Boole).

   (d) $|\mathbb{P}(A) - \mathbb{P}(B)| \leqslant \mathbb{P}(A \Delta B)$ (desigualdad triangular), con $A \Delta B = (A \cap B^c) \cup (B \cap A^c)$.

   (e) $\mathbb{P}(A \cap B) \leqslant \mathbb{P}(A) \leqslant \mathbb{P}(A \cup B) \leqslant \mathbb{P}(A) + \mathbb{P}(B)$.

   (f) $\mathbb{P}[(A \cap B^c) \mid C] = \mathbb{P}(A \mid C) - \mathbb{P}[(A \cap B) \mid C]$.

   (g) $\mathbb{P}(A) < \mathbb{P}(A \cup B) \Rightarrow \mathbb{P}(A \cap B) < \mathbb{P}(B)$.

9. Probar que la función dada es una función de probabilidad:

   (a) $\mathbb{P}(A) = \sum_{x \in A} p(1-p)^{x-1}$, $0 < p < 1$, con $A \subset \Omega = \{1, 2, \cdots\}$.

   (b) $\mathbb{P}(A) = \dfrac{\#(A)}{\#(\Omega)}$, $A \subset \Omega$.

   (c) $\mathbb{P}(A) = \displaystyle\int_{x \in A} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}x^2\right\} dx$, con $\Omega = \mathbb{R}$.

10. Mostrar que la probabilidad de que ocurra exactamente uno de los eventos A o B es igual a $\mathbb{P}(A) + \mathbb{P}(B) - 2\,\mathbb{P}(A \cap B)$.

11. Un experimento aleatorio consiste en ubicar r bolitas numeradas del 1 al r en n celdas distintas numeradas del 1 al n. Este experimento permite ubicar incluso todas las bolitas en una misma celda.

    (a) Describir el espacio muestral y determinar su cardinalidad.

(b) Determinar la probabilidad de que cada bolita sea ubicada en solo una celda.

(c) Determinar la probabilidad de que una celda específica tenga k bolitas (con $k \leqslant n$).

12. Ocho amigos se fueron de vacaciones a Europa. Ellos tomaron vuelos distintos y se pusieron de acuerdo para encontrarse en el Grand Hotel de París. Pero resulta que en esa ciudad hay ocho hoteles con ese nombre. ¿Cuál es la probabilidad de que todos ellos escojan hoteles diferentes?

13. La noche anterior a la prueba de la asignatura de Probabilidades se reúnen 4 alumnos para estudiar (esta situación no es aconsejable). Encontrar la probabilidad de que por lo menos 2 estudiantes hayan nacido el mismo mes.

14. Se ha lanzado un dado legal 4 veces. ¿Cuál es la probabilidad de que se obtenga el número 3 en 2 de las 4 tiradas?

15. Una moneda es lanzada 3 veces. Considerar el evento definido por A_i: *sale cara en el i-ésimo lanzamiento*, para $i = 1, 2, 3$. Determinar los siguientes eventos:

(a) $A_1^c \cap A_2$.

(b) $A_1^c \cup A_2$.

(c) $(A_1^c \cap A_2^c)^c$.

(d) $A_1 \cap (A_2 \cup A_3)$.

16. Considerar el espacio muestral formado por las 24 permutaciones de los números 1, 2, 3 y 4, todos equiprobables. Se define el evento A_i: *el número i aparece en el i-ésimo lugar*, $i = 1, 2, 3, 4$. Calcular $\mathbb{P}(A_1 \cap A_2 \cap A_3 \cap A_4)$.

17. A, B y C son eventos de un mismo espacio muestral, tal que $\mathbb{P}(B) = 0{,}5$, $\mathbb{P}(C) = 0{,}3$, $\mathbb{P}(B|C) = 0{,}4$ y $\mathbb{P}(A|(B \cap C)) = 0{,}5$. Calcular $\mathbb{P}(A \cap B \cap C)$.

18. Sean 2 eventos A y B de un mismo espacio muestral y tales que $\mathbb{P}(A) = \frac{2}{3}$ y $\mathbb{P}(B) = \frac{4}{9}$. Mostrar los siguientes resultados:

    (a) $\mathbb{P}(A \cup B) \geqslant \frac{2}{3}$.

    (b) $\frac{2}{9} \leqslant \mathbb{P}(A \cap B^c) \leqslant \frac{5}{9}$.

    (c) $\frac{1}{9} \leqslant \mathbb{P}(A \cap B) \leqslant \frac{4}{9}$.

19. En una fiesta de cumpleaños, donde hay n niños se reparten d dulces.

    (a) Obtener el número de posibilidades para repartir los d dulces.

    (b) Calcular la probabilidad de que al repartir los d dulces, al niño de cumpleaños le corresponda al menos uno.

    (c) Si la cantidad de niños es igual a la cantidad de dulces, calcular nuevamente (b).

20. Un vendedor de autos nuevos ha comprobado que los clientes solicitan en especial algunos de los siguientes extras: transmisión automática (A), neumáticos pantaneros (B) y/o radio (C). Si el 70 % de los clientes solicita A, el 75 % solicita B, el 80 % solicita C, el 80 % requiere A o B, el 85 % requiere A o C, el 90 % requiere B o C, y el 95 % requiere A o B o C. Calcular la probabilidad de que:

    (a) El próximo cliente solicite a lo menos una de las tres opciones.

    (b) El próximo cliente solicite solo una radio.

    (c) El próximo cliente solicite solo una de las tres opciones

21. Un detector de mentiras muestra una lectura positiva (indica una mentira) el 10 % de las veces cuando una persona está diciendo la verdad y el 95 % de las veces cuando la persona está mintiendo. Si dos personas son sospechosas de un mismo crimen y una de ellas es culpable, determinar la probabilidad de que el detector muestre:

    (a) Una lectura positiva para ambos sospechosos.

    (b) Una lectura positiva para el culpable y negativa para el inocente.

    (c) Un resultado completamente equivocado.

    (d) Una lectura positiva para cualquiera de los sospechosos.

22. Un sistema de propulsión está formado por un motor y dos calderas. El sistema funciona cuando está operando el motor y al menos una caldera. La probabilidad de que el sistema funcione es de 0,7, la probabilidad de que funcione el motor y solo la caldera 1 es de 0,4, mientras que la probabilidad de que funcione el motor y solo la caldera 2 es de 0,5. Determinar la probabilidad de que el sistema funcione con el motor y ambas calderas.

23. Una muestra de tamaño 3 es seleccionada sin reemplazo desde una urna que contiene 6 fichas, de las cuales 4 son azules y el resto, verdes. Determinar la probabilidad de que:

    (a) La primera y la segunda fichas extraídas sean azules y la otra verde.

    (b) La muestra contenga exactamente 2 fichas azules.

24. En sus ratos libres un profesor se dedica a la pesca. Cierto día este profesor ha pescado 10 peces, pero 2 son más pequeños que el tamaño mínimo legal. Ese día no fue del todo bueno, ya que un inspector de pesca revisó su canasta. El inspector seleccionó al azar (sin reemplazo) 2 peces del total de 10. ¿Cuál es la probabilidad de que no seleccione ninguno de los pescados de tamaño ilegal?

25. En cada una de las 10 urnas $U_1, U_2, ..., U_{10}$ hay 10 fichas. La urna con número n (para $1 \leqslant n \leqslant 10$) contiene n fichas blancas y $10 - n$ negras. Si se extrae una urna al azar y de esta se saca una ficha, la cual se sabe que es blanca, ¿cuál es la probabilidad de que la ficha provenga de la urna 4?

26. Una caja A contiene 12 fichas negras y 8 fichas rojas. Otra caja idéntica B contiene 10 fichas negras y 20 rojas.

    (a) Se toma una caja al azar y se saca una ficha. Encontrar la probabilidad de que la ficha seleccionada sea negra.

    (b) Se toma una caja al azar y se saca una ficha que resulta ser negra. ¿Cuál es la probabilidad de que provenga de la caja B?

27. Una caja contiene n fichas de las cuales m son blancas. Se extraen 2 fichas y se define el evento A_i: *la i-ésima ficha seleccionada es blanca*, $i = 1, 2$.

    (a) Obtener $\mathbb{P}(A_2|A_1)/\mathbb{P}(A_2)$.

    (b) Verificar que la expresión obtenida en (a) tiende a 1 cuando n y m tienden a infinito.

28. Una compañía de seguros divide a las personas en dos clases: quienes son propensos a accidentes y quienes no lo son. Sus estadísticas muestran que una persona propensa a accidentes tendrá, en no más de un año, un accidente con probabilidad 0,4, mientras que esta probabilidad decrece a 0,2 para personas no propensas a accidentes. Si pensamos que un 30 % de la población es propensa a accidentes:

    (a) ¿Cuál es la probabilidad de que una persona que compra una nueva póliza tenga un accidente en no más de un año?

    (b) Suponer que un nuevo asegurado ha tenido un accidente a no más de un año de haber comprado su póliza. ¿Cuál es la probabilidad de que sea propenso a accidentes?

29. Un análisis para detectar un escape en una planta de energía nuclear indica que no hay escape en el 95 % de los casos cuando la planta está funcionando correctamente, y señala la falla en el 99 % de los casos cuando realmente hay un escape. De todas las plantas existentes en un país se sabe que el 1 % presenta fallas de escape.

    (a) Si se analiza una planta cualquiera y el resultado del análisis indica que hay un escape. ¿Cuál es la probabilidad de que el resultado esté errado?

    (b) ¿Cuál es la probabilidad de que en una planta elegida al azar el análisis indique que no hay un escape?

    (c) ¿Cuál es la probabilidad de error del análisis?

30. Como parte de su proceso de selección una empresa ofrece a los postulantes un curso de entrenamiento de dos semanas. Al cabo de ese periodo los postulantes son clasificados como *buenos*, *satisfactorios*

y *malos*; la experiencia indica que los postulantes se distribuyen en 25 %, 50 % y 25 %, respectivamente.

Se desea aplicar un test para reemplazar el entrenamiento, el que tiene como requisito que asegure que un postulante reprobado en el test es categorizado como *malo* en el caso de que haya realizado el curso. Con este objetivo, al iniciar un nuevo proceso de selección, a los postulantes se les aplica el test y según el resultado son calificados como *aprobados* y *reprobados*; luego de terminado el curso se obtiene que de los clasificados como buenos, satisfactorios y malos, el 80 %, 50 % y 20 %, respectivamente, habían sido calificados como aprobados en el test.

Con estos resultados, determinar la probabilidad de que un postulante aprobado en el test sea considerado malo si realizó el curso.

31. Dos máquinas de una planta elaboran el total de la producción de cierto artículo. La primera máquina elabora el 10 % de la producción total. La probabilidad de producir un artículo defectuoso con dichas máquinas es de 0,01 la primera y de 0,05 la segunda. ¿Cuál es la probabilidad de que un artículo tomado al azar de la producción de un día haya sido producido con la primera máquina, sabiendo que es defectuoso?

32. Suponer que en el 40 % de los accidentes en autopistas participa la velocidad excesiva de por lo menos uno de los conductores, y en el 30 % el consumo de bebidas alcohólicas también en al menos uno de los conductores. En el caso de que el conductor haya consumido alcohol, existe una probabilidad del 60 % de que conduzca a exceso de velocidad, mientras que en el caso contrario esta probabilidad es de apenas un 10 %. Ocurre un accidente con participación de exceso de velocidad, ¿cuál es la probabilidad de que haya habido consumo de bebidas alcohólicas?

33. Un análisis de sangre de laboratorio es efectivo en el 95 % de los casos para detectar una determinada enfermedad cuando, en efecto, está presente. Sin embargo, la prueba también arroja un resultado falso positivo para el 1 % de las personas sanas testeadas (es

decir, si se realiza una prueba a una persona sana, entonces, con una probabilidad de 0,01, el resultado de la prueba indicará que tiene la enfermedad). Si el 0,5 % de la población realmente tiene la enfermedad, ¿cuál es la probabilidad de que una persona tenga la enfermedad dado que el resultado de la prueba es positivo?

34. Supongamos que tenemos 3 cartas de idéntica forma, excepto que ambos lados de la primera carta son de color rojo, ambos lados de la segunda carta son de color negro, y un lado de la tercera carta es de color rojo y el otro lado negro. Las 3 cartas se mezclan en un sombrero y se selecciona una al azar. Si el lado superior de la carta elegida (al ponerla sobre una mesa) es de color rojo, ¿cuál es la probabilidad de que el otro lado sea de color negro?

35. En un sistema de alarma, la probabilidad de que se produzca un peligro es de 0,1. La probabilidad de que la alarma funcione sabiendo que se produce peligro es de 0,95. La probabilidad de que funcione la alarma dado que no hay peligro es de 0,03. Hallar la probabilidad de que:

    (a) No haya habido peligro sabiendo que la alarma funcionó.

    (b) Haya un peligro o la alarma no funcione.

    (c) No funcione la alarma.

Problemas adicionales

A continuación, se presentan algunos problemas que tienen el carácter de desafío y que han sido trabajados ampliamente por diversos autores. Sus soluciones están disponibles en textos de probabilidades (se puede realizar la búsqueda por el nombre del problema).

1. (Problema de los cumpleaños). En un grupo de r personas, ¿cuál es la probabilidad de que por lo menos dos de ellas estén de cumpleaños en el mismo día? ¿Cuál es el número r de personas para el que se obtiene una probabilidad de ocurrencia mayor a 0,5?

2. (Paradoja de Bertrand). Un círculo unitario, como se presenta en la figura 1.5(a), tiene inscrito un triángulo equilátero de lado igual a

$\sqrt{3}$. ¿Cuál es la probabilidad de que una cuerda de ese círculo, escogida al azar, tenga una longitud mayor que el lado de ese triángulo?

3. (Muestras con y sin reposición). En un lote de n piezas existen m defectuosas. Una muestra sin reposición de r de esas piezas, $r < n$, es seleccionada al azar y se desea saber cuál es la probabilidad de obtener k piezas defectuosas, $k \leqslant \text{mín}(r, m)$.

4. (Problema de la elección y principio de la reflexión). Suponer que dos candidatos I y II disputan una elección recibiendo, respectivamente, a y b votos. Asumiendo que $a > b$, calcular la probabilidad de que el candidato I esté siempre adelante en el conteo de votos.

5. (Problema del juego de una feria). El juego de una feria consiste en lanzar monedas de radio r sobre un tablero cuadriculado, donde el lado de cada cuadrado mide a unidades. Un jugador se hace acreedor de un premio si la moneda lanzada no toca ninguna de las líneas. ¿De qué tamaño deben ser a y r para que la probabilidad de ganar en este juego sea menor a $\frac{1}{4}$?

6. (Problema de los dos amigos). Dos amigos deciden encontrarse en cierto lugar pero olvidan la hora exacta de la cita, únicamente recuerdan que era entre las 12:00 y las 13:00 horas. Cada uno de ellos decide llegar al azar en ese lapso y esperar solo 10 minutos. ¿Cuál es la probabilidad de que se encuentren?

7. (Problema de la aguja de Buffon). Considerar un conjunto infinito de líneas horizontales paralelas sobre una superficie plana, donde la distancia entre una línea y otra es L. Se deja caer una aguja de longitud l sobre la superficie. Suponga $l \leqslant L$. Determinar la probabilidad de que la aguja cruce alguna línea.

8. (Problema de la varilla de metal). Una varilla de metal de longitud l se rompe en dos puntos distintos escogidos al azar. ¿Cuál es la probabilidad de que los tres segmentos así obtenidos formen un triángulo?

9. (Las cajas de cerillos de Banach). Una persona tiene dos cajas de cerillos, cada una de las cuales contiene n cerillos. La persona pone

una caja en cada uno de sus bolsillos (izquierdo y derecho) y cada vez que requiere un cerillo elije una caja al azar de uno de sus bolsillos y toma de ella un cerillo. ¿Cuál es la probabilidad de que en el momento en que la persona se da cuenta de que la caja del bolsillo izquierdo está vacía en la caja del bolsillo derecho haya exactamente r cerillos?

10. (La urna de Polya). En una urna se tienen a fichas blancas y b fichas azules. Un ensayo consiste en tomar una bola al azar y regresarla a la urna junto con c fichas del mismo color. Se repite este ensayo varias veces y se define el evento *obtiene una ficha blanca en la n-ésima extracción*. Determinar la probabilidad de ocurrencia de dicho evento.

11. (El tablero de Galton). Considere el tablero vertical donde se han puesto 5 filas de clavos en forma triangular (ordenados de tal manera que 3 clavos, 1 en la fila superior y 2 consecutivos de la fila inferior, representan los vértices de un triángulo equilátero). Una bola que se deja caer desde la parte superior choca contra el primer clavo y baja al clavo inferior izquierdo con una probabilidad de 0,5 o baja al clavo inferior derecho con una probabilidad de 0,5, y así sucesivamente hasta caer en alguna de las 6 urnas que se encuentran en la parte inferior (como se muestra en la figura 1.5(b)):

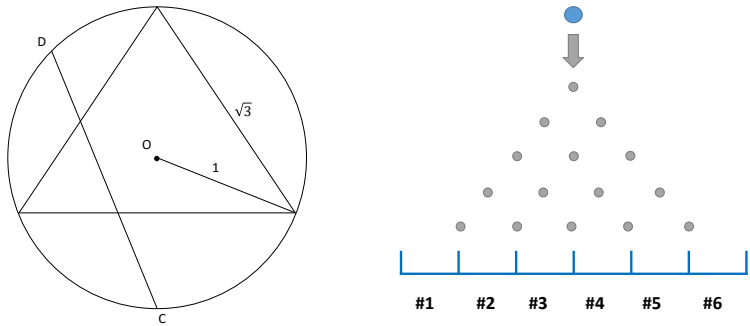
    (a) Determinar el número total de trayectorias distintas que la bola puede tomar. ¿Son igualmente probables estas trayectorias?

    (b) Determinar el número de trayectorias que llevan a cada una de las urnas.

    (c) Calcular la probabilidad de que la bola caiga en cada una de las urnas.

    (d) Resolver los tres puntos anteriores cuando se tienen n filas de clavos y por lo tanto $n + 1$ urnas y la probabilidad de que la bola caiga a la izquierda es $1 - p$ y de que caiga a la derecha es p.

Figura 1.5: Esquemas de la paradoja de Betrand y del tablero de Galton



(a) Paradoja de Bertrand

(b) El tablero de Galton

Fuente: elaboración propia.

Capítulo 2

Variables aleatorias

2.1. Variables aleatorias discretas y continuas

Considérese el contexto de la realización de un experimento aleatorio. En este capítulo se definirá una función que relaciona eventos pertenecientes al espacio muestral con valores definidos en un conjunto numérico (por ejemplo, los enteros o los reales). Esta función que atribuye valores numéricos según sea el interés respecto al fenómeno, será llamada variable aleatoria y por medio ella se podrá asignar una medida de ocurrencia a los intervalos numéricos obtenidos.

Con esto, se define una variable aleatoria como una función que va del espacio muestral Ω a los reales. Es decir:

$$X : \Omega \to \mathbb{R}$$
$$\omega \in \Omega \mapsto x = X(\omega) \in \mathbb{R}.$$

En seguida, se puede notar que, según la naturaleza del conjunto dado por el recorrido de la variable aleatoria, se dirá que es de carácter discreto o continuo, con lo cual se puede calcular la probabilidad de ocurrencia de un rango de valores definidos en el recorrido de la variable aleatoria (que pueden corresponder a eventos).

A continuación, se define formalmente el concepto de variable aleatoria y luego los elementos que permiten calcular probabilidades.

Definición **2.1** Una variable aleatoria X en un espacio de probabilidad

$(\Omega, \mathcal{F}, \mathbb{P})$ es una función de valores reales definidos en Ω, tal que:

$$\{X \leqslant x\} = \{\omega \in \Omega | X(\omega) \leqslant x\} \in \mathcal{F}; \quad \forall x \in \mathbb{R}.$$

A continuación, se presenta un ejemplo para graficar este resultado.

Ejemplo **2.1** Sea el experimento aleatorio: lanzar una moneda 3 veces y observar el resultado (c: cara, s: sello), y sea X una variable que mide *el número de caras que aparecen*. Para este experimento se pide definir: el espacio muestral Ω, los valores que asume la variable aleatoria $X(\omega_i)$ y con este resultado definir su recorrido Rec_X.

Solución: Los valores están dados respectivamente por:

$$\Omega = \{(c,c,c), (c,c,s), (c,s,c), (s,c,c), (c,s,s), (s,c,s), (s,s,c), (s,s,s)\}.$$

$X(\{ccc\}) = 3, X(\{c,c,s\}) = 2, \cdots, X(\{sss\}) = 0. \Rightarrow Rec_X = \{0,1,2,3\}.\square$

Ejemplo **2.2** Un experimento aleatorio consiste en poner en funcionamiento un componente electrónico hasta que falla, luego, se define la variable aleatoria X: *tiempo medio de falla de un componente electrónico*. Para este experimento se pide definir: el espacio muestral Ω, el recorrido Rec_X y los valores que asume la variable aleatoria $X(\omega_i)$.

Solución: Sean c_i y t_i, $i = 1, 2, \cdots, n$ los componentes medidos y los tiempos de falla observados, respectivamente, luego los elemento pedidos son, respectivamente:

$$\Omega = \{c_1, c_2, \cdots, c_n\}.$$

$X(\{c_1\}) = t_1, X(\{c_2\}) = t_2, \cdots, X(\{c_n\}) = t_n \Rightarrow Rec_X = \{t_1, t_2, \cdots, t_n\}.$
$\square$

Definición **2.2** Si la variable aleatoria X asume valores en un conjunto finito o infinito numerable es llamada *variable aleatoria discreta*. En caso de asumir valores en un intervalo de la recta real es llamada *variable aleatoria continua*.

Ejemplo **2.3** El ejemplo 2.1 es un caso en que la variable aleatoria es discreta, pues asume valores finitos. Por su parte, el ejemplo 2.2 es un caso en que la variable aleatoria es continua, pues los tiempos son medidos en un intervalo de recta real. $\square$

A continuación, el interés es obtener la probabilidad de ocurrencia de algún evento para la variable aleatoria. Por ejemplo, la probabilidad de observar menos de 2 caras al realizar el experimento aleatorio del ejemplo 2.1 o la probabilidad de obtener un tiempo de falla superior a un tiempo t_k, con $t_k < \max t_i$, al realizar el experimento aleatorio del ejemplo 2.2. Para tal objeto, se define la que se llamará función de distribución acumulada (o simplemente función de distribución).

2.2. Función de distribución

Para obtener el cálculo de probabilidad de $\{X \leqslant x\}$, es decir, $\mathbb{P}(\{X \leqslant x\})$, se define la función de distribución de la variable aleatoria X.

Definición **2.3** La función de distribución para la variable aleatoria X es una función $F = F_X$ definida por:

$$F(x) = \mathbb{P}(X \leqslant x) = \mathbb{P}(\{\omega \in \Omega | X(\omega) \leqslant x\}),\ x \in \mathbb{R}.$$

Propiedad **2.1** Propiedades de la función de distribución:

(P1) $\lim\limits_{x \to -\infty} F_X = 0$ y $\lim\limits_{x \to +\infty} F_X = 1$.

(P2) F_X es no decreciente: $x_1 < x_2 \Rightarrow F(x_1) \leqslant F(x_2)\ \ \forall x_1, x_2 \in \mathbb{R}$.

(P3) F_X es continua por la derecha: $\lim\limits_{x \to a^+} F(x) = F(a)\ \ a \in \mathbb{R}$.

Como consecuencia de las propiedades del resultado anterior, están los siguientes resultados:

1. $0 \leqslant F_X \leqslant 1$.
2. $\mathbb{P}(a < X \leqslant b) = \mathbb{P}(X \leqslant b) - \mathbb{P}(X \leqslant a) = F(b) - F(a)$.
3. Para cualquier $x \in \mathbb{R}$, $\mathbb{P}(X = x) = F(x) - \mathbb{P}(X < x)$.
4. El conjunto de puntos de discontinuidad de F es finito o numerable.

Ejemplo **2.4** Para el caso del ejemplo 2.1, la función de distribución está dada por:
$F(a) = 0,$ si $a < 0;\ F(0) = \frac{1}{8};\ F(1) = \frac{1}{2};\ F(2) = \frac{7}{8};\ F(b) = 1,$ si $b \geqslant 1$.
□

Variables aleatorias discretas

En el caso de una variable aleatoria discreta, es posible definir una función que represente la probabilidad para cualquier valor de una variable aleatoria X, esto es, $p(x) = \mathbb{P}(X = x)$, la cual es llamada función de probabilidad o función de cuantía.

Ejemplo **2.5** Considerar el experimento aleatorio de lanzar 7 monedas al aire y observar el resultado. Sea X: número de caras que se aparecen al lanzar las 7 monedas. Calcular la probabilidad de que aparezcan x caras al realizar el experimento aleatorio descrito.

Solución: El espacio muestral se define por:

$$\Omega = \{(x_1, \cdots, x_7) \mid x_i = 2 \text{ posibilidades}, i = 1, \cdots, 7\} \Rightarrow \#(\Omega) = 2^7.$$

Por su parte, se tiene que $Rec_X = \{0, 1, \cdots, 7\}$. Luego, la probabilidad para los valores $x_0 \in Rec_X$: $\mathbb{P}(X = x_0) = p(x_0)$ es:

$$\begin{aligned}
p(0) =& \mathbb{P}(\{s, \cdots, s\}) = \frac{1}{2^7} = \frac{\binom{7}{0}}{2^7}.\\
p(1) =& \mathbb{P}(\{c, s, s, \cdots, s\}, \{s, c, s, \cdots, s\}, \cdots, \{s, s, \cdots, c\}) = \frac{7}{2^7} = \frac{\binom{7}{1}}{2^7}.\\
&\vdots\\
p(7) =& \mathbb{P}(\{c, \cdots, c\}) = \frac{1}{2^7} = \frac{\binom{7}{7}}{2^7}.
\end{aligned}$$

Generalizando, para $x \in Rec_X$ se obtiene la expresión:

$$\mathbb{P}(X = x) = p(x) = \mathbb{P}((\underbrace{c, \cdots}_{x}, \underbrace{s, \cdots}_{7-x})) = \frac{\binom{7}{x}}{2^7}, \text{ para } x = 0, 1, \cdots, 7. \square$$

A partir de este ejemplo, se establece la propiedad siguiente, la cual, dada una función que asigne probabilidad de ocurrencia para todos los valores de una variable aleatoria discreta, establece si corresponde a una función de cuantía.

Propiedad **2.2** Sean $\{x_1, x_2, \cdots, x_n\}$ valores del recorrido de una variable aleatoria discreta X y sean $\{p_1, p_2, \cdots, p_n\}$ las respectivas probabilidades de cada valor de la variable aleatoria. Si $p_i = p(x_i)$ cumple con las siguientes condiciones, se llamará función de cuantía:

(1) $p_i > 0$, $\forall i$, $1 \leqslant i \leqslant n$.

(2) $\sum_{i=1}^{\infty} p_i = 1$.

Demostración: (1) Como $p_i = \mathbb{P}(X = x_1) = \mathbb{P}(A) \geqslant 0$ y si $A \neq \phi$ se cumplirá que $p(i) \geqslant 0$.

(2) $F(x) = \sum_{i=1}^{x} p_i$, además $\sum_{i=1}^{n} p_i = 1$ y $p_i = 0$ para $i \geqslant n$. Como $\lim_{x\to\infty} F(x) = 1$, entonces $\sum_{i=1}^{\infty} p_i = 1$.

Ejemplo **2.6** Considerando la función de probabilidad obtenida en el ejemplo 2.5, probar que es efectivamente una función de cuantía.

Solución: En el ejemplo mencionado, la función es $p(x) = \frac{\binom{7}{x}}{2^7}$. Luego, es inmediato que $p(x) > 0$, además:

$$\sum_{x=0}^{7} \frac{\binom{7}{x}}{2^7} = \frac{1}{2^7} \sum_{x=0}^{7} \binom{7}{x} = 1.$$

Con esto, se prueba que $p(x)$ es función de probabilidad para la variable aleatoria X definida. □

Utilizando la función de probabilidad, es posible hacer la relación con la función de distribución para el cálculo de ocurrencia de un conjunto $B \in \mathbb{R}$, en efecto: si los posibles valores de X variable aleatoria discreta son $\{x_1, x_2, \cdots, x_n\}$, $n > 7$ y $B = \{x_1, x_2, \cdots, x_6\}$, luego:

$$\begin{aligned} F(x_6) =& \mathbb{P}(X \leqslant x_6) = \mathbb{P}(X = x_1) + \mathbb{P}(X = x_2) + \cdots + \mathbb{P}(X = x_6) \\ =& p(x_1) + p(x_2) + \cdots + p(x_6). \end{aligned}$$

Ejemplo **2.7** Sea la variable aleatoria X discreta con función de cuantía:

$$p(x) = c \cdot 0{,}8^x, \quad x = 0, 1, 2, \cdots$$

(a) Determinar el valor de c que hace que $p(x)$ sea una función de cuantía.

(b) Obtener la función de distribución para X.

(c) Calcular $\mathbb{P}(10 \leqslant X \leqslant 25 \mid 15 \leqslant X \leqslant 30)$.

Solución:
(a) Como $\sum_{x=0}^{\infty} p(x) = 1 \Rightarrow \sum_{x=0}^{\infty} c \cdot 0{,}8^x = \dfrac{c}{1-0{,}8} = 1 \Rightarrow c = 0{,}2.$
Por tanto, $p(x) = 0{,}2 \cdot 0{,}8^x$, $x = 0, 1, 2, \cdots$
Se usó la aproximación en serie dada por: $\sum_{x=0}^{\infty} ab^x = \dfrac{a}{1-b}$, $|b| < 1$.
(b) La función de distribución se obtiene por:

$$\begin{aligned} F(x) =& \mathbb{P}(X \leqslant x) = \sum_{k=0}^{x} p(k) = 1 - \mathbb{P}(X > x) = 1 - \sum_{k=x+1}^{\infty} 0{,}2 \cdot 0{,}8^k \\ =& 1 - \sum_{t=0}^{\infty} 0{,}2 \cdot 0{,}8^{t+(x+1)}, \text{ siendo } t = k - (x+1) \\ =& 1 - 0{,}2 \cdot 0{,}8^{x+1} \sum_{t=0}^{\infty} 0{,}8^t \\ =& 1 - 0{,}2 \cdot 0{,}8^{x+1} \frac{1}{0{,}2} = 1 - 0{,}8^{x+1}. \end{aligned}$$

Por lo tanto, $F(x) = 1 - 0{,}8^{x+1}$, $x = 0, 1, 2, \cdots$. En el gráfico de la figura 2.1 se aprecia la función de cuantía, $p(x)$, y la función de probabilidad acumulada, $F(x)$.
(c) Sea $B = 10 \leqslant X \leqslant 25 \mid 15 \leqslant X \leqslant 30$:

$$\begin{aligned} \mathbb{P}(B) =& \frac{\mathbb{P}(15 \leqslant X \leqslant 25)}{\mathbb{P}(15 \leqslant X \leqslant 30)} = \frac{\mathbb{P}(X \leqslant 25) - \mathbb{P}(X \leqslant 14)}{\mathbb{P}(X \leqslant 30) - \mathbb{P}(X \leqslant 14)} \\ =& \frac{F(25) - F(14)}{F(30) - F(14)} = \frac{0{,}8^{15} - 0{,}8^{26}}{0{,}8^{15} - 0{,}8^{31}} = 0{,}941. \square \end{aligned}$$

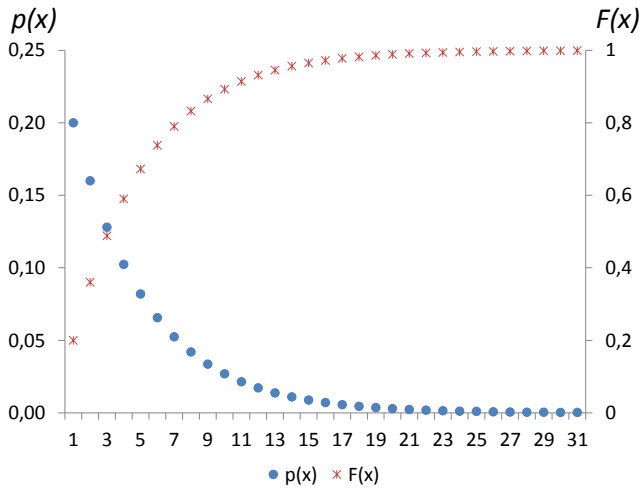
Ejemplo **2.8** Obtener la función de probabilidad acumulada para la función de probabilidad (del ejemplo 2.12) dada por:

	Recorrido de X			
	$x = 0$	$x = 1$	$x = 2$	$x = 3$
$\mathbb{P}(X = x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Solución: La función de distribución está dada por:

$$F(x) = \mathbb{P}(X \leqslant x) = \begin{cases} 0 & \text{si} \quad x < 0 \\ \frac{1}{8} & \text{si} \quad 0 \leqslant x < 1 \\ \frac{1}{2} & \text{si} \quad 1 \leqslant x < 2 \\ \frac{7}{8} & \text{si} \quad 2 \leqslant x < 3 \\ 1 & \text{si} \quad x \geqslant 3. \end{cases} \square$$

Figura 2.1: Gráfico de la función de densidad y la función de probabilidad



Fuente: elaboración propia.

Ejercicio propuesto **2.1** Resolver los siguientes puntos:

(1) Graficar la función del ejemplo 2.8.

(2) Considerar la función de cuantía:

$$p(x) = \frac{4!}{(4-x)!x!}\left(\frac{1}{3}\right)^x\left(\frac{2}{3}\right)^{4-x}, \quad x = 0, 1, 2, 3, 4$$

(a) Probar que $p(x)$ es realmente una función de cuantía.

(b) Determinar la función de distribución y graficar.

Variables aleatorias continuas

Para una variable aleatoria continua, se tendrá la función $f(x)$, llamada función de densidad de X, dada por:

$$F_X(x) = \int_{-\infty}^{x} f(t)\,dt, \quad \forall x \in \mathbb{R}.$$

Propiedad **2.3** Si $f(x)$ es la función de densidad de una variable aleatoria X, entonces cumplirá que:

(1) $f(x) \geqslant 0$ para todo $x \in Rec_X$.

(2) $\displaystyle\int_{x \in Rec_X} f(x)\, dx = 1.$

Demostración: (1) Es inmediato por la definición de función de densidad. (2) Por la propiedad 2.1:

$$\lim_{x \to \infty} F(x) = 1 \Rightarrow \int_{-\infty}^{\infty} f(t)\, dt = 1.$$

Ejemplo **2.9** Considerar la función de densidad de la variable aleatoria continua X dada por $f(x) = \exp\{-x\}$, $x > 0$. Obtener los siguientes resultados:

(a) Verificar que $f(x)$ es efectivamente la función de densidad para la variable aleatoria X.

(b) Calcular la función de distribución.

(c) Utilizar la función de distribución para calcular $\mathbb{P}(1{,}0 \leqslant X < 3{,}0 \mid 1{,}5 \leqslant X \leqslant 4{,}0)$.

Solución: (a) Se observa que $f(x) > 0$, pues $e^{-x} > 0$, $\forall x \in \mathbb{R}^+$. Además:

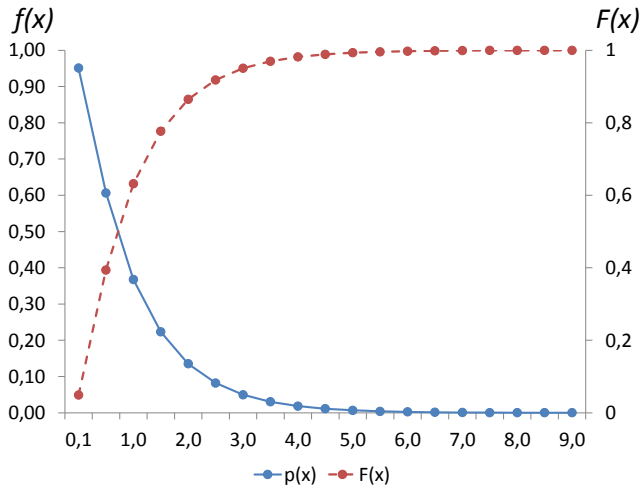
$$\int_0^{\infty} e^{-x}\, dx = \int_0^{\infty} x^0 e^{-x}\, dx = 0! = 1.$$

Por tanto, $f(x)$ es la función de densidad. El resultado anterior se obtuvo usando la *función gamma*: $\displaystyle\int_0^{\infty} x^{n-1} e^{-x}\, dx = \Gamma(n) = (n-1)!$, $n \in \mathbb{N}$.

(b) La función de distribución, y el gráfico (véase figura 2.2) para esta, y la función de densidad son:

$$F(x) = \mathbb{P}(X \leqslant x) = \int_0^{x} e^{-t}\, dt = 1 - e^{-x}, \;\; x > 0.$$

Figura 2.2: Gráfico de la función de densidad y la función de distribución



Fuente: elaboración propia.

(c) Para calcular la probabilidad condicional, se debe reescribir la expresión con la forma de la función de distribución. En efecto:

$$\begin{aligned}\mathbb{P}(1{,}0 \leqslant X < 3{,}0 \mid 1{,}5 \leqslant X \leqslant 4{,}0) &= \frac{\mathbb{P}(1{,}5 \leqslant X < 3{,}0)}{\mathbb{P}(1{,}5 \leqslant X \leqslant 4{,}0)} \\ &= \frac{F(3{,}0) - F(1{,}5)}{F(4{,}0) - F(1{,}5)} = \frac{e^{-1{,}5} - e^{-3{,}0}}{e^{-1{,}5} - e^{-4{,}0}}.\square\end{aligned}$$

Ejemplo **2.10** Sea $f(x) = \frac{1}{3}x^2$ donde $-1 \leqslant x \leqslant 2$. A partir de la función de distribución, obtener $\mathbb{P}(0 < X \leqslant 1)$.

Solución: Se puede verificar que $f(x)$ es la función de densidad. Luego, la función de distribución está determinada por:

$$F(x) = \mathbb{P}(X \leqslant x) = \begin{cases} 0 & \text{si} \quad x < -1 \\ \frac{1}{9}(x^3 + 1) & \text{si} \quad -1 \leqslant x < 2 \\ 1 & \text{si} \quad x \geqslant 2. \end{cases}$$

Así, se puede reexpresar la probabilidad para aplicar la función de distribución antes obtenida. En efecto:
$\mathbb{P}(0 < X \leqslant 1) = \mathbb{P}(X \leqslant 1) - \mathbb{P}(X \leqslant 0) = F(1) - F(0) = \frac{2}{9} - \frac{1}{9} = \frac{1}{9}$. $\square$

Observación **2.1** Para X variable aleatoria continua se tiene que:

$$F(x) = \mathbb{P}(X \leqslant x) = \mathbb{P}(X < x).$$

Cálculo de probabilidad para una variable aleatoria

Luego de la definición y de haber identificado el espacio muestral y el recorrido de la variable aleatoria, interesa obtener la probabilidad de ocurrencia de un evento constituido por un conjunto de elementos del recorrido.

Sea B el conjunto de todos los intervalos en los reales abiertos, semiabiertos y cerrados. Luego, cualquier conjunto $B \in \mathbb{R}$, se tiene que:

$$\mathbb{P}(X \in B) = \begin{cases} \displaystyle\sum_{i|x_i \in B} p(x_i) & \text{si } X \text{ es una variable aleatoria discreta} \\ \displaystyle\int_B f(x)\,dx & \text{si } X \text{ es una variable aleatoria continua.} \end{cases}$$

En efecto, sea X una variable aleatoria con función de cuantía $p(x)$ (caso discreto) o función de densidad $f(x)$ (caso continuo) y la probabilidad de ocurrencia de un intervalo en los reales, se tiene:

$$\mathbb{P}(a < X \leqslant b) = \mathbb{P}(X \leqslant b) - \mathbb{P}(X \leqslant a) = F_X(b) - F_X(a).$$

Si se considera $B_1 = \{X \leqslant b\}$ y $B_2 = \{X \leqslant a\}$, se tiene que:

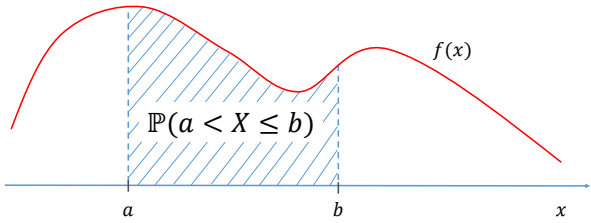
$$\mathbb{P}(a < X \leqslant b) = \begin{cases} \displaystyle\sum_{B_1} p(x) - \sum_{B_2} p(x) = \sum_{x=a}^{b} p(x), & \text{caso discreto} \\ \displaystyle\int_{B_1} f(x)\,dx - \int_{B_2} f(x)\,dx = \int_a^b f(x)\,dx, & \text{caso continuo.} \end{cases}$$

Nótese que para el caso de una variable aleatoria continua, la probabilidad de que X tome valores en el intervalo (a, b) corresponde al área bajo la curva de la función de densidad entre a y b, tal como en la figura 2.3.

Ejemplo **2.11** Considerando el ejemplo 2.1:

$$\begin{aligned} p(0) =& \mathbb{P}(X = 0) = \frac{1}{8} \\ p(1) =& \mathbb{P}(X = 1) = \frac{3}{8} \\ p(2) =& \mathbb{P}(X = 2) = \frac{3}{8} \\ p(3) =& \mathbb{P}(X = 3) = \frac{1}{8}. \square \end{aligned}$$

Figura 2.3: Gráfico de la probabilidad de que X esté en (a, b)



Fuente: elaboración propia.

Ejemplo **2.12** En función del espacio muestral y la variable aleatoria X definida en el ejemplo 2.1, hallar la probabilidad para cada elemento del recorrido de X.

Solución: Las probabilidades respectivas están en la tabla siguiente:

	Recorrido de X			
	$x=0$	$x=1$	$x=2$	$x=3$
$\mathbb{P}(X=x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

□

Ejemplo **2.13** Sea X una variable aleatoria discreta con función de cuantía:

$$p(x) = \mathbb{P}(X = x) = \frac{5^x e^{-5}}{x!}, \quad x = 0, 1, \cdots .$$

(a) Verificar que $p(x)$ es función de cuantía, (b) calcular $\mathbb{P}(0 < X \leqslant 3)$.

Solución: (a) $p(x)$ es función de cuantía, pues:

(1) $p(x) \geqslant 0$, para todo $x \in Rec_X$, y

$$(2)\ \sum_{x=0}^{\infty} \mathbb{P}\left(X = x\right) = \sum_{x=0}^{\infty} \frac{5^x e^{-5}}{x!} = e^{-5} \sum_{x=0}^{\infty} \frac{5^x}{x!} = e^{-5} e^{5} = 1.$$

(b) La probabilidad pedida se obtiene mediante:

$$\mathbb{P}(0 < X \leqslant 3) = \sum_{x=0}^{3} \mathbb{P}\left(X = x\right) = \mathbb{P}(X = 1) + \mathbb{P}(X = 2) + \mathbb{P}(X = 3)$$
$$= e^{-5} \cdot \left(5 + \frac{25}{2} + \frac{125}{6}\right) = 0{,}258. \square$$

Ejercicio propuesto **2.2** Considerar el experimento aleatorio de lanzar una moneda n veces y sea la variable aleatoria X: *cantidad de caras que aparecen* al realizar el experimento. Luego:

(a) Determinar $p(x)$ y verificar si es función de cuantía.

(b) Calcular las siguientes probabilidades: (b.1) $\mathbb{P}(X \leqslant 1)$, (b.2) $\mathbb{P}(X \geqslant 2)$, (b.3) $\mathbb{P}(2 \leqslant X \leqslant 5)$, (b.4) $\mathbb{P}(X \leqslant 4 \mid X > 1)$.

(c) Se redefine el experimento aleatorio utilizando ahora una moneda *cargada* la cual es lanzada n veces, donde la probabilidad de obtener cara es de 60 %. Considerar nuevamente la variable aleatoria X: *cantidad de caras que aparecen* y obtener (a).

Ejemplo **2.14** Sea $f(x) = \beta^{-1} \exp\left\{-\beta^{-1}x\right\}$, $x > 0$, β es una constante positiva. La función de esta variable aleatoria se muestra en la figura siguiente para tres valores del parámetro β. (a) Verificar que $f(x)$ es función de densidad y (b) calcular $\mathbb{P}(0{,}8 < X < 1{,}2)$ y $\mathbb{P}(X > 1{,}2 \mid X > 0{,}8)$.

Solución: (a) $f(x)$ es función de densidad, pues:

(i) $f(x) \geqslant 0$, pues $\beta > 0$ y (ii) la integral para $x \in \mathrm{Rec}_X$ de f es:

$$\int_0^{\infty} \frac{1}{\beta} e^{-\frac{x}{\beta}}\, dx = \frac{1}{\beta} \int_0^{\infty} e^{-\frac{x}{\beta}}\, dx = \frac{1}{\beta}(-\beta e^{-\frac{x}{\beta}})\Big|_0^{\infty} = 1 - \lim_{x\to\infty} e^{-\frac{x}{\beta}} = 1.$$

(b) Cálculo de las probabilidades pedidas:

$$\mathbb{P}(0{,}8 < X < 1{,}2) = \int_{0{,}8}^{1{,}2} \frac{1}{\beta} e^{-\frac{x}{\beta}}\, dx = e^{-\frac{0{,}8}{\beta}} - e^{-\frac{1{,}2}{\beta}}.$$

$$\mathbb{P}(X > 1{,}2 \mid X > 0{,}8) = \frac{\mathbb{P}(X > 1{,}2 \cap X > 0{,}8)}{\mathbb{P}(X > 0{,}8)} = \frac{\mathbb{P}(X > 1{,}2)}{\mathbb{P}(X > 0{,}8)}$$
$$= e^{\frac{-0{,}4}{\beta}}. \square$$

Ejercicio propuesto **2.3** Sea $f(x) = \dfrac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\dfrac{1}{2\sigma^2}(x-\mu)^2\right\}$, $x \in \mathbb{R}$, $\mu \in \mathbb{R}$ y $\sigma > 0$. Verificar que $f(x)$ es la función de densidad.

Indicación: Nótese que la función de densidad propuesta es simétrica en torno a μ: $f(\mu - x_0) = f(\mu + x_0)$, por lo que el cálculo de la integral de $f(x)$ en el intervalo $(-\infty, +\infty)$ es equivalente a 2 veces la integral de $f(x)$ en el intervalo $(\mu, +\infty)$. Luego, la sustitución $z = \frac{x-\mu}{\sigma}$ lleva a una integral de tipo función gamma.

Valor esperado y varianza de una variable aleatoria

La esperanza matemática de una variable aleatoria es una generalización del concepto de media aritmética. Dada una muestra de valores observados de una variable X: x_i, con sus correspondientes frecuencias: f_i, $i = 1, 2, \cdots, n$, donde: $\sum_{i=1}^{n} f_i = n$. La media muestral es:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} f_i \cdot x_i = \sum_{i=1}^{n} x_i \frac{f_i}{n}, \text{ donde } \frac{f_i}{n} \text{ es la frecuencia relativa.}$$

La frecuencia relativa se puede considerar como la probabilidad que tiene el valor x_i de presentarse en la muestra total de tamaño n. Por tanto:

$$\mathbb{P}(X = x_i) = \frac{f_i}{n} \Rightarrow \bar{x} = \sum_{i=1}^{n} x_i \mathbb{P}(X = x_i).$$

Esta forma de expresar la media de la muestra sugiere la definición de *esperanza* en el caso discreto.

Definición **2.4** La definición está dada de acuerdo con la clasificación de la variable aleatoria:

(1) X es una variable aleatoria discreta con función de probabilidad $\mathbb{P}(X = x_i)$, $i = 0, 1, \cdots, n$, se define la esperanza o el valor esperado de la variable aleatoria X, denotado por $\mu_X = E[X]$, por:

$$\mu_X = E[X] = \sum_{k=x_1}^{x_n} k \cdot \mathbb{P}(X = k) = \sum_{k=x_1}^{x_n} k \cdot p(k).$$

Si se piensa la suma como una serie, se exige que sea absolutamente convergente.

(2) Si X es una variable aleatoria continua con función de densidad $f(x)$, $x_1 < x_2$, entonces el valor esperado o esperanza es:

$$\mu_X = E[X] = \int_{x_1}^{x_2} x \cdot f(x)\, dx.$$

Siempre y cuando la integral exista; en caso contrario se dice que la variable aleatoria no tiene asociado un valor esperado.

La esperanza o valor esperado de una variable aleatoria se podría interpretar conceptualmente como:

1. El valor medio teórico de todos los valores que puede tomar la variable, representada como una medida de centralización.

2. El centro de gravedad de los puntos que corresponden a los valores de la variable, asignándoles una cantidad de masa proporcional a la función de densidad en cada punto.

3. Si la variable es la ganancia o pérdida en un determinado juego de azar, la esperanza representaría la ganancia media por jugada. Un juego es equitativo si su esperanza es cero.

Ejemplo **2.15** Calcular el valor esperado asociado a la variable aleatoria discreta X con función de cuantía:

$$p(x) = \binom{n}{x}\alpha^x(1-\alpha)^{n-x};\ \ x = 0, 1, \cdots, n;\ \ 0 < \alpha < 1\ (\alpha \text{ constante}).$$

Solución: Aplicando la definición de esperanza:

$$\begin{aligned}
E[X] &= \sum_{x=0}^{n} x \binom{n}{x} \alpha^x (1-\alpha)^{n-x} \\
&= \sum_{x=0}^{n} x \frac{n!}{x!(n-x)!} \alpha^x (1-\alpha)^{n-x} \\
&= \sum_{x=1}^{n} \cancel{x} \frac{n!}{\cancel{x}\,(x-1)!\,(n-x)!} \alpha^x (1-\alpha)^{n-x} \\
&= \sum_{x=1}^{n} \frac{n!}{(x-1)!(n-x)!} \alpha^x (1-\alpha)^{n-x} \\
&= n\,\alpha \sum_{x=0}^{n} \binom{n-1}{x-1} \alpha^x (1-\alpha)^{n-1-x} \\
&= n\,\alpha\,(\cancel{\alpha} + (1-\cancel{\alpha}))^{n-1} = n\,\alpha\,(1)^{n-1} = n\,\alpha. \square
\end{aligned}$$

Ejemplo **2.16** Sea X una variable aleatoria cuya función de densidad está dada por $f(x) = \frac{1}{3}x^2$, $-1 \leqslant x \leqslant 2$. Determinar el valor esperado de X.

Solución: Dado que la variable aleatoria es continua, entonces el valor esperado corresponde a:

$$E[X] = \int_{-1}^{2} x \cdot f(x)\,dx = \int_{-1}^{2} \frac{x^3}{3}\,dx = \left.\frac{x^4}{12}\right|_{-1}^{2} = \frac{1}{12}\left(2^4 - (-1)^4\right) = \frac{15}{12}. \square$$

Ejemplo **2.17** Considerar la función de probabilidad dada en el ejemplo 2.1 y obtener el valor esperado.

Solución: La variable aleatoria del ejemplo 2.1 está dada por:

	Recorrido de X			
	$x=0$	$x=1$	$x=2$	$x=3$
$\mathbb{P}(X=x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Luego, $E[X] = 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} = \frac{3}{2} = 1{,}5$. □

Existen ciertas situaciones problemáticas en que el valor esperado de la variable aleatoria no está definido, pues la sumatoria o integral no es convergente en el recorrido de la variable aleatoria. A continuación, se presentan dos ejemplos en que ocurre esta situación.

Ejemplo **2.18** Resolver los siguientes problemas asociados al calculo del valor esperado de una variable aleatoria.

(1) Considerar el experimento asociado a una variable aleatoria con función de densidad dada por $f(x) = \big(\pi(1+x^2)\big)^{-1}$, $-\infty < x < +\infty$. Determinar el valor esperado de X.

(2) Considerar el experimento aleatorio de lanzar una moneda honesta hasta obtener una cara. Si la primera cara ocurre en el k-ésimo intento, el apostador gana \$ 2^k. Calcular la ganancia esperada por un apostador que se somete a este experimento.

Solución: (1) En este caso, la integral mediante la cual se obtiene el valor esperado es divergente en el intervalo en que está definido el recorrido, por lo tanto, se dice que el valor esperado de la variable aleatoria X es indeterminado.

(2) Sea X: *ganancia obtenida*, con lo cual $Rec_X = \{2, 2^2, 2^3, \cdots\}$. La probabilidad de obtener cara en el k-ésimo intento (función de probabilidad o cuantía de X) y el valor esperado de X, se obtienen respectivamente por:

$$\mathbb{P}\left(X = 2^k\right) = \left(\frac{1}{2}\right)^{k-1}\left(\frac{1}{2}\right) = \left(\frac{1}{2}\right)^{k},\ k = 1, 2, \cdots .$$

$$E\left[X\right] = \sum_{k=1}^{\infty} k \cdot 2^k = 2\frac{1}{2} + 2^2\frac{1}{2} + 2^3\frac{1}{2} + \cdots = +\infty.\square$$

Valor esperado de una función de una variable aleatoria

En ciertas oportunidades, la variable aleatoria a la que se desea calcular la esperanza, X, está dada o se puede expresar como función de otra variable aleatoria Y, es decir, $X = G(Y)$.

Definición **2.5** Se considera la variable aleatoria X con función de cuantía $p(x)$ o función de densidad $f(x)$. El valor esperado de una función de la variable aleatoria, $g(X)$, es denotado con $E\left[g(X)\right]$ y está dado por:

$$E\left[g(X)\right] = \begin{cases} \displaystyle\sum_{x \in Rec_X} g(x)p(x), & \text{caso discreto} \\ \displaystyle\int_{x \in Rec_X} g(x)f(x)\,dx, & \text{caso continuo.} \end{cases}$$

Casos particulares:

(a) Si $g(X) = X$, entonces se obtiene el valor esperado de X: $E\left[X\right]$.

(b) Si $g(X) = X^k$, $k \in \mathbb{N}$; entonces $E\left[X^k\right] = \mu_k$ es el momento de orden k de la variable aleatoria X.

(c) Si $g(X) = \left(X - E\left[X\right]\right)^k$, $k \in \mathbb{N}$; entonces $E\left[\left(X - E\left[X\right]\right)^k\right]$ es el momento de orden k alrededor del valor esperado de la variable aleatoria X.

(d) Si $g(X) = e^{t\,X}$, $t \in \mathbb{R}^+$; entonces $E\left[e^{tX}\right] = M_X(t) = M(t)$ es la función generadora de momentos (fgm) de la variable aleatoria X.

Propiedad **2.4** Sea X una variable aleatoria (discreta o continua) y $c \in \mathbb{R}$, una constante.

(a) $E\left[c\right] = c$.

(b) $E\left[cX\right] = c \cdot E\left[X\right]$.

Demostración: Sea X una variable aleatoria discreta (los resultados son análogos para el caso continuo):

(a) $E\left[c\right] = \sum c \cdot p(x) = c \sum p(x) = c$, pues $\sum_x p(x) = 1$.

(b) $E\left[cX\right] = \sum c \cdot x \cdot p(x) = c \sum_x x \cdot p(x) = c \cdot E\left[X\right]$.

Ejemplo **2.19** La variable aleatoria discreta X tiene función de cuantía: $p(x) = p^x(1-p)^{1-x}$, con $x = 0, 1$, y donde $0 < p < 1$. Obtener:

(a) Valor esperado.

(b) Momento de orden k.

(c) Momento de orden k en torno al valor esperado.

(d) Función generadora de momentos.

Solución: Considerando que la variable aleatoria X es discreta, se tiene que:

(a) El valor esperado de la variable aleatoria X es:

$$E\left[X\right] = \sum_{x=0}^{1} x \cdot p(x) = \sum_{x=0}^{1} x \cdot p^x(1-p)^{1-x} = p.$$

(b) El momento de orden k es:

$$\begin{aligned} E\left[X^k\right] = \sum_{x=0}^{1} x^k \cdot p(x) = \sum_{0}^{1} x^k \cdot p^x(1-p)^{1-x} = \\ = 0^k p^0 (1-p)^{1-0} + 1^k p^1 (1-p)^{1-1} = p. \end{aligned}$$

(c) El momento de orden k en torno al valor esperado de X es:

$$\begin{aligned} E[(X-p)^k] = \sum_{x=0}^{1} (x-p)^k p(x) = \sum_{x=0}^{1} (x-p)^k p^x (1-p)^{1-x} = \\ = (-p)^k (1-p) + (1-p)^k p. \end{aligned}$$

(d) La función generadora de momentos de X es: $M(t) = E\left[e^{tX}\right] = \sum_{x=0}^{1} e^{tx} p(x) = (1-p) + p \cdot e^t$. □

La función generadora de momentos, al igual que la función de probabilidad (de densidad o cuantía), está asociada a la variable aleatoria y, como su nombre lo indica, permite obtener el momento de orden k-ésimo y, a través de este, el valor esperado de la variable aleatoria (para $k = 1$).

Propiedad **2.5** Sea X una variable aleatoria, la derivada de orden k de la fgm evaluada en $t = 0$ corresponde al momento de orden k. Es decir:

$$\frac{d^k}{dt^k} M_X(t)\Big|_{t=0} = E\left[X^k\right] \;\; ; \quad k = 1, 2, \cdots .$$

Demostración: Se presenta un esquema de demostración considerando que la variable aleatoria X es discreta y que la función de cuantía $p(x)$ está definida para $x = x_1, x_2, \cdots$ (el resultado es análogo para el caso continuo).

A partir de la definición 2.5(d) de fgm se tiene que:

$$\frac{d}{dt} M(t) = \sum_{x=x_1}^{\infty} x \cdot e^{tx} \cdot p(x) \to \frac{d}{dt} M(t)\Big|_{t=0} = \sum_{x=x_1}^{\infty} x \cdot p(x) = E\left[X\right]$$

$$\frac{d^2}{dt^2} M(t) = \sum_{x=x_1}^{\infty} x^2 \cdot e^{tx} \cdot p(x) \to \frac{d^2}{dt^2} M(t)\Big|_{t=0} = \sum_{x=x_1}^{\infty} x^2 \cdot p(x) = E\left[X^2\right]$$

$$\vdots$$

$$\frac{d^k}{dt^k} M_X(t) = \sum_{x=x_1}^{\infty} x^k \cdot e^{tx} \cdot p(x) \to \frac{d^k}{dt^k} M(t)\Big|_{t=0} = \sum_{x=x_1}^{\infty} x^k\, p(x) = E\left[X^k\right].$$

Observación **2.2** Una demostración formal de la propiedad 2.5 se obtiene considerando la expansión en serie de Taylor para $\exp\{t\,X\}$. Este desarrollo aparece mencionado en la sección de ejercicios propuestos.

Propiedad **2.6** Sea X una variable aleatoria y $g(X) = g_1(X) + g_2(X) + \cdots + g_n(X)$, una función real valuada de dicha variable aleatoria. Entonces (para el caso discreto):

$$\begin{aligned} E\left[g(X)\right] &= \sum (g_1(x) + g_2(x) + \cdots + g_n(x))p(x) \\ &= \sum g_1(x)p(x) + \sum g_2(x)p(x) + \cdots + \sum g_n(x)p(x) \\ &= E\left[g_1(X)\right] + E\left[g_2(X)\right] + \cdots + E\left[g_n(X)\right] \end{aligned}$$

Es decir, $E\left[\sum_{i=1}^{n} g_i(X)\right] = \sum_{i=1}^{n} E\left[g_i(X)\right]$.

Corolario **2.1** Sea X una variable aleatoria y c_1 y c_2 constantes. La función de la variable aleatoria X, $g(X) = c_1 + c_2X$, cumple con:

(a) $E[g(X)] = E[c_1 + c_2X] = E[c_1] + E[c_2X] = c_1 + c_2E[X]$.

(b) La fgm de $g(X)$ está dada por:

$M_{g(X)}(t) = E[e^{t\,g(X)}] = E[e^{t\,(c_1+c_2X)}] = e^{t\,c_1}E[e^{t\,c_2X}] = e^{t\,c_1}M_X(t^*)$; $t^* = t\,c_2$.

Medidas de localización

Este concepto permite tener una idea global de dónde se encuentra localizada la distribución de la variable aleatoria X. Existen dos medidas bastante utilizadas: la media o valor esperado, estudiado recientemente, y la mediana, en la cual se profundizará brevemente a continuación.

Definición **2.6** La mediana de una variable aleatoria X es algún número m tal que:

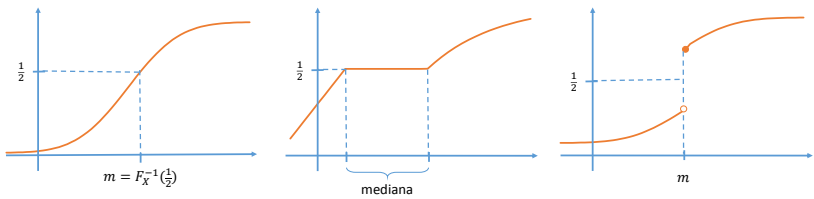
$$\mathbb{P}(X \geqslant m) \geqslant \frac{1}{2} \text{ y } \mathbb{P}(X \leqslant m) \geqslant \frac{1}{2}.$$

Como se observa en la figura 2.4, la mediana no es necesariamente única y no siempre se puede calcular haciendo uso de su definición, lo cual es un problema para describir la localización de la distribución de la variable aleatoria.

Como recomendación, se puede tener que, si se sabe que la distribución de la variable aleatoria es simétrica, entonces el valor esperado coincide con la mediana; en el caso de distribuciones asimétricas, la mediana puede ser una mejor medida de localización.

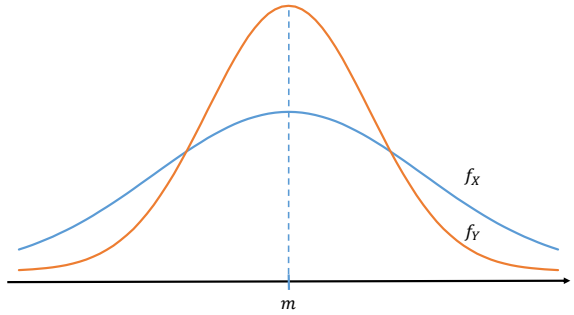
Cuando se quiere describir una distribución de una variable aleatoria se utilizan las medidas de localización. Sin embargo, no son suficientes y es necesario agregar las llamadas medidas de variabilidad, las cuales vienen a solucionar un problema: suponer que se tienen dos variables aleatorias con la misma medida de posicionamiento no asegura que tengan la misma distribución; por el contrario, pueden ser muy diferentes (como es el caso de la figura 2.5), por lo que surge la necesidad de agregar otra medida para caracterizar la distribución.

Figura 2.4: La mediana para tres tipos de variable aleatoria



Fuente: elaboración propia.

Figura 2.5: La mediana para dos funciones de densidad



Fuente: elaboración propia.

Varianza de una variable aleatoria

La varianza indica la dispersión de la distribución de la variable aleatoria, lo que permite contar con una medida de cuán homogénea (menor dispersión) o heterogénea (mayor dispersión) es la variable.

Definición **2.7** Sea la variable aleatoria X, se define la varianza de X, denotada con $Var[X]$ o $V[X]$, como:

$$Var[X] = E[(X - E[X])^2].$$

A partir de esta definición, se puede obtener una representación alter-

nativa que simplifique el cálculo. Esto es:

$$\begin{aligned} Var[X] &= E[(X - E[X])^2] = E\{(X^2) - 2X\,E[X] + [E[X]]^2\} = \\ &= E[X^2] - 2E[X]\,E[X] + (E[X])^2 = \\ &= E[X^2] - (E[X])^2 = \\ &= m_2 - m_1^2. \end{aligned}$$

Es decir, la varianza se puede expresar como la diferencia entre el momento de orden 2 y el momento de orden 1 (esperanza) al cuadrado.

Propiedad **2.7** Sea la variable aleatoria X, $k \in \mathbb{R}$ una constante, bajo el supuesto de que existe la esperanza de X, se cumple que:

(a) $Var[k] = 0$.

(b) $Var[kX] = k^2\,Var[X]$.

Demostración: (a) $Var[k] = E[k^2] - (E[k])^2 = k^2 - k^2 = 0$. (b) Desde la definición 2.7, se tiene que:

$$\begin{aligned} Var[kX] &= E[(kX)^2] - (E[kX])^2 = k^2E[X^2] - k^2(E[X])^2 = \\ &= k^2(E[X^2] - (E[X])^2) = k^2Var[X]. \end{aligned}$$

Ejemplo **2.20** Sea la función de densidad de la variable aleatoria X:

$$f(x) = (2\beta^3)^{-1}x^2 \exp\left\{-\frac{x}{\beta}\right\}, \quad x > 0, \quad \beta > 0 \ \text{ (con } \beta \text{ constante)}.$$

(a) Determinar la fgm de X.

(b) Calcular el valor esperado y la varianza.

(c) Calcular la esperanza y la varianza de $g(X) = 2 + 3X$.

Solución: (a) La fgm de X está dada por:

$$\begin{aligned} M_X(t) = E[e^{tX}] &= \int_0^\infty e^{t\,x} f(x)\,dx = \int_0^\infty (2\beta^3)^{-1}x^2 e^{tx} e^{-\frac{x}{\beta}}\,dx \\ &= (2\beta^3)^{-1} \int_0^\infty x^2 \exp\left\{-x(\beta^{-1} - t)\right\}\,dx \end{aligned}$$

Sea $z = x(\beta^{-1} - t)$, luego $dx = (\beta^{-1} - t)^{-1}$ y se mantienen los límites de integración. Así:

$$\begin{aligned} &= (2\beta^3)^{-1}(\beta^{-1} - t)^{-3} \int_0^\infty z^2 e^{-z}\, dz \\ &= (1 - \beta t)^{-3}. \end{aligned}$$

(b) El valor esperado y la varianza se obtienen a través de la fmg:

$$\frac{d}{dt} M_X(t) = \frac{d}{dt}(1 - \beta\, t)^{-3} = 3\beta(1 - \beta\, t)^{-4}$$

Haciendo $t = 0$, $E[X] = 3\beta$. Para calcular la varianza se necesita el momento de orden 2:

$$\frac{d^2}{dt^2} M_X(t) = \frac{d}{dt}\{3\beta(1 - \beta\, t)^{-4}\} = 12\beta^2(1 - \beta\, t)^{-5}$$

Haciendo $t = 0$, $E[X^2] = 12\beta^2$. Por lo tanto:

$$Var[X] = 12\beta^2 - (3\beta)^2 = 3\beta^2.$$

(c) Por los resultados obtenidos anteriormente, se sabe que $E[X] = 3\beta$ y $Var[X] = 3\beta^2$. Así:

$$\begin{aligned} E[g(X)] &= 2 + 3E[X] = 2 + 9\beta \\ Var[g(X)] &= 27\beta^2. \end{aligned}$$

□

Ejemplo **2.21** Sea X una variable aleatoria continua con función de densidad:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}},\ -\infty < x < \infty.$$

Obtener expresiones para:

(a) La fgm de X.

(b) La fgm de $g(X) = \sigma X + \mu$ ($\sigma > 0$, $\mu \in \mathbb{R}$).

(c) El valor esperado y la varianza de $g(X)$.

Solución:

(a) Dado que la fgm es $M_X(t) = E[e^{tX}]$, se tiene que:

$$M_X(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx} e^{-\frac{x^2}{2}}\, dx = \frac{1}{\sqrt{2\pi}} e^{\frac{t^2}{2}} \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2}(x-t)^2\right\}\, dx$$

La función $\exp\left\{-\frac{1}{2}(x-t)^2\right\}$ es simétrica en torno a t, luego:

$$= \frac{2e^{\frac{t^2}{2}}}{\sqrt{2\pi}} \int_{t}^{\infty} \exp\left\{-\frac{1}{2}(x-t)^2\right\}\, dx.$$

Haciendo la sustitución $2z = (x-t)^2$, se tiene que:

$$= \frac{1}{\sqrt{2\pi}} e^{\frac{t^2}{2}} \int_{0}^{\infty} z^{-\frac{z}{2}} e^{-z}\, dz.$$

Resolviendo la función gamma, finalmente, se obtiene:

$$M_X(t) = \exp\left\{\frac{t^2}{2}\right\}.$$

(b) $M_{g(X)}(t) = e^{\mu t} M_X(\sigma t) = e^{\mu t} e^{\sigma^2 \frac{t^2}{2}} = e^{\mu t + t^2 \frac{\sigma^2}{2}}$.

(c) $E[g(X)] = \sigma E[X] + \mu = \mu$ y $Var[g(X)] = \sigma^2 Var[X] = \sigma^2$. El valor esperado y la varianza de $g(X)$ se obtienen a partir de la fgm de X (quedan de ejercicio para el lector). □

2.3. Transformaciones de variables aleatorias

Anteriormente, se ha trabajado obteniendo expresiones para la función generadora de momentos, valor esperado y varianza de funciones de variables aleatorias, $Y = g(X)$. En esta sección, se estudiará brevemente cómo obtener la función de densidad de estas variables aleatorias a partir de otra variable aleatoria.

Propiedad **2.8** Sea la variable aleatoria X con función de distribución $F_X(x)$. La nueva variable aleatoria $T = g(X)$ tiene función de distribución:

$$F_T(t) = \begin{cases} F_X(g^{-1}(t)), & \text{si } g(X) \text{ es función creciente} \\ 1 - F_X(g^{-1}(t)), & \text{si } g(X) \text{ es función decreciente.} \end{cases}$$

Luego, la función de densidad de T está dada por:

$$f_T(t) = \frac{d}{dt}F_T(t) = f_X\left(g^{-1}(t)\right)\left|\frac{d}{dt}g^{-1}(t)\right|.$$

Demostración: Sea $F_T(t)$ la función de distribución de la variable aleatoria T y $g(X)$ (monótona) creciente, entonces:

$$\begin{aligned} F_T(t) &= \mathbb{P}(T \leqslant t) = \mathbb{P}(g(X) \leqslant t) \\ &= \mathbb{P}(X \leqslant g^{-1}(t)) = F_X(g^{-1}(t)). \end{aligned}$$

$$\Rightarrow f_T(t) = \frac{d}{dt}F_X\left(g^{-1}(t)\right) = f_X\left(g^{-1}(t)\right)\frac{d}{dt}g^{-1}(t).$$

Ahora, si $g(X)$ (monótona) decreciente, entonces:

$$\begin{aligned} F_T(t) &= \mathbb{P}(T \leqslant t) = \mathbb{P}(g(X) \leqslant t) \\ &= \mathbb{P}(X > g^{-1}(t)) = 1 - F_X(g^{-1}(t)). \end{aligned}$$

$$\Rightarrow f_T(t) = \frac{d}{dt}\left[1 - F_X\left(g^{-1}(t)\right)\right] = -f_X\left(g^{-1}(t)\right)\frac{d}{dt}g^{-1}(t).$$

Finalmente, la función de densidad de T, para $g(X)$ creciente o decreciente, es:

$$\Rightarrow f_T(t) = f_X\left(g^{-1}(t)\right)\left|\frac{d}{dt}g^{-1}(t)\right|.$$

Ejemplo **2.22** Sea la variable aleatoria X con función de densidad:

$$f(x) = \frac{1}{2\beta^3}x^2\exp\left\{-\frac{x}{\beta}\right\}, \quad x > 0, \quad \beta > 0 \ \text{ (con } \beta \text{ constante)}.$$

Obtener la función de densidad de la variable aleatoria Y dada por $Y = 2 + 3X$.

Solución: Como la nueva variable aleatoria Y es función de X, entonces, usando la propiedad 2.8:

$$\begin{aligned} f_Y(y) &= \frac{d}{dy}F_X\left(\frac{1}{3}(y-2)\right) = \frac{1}{3}f_X\left(\frac{1}{3}(y-2)\right) \\ &= \frac{1}{3}(2\beta^3)^{-1}\left(\frac{1}{3}(y-2)\right)^2\exp\left\{-\frac{y-2}{3\beta}\right\} \\ &= (54\beta^3)^{-1}(y-2)^2\exp\left\{-\frac{y-2}{3\beta}\right\}, \quad y > 2. \end{aligned}$$

□

Ejercicio propuesto **2.4** Sea X una variable aleatoria y sea la función $f(x) = k(x^3 + 1)$, $x \in (0, 2)$.

(a) Determinar k, tal que $f(x)$ sea función de densidad para X.

(b) Determinar F_X y M_X.

(c) Calcular el valor esperado y la varianza de X.

(d) Sea la nueva variable aleatoria $Y = 2X - 1$, obtener la función de densidad de Y y M_Y.

2.4. Ejercicios resueltos

1. Considerar la variable aleatoria discreta X con función de cuantía:

$$p(x) = \frac{q(1-q)^x}{1-(1-q)^{m+1}}, \quad x = 0, 1, \cdots, m; \quad 0 < q < 1.$$

   (a) Probar que $p(x)$ es una función de cuantía.

   (b) Determinar el valor esperado de la función $H(X) = s^X$.

   (c) Utilizar el resultado de (b) para obtener una expresión para el valor esperado de X.

   Solución: (a) Función de cuantía: $f(x) > 0$, y la sumatoria en el recorrido de la variable se obtiene:

$$\begin{aligned}\sum_{x=0}^{m} p(x) &= \sum_{x=0}^{m} \frac{q(1-q)^x}{1-(1-q)^{m+1}} \\ &= \frac{q}{1-(1-q)^{m+1}} \left(\sum_{x=0}^{\infty} (1-q)^x - \sum_{x=m+1}^{\infty} (1-q)^x \right)\end{aligned}$$

   Como $0 < 1 - q < 1$, se garantiza la convergencia de la serie y, haciendo la sustitución $t = x - (m + 1)$, se obtiene una serie de tipo geométrica:

$$\begin{aligned}&= \frac{q}{1-(1-q)^{m+1}} \left(\frac{1}{1-(1-q)} - \frac{(1-q)^{m+1}}{q} \right) \\ &= \frac{\cancel{q}}{\cancel{1-(1-q)^{m+1}}} \frac{1}{\cancel{q}} \cancel{(1-(1-q)^{m+1})} = 1.\end{aligned}$$

(b) El valor esperado de s^X está dado por:

$$E\left[s^X\right] = \sum_{x=0}^{m} s^x p(x) = \sum_{x=0}^{m} s^x \frac{q(1-q)^x}{1-(1-q)^{m+1}}$$
$$= \frac{q}{1-(1-q)^{m+1}} \sum_{x=0}^{m} (s(1-q))^x.$$

La convergencia se cumple para $0 < s < 1$. Además, obsérvese que la sumatoria tiene la misma forma que la del punto (a), por lo tanto, la sustitución $t = x - (m+1)$ lleva a obtener una serie de tipo geométrica. En efecto:

$$= \frac{q}{1-(1-q)^{m+1}} \left(\frac{1}{1-s(1-q)} - \frac{(s(1-q))^{m+1}}{1-s(1-q)} \right)$$
$$= \frac{q}{1-(1-q)^{m+1}} \frac{1}{1-s(1-q)} (1-(s(1-q))^{m+1}).$$

(c) Si se considera $s = e^t$, se tiene que $E[s^X] = E\left[e^{tX}\right] = M_X(t)$, con lo cual se puede obtener el valor esperado (véase propiedad 2.5). En efecto:

$$M_X(t) = \frac{q}{1-(1-q)^{m+1}} \frac{1}{1-e^t(1-q)} (1-(e^t(1-q))^{m+1}).$$
$$\frac{d}{dt} M_X(t) = \frac{(e^t(1-q))(1-(e^t(1-q))^{m+1})}{(1-(1-q)^{m+1})^2} - \frac{(m+1)(e^t(1-q))^{m+1}}{1-e^t(1-q)}.$$

Finalmente, el valor esperado se obtiene con la derivada de fgm respecto a t y evaluada en $t = 0$:

$$E[X] = \frac{1}{(1-(1-q)^{m+1})^2} (1-(1-q)^{m+1}) - \frac{(m+1)(1-q)^{m+1}}{1-(1-q)}$$
$$= \frac{(m+1)(1-q)}{(1-(1-q)^{m+1})^2} (1-(1-q)^{m+1}) - \frac{(m+1)(1-q)^{m+1}}{q}.$$

2. Considerar la variable aleatoria discreta X con función de cuantía $p(x) = \dfrac{e^{-\lambda}\lambda^x}{x!}$; $x = 0, 1, 2, \cdots$; $\lambda > 0$.

   (a) Probar que $p(x)$ es una función de cuantía.

   (b) Determinar la función generadora de momentos de X.

(c) Calcular el valor esperado y la varianza de X.

Solución: (a) Es inmediato que $p(x) > 0$, además:

$$\sum_{x=0}^{\infty} p(x) = \sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = e^{-\lambda}e^{\lambda} = 1.$$

(b) La función generadora de momentos se obtiene mediante:

$$\begin{aligned} M_X(t) =& E\left[e^{tX}\right] = \sum_{x=0}^{\infty} e^{tx}p(x) = \sum_{x=0}^{\infty} e^{tx}\frac{e^{-\lambda}\lambda^x}{x!} = \\ =& e^{-\lambda} \sum_{x=0}^{\infty} \frac{(e^t\lambda)^x}{x!} = e^{-\lambda}e^{e^t\lambda} = e^{-\lambda(1-e^t)}. \end{aligned}$$

(c) El valor esperado se obtiene a partir de la fgm obtenida anteriormente:

$$\frac{d}{dt}M_X(t) = \lambda e^t e^{-\lambda}e^{e^t\lambda} \Rightarrow \frac{d}{dt}M_X(t)\Big|_{t=0} = \lambda e^{-\lambda}e^{\lambda} = \lambda = E[X].$$

El momento de orden 2 se obtiene por:

$$\begin{aligned} \frac{d^2}{dt^2}M_X(t) &= \lambda e^t e^{-\lambda}e^{e^t\lambda} + (\lambda e^t)^2 e^{-\lambda}e^{e^t\lambda}. \\ \frac{d^2}{dt^2}M_X(t)\Big|_{t=0} &= \lambda e^{-\lambda}e^{\lambda} + (\lambda)^2 e^{-\lambda}e^{\lambda} = \lambda + \lambda^2 = E[X^2]. \end{aligned}$$

Finalmente, $Var[X] = E[X^2] - (E[X])^2 = \lambda + \lambda^2 - \lambda^2 = \lambda$.

3. Sea X una variable aleatoria continua con una función de densidad $f(x) = 2x$ y $Rec_X = \{x \mid 0 < x < 1\}$. Se verificará que $f(x)$ es efectivamente una función de densidad.

   Solución: $f(x)$ es una función de densidad cumpliendo que:

   (1) $f(x) = 2x \geqslant 0$, pues $x > 0$

   (2) $\displaystyle\int_0^1 2x\,dx = x^2\Big|_0^1 = 1.$

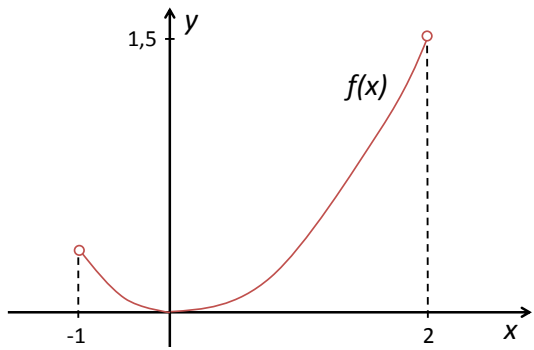
   Por lo tanto, $f(x)$ es una función de densidad.

4. Probar que la siguiente función es de densidad:

$$f(x) = \begin{cases} \frac{1}{3}x^2, & \text{si} \quad -1 < x < 2 \\ 0, & \text{otro caso.} \end{cases}$$

Solución: En el gráfico de la figura 2.6 se muestra la función $f(x)$ definida en el intervalo $(-1, 2)$.

Figura 2.6: Gráfico de la función de densidad de X



Fuente: elaboración propia.

$f(x)$ es una función de densidad, pues cumple que:

(1) $\frac{1}{3}x^2 \geqslant 0$, pues $\frac{1}{3} > 0$ y $x^2 \geqslant 0$.

(2) $\displaystyle\int_{-1}^{2} \frac{1}{3}x^2\,dx = \left.\frac{x^3}{9}\right|_{-1}^{2} = 1.$

5. Sea la variable aleatoria continua X con una función de densidad:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad -\infty < x < +\infty.$$

(a) Probar que $f(x)$ es una función de densidad.

(b) Determinar la fgm de la variable aleatoria X.

(c) A partir de (b), calcular la esperanza y la varianza de X.

Solución: (a) Para verificar si la función propuesta es de densidad, se deben cumplir los supuestos: se tiene que $f(x) > 0$, el otro supuesto se verifica mediante:

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}}\,dx.$$

La función $h(x) = \exp\left\{-\frac{x^2}{2}\right\}$ es una función par (simétrica en torno a cero), por tanto, es válido que $\int_{-\infty}^{\infty} h(x)\,dx = 2\int_{0}^{\infty} h(x)\,dx$. Luego:

$$= \frac{1}{\sqrt{2\pi}} 2 \int_{0}^{\infty} e^{-\frac{x^2}{2}}\,dx.$$

Es posible construir una función gamma mediante el cambio de variable $t = \frac{x^2}{2}$, con lo que se obtiene:

$$= \frac{1}{\sqrt{\pi}} \int_{0}^{\infty} t^{-\frac{1}{2}} e^{-t}\,dt = \frac{1}{\sqrt{\pi}} \Gamma\left(\frac{1}{2}\right) = 1.$$

Con lo que se prueba que $f(x)$ es una función de densidad.

(b) Resuelto anteriormente, véase ejemplo 2.21(a).

(c) A partir de la función generadora de momentos, se obtiene la esperanza y el momento de orden 2 y, con estos dos resultados, se construye la varianza. En efecto:

$$\frac{d}{dt} M_X(t) = t\, e^{\frac{t^2}{2}} \Rightarrow \left.\frac{d}{dt} M_X(t)\right|_{t=0} = E[X] = 0$$

$$\frac{d^2}{dt^2} M_X(t) = e^{\frac{t^2}{2}} + t^2 e^{\frac{t^2}{2}} \Rightarrow \left.\frac{d^2}{dt^2} M_X(t)\right|_{t=0} = E[X^2] = 1.$$

Con estos resultados, $Var[X] = E[X^2] - (E[X])^2 = 1$.

6. Sea la función asociada a la variable aleatoria continua X: $f(x) = c \cdot \theta\, e^{-\theta x}$, $\theta > 0$, $x > 0$. Determinar:

   (a) El valor de c para que $f(x)$ sea una función de densidad.

   (b) La función de probabilidad acumulada de X, es decir, $F_X(t)$.

   (c) La fgm de la variable aleatoria X.

(d) $E[X^n]$ y luego usar para calcular la varianza.

Solución: (a) Para que la función $f(x)$ sea de densidad debe cumplir los supuestos: (1) se tiene que $f(x) > 0 \Rightarrow c > 0$, (2) con el otro supuesto se obtiene el valor exacto de c:

$$\int_0^\infty f(x)\,dx = c\int_0^\infty \theta e^{-\theta x}\,dx = ce^{-\theta x}\Big|_0^\infty = 1.$$

Luego, para $c = 1$, $f(x)$ es una función de densidad.

(b) La probabilidad acumulada de X:

$$F_X(x) = \mathbb{P}(X \leqslant t) = \int_0^t f(x)\,dx = \int_0^t \theta e^{-\theta x}\,dx = e^{-\theta x}\Big|_0^t = 1-e^{-\theta t}.$$

(c) La función generadora de momentos de la variable aleatoria X se obtiene por:

$$\begin{aligned} M_X(t) = E[e^{tX}] &= \int_0^\infty e^{tx}\theta e^{-\theta x}\,dx = \int_0^\infty \theta e^{-(\theta-t)x}\,dx \\ &= -\frac{\theta}{\theta-t}e^{-(\theta-t)x}\Big|_0^\infty = \frac{\theta}{\theta-t} = \left(1-\frac{t}{\theta}\right)^{-1}. \end{aligned}$$

(d) $E[X^n]$ se obtiene aplicando la definición y usando la transformación $z = \theta x$. En efecto:

$$E[X^n] = \int_0^\infty x^n\theta e^{-\theta x}\,dx = \theta\int_0^\infty x^n e^{-\theta x}\,dx = \frac{1}{\theta^n}\,\Gamma(n+1) = \frac{n!}{\theta^n}.$$

Usando el resultado anterior, la esperanza ($n = 1$) y la varianza de X son:

$$E[X] = \frac{1!}{\theta^1} = \theta^{-1}.$$

$$Var[X] = E[X^2] - (E[X])^2 = \frac{2!}{\theta^2} - (\theta^{-1})^2 = \theta^{-2}.$$

7. Sea X una variable aleatoria continua con una función de densidad $f(x) = \frac{1}{2}e^{-|x|}$, para $-1 < x < 1$. Comprobar que: (a) $E[X] = 0$; (b) $Var[X] = 2 - 5e^{-1}$.

Solución: Valor esperado y varianza de X se calcula como:

$$\begin{aligned}E[X] &= \int_{-1}^{1} x\frac{1}{2}e^{-|x|}\,dx = \frac{1}{2}\left(\int_{-1}^{0} x\,e^{x}\,dx + \int_{0}^{1} xe^{-x}\,dx\right)\\ &= \frac{1}{2}\left(xe^{x} - e^{x}\big|_{-1}^{0} - xe^{-x} - e^{-x}\big|_{0}^{1}\right)\\ &= \frac{1}{2}\left(0 - e^{0} - (-1e^{-1} - e^{-1}) - 1e^{-1} - e^{-1} - (-0e^{0} - e^{0})\right)\\ &= \frac{1}{2}\left(-1 + e^{-1} + e^{-1} - e^{-1} - e^{-1} + 1\right) = 0.\end{aligned}$$

Según el resultado anterior, $Var[X] = E[X^2]$, entonces:

$$\begin{aligned}Var[X] = \int_{-1}^{1} x^2\frac{1}{2}e^{-|x|}\,dx &= \frac{1}{2}\left(\int_{-1}^{0} x^2e^{x}\,dx + \int_{0}^{1} x^2e^{-x}\,dx\right)\\ &= \frac{1}{2}\left(x^2e^{x} - 2xe^{x} + 2e^{x}\big|_{-1}^{0} - x^2e^{-x} - 2xe^{-x} - 2e^{-x}\big|_{0}^{1}\right)\\ &= \frac{1}{2}\left[0e^{0} - 0e^{0} + 2e^{0} - (1e^{-1} + 2e^{-1} + 2e^{-1}) - 1e^{-1} -\right.\\ &\quad \left. -2e^{-1} - 2e^{-1} - (-0e^{-0} - 0e^{-0} - 2e^{-0})\right]\\ &= \frac{1}{2}(2 - e^{-1} - 2e^{-1} - 2e^{-1} - e^{-1} - 2e^{-1} - 2e^{-1} + 2)\\ &= \frac{1}{2}(4 - 10e^{-1}) = 2 - 5e^{-1}.\end{aligned}$$

8. Sea $f_X(x) = \dfrac{1}{\sqrt{2\pi\theta^2}}\exp\left\{-\dfrac{1}{2\theta^2}(x-\alpha)^2\right\}$ con $-\infty < x < \infty$, $-\infty < \alpha < \infty$ y $\theta > 0$ (α y θ constantes).

   (a) Probar que la función de densidad de $Z = \dfrac{X-\alpha}{\theta}$ es:

   $$f_Z(z) = \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}z^2\right\},\ \text{con}\ -\infty < z < \infty.$$

   (b) Probar que la función de densidad de $Y = Z^2$ es:

   $$f_Y(y) = \frac{1}{\sqrt{2\pi}}y^{-\frac{1}{2}}e^{-\frac{y}{2}},\ y > 0.$$

Solución: (a) Si se obtiene la función de distribución acumulada de Z se puede obtener su densidad, pues $\frac{d}{dz}F_Z(z) = f_Z(z)$. Así:

$$\begin{aligned}F_Z(z) = \mathbb{P}(Z \leqslant z) &= \mathbb{P}\left(\frac{X-\alpha}{\theta} \leqslant z\right)\\ &= \mathbb{P}(X \leqslant z\theta + \alpha) = F_X(z\theta + \alpha).\end{aligned}$$

Derivando (usando la regla de la cadena), la función de densidad es:

$$f_Z(z) = f_X(z\theta + \alpha) \cdot \theta = \frac{1}{\sqrt{2\pi}\cancel{\theta}} e^{-\frac{z^2}{2}} \cancel{\theta} = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}.$$

El intervalo en que se encuentra z es:

$$-\infty < x < \infty \Rightarrow -\infty < z\theta + \alpha < \infty \Rightarrow -\infty < z < \infty.$$

(b) Se procede de igual manera que en el caso anterior. En efecto:

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \leqslant y) = \mathbb{P}(Z^2 \leqslant y) = \mathbb{P}(|Z| \leqslant \sqrt{y}) \\ &= \mathbb{P}(-\sqrt{y} \leqslant Z \leqslant \sqrt{y}) = \mathbb{P}(Z \leqslant \sqrt{y}) - \mathbb{P}(Z \leqslant -\sqrt{y}) \\ &= F_Z(\sqrt{y}) - F_Z(-\sqrt{y}). \end{aligned}$$

Derivando (aplicando regla de la cadena), la función de densidad es:

$$\begin{aligned} f_Y(y) &= f_Z(\sqrt{y}) \cdot \frac{1}{2\sqrt{y}} + f_Z(-\sqrt{y}) \frac{1}{2\sqrt{y}} \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{(\sqrt{y})^2}{2}} \frac{1}{2\sqrt{y}} + \frac{1}{\sqrt{2\pi}} e^{-\frac{(-\sqrt{y})^2}{2}} \frac{1}{2\sqrt{y}} \\ &= \frac{\cancel{2}}{\sqrt{2\pi}} e^{-\frac{y}{2}} \frac{1}{\cancel{2}\sqrt{y}} = \frac{1}{\sqrt{2\pi}} y^{-\frac{1}{2}} e^{-\frac{y}{2}}. \end{aligned}$$

El intervalo en que se encuentra y se obtiene considerando que: $Y = Z^2 \Rightarrow y > 0$.

9. Sea X una variable aleatoria que representa la concentración, en partes por millón (ppm), de un cierto contaminante del aire. Esta variable tiene una función de densidad:

$$f(x) = (2\alpha)^{-1} \exp\left\{ -\frac{|x - \beta|}{\alpha} \right\}, \ -\infty < x < \infty, \ \alpha > 0, \ \beta \in \mathbb{R}.$$

(a) Encontrar una expresión para $F_X(x)$ (FDA de X). Luego, considerar $\alpha = 1$ y $\beta = 2$ y utilizar F_x para calcular el valor numérico de $\mathbb{P}(X > 4 \mid X > 2)$.

(b) Determinar e interpretar el valor esperado de X.

Solución: (a) La función de densidad presenta un cambio de signo en $x = \beta$, por tanto, la FDA se calcula en dos intervalos reales divididos por el punto de cambio. En efecto, para $x \leqslant \beta$ se tiene:

$$\begin{aligned} F_X(x) &= \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{x} (2\alpha)^{-1} \exp\left\{\frac{t-\beta}{\alpha}\right\} dt \\ &= (2\alpha)^{-1} \exp\left\{\frac{t-\beta}{\alpha}\right\} \alpha \Big|_{-\infty}^{x} = \frac{1}{2} \exp\left\{\frac{x-\beta}{\alpha}\right\}. \end{aligned}$$

Para $x > \beta$ se tiene:

$$\begin{aligned} F_X(x) &= \int_{-\infty}^{\beta} f(t)\,dt + \int_{\beta}^{x} -(2\alpha)^{-1} \exp\left\{-\frac{t-\beta}{\alpha}\right\} dt \\ &= \frac{1}{2} - \frac{1}{2} \exp\left\{-\frac{t-\beta}{\alpha}\right\} \Big|_{\beta}^{x} = 1 - \frac{1}{2} \exp\left\{-\frac{x-\beta}{\alpha}\right\}. \end{aligned}$$

Para $\alpha = 1$ y $\beta = 2$, la *FDA* es:

$$F_X(x) = \begin{cases} \frac{1}{2} \exp\{x-2\}, & \text{si } -\infty < x \leqslant 2 \\ 1 - \frac{1}{2} \exp\{-(x-2)\}, & \text{si } x > 2. \end{cases}$$

Luego, la probabilidad que se pide se calcula por:

$$\begin{aligned} \mathbb{P}(X > 4 \mid X > 2) &= \frac{\mathbb{P}(X > 4 \cap X > 2)}{\mathbb{P}(X > 2)} = \frac{\mathbb{P}(X < 4)}{\mathbb{P}(X > 2)} = \frac{1 - F_X(4)}{1 - F_X(2)} \\ &= \frac{1 - 1 + \frac{1}{2}e^{-(4-2)}}{1 - \frac{1}{2}} = \frac{\frac{1}{2}e^{-2}}{\frac{1}{2}} = 2\frac{1}{2}e^{-2} = 0{,}135. \end{aligned}$$

(b) El valor esperado (usando $\alpha = 1$ y $\beta = 2$):

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} \frac{1}{2} x\, e^{-|x-2|}\,dx \\ &= \frac{1}{2}\left(\int_{-\infty}^{2} x\, e^{x-2}\,dx + \int_{2}^{\infty} x e^{-(x-2)}\,dx \right) \\ &= \frac{1}{2}\left(x e^{x-2} - e^{x-2} \Big|_{-\infty}^{2} - x e^{-(x-2)} - e^{-(x-2)} \Big|_{2}^{\infty} \right) \\ &= \frac{1}{2}\left(2e^0 - e^0 - 0 + 0 + 2e^0 + e^0\right) = \frac{1}{2}(2 - 1 + 2 + 1) = 2. \end{aligned}$$

Es decir, la concentración esperada del contaminante en el aire es de 2 *ppm*.

10. El tiempo de sobrevida en horas de un aparato electrónico es una variable aleatoria continua X con una función de densidad:

$$f(x) = \begin{cases} \frac{4}{5}x^3, & \text{si } 0 \leqslant x < 1 \\ \frac{4}{5}\exp\{1-x\}, & \text{si } x \geqslant 1. \end{cases}$$

Luego, la *FDA* está dada por la expresión:

$$F_X(x) = \begin{cases} 0, & \text{si } x < 0 \\ \frac{1}{5}x^4, & \text{si } 0 \leqslant x < 1 \\ \frac{4}{5}(1-\exp\{1-x\}), & \text{si } x \geqslant 1. \end{cases}$$

2.5. Ejercicios propuestos

1. Considerar la variable aleatoria discreta X con una función de cuantía

$$p(x) = \binom{n}{x} k^x (1-k)^{n-x}, \quad 0 < k < 1 \quad x = 0, 1, 2, ..., n.$$

   (a) Probar que $p(x)$ es una función de cuantía.

   (b) Determinar la función generadora de momentos de X.

   (c) Calcular el valor esperado y la varianza de X.

2. Para el problema 27 (ejercicios propuestos del capítulo 1), suponer que la caja contiene 15 fichas en total, de las cuales 6 son blancas, y que se efectúa el experimento aleatorio de seleccionar 2 fichas. Ahora se define la variable aleatoria X: *número de fichas blancas obtenidas en las 2 extracciones*.

   (a) Obtener el recorrido de la variable aleatoria X.

   (b) Obtener para cada valor del recorrido de X la respectiva probabilidad de ocurrencia.

3. Sea la variable aleatoria continua X con una función de densidad:

$$f(x) = \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, \quad n \in \mathbb{N} \quad x > 0.$$

   (a) Probar que $f(x)$ es una función de densidad.

   (b) Obtener la función de distribución de X.

(c) Determinar la fgm de la variable aleatoria X.

(d) Calcular $E[X^k]$ y usar este resultado para calcular la varianza.

4. La función de distribución de la variable aleatoria X es:

$$F(x) = \begin{cases} 0 & \text{si } x < -1 \\ \frac{1}{3}(x+1) + \frac{2}{3}(x+1)^2 & \text{si } -1 \leqslant x \leqslant 0 \\ 1 & \text{si } x > 0. \end{cases}$$

Determinar: (a) $\mathbb{P}(X > -\frac{1}{3})$; (b) $\mathbb{P}(|X| > 0{,}5)$; (c) $\mathbb{P}(\left|X - \frac{1}{3}\right| < 1)$.

5. Sea X una variable aleatoria continua con:

$$f(x) = \begin{cases} \frac{x}{25} & \text{si} \quad 0 \leqslant x < 5 \\ \frac{10-x}{25} & \text{si} \quad 5 \leqslant x < 10 \\ 0 & \text{otro caso.} \end{cases}$$

(a) Verificar que $f(x)$ es una función de densidad.

(b) Encontrar $F(x)$.

(c) Determinar el valor esperado de X.

6. Sea $F(x)$ la función de distribución de una variable aleatoria X:

$$F(x) = \begin{cases} ae^x, & \text{si } x \leqslant 0 \\ -\frac{1}{2}e^{-x} + b, & \text{si } x > 0. \end{cases}$$

(a) Determinar el (los) valor(es) de a y b, tal que se cumplan las propiedades de la función de distribución.

(b) Determinar la función de densidad de X y (re)definir las condiciones para a y b.

7. Sea $F(x)$ la función de distribución de una variable aleatoria X:

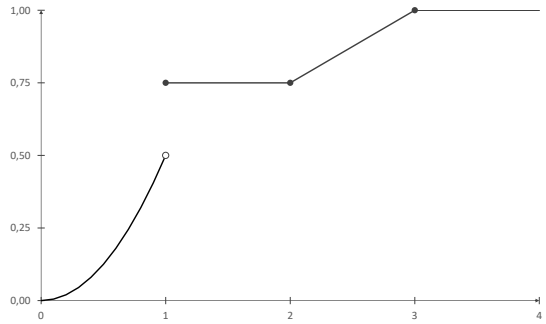
$$F(x) = \begin{cases} 0, & \text{si } x < 0 \\ 2x - x^2, & \text{si } 0 \leqslant x < 1 \\ 1. & \text{si } x > 1. \end{cases}$$

(a) Mostrar que $f(x) = 2 - 2x$, $0 < x < 1$ y verificar que es una función de densidad.

(b) Mostrar que $E[X] = \frac{1}{3}$.

8. El gráfico de la figura 2.7 representa la función de distribución de una variable aleatoria X, la cual en $[0, 1]$ es una función cuadrática y en otro caso es una recta.

Figura 2.7: Gráfico de la función de distribución de X



Fuente: elaboración propia.

(a) Obtener una expresión para la función de distribución.

(b) Obtener, a partir de (a), la función de densidad de X.

(c) Calcular (1) $\mathbb{P}(X < 1)$, (2) $\mathbb{P}(X \leqslant 1)$, (3) $\mathbb{P}(X \geqslant 1)$, (4) $\mathbb{P}(\frac{1}{2} < X < \frac{5}{2})$ y (5) $\mathbb{P}(1 \leqslant X < 4)$.

9. Sea X una variable aleatoria cuya función de densidad se representa en el gráfico de la figura 2.8 (siendo $0 < k < 2$), donde los segmentos de recta unen los puntos $(0, 0)$ y (k, a), y los puntos (k, a) y $(2, 0)$.

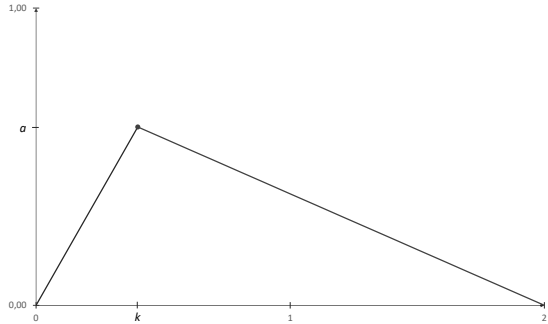
(a) Obtener una expresión para la función de densidad.

(b) Obtener, a partir de (a), el valor esperado y la varianza de X.

10. La variable aleatoria X tiene una función de densidad $f(x) = ax + bx^2$, $0 < x < 1$. Si $E[X] = 0{,}6$, entonces:

(a) Encontrar los valores de a y b, tal que $f(x)$ sea efectivamente una función de densidad.

Figura 2.8: Gráfico de la función de densidad de X



Fuente: elaboración propia.

(b) Calcular $\mathbb{P}(X \leqslant \frac{1}{2})$.

(c) Obtener la varianza de X.

11. Resolver cada una de las siguientes situaciones para determinar la fgm y, a partir de este resultado, obtener la esperanza y la varianza de la variable aleatoria X.

(a) Sea X una variable aleatoria discreta con una función de cuantía $p(x) = 0{,}2 \cdot \frac{5^x e^{-5}}{x!} + 0{,}8\,\alpha(1-\alpha)^x$, $x = 0, 1, 2, \cdots$, $(0 < \alpha < 1$, constante).

(b) Sea X una variable aleatoria continua con una función de densidad $f(x) = 2 \cdot e^{-2x}$, $x > 0$.

12. Sea $f_X(x) = \frac{|x|}{16}$, si $|x| \leqslant 4$ la función de densidad de la variable aleatoria X.

(a) Mostrar que si $Y = X^2$, entonces $f_Y(y) = \frac{1}{16}$ con $y \in (0, 16)$.

(b) Obtener $f_Z(z)$ y $F_Z(z)$, donde $Z = |X|$.

13. En la propiedad 2.5 se realizó un esquema de demostración para el resultado:

$$E[X^k] = \frac{\partial^k}{\partial t^k} M_X(t)\Big|_{t=0}.$$

Ahora, se pide demostrar este resultado considerando la expansión en serie de Taylor de e^{tX} y, con este resultado, obtener las derivadas respectivas y evaluar en $t = 0$; la expresión (convergente) es:

$$e^{tX} = \sum_{i=0}^{\infty} \frac{(tX)^i}{i!}.$$

14. Sea X una variable aleatoria con función de densidad $f(x) = \frac{1}{2},\ -1 < x < 1$. Determinar:

    (a) La función de distribución acumulada de X.

    (b) La función de densidad de la nueva variable aleatoria $Y = X^2$.

15. Sea X una variable aleatoria con función de densidad $f(x) = a(1 - x)$, $0 < x < 1$.

    (a) Determinar a tal que $f(x)$ sea una función de densidad.

    (b) Encontrar la función de densidad de la nueva variable aleatoria $T = \frac{1}{X-1}$.

Capítulo 3

Distribuciones

3.1. Introducción

Considerando que una distribución de probabilidad es un listado de todas las posibilidades de resultados de una variable aleatoria junto con la probabilidad asociada a cada uno de estos resultados, para ciertas variables aleatorias (discretas y continuas) se puede determinar la probabilidad de ocurrencia de los valores determinados por el recorrido mediante una función, la cual es llamada función de distribución.

Las expresiones dadas por las funciones de probabilidad de las distribuciones de variables aleatorias van a depender de ciertas constantes, que se llamarán parámetros de la distribución y surgen de la naturaleza e interés de los experimentos de estudio. Habitualmente los parámetros determinan tres medidas que describen el comportamiento de la población para la característica de interés X. Estas son: posición, escala o dispersión y forma o asimetría.

Es decir, las distribuciones de probabilidad son aproximaciones teóricas a la realidad poblacional, teniendo en cuenta que la realidad es compleja y puede asumir múltiples formas que son muy difíciles de representar mediante un modelo matemático.

Además, como un contenido optativo, este capítulo finaliza con un apartado dedicado a trabajar las distribuciones de variables aleatorias con MS-Excel. Dicha herramienta entrega la oportunidad de realizar cálculos de probabilidades y de densidades.

Definición **3.1** Como se ha anticipado, las distribuciones de variables aleatorias que se estudiarán a continuación responden a situaciones particulares que habitualmente reciben un nombre propio, por lo que adoptaremos una notación para resumir la distribución de una variable aleatoria X:

$$X \sim \text{Distribución}(\boldsymbol{\theta}),$$

donde X es la variable aleatoria, $\sim$ es un símbolo que se lee como *distribución* o *distribuye* y $\boldsymbol{\theta}$ es un vector que contiene los parámetros que especifican la distribución de probabilidad de la variable aleatoria.

Si se sabe que la variable aleatoria X sigue una cierta distribución, entonces esta tiene asociados algunos elementos: función de cuantía o de densidad, la función de distribución, función generadora de momentos, valor esperado (si existe) y varianza. Al contrario, se puede identificar una distribución para una variable aleatoria X si se conoce la función de probabilidad o la función generadora de momentos; no así el valor esperado y la varianza, que no son suficientes para identificar una distribución específica.

En resumen:

$$\begin{aligned} X \sim \text{Distribución}(\boldsymbol{\theta}) &\Leftrightarrow p(x), f(x); F(x); M_X(t) \\ &\Rightarrow E[X], Var[X]. \end{aligned}$$

3.2. Distribuciones de variables aleatorias discretas

Cuando el recorrido de la variable derivado del experimento es de carácter discreto (toma valores finitos o infinitos numerables), es decir, el recorrido está contenido en el conjunto de los enteros, se habla de una distribución de variables aleatorias discretas. A continuación, se hace un recorrido por algunas de las distribuciones usuales, para las cuales se especifica el experimento asociado y la función de cuantía, a partir de la cual se deriva la función generadora de momentos para obtener la esperanza y la varianza.

Distribución de Bernoulli

Definición **3.2** Un experimento se dice que es de tipo Bernoulli (o ensayo de Bernoulli) si solo consta de dos resultados posibles mutuamente excluyentes, éxito y fracaso, con una cierta probabilidad en la población de estudio. Por tratarse de eventos complementarios, éxito y fracaso, la suma de las probabilidades asociadas a cada uno de estos eventos debe ser 1.

En efecto, supongamos que la probabilidad de éxito de un experimento realizado bajo las condiciones de un ensayo de Bernoulli es p, y la probabilidad de fracaso es $q = 1 - p$. Es decir:

$$p(x) = \begin{cases} p & \text{si} \quad x = 1 \quad \text{(éxito)} \\ 1-p & \text{si} \quad x = 0 \quad \text{(fracaso).} \end{cases}$$

Luego, $p(x) = p^x(1-p)^{1-x}$, $x = 0, 1$ es la función de cuantía asociada a la variable aleatoria discreta X que distribuye Bernoulli de parámetro p. Esto se denota con $X \sim \text{Bernoulli}(p)$.

A partir de la función de cuantía se puede obtener la función generadora de momentos y, de ese resultado, obtener el valor esperado y la varianza:

$$\begin{aligned} M_X(t) &= \sum_{x=0}^{1} e^{tx} p(x) = e^{t\cdot 0}p(0) + e^{t\cdot 1}p(1) = 1 - p + pe^t = q + pe^t. \\ E[X] &= \frac{\partial}{\partial t} M_X(t)\Big|_{t=0} = p. \\ E[X^2] &= \frac{\partial^2}{\partial t^2} M_X(t)\Big|_{t=0} = p. \\ Var[X] &= p - p^2 = p \cdot (1-p) = p \cdot q. \end{aligned}$$

Distribución binomial

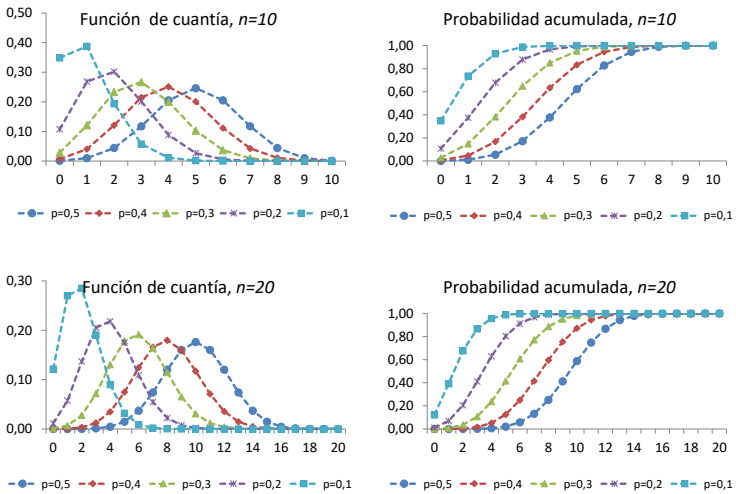
La distribución binomial consiste en la repetición de n ensayos de Bernoulli, es decir, cada una de las n repeticiones del experimento tiene dos resultados posibles, mutuamente excluyentes, éxito y fracaso. Además, las repeticiones son independientes entre sí y tienen una probabilidad de éxito, p, constante de ensayo en ensayo.

Definición **3.3** La variable aleatoria X, que mide el número de éxitos que aparecen en los n ensayos, pertenece a la familia de las distribuciones discretas. Entonces, X sigue una distribución binomial si es el resultado de n experimentos de Bernoulli; la notación y la función de cuantía son:

$$X \sim \text{Binomial}(n, p),\ n > 0,\ 0 < p < 1$$
$$\Leftrightarrow p(x) = \binom{n}{x} p^x (1-p)^{n-x},\ x = 0, 1, 2, \cdots, n.$$

La distribución depende de la cantidad de repeticiones del ensayo de Bernoulli, n, y del parámetro que indica la proporción de éxitos, p. El gráfico de la figura 3.1 muestra, para $n = 10$ y $n = 20$, la función de cuantía y la función de probabilidad acumulada considerando p igual a 0,1; 0,2; 0,3; 0,4; 0,5.

Figura 3.1: Función de cuantía y función de distribución



Fuente: elaboración propia.

Ejemplo **3.1** En este ejemplo se obtendrá la función de cuantía para la distribución binomial a partir de la definición del experimento aleatorio. La variable aleatoria considera la probabilidad de ocurrencia de x éxitos

entre n ensayos. Para obtener la función de cuantía, se obtendrá primero la probabilidad de ocurrencia de x éxitos en los primeros ensayos, entonces si se tienen n ensayos y los primeros x son éxitos, la probabilidad de ocurrencia es:

$$\mathbb{P}(X = x) = p^x(1-p)^{n-x}.$$

Luego, todas las posibles combinaciones para obtener x éxitos entre los n ensayos, se obtienen por:

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}.$$

Así, la probabilidad de ocurrencia de x éxitos entre los n ensayos Bernoulli con probabilidad de éxito p, que corresponde a la función de cuantía de una variable aleatoria X de distribución binomial, es:

$$\mathbb{P}(X = x) = p(x) = \binom{n}{x} p^x(1-p)^{n-x}, \quad x = 0, 1, 2, \cdots, n. \square$$

Propiedad **3.1** La función generadora de momentos de la distribución binomial, se obtiene a partir de la función de cuantía como sigue:

$$\begin{aligned} M_X(t) = E(e^{tX}) &= \sum_{x=0}^{n} e^{tx} p(x) \\ &= \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x(1-p)^{n-x} \\ &= \sum_{x=0}^{n} \binom{n}{x} (pe^t)^x (1-p)^{n-x}. \end{aligned}$$

Esta expresión corresponde a la forma del binomio de Newton para $a = pe^t$ y $b = (1-p)$, entonces:

$$M_X(t) = (1 - p + pe^t)^n = (q + pe^t)^n.$$

Por su parte, de la función generadora de momentos se puede obtener el valor esperado y la varianza:

$$\begin{aligned} E[X] &= \frac{d}{dt} M_X(t)\Big|_{t=0} = np. \\ E[X^2] &= \frac{d^2}{dt^2} M_X(t)\Big|_{t=0} = n(n-1)p^2 + np \\ Var[X] &= E[X^2] - \{E[X]\}^2 = np(1-p) = n \cdot p \cdot q. \end{aligned}$$

Ejemplo **3.2** Suponer que un estudiante rinde una prueba de 10 preguntas de verdadero y falso y que además este estudiante no está preparado y adivina las respuestas (es decir, selecciona las respuestas al azar). Considerando la variable aleatoria X: *número de respuestas contestadas correctamente por el estudiante*, determinar:

(a) La distribución de probabilidad de la variable aleatoria X.

(b) La probabilidad de que conteste correctamente todas las preguntas.

(c) La probabilidad de que conteste correctamente 5 preguntas.

(d) La probabilidad de que apruebe el examen, si este se aprueba contestando al menos 7 preguntas correctamente.

(e) El número esperado de preguntas contestadas correctamente y la variabilidad.

(f) La probabilidad de que obtenga un puntaje de al menos el 40 %.

Solución: (a) Cada pregunta de la prueba es un ensayo de Bernoulli, pues tiene dos respuestas posibles: *éxito* (respuesta correcta) y *fracaso* (respuesta incorrecta). Por la construcción de la prueba, la probabilidad de éxito y fracaso debe ser igual y constante para cada una de las 10 preguntas. Si X: *número de respuestas correctas*, entonces: $X \sim \text{Binomial}(n = 10, p = 0{,}5)$ con:

$$p(x) = \binom{10}{x} \cdot 0{,}5^x \cdot 0{,}5^{10-x} = \binom{10}{x} \cdot 0{,}5^{10}.$$

(b) $p(10) = \binom{10}{10} \cdot 0{,}5^{10} = 0{,}5^{10} = 0{,}00098$.
(c) $p(5) = \binom{10}{5} \cdot 0{,}5^{10} = 0{,}246$.
(d) $\mathbb{P}(X \geqslant 7) = \sum_{x=7}^{10} \cdot \binom{10}{x} \cdot 0{,}5^{10} = 0{,}17188$.
(e) $E[X] = n \cdot p = 10 \cdot 0{,}5 = 5$ y $V(X) = n \cdot p \cdot q = 10 \cdot 0{,}5 \cdot 0{,}5 = 2{,}5$.
(f) $\mathbb{P}(X \geqslant 4) = 0{,}828$. □

Distribución geométrica

Sea el experimento de realizar repeticiones de ensayos tipo Bernoulli hasta obtener un éxito, bajo las condiciones de independencia entre los eventos y

de probabilidad de éxito constante de ensayo en ensayo. La variable aleatoria que mide este tipo de experimentos se dice que sigue una distribución geométrica.

Definición **3.4** $X \sim \text{Geométrica}(p) \Leftrightarrow p(x) = (1-p)^{x-1} \cdot p,\ x = 1, 2, \cdots$

Ejemplo **3.3** En este ejemplo se obtendrá la función de cuantía para la distribución geométrica a partir de la definición del experimento aleatorio. La función de cuantía se puede derivar obteniendo resultados para algunos casos particulares y generalizando para x. Para esto hay que tener en cuenta que $\mathbb{P}(X = i)$ es la probabilidad de que en el i-ésimo ensayo se produzca el primer éxito y que la probabilidad de éxito es p. Por tanto:

$$\begin{aligned}
&\mathbb{P}(X = 1) = p \\
&\mathbb{P}(X = 2) = p(1-p) \\
&\mathbb{P}(X = 3) = p(1-p)(1-p) \\
&\qquad\vdots \\
&\mathbb{P}(X = x) = p\underbrace{(1-p)\cdots(1-p)}_{x-1 \text{ veces}} = p\,(1-p)^{x-1}.
\end{aligned}$$

□

Propiedad **3.2** A partir de la función de cuantía se puede obtener la función generadora de momentos y, a su vez, el valor esperado y la varianza de la variable aleatoria con distribución geométrica.

(a) $M_X(t) = pe^t(1 - qe^t)^{-1}$.

(b) $E[X] = p^{-1}$.

(c) $Var[X] = (1-p)p^{-2}$.

Demostración: (a) La función generadora de momentos:

$$M_X(t) = E(e^{tX}) = \sum_{x=1}^{\infty} e^{tx} pq^x = p \sum_{x=0}^{\infty} (qe^t)^x + \frac{p}{q},$$

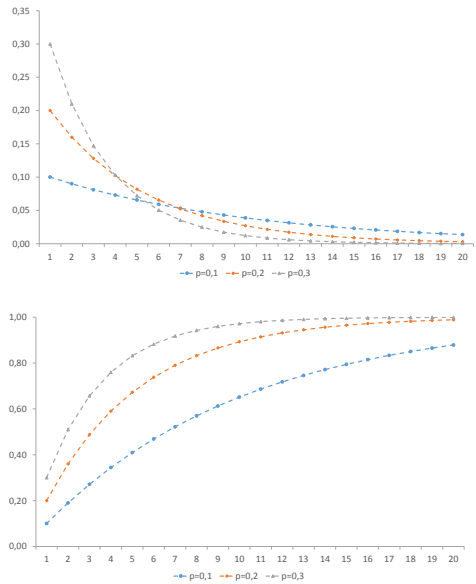
la convergencia se garantiza para $q \cdot e^t < 1$:

$$= \frac{p}{q} \frac{1}{(1 - qe^t) + 1} = \frac{p}{\not{q}} \frac{\not{q}e^t}{(1 - qe^t)}$$
$$= \frac{pe^t}{1 - qe^t}.$$

(b) y (c) Valor esperado y varianza quedan de ejercicio para el lector. Indicación: véase propiedad 2.5 y la definición 2.7.

Ejercicio propuesto **3.1** Construir gráficos de una variable aleatoria X con distribución geométrica, con $p = 0{,}1$, $p = 0{,}2$ y $p = 0{,}3$, para la función de cuantía y la función de distribución acumulada. Esbozar los gráficos y luego utilizar MS-Excel para el gráfico exacto (este último se aprecia en la figura 3.2).

Figura 3.2: Función de cuantía y función de distribución de X



Fuente: Elaboración propia.

Ejemplo **3.4** Se sabe que en un cierto proceso de fabricación, en promedio, 1 de cada 100 artículos está defectuoso. ¿Cuál es la probabilidad de que

el quinto artículo que se inspecciona sea el primero defectuoso que se encuentra?

Solución: Se define la variable aleatoria X: *número de inspecciones necesarias hasta obtener el primero defectuoso*, donde la probabilidad de encontrar un artículo defectuoso es de 0,01. Por tanto: $X \sim$ Geométrica$(p = 0{,}01)$.

Finalmente, la probabilidad de que el quinto artículo que se inspecciona sea el primero defectuoso que se encuentra es:
$\mathbb{P}(X = 5) = 0{,}01(1 - 0{,}01)^{5-1} = 0{,}0096 = 0{,}1\,\%$. □

Distribución binomial negativa

En experimentos donde es de interés determinar la cantidad de ensayos (independientes) necesarios para observar el k-ésimo éxito, se dice que los valores observados tienen o siguen una distribución de probabilidad binomial negativa.

Definición **3.5** En experimentos donde es de interés determinar el número de ensayos de tipo Bernoulli independientes que son necesarios para obtener el k-ésimo éxito, con probabilidad p (y probabilidad de fracaso $q = 1-p$), $k = 2, 3, \cdots$, la variable aleatoria sigue una distribución llamada binomial negativa ($Bneg$) con parámetros k y p, es decir: $X \sim \text{Bneg}(k, p)$.

Definición **3.6** $X \sim \text{Bneg}(k, p) \Leftrightarrow p(x) = \binom{x-1}{k-1} p^k q^{x-k}$, $x = k, k + 1, ...,$ $k = 1, 2, \cdots$.

Ejemplo **3.5** Obtener la función de cuantía para la distribución binomial negativa a partir de la definición del experimento aleatorio.

Solución: De acuerdo con la definición del experimiento aleatorio para esta distribución, el recorrido de la variable aleatoria X es: $k, k + 1, \cdots$ y, en cada realización del experimento, la probabilidad de éxito es p. Con estos datos se obtiene la probabilidad para los valores del recorrido, los cuales son:

$X = k$: necesariamente cada uno de los primeros k ensayos son éxitos, entonces $\mathbb{P}(X = k) = p(k) = p^k$.

$X = k+1$: el k-ésimo éxito debe ocurrir en el ensayo $k+1$, por lo que debe haber exactamente $k-1$ éxitos en los primeros ensayos. Entonces $\mathbb{P}(X = k+1) = p(k+1) = \binom{k}{k-1} p^k \cdot q$.

$X = x$: el k-ésimo éxito debe ocurrir en el ensayo x y debe haber exactamente $k-1$ éxitos en los primeros $x-1$ ensayos. De este análisis se obtiene la función de cuantía de una variable aleatoria $X \sim \text{Bneg}(k, p)$:

$$\mathbb{P}(X = x) = p(x) = \binom{x-1}{k-1} p^k q^{x-k}, \quad x = k, k+1, \cdots . \ \square$$

Propiedad **3.3** La función generadora de momentos, el valor esperado y la varianza de X están dados, respectivamente, por:

(a) $M_X(t) = \left(p\, e^t (1 - q\, e^t)^{-1}\right)^k$.

(b) $E[X] = \frac{k}{p}$, $k = 2, 3, \cdots$.

(c) $Var[X] = k \cdot q \cdot p^{-2}$, $k = 2, 3, \cdots$.

Ejercicio propuesto **3.2** Demostrar los resultados para la propiedad 3.3.

Indicación: Usar el siguiente resultado:

$$(1-t)^{-n} = \sum_{i=0}^{\infty} \binom{i+n-1}{i} t^i.$$

Observación **3.1** Si $k = 1$, entonces la distribución binomial negativa es equivalente a la distribución geométrica. Este resultado es claro, pues la variable aleatoria con distribución binomial negativa, en este caso, mide el número de prueba hasta que ocurre el primer ($k = 1$) éxito. Para evitar esta situación, algunos textos definen la distribución binomial negativa para valores de k superiores a 1, es decir, $k = 2, 3, \cdots$.

Ejemplo **3.6** Si la probabilidad de que un niño expuesto a una enfermedad contagiosa la contraiga es de 0,4, ¿cuál es la probabilidad de que el décimo niño expuesto a la enfermedad sea el tercero en contraerla?

Solución: Si se considera la variable aleatoria X: *número de niños expuestos a una enfermedad contagiosa hasta que ocurra el tercer contagiado*, por tanto:

$$X \sim \text{Bneg}(k = 3; p = 0{,}4) \Rightarrow p(x) = \binom{x-1}{2} \cdot 0{,}4^3 \cdot 0{,}6^{x-3}, \quad x = 3, 4, \cdots$$

Así, $\mathbb{P}(X = 10) = \binom{9}{2} \cdot 0{,}4^3 \cdot 0{,}6^7 = 0{,}065$.

Distribución hipergeométrica

Supóngase que para cierta característica se puede dividir al total de la población de estudio, N, en dos grupos. El *grupo I* tiene m elementos y, por su parte, el *grupo II* tiene los restantes $N - m$ elementos. El experimento consiste en extraer n elementos de la población (uno a uno y sin reemplazo) y enseguida contar la cantidad de éxitos encontrados (si un elemento pertenece al *grupo I* se considera un éxito; por tanto, si es del *grupo II* se considera un fracaso).

Definición **3.7** La variable aleatoria X, que cuenta el número de éxitos entre n objetos elegidos sin reemplazo de una población antes caracterizada, sigue una distribución llamada *hipergeométrica*, cuyos parámetros son el tamaño de la población: N, el número de elementos del *grupo I* de la población: m, y la cantidad de elementos que se seleccionan de la población: n. La siguiente es la función de cuantía y se presentan gráficos para distintos valores de los tamaños de población, población de éxito y muestra:

$$X \sim \text{Hiperg}(N, m, n) \Leftrightarrow \mathbb{P}(X = x) = \frac{\binom{m}{x}\binom{N-m}{n-x}}{\binom{N}{n}}, \quad x = 0, \cdots, \text{mín}(m, n).$$

Propiedad **3.4** El valor esperado y la varianza de X son, respectivamente:

(a) $E[X] = n \cdot p$, donde $p = \frac{m}{N}$.

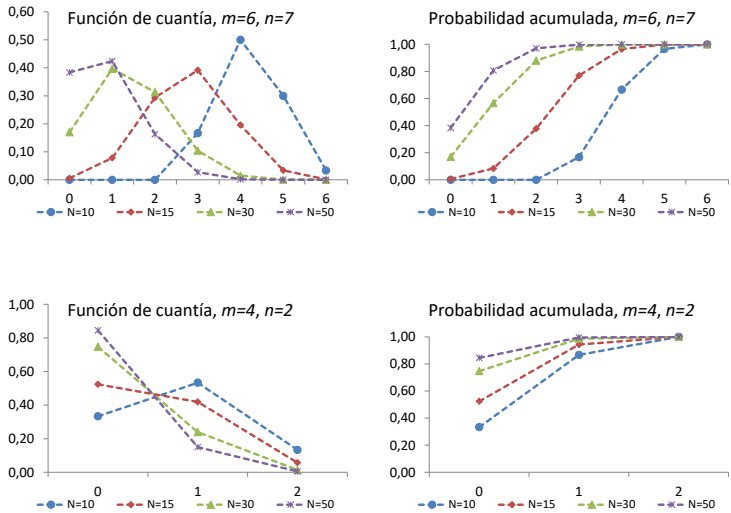
(b) $Var[X] = n \cdot p \cdot q\frac{N-n}{N-1}$, donde $p = \frac{m}{N}$ y $q = 1 - \frac{m}{N}$.

Los gráficos de la figura 3.3 muestran el comportamiento de la función de cuantía $f(x)$ y la función de distribución $F(x)$ para distintos valores de los parámetros de la distribución, de tamaños muestrales y poblacionales.

Ejemplo **3.7** Un lote de 20 artículos es aceptado o rechazado basándose en una inspección de 4 extracciones al azar. Se acepta el lote si a lo más hay un producto defectuoso. ¿Cuál es la probabilidad de que el lote sea rechazado si se sabe que 2 de los 20 productos están en mal estado?

Solución: Para este ejemplo, se mide la variable aleatoria X: *número*

Figura 3.3: Función de cuantía y función de distribución de $X \sim$ Hiperg



Fuente: elaboración propia.

de artículos defectuosos en las 4 elecciones. Por tanto:

$$X \sim \text{Hiperg}(N = 20, m = 2, n = 4)$$
$$\Rightarrow \mathbb{P}(X = x) = p(x) = \frac{\binom{2}{x}\binom{18}{4-x}}{\binom{20}{4}}, \quad x = 0, 1, 2.$$

En este caso, $\mathbb{P}(X > 1) = \mathbb{P}(X = 2) = p(2) = \frac{\binom{2}{2}\binom{18}{2}}{\binom{20}{4}} = 0{,}0316.$ □

Distribución de Poisson

Existen experimentos en los que interesa determinar con qué probabilidad ocurren eventos como el número de averías de una cierta máquina en una jornada de trabajo, el número de partículas emitidas por un átomo radiactivo en t segundos, el número de errores tipográficos en una revista, el número de llamadas que ingresan a una central telefónica en un periodo de tiempo determinado, etc. Es decir, en este tipo de experimentos interesa asignarle una probabilidad de ocurrencia a un evento en un periodo de

tiempo fijo o en una región determinada.

Aquellos experimentos que cumplen las condiciones anteriores se ajustan a un proceso de Poisson, que se denota con $X \sim \text{Poisson}(\lambda)$.

Definición **3.8** Una variable aleatoria X sigue una distribución de Poisson de parámetro λ ($\lambda > 0$) si su función de densidad es:

$$p(x) = \mathbb{P}(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2, \cdots$$

Propiedad **3.5** La función generadora de momentos, el valor esperado y la varianza de X son, respectivamente:

(a) $M_X(t) = e^{\lambda(e^t-1)}$.

(b) $E[X] = \lambda$.

(c) $Var[X] = \lambda$.

Demostración: Sea $X \sim \text{Poisson}(\lambda)$:

(a) $M(t) = E\left[e^{tX}\right] = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!} = e^{-\lambda} e^{\lambda e^t} = e^{\lambda(e^t-1)}$.

(b) $E[X] = \left.\frac{d}{dt} M(t)\right|_{t=0} = \lambda$.

(c) $E[X^2] = \left.\frac{d^2}{dt^2} M(t)\right|_{t=0} = \lambda^2 + \lambda$. Así, $Var[X] = \lambda$.

Ejemplo **3.8** Suponer que los clientes que llegan a una fila de espera lo hacen en una tasa de 4 por minuto y que este proceso de llegada ocurre de acuerdo con un proceso de Poisson. Determinar la probabilidad de que al menos una persona llegue a la fila en un intervalo de medio minuto.

Solución: Considerar la variable aleatoria X: *número de personas que llegan a una fila en medio minuto*. De acuerdo con la definición de la variable aleatoria, se debe determinar la tasa de personas que llegan a la fila en medio minuto; considerando que es proporcional y constante, entonces la tasa de personas que llegan a la fila en medio minuto es de 2. Por tanto, $X \sim \text{Poisson}(\lambda = 2)$ y la función de cuantía está dada por:

$$p(x) = \mathbb{P}(X = x) = \frac{e^{-2}2^x}{x!}, \quad x = 0, 1, 2, \cdots,$$

de donde la probabilidad de que llegue al menos una persona a la fila en el intervalo de medio minuto es:

$$\sum_{x=1}^{\infty} p(x) = \mathbb{P}(X \geqslant 1) = 1 - \mathbb{P}(X = 0) = 1 - \frac{e^{-2}2^0}{0!} = 1 - e^{-2} = 0{,}865.\square$$

Aproximación a la distribución binomial por la distribución de Poisson

Es conveniente realizar esta aproximación cuando se tiene una gran cantidad de ensayos de Bernoulli y una probabilidad de éxito en cada ensayo muy baja (cercana a cero). En ese caso, la distribución binomial se aproxima bien mediante la distribución de Poisson, donde λ es el producto de n y p, la cantidad de ensayos y la probabilidad de éxito de cada uno. Es decir:

Sea $X \sim \text{Binomial}(n, p)$ donde $n \to \infty$ y $p \to 0$, entonces: $X \dot{\sim} \text{Poisson}(\lambda = np)$, donde $\dot{\sim}$ denota que es una distribución aproximada.

Observación **3.2** En la práctica, se puede obtener una buena aproximación para $n > 30$ y $p \leqslant 0{,}1$. En el mejor de los escenarios, si $n > 100$ y $p \leqslant 0{,}1$, se obtiene una aproximación muy buena.

Ejemplo **3.9** En un sistema de control de la calidad de un producto terminado, un experto estima que hay una probabilidad de 0,001 de encontrar un artículo defectuoso durante un periodo de tiempo de 5 minutos en una estación de la cadena de producción continua. Es de interés determinar la probabilidad de: encontrar exactamente ningún artículo defectuoso y de encontrar un artículo defectuoso, en un experimento en que se revisa la aparición de artículos defectuosos en 100 periodos de 5 minutos.

Solución: Sea X: *número de artículos defectuosos observados en 100 periodos al azar de 5 minutos*, entonces $X \sim \text{Binomial}(n = 100; p = 0{,}001)$, con lo cual se obtiene que:

$\mathbb{P}(X = 0) = 0{,}9048$: la probabilidad de no encontrar artículos defectuosos es de 90,48 %, y

$\mathbb{P}(X = 1) = 0{,}0906$: la probabilidad de encontrar un artículo defectuoso es de 9,06 %.

Nótese que en este ejemplo n es grande y p pequeño, usando la distribución de Poisson dada por $X \sim \text{Poisson}(\lambda = np = 0{,}1)$, entonces la probabilidad de *no encontrar artículos defectuosos* y de *encontrar un artículo defectuoso* es, respectivamente:

$\mathbb{P}(X = 0) = 0{,}9048$ y $\mathbb{P}(X = 1) = 0{,}0905$. Es decir, 90,48 % y 9,05 %, respectivamente, que corresponden a valores cercanos a los obtenidos utilizando la distribución binomial. □

3.3. Distribuciones de variables aleatorias continuas

Distribución normal

La distribución normal constituye un tipo de distribución de probabilidad de más amplio uso práctico y estudio teórico durante el desarrollo de la estadística y la probabilidad. Esta amplitud de uso está dada porque muchos fenómenos reales se pueden aproximar mediante esta distribución; de ahí que su estudio se masificara y sea hoy fuente obligada para adentrarse (o al menos para iniciarse) en la inferencia estadística y en el estudio de la modelación estadística.

Desde el punto de vista gráfico, la distribución es simétrica en torno a un punto (llamado parámetro de localización) donde se localiza la media, la mediana y la moda. Por esta razón la forma es de campana (el gráfico de la distribución normal es llamado *campana de Gauss*). La extensión o amplitud de la distribución hacia ambos extremos es similar y se determina por un parámetro llamado de escala o de dispersión. Estos dos parámetros, de posición y de escala, son los que determinan la distribución.

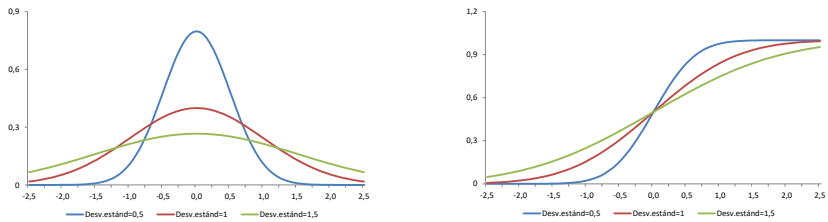
Definición **3.9** Sea X una variable aleatoria continua que toma valores en todos los reales y sean μ y σ^2 los parámetros de posición y escala, respectivamente, entonces:

$$X \sim N(\mu, \sigma^2) \Leftrightarrow f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\},$$
$$\text{donde} -\infty < x < \infty,\ \mu \in \mathbb{R} \text{ y } \ \sigma^2 > 0.$$

La densidad $f(x)$ es una función simétrica en torno a μ, por tanto, $\mathbb{P}(X < \mu) = 0{,}5$ y $\mathbb{P}(X > \mu) = 0{,}5$. En la figura 3.4 se aprecia el

gráfico de una variable aleatoria con distribución normal media 0 y la desviación estándar para tres valores. Aquí es posible observar que valores más grandes de la desviación estándar implican mayor dispersión de los datos; algo parecido se aprecia respecto del aumento de la curva para la función de distribución acumulada.

Figura 3.4: Función de densidad y función de distribución $X \sim N(0, \sigma_0^2)$



Fuente: elaboración propia.

A continuación, se entregan algunos resultados para caracterizar la distribución normal asociados a la función generadora de momentos y los parámetros de posición y de escala.

Propiedad **3.6** La función generadora de momentos, el valor esperado y la varianza son:

(a) $M_X(t) = \exp\left\{t\mu + t^2\,\frac{\sigma^2}{2}\right\}$.

(b) $E[X] = \mu$.

(c) $Var[X] = \sigma^2$.

Demostración: (a) Para obtener la función generadora de momentos de la variable aleatoria X, se aplica la definición y se resuelve la integral. En efecto, se tiene que:

$$M_X(t) = E\left[e^{t\,X}\right] = \int_{-\infty}^{\infty} \exp\{t\,x\}\,f(x)\,dx. \tag{3.1}$$

Resolviendo la ecuación 3.1 se puede obtener la función generadora de momentos de X.

(b) y (c) A partir de la primera y la segunda derivada parcial de la función generadora de momentos de X respecto a t y evaluadas en $t = 0$, se puede obtener una expresión para el valor esperado y la varianza. En efecto:

$$\frac{d}{dt}M_X(t) = \left(\mu + t\sigma^2\right) \cdot \exp\left\{t\mu + t^2\frac{\sigma^2}{2}\right\}.$$
$$\frac{d^2}{dt^2}M_X(t) = \left[\mu\left(\mu + t\sigma^2\right) + \sigma^2 + t\sigma^2\left(\mu + t\sigma^2\right)\right] \cdot \exp\left\{t\mu + t^2\frac{\sigma^2}{2}\right\}.$$

Evaluando ambas expresiones obtenidas anteriormente en $t = 0$, se tiene que el valor esperado y la varianza son:

$$E[X] = \mu \text{ y } Var[X] = \sigma^2.$$

Distribución normal estándar

Existe un caso particular de la distribución normal en que $\mu = 0$ y $\sigma^2 = 1$. Esta es la llamada distribución normal estándar, la cual presenta ventajas particularmente para el cálculo de probabilidades.

Propiedad **3.7** Si $X \sim \mathrm{N}(\mu, \sigma^2)$, entonces $Z = g(X) = \frac{X-\mu}{\sigma} \sim \mathrm{N}(0, 1)$. Es decir, la nueva variable aleatoria Z sigue una distribución *normal estándar* (con media 0 y varianza 1). Nótese que la función de densidad de Z es:

$$f(z) = \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}z^2\right\} \quad -\infty < z < +\infty.$$

Demostración: La idea será obtener la función generadora de momentos de la variable aleatoria Z y ver que tenga la forma de la distribución normal estándar, de esta manera:

$$\begin{aligned} M_Z(t) =& E[e^{tZ}] = E\left[\exp\left\{t\,\frac{X-\mu}{\sigma}\right\}\right] = \exp\left\{-\frac{t\mu}{\sigma}\right\}E\left[e^{\frac{t}{\sigma}}\right] = \\ &= \exp\left\{-\frac{t\mu}{\sigma}\right\}M_X(\tfrac{t}{\sigma}) = \cancel{\exp\left\{-\frac{t\mu}{\sigma}\right\}}\exp\left\{\cancel{\frac{t}{\sigma}}\mu + \frac{t^2}{\cancel{\sigma^2}}\cdot\frac{\cancel{\sigma^2}}{2}\right\} = \\ &= \exp\left\{\frac{t^2}{2}\right\}. \end{aligned}$$

Con este resultado y el hecho de que en la propiedad 3.6(a) si $\mu = 0$ y $\sigma^2 = 1$, entonces se obtiene que $Z \sim N(0, 1)$.

Tal como se anticipó anteriormente y después del resultado de la propiedad 3.7 queda de manifiesto, dado que cualquiera sea el valor numérico de los parámetros μ y σ^2, la transformación lleva a una variable aleatoria con parámetros 0 y 1, respectivamente; luego el cálculo de probabilidades para esta distribución puede realizarse a través de $F_Z(z)$. La siguiente definición permite obtener valores para la probabilidad de Z.

Definición **3.10** Se define la función $\Phi(\cdot)$ como:

$$\Phi(z) = F_Z(z) = \mathbb{P}(Z \leqslant z).$$

Además, se puede mostrar que $\Phi(z)$ es una función que admite inversa (es 1 a 1), con lo cual se cumple que:

$$\Phi(z) = \alpha \Leftrightarrow z = z_\alpha = \Phi^{-1}(\alpha).$$

Definición **3.11** Sea la variable aleatoria $X \sim N(\mu, \sigma^2)$, el resultado para Z de la propiedad 3.7 y la definición 3.10, se obtiene:

$$\begin{aligned} F_X(x) = \mathbb{P}(X \leqslant x) &= \mathbb{P}\left(\frac{X-\mu}{\sigma} \leqslant \frac{x-\mu}{\sigma}\right) \\ &= \mathbb{P}\left(Z \leqslant \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right). \end{aligned}$$

Es decir, a través de la función $\Phi(\cdot)$ se permite el cálculo de probabilidades para una variable aleatoria con distribución normal con cualquier valor de μ y σ^2, utilizando la distribución normal estándar.

La importancia del resultado de la definición 3.11 es que, si se cuenta con los valores de la imagen de Φ para todos los valores de z (argumento de la función), se tendrían los valores tabulados para una variable aleatoria con distribución normal estándar. Además, como la función $\Phi(\cdot)$ admite inversa, es posible obtener un valor dentro del recorrido de la variable aleatoria Z dado por una probabilidad.

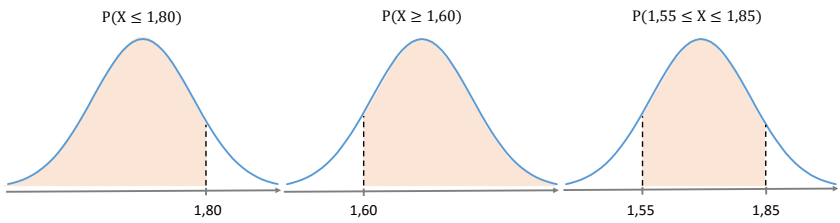
Este valor puede ser llamado percentil o puntaje z, por ejemplo, el percentil 85 está definido por $z = 1{,}03$, o sea $\Phi(1{,}03) = 0{,}85$. Además, como la función Φ admite inversa, entonces, $\Phi^{-1}(0{,}85) = 1{,}03$. Nótese que otra forma de escribir el resultado anterior es $z_{0{,}85} = \Phi^{-1}(0{,}85) = 1{,}03$, que simplifica la escritura reemplazando la función inversa por la variable aleatoria y la probabilidad acumulada como subíndice.

Ejemplo **3.10** Sea X una variable aleatoria $X \sim N(\mu = 1{,}75 \text{ y } \sigma^2 = 0{,}5^2)$. Se desea calcular (a) $\mathbb{P}(X \leqslant 1{,}80)$, (b) $\mathbb{P}(X \geqslant 1{,}60)$ y (c) $\mathbb{P}(1{,}55 \leqslant X \leqslant 1{,}85)$.

Solución: Según el gráfico (véase figura 3.5) y usando la función $\Phi(\cdot)$:

$$\text{(a) } \mathbb{P}(X \leqslant 1{,}80) = \mathbb{P}\left(Z \leqslant \frac{1{,}80 - 1{,}75}{0{,}5}\right) =$$
$$= \mathbb{P}(Z \leqslant 0{,}1) = \Phi(0{,}1) = 0{,}5398.$$
$$\text{(b) } \mathbb{P}(X \geqslant 1{,}60) = 1 - \mathbb{P}\left(Z \leqslant \frac{1{,}60 - 1{,}75}{0{,}5}\right) = 1 - \mathbb{P}(Z \leqslant -0{,}3)$$
$$= 1 - \Phi(-0{,}3) = 1 - 0{,}3821 = 0{,}6179.$$
$$\text{(c) } \mathbb{P}(1{,}55 \leqslant X \leqslant 1{,}85) = \mathbb{P}\left(\frac{1{,}55 - 1{,}75}{0{,}5} \leqslant Z \leqslant \frac{1{,}85 - 1{,}75}{0{,}5}\right)$$
$$= \mathbb{P}(-0{,}4 \leqslant Z \leqslant 0{,}2) = \Phi(0{,}2) - \Phi(-0{,}4)$$
$$= 0{,}5793 - 0{,}3446 = 0{,}2347.$$

Figura 3.5: Gráfico para las probabilidades a calcular



Fuente: elaboración propia.

Ejemplo **3.11** (1) Obtener el percentil de orden 97,5 % de una variable aleatoria que sigue una distribución normal estándar $Z \sim N(0, 1)$.

(2) Sea $k \in \mathbb{R}^+$ y $Z \sim N(0, 1)$. En la ecuación $\mathbb{P}(-k \leqslant Z \leqslant k) = 0{,}95$, obtener el valor de k.

Solución: (1) El percentil de orden 97,5 % de una variable aleatoria que sigue una distribución normal estándar está dado por:

$$z_{0{,}975} = \Phi^{-1}(0{,}975) = 1{,}96.$$

Es decir, si $Z \sim N(0,1)$, entonces se tiene que en esta variable aleatoria bajo el valor $z = 1{,}96$ se encuentra el 97,5 % de las observaciones.

(2) Del gráfico de la figura 3.6(a) se tiene que:

$$\begin{aligned}\mathbb{P}(-k \leqslant Z \leqslant k) =& \mathbb{P}(Z \leqslant k) - \mathbb{P}(Z < -k)\\ =& \mathbb{P}(Z \leqslant k) - (1 - \mathbb{P}(Z \leqslant k))\\ =& 2 \cdot \mathbb{P}(Z \leqslant k) - 1 = 0{,}95.\end{aligned}$$

Aplicando la condición inicial, se tiene que:

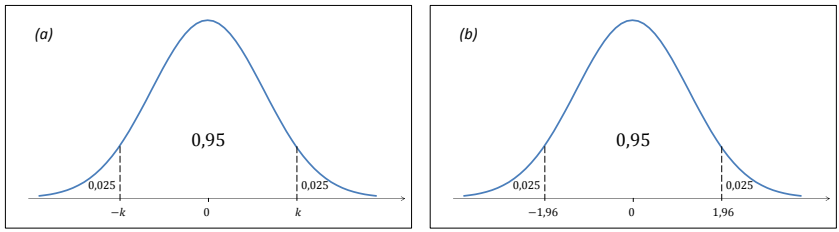
$$2 \cdot \mathbb{P}(Z \leqslant k) - 1 = 0{,}95 \Rightarrow \mathbb{P}(Z \leqslant k) = 0{,}975.$$

Aplicando la función $\Phi(\cdot)$ y la función $\Phi^{-1}(\cdot)$:

$$\Phi(k) = 0{,}975 \Rightarrow k = \Phi^{-1}(0{,}975) = z_{0{,}975} = 1{,}96.$$

Es decir, $\mathbb{P}(-1{,}96 \leqslant Z \leqslant 1{,}96) = 0{,}95$ (véase el gráfico de la figura 3.6(b)). □

Figura 3.6: Gráfico de cuantiles para la distribución normal estándar



(a) Valores $-k$ y k (b) Valores $-z_{0{,}975}$ y $z_{0{,}975}$

Fuente: elaboración propia.

Distribución uniforme

Existen situaciones experimentales en que cada valor en el recorrido de la variable aleatoria tiene la misma densidad y está definido para un intervalo específico; por esto los parámetros de la distribución están dados por los límites del intervalo en que está definida. Nótese que en caso de que el intervalo esté especificado, por ejemplo $(0, 1)$, la distribución estará

completamente especificada. Por las características dadas, es llamada distribución uniforme.

Definición **3.12** La variable aleatoria X sigue una distribución uniforme en el intervalo (a, b), si:

$$X \sim \text{Uniforme}(a, b) \Leftrightarrow f(x) = \frac{1}{b-a}, \quad a < x < b.$$

Dado que la distribución está definida para un intervalo acotado en los reales, para cualquier otro intervalo de valores de $x \in \mathbb{R}$ que no intersecte al intervalo (a, b), la densidad estará definida como cero. Habitualmente esto se deja de manifiesto en la definición de la función de densidad asociada a la variable aleatoria; sin embargo, para simplificar, en adelante se omite, teniendo en cuenta esta acotación.

Propiedad **3.8** La función generadora de momentos, el valor esperado y la varianza de X están dados, respectivamente, por:

(a) $M_X(t) = \frac{e^{tb}-e^{ta}}{t(b-a)}$.

(b) $E[X] = \frac{a+b}{2}$.

(c) $Var[X] = \frac{(b-a)^2}{12}$.

Demostración:

(a) $M_X(t) = E[e^{tX}] = \displaystyle\int_a^b e^{xt}\frac{1}{b-a}\,dx = \frac{1}{t(b-a)}e^{xt}\Big|_a^b = \frac{e^{tb}-e^{ta}}{t(b-a)}$.

(b) $E[X] = \displaystyle\int_a^b x\frac{1}{b-a}\,dx = \frac{x^2}{2}\frac{1}{b-a}\Big|_a^b = \frac{1}{2}(b+a)$.

(c) $E[X^2] = \displaystyle\int_a^b x^2\frac{1}{b-a}\,dx = \frac{1}{b-a}\cdot\frac{1}{3}x^3\Big|_a^b = \frac{1}{3}\cdot(b^2+ab+a^2)$

Luego, la varianza es:

$$\begin{aligned} Var[X] =& E[X^2] - (E[X])^2 \\ =& \frac{1}{12}\left(4b^2+4ab+4a^2-3b^2-6ab-3a^2\right) = \frac{1}{12}(b-a)^2. \end{aligned}$$

Nótese que los resultados obtenidos en (b) y (c) también se podrían obtener utilizando la función generadora de momentos.

Distribución gamma

Esta distribución (al igual que la distribución exponencial) es utilizada principalmente para modelar variables que describen los tiempos de sobrevida o tiempos hasta que se produce β veces un determinado suceso.

Desde un punto de vista teórico, es utilizada para modelar variables aleatorias con asimetría hacia los valores bajos del recorrido de distribución o a la izquierda de la media. La distribución queda especificada por dos parámetros, α y β, donde α da la forma a la distribución y también determina la convergencia de la distribución. Por su parte, β determina el alcance de la asimetría de la densidad y se puede interpretar como el *tiempo promedio entre la ocurrencia de un suceso.*

Definición **3.13** La variable aleatoria X sigue una distribución gamma si:

$$f(x) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-\frac{x}{\beta}}, \;\; x > 0, \; \alpha > 0, \; \beta > 0 \; (\alpha, \beta \text{ valores fijos}),$$

donde $\Gamma(\alpha) = \displaystyle\int_0^{\infty} t^{\alpha-1} e^{-t}\, dt$ es la llamada función gamma.

Propiedad **3.9** La función generadora de momentos, el valor esperado y la varianza de X están dados, respectivamente, por:

(a) $M_X(t) = (1-\beta t)^{-\alpha}$.

(b) $E[X] = \alpha\beta$.

(c) $Var[X] = \alpha\beta^2$.

Demostración: (a) La función generadora de momentos de X:

$$\begin{aligned}
M_X(t) = E[e^{tX}] &= \int_0^{\infty} e^{t\,x} f(x)\, dx \\
&= \int_0^{\infty} \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} \exp\left\{-\frac{x}{\beta} + tx\right\} dx \\
&= \int_0^{\infty} \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} \exp\left\{-x(\frac{1}{\beta} - t)\right\} dx \\
&= \frac{1}{\Gamma(\alpha)(1-\beta t)^{\alpha}} \int_0^{\infty} z^{\alpha-1} e^{-z}\, dz, \text{ con } z = x(\tfrac{1}{\beta} + t) \\
&= \frac{1}{\cancel{\Gamma(\alpha)}(1-\beta t)^{\alpha}} \cancel{\Gamma(\alpha)} \\
&= (1-\beta t)^{-\alpha}.
\end{aligned}$$

(b) Valor esperado:

$$E[X] = \frac{d}{dt}M_X(t)\Big|_{t=0} = \alpha\beta(1-\beta t)^{-\alpha-1}\big|_{t=0} = \alpha\beta.$$

(c) $Var[X] = E[X^2] - (E[X])^2$, entonces:

$$\begin{aligned} E[X^2] &= \frac{d^2}{dt^2}M_X(t)\Big|_{t=0} = \frac{d}{dt}\alpha\beta(1-\beta t)^{-\alpha-1}\Big|_{t=0} \\ &= \alpha\beta^2(\alpha+1)(1-\beta t)^{-\alpha-2}\big|_{t=0} \\ &= \alpha\beta^2(\alpha+1). \end{aligned}$$

Finalmente, se obtiene: $Var[X] = \alpha\beta^2(\alpha+1) - (\alpha\beta)^2 = \alpha\beta^2$.

Propiedad **3.10** Si $X \sim \text{Gamma}(\alpha = 1, \beta)$ entonces $X \sim \text{Exp}(\beta)$ es la distribución llamada *exponencial* de parámetro β, para la cual:

(a) $X \sim \text{Exp}(\beta) \Leftrightarrow f(x) = \frac{1}{\beta}e^{-\frac{x}{\beta}}, \;\; x > 0, \;\; (\beta > 0 \text{ constante}).$

(b) La función generadora de momentos es $M_X(t) = (1-\beta t)^{-1}$.

(c) El valor esperado es: $E[X] = \beta$.

(d) La varianza es: $Var[X] = \beta^2$.

Demostración: Estos resultados son inmediatos si en la propiedad 3.9 $\alpha = 1$.

Una notación alternativa para la distribución exponencial es utilizar el parámetro λ en lugar de β, con $\lambda = \beta^{-1}$.

Propiedad **3.11** Si $X \sim \text{Gamma}(\alpha = \frac{n}{2}, \beta = 2)$, entonces: $X \sim \chi^2(n)$, llamada distribución *chi-cuadrado*, donde n es denominado *grados de libertad*.

(a) La función de densidad es:

$$f(x) = \frac{1}{\Gamma\left(\frac{n}{2}\right)2^{\frac{n}{2}}}x^{\frac{n}{2}-1}\exp\left\{-\frac{x}{2}\right\}, \; x > 0.$$

(b) El valor esperado y la varianza son $E[X] = n$ y $Var[X] = 2n$, respectivamente.

Demostración: Usar en la propiedad 3.9: $\alpha = \frac{n}{2}$ y $\beta = 2$.

Distribución beta

Esta distribución se ajusta a diseños experimentales donde la variable aleatoria toma valores en el intervalo $(0, 1)$, por lo que es posible utilizarla para modelar proporciones y se usa en inferencia como un paso previo a la distribución binomial. Este último dato hace referencia a la llamada *inferencia bayesiana*, donde la distribución beta es utilizada frecuentemente como distribución *a priori* cuando se trabaja con observaciones de tipo binomial.

Definición **3.14** La variable aleatoria $X \sim \text{Beta}(\alpha, \beta)$ si su función de densidad es:

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \quad 0 < x < 1 \quad \alpha > 0 \quad \beta > 0. \tag{3.2}$$

Una forma alternativa de definir la función de densidad de distribución beta es por medio de la utilización de la función beta, que se define como $\text{B}(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$. Con esto se tiene que la expresión 3.2 se puede reescribir como:

$$f(x) = \frac{1}{\text{B}(\alpha, \beta)} x^{\alpha-1}(1-x)^{\beta-1}, \quad 0 < x < 1, \quad \alpha > 0, \quad \beta > 0.$$

Propiedad **3.12** El valor esperado, la varianza y el momento de orden k de $X \sim \text{Beta}(\alpha, \beta)$ son, respectivamente:

(a) $E[X] = \alpha \cdot (\alpha+\beta)^{-1}$.

(b) $Var[X] = \alpha\beta \cdot \left[(\alpha+\beta)^2(\alpha+\beta+1)\right]^{-1}$.

(c) $E[X^k] = \Gamma(\alpha+\beta)\,\Gamma(\alpha+k) \cdot (\Gamma(\alpha)\Gamma(\alpha+\beta+k))^{-1}$.

Demostración: Se obtendrá el momento de orden k-ésimo y, a partir de ese resultado, el valor esperado y la varianza. En efecto:

$$\begin{aligned} E[X^k] =& \text{B}(\alpha, \beta)^{-1} \cdot \int_0^1 x^{(\alpha+k)-1}(1-x)^{\beta-1}\, dx \\ =& \text{B}(\alpha, \beta)^{-1} \cdot \text{B}(\alpha+k, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\cancel{\Gamma(\beta)}} \cdot \frac{\Gamma(\alpha+k)\cancel{\Gamma(\beta)}}{\Gamma(\alpha+k+\beta)} \\ =& \frac{\Gamma(\alpha+\beta)\Gamma(\alpha+k)}{\Gamma(\alpha)\Gamma(\alpha+k+\beta)}. \end{aligned}$$

Con $k = 1$ se tiene $E[X]$ el valor esperado y con $k = 2$ se tiene $E[X^2]$, que se utiliza para calcular la varianza:

$$\begin{aligned} E[X] =& \frac{\Gamma(\alpha+\beta)\Gamma(\alpha+1)}{\Gamma(\alpha)\Gamma(\alpha+1+\beta)} = \frac{\alpha\cancel{\Gamma(\alpha+\beta)}\cancel{\Gamma(\alpha)}}{(\alpha+\beta)\cancel{\Gamma(\alpha)}\cancel{\Gamma(\alpha+\beta)}} \\ =& \frac{\alpha}{\alpha+\beta}. \\ E[X^2] =& \frac{\Gamma(\alpha+\beta)\Gamma(\alpha+2)}{\Gamma(\alpha)\Gamma(\alpha+2+\beta)} = \frac{\alpha(\alpha+1)\cancel{\Gamma(\alpha+\beta)}\cancel{\Gamma(\alpha)}}{(\alpha+\beta)(\alpha+\beta+1)\cancel{\Gamma(\alpha)}\cancel{\Gamma(\alpha+\beta)}} \\ =& \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}. \\ Var[X] =& \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)} - \left(\frac{\alpha}{\alpha+\beta}\right)^2 \\ =& \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}. \end{aligned}$$

Propiedad **3.13** Si $X \sim \text{Beta}(1,1)$, es decir, los parámetros son $\alpha = \beta = 1$, se obtiene que $X \sim \text{Uniforme}(0,1)$.

Demostración: es inmediata al evaluar la función de densidad de una variable aleatoria con distribución beta en $\alpha = \beta = 1$, con lo que se obtiene la función de densidad de una distribución uniforme con $a = 1$ y $b = 1$.

Otras distribuciones de variables aleatorias continuas

Las siguientes variables aleatorias tienen importantes aplicaciones en la modelación de distintos fenómenos. Sin embargo, para los alcances de este texto se presentarán solo las distribuciones y las principales propiedades.

Distribución de Weibull

La variable aleatoria continua X sigue una distribución de Weibull si:

$$X \sim \text{W}(\alpha,\theta) \Leftrightarrow f(x) = \frac{\alpha}{\theta^\alpha} x^{\alpha-1} \exp\left\{-\left(\frac{x}{\theta}\right)^\alpha\right\}, \quad x > 0,\ \alpha > 0,\ \theta > 0.$$

Propiedad **3.14** El valor esperado y la varianza para la distribución de Weibull son:

$$E[X] = \theta\,\Gamma(1+\alpha^{-1}) \text{ y } Var[X] = \theta^2\left[\Gamma(1+2\alpha^{-1}) - \Gamma^2(1+\alpha^{-1})\right].$$

Distribución log-normal

La variable aleatoria continua X se dice que tiene distribución log-normal, con notación $X \sim \text{LogN}(\mu, \sigma^2)$, si y solo si:

$$f(x) = \frac{1}{\sqrt{2\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}\left(\log x - \mu\right)^2\right\}, \quad x > 0,\ \mu \in \mathbb{R},\ \sigma > 0.$$

Propiedad **3.15** El valor esperado y la varianza son:
$E[X] = \exp\left\{\mu + \frac{\sigma^2}{2}\right\}$ y $Var[X] = \exp\left\{2\mu + \sigma^2\right\}$.

3.4. Distribuciones de variables aleatorias con MS-Excel

Si X es una variable aleatoria que sigue una distribución conocida (ya sea de naturaleza discreta o continua), es posible obtener algunos resultados gráficos y de cálculo de probabilidades asociados a ella. En este contexto, resulta útil apoyarse en una herramienta informática para procesar la información. En este caso, se utilizará MS-Excel, básicamente porque se trata de un *software* masificado entre los usuarios de MS y, además, tiene asociada la posibilidad de graficar y realizar cálculos mediante el ingreso de fórmulas, y cuenta con funciones para las distribuciones de uso común.

A continuación, se presenta la representación de algunas de las distribuciones estudiadas en este texto mediante el uso de MS-Excel.

Distribución	Probabilidad acumulada ($F(x)$)
Binomial	`DISTR.BINOM.N(Número de éxitos;Núm. de ensayos; Probabilidad de éxito;Acumulado)`
Bernoulli	`DISTR.BINOM.N(Núm de éxitos;1;Probabilidad de éxito; VERDADERO)`
Binomial negativa	`NEGBINOM.DIST(Núm_fracasos;Núm_éxitos;Prob_éxito; Acumulado)`
Hipergeométrica	`DISTR.HIPERGEOM.N(Muestra_éxito;Núm_de_muestra; Población_éxito;Núm_de_población;Acumulado)`
Poisson	`POISSON.DIST(x;Media;Acumulado)`
Normal	`DISTR.NORM.N(x;Media;Desv_estándar;Acumulado)`
Normal estándar	`DISTR.NORM.ESTAND.N(z;Acumulado)`
Gamma	`DISTR.GAMMA.N(x;Alpha;Beta;Acumulado)`
Beta	`DISTR.BETA.N(x;Alpha;Beta;Acumulado)`
Weibull	`DISTR.WEIBULL(x;Alpha;Beta;Acumulado)`
Log-normal	`DISTR.LOGNORM(x;Media;Desv_estándar;Acumulado)`

En todos los casos, el argumento de las funciones `Acumulado` tiene dos

opciones: `VERDADERO`, que entrega la probabilidad acumulada hasta x o hasta el número de éxito, y `FALSO`, que arroja la función de cuantía para x o el número de éxitos.

3.5. Ejercicios resueltos

1. Ciertos ítems producidos por una máquina son clasificados como defectuosos o no defectuosos; la probabilidad de que ocurra un ítem defectuoso es de 0,05 y, además, son independientes.

   (a) Suponer que se extraen 10 ítems para ser inspeccionados, ¿cuál es la probabilidad de que uno o más de un ítem sea defectuoso?

   (b) Estos ítems posteriormente son embalados en cajas de 20 unidades para ser distribuidos, con la especificación de que un 15 % de los ítems por caja son defectuosos. Un inspector rechazará una caja si al revisar 2 ítems encuentra a lo menos 1 defectuoso. Determinar la probabilidad de que la caja sea rechazada.

   (c) Suponer que se agregó un componente nuevo a dicha máquina para mejorar la calidad, produciendo 100 ítems por hora, de los cuales 1 es defectuoso. ¿Cuál es la probabilidad de que produzca menos de 3 ítems defectuosos en una producción de 1000 unidades? ¿Cuál es el número esperado de ítems defectuosos en una jornada laboral?

   Solución: (a) El experimento es una repetición de ensayos de Bernoulli, por tanto, se trata de una variable aleatoria binomial con $n = 10$ y $p = 0,05$. Si se llama a la variable aleatoria X, entonces la respuesta a la pregunta es:

   $\mathbb{P}(X \geqslant 1) = 1 - \mathbb{P}(X = 0) = 1 - \binom{10}{0}0{,}05^0(1 - 0{,}05)^{10-0} = 1 - 0{,}95^{10} = 0{,}401.$

   (b) El experimento se trata de una distribución hipergeométrica con $N = 20$, $m = 3$ ($m = 20 \cdot 15\,\%$) y $n = 2$, entonces:

   $\mathbb{P}(X \geqslant 1) = 1 - \mathbb{P}(X = 0) = 1 - \frac{\binom{3}{0}\binom{17}{2}}{\binom{20}{2}} = 1 - \frac{136}{190} = 0{,}284.$

   (c) Se trata de un experimento binomial, pero como n es grande y p es pequeño se puede obtener una buena aproximación a través de

la distribución de Poisson de parámetro $\lambda = np = 10$. Así $f(x) = \frac{e^{-10}10^x}{x!}$. Luego, la respuesta a la pregunta es:

$$\begin{aligned}\mathbb{P}(X < 3) =& \mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \mathbb{P}(X = 2)\\ =& e^{-10}(\frac{10^0}{0!} + \frac{10^1}{1!} + \frac{10^2}{2!}) = 0{,}0028.\end{aligned}$$

En una jornada laboral de 8 horas se esperan 8 ítems defectuosos.

2. La variable aleatoria X tiene una fgm dada por $(\frac{5}{8} + \frac{3}{8}e^t)^5$.

   (a) Probar que $X \sim \text{Binomial}(n = 5, p = \frac{3}{8})$.

   (b) Obtener la fgm de la variable aleatoria $g(X) = 2X + 3$.

   (c) Obtener la probabilidad de que $g(X) \leqslant 5$.

   Solución: El problema se basa en identificar la distribución de probabilidad a partir de la función generadora de momentos (en el capítulo 2 se vio que la fgm identifica a una distribución de probabilidad y viceversa) y la aplicación de la propiedad de la función de una variable aleatoria.

   (a) Si $X \sim \text{Binomial}(n, p) \Rightarrow M_X(t) = (1 - p + pe^t)^n = (q + pe^t)^n$. En el caso propuesto, si $n = 5$ y $p = \frac{3}{8}$, se tiene la distribución binomial con dichos parámetros.

   (b) La fgm de $g(X) = 2X + 3$ es:

$$\begin{aligned}M_{g(X)}(t) =& E\left[e^{t\cdot g(X)}\right] = E\left[e^{t(2X+3)}\right]\\ =& E\left[e^{2tX}\cdot e^{3t}\right] = e^{3t}\cdot E\left[e^{2tX}\right]\\ =& e^{3t}\cdot M_X(2t) = e^{3t}\cdot\left(\frac{5}{8} + \frac{3}{8}e^{2t}\right)^5.\end{aligned}$$

   (c) $\mathbb{P}\left(2X + 3 \leqslant 5\right) = \mathbb{P}\left(X \leqslant 1\right) = p(0) + p(1) = 4\left(\frac{5}{8}\right)^5$.

3. Un jugador de básquetbol realiza repetidos lanzamientos desde la línea de tiros libres. Suponer que los lanzamientos son ensayos de Bernoulli independientes con $p = 0{,}7$.

   (a) Calcular la probabilidad de que al jugador le tome menos de 5 lanzamientos lograr su primer éxito.

(b) Calcular la probabilidad de que al jugador le tome menos de 5 lanzamientos lograr su segundo acierto.

(c) Obtener el número esperado de lanzamientos para que el jugador logre su cuarto acierto.

Solución: (a) Sea la variable aleatoria X_1: *el número de lanzamientos hasta el primer acierto*, entonces $X \sim$ Geométrica$(p = 0{,}7)$. Luego, la probabilidad de que requiera menos de 5 lanzamientos para acertar por primera vez es:

$$\mathbb{P}(X_1 < 5) = \mathbb{P}(X_1 \leqslant 4) = 1 - 0{,}3^4 = 0{,}992.$$

(b) Sea X_2: *el número de lanzamientos hasta el segundo acierto*, entonces $X \sim \text{Bneg}(k = 2; p = 0{,}7)$. Luego, la probabilidad de que realice menos de 5 lanzamientos hasta el segundo acierto es:

$$\mathbb{P}(X_2 \leqslant 4) = \sum_{x=2}^{4} \binom{x-1}{1} 0{,}7^2 0{,}3^{x-2} = 0{,}916.$$

(c) Sea X_3: *el número de lanzamientos hasta efectuar el cuarto acierto*, entonces $X_3 \sim \text{Bneg}(k = 4; p = 0{,}7)$. Luego, el número esperado de lanzamientos hasta el cuarto acierto es:

$$E[X_3] = \frac{4}{0{,}7} = 5{,}71.$$

4. Un ingeniero va todos los días de su casa a su oficina ubicada en el centro de la ciudad. El tiempo promedio para su viaje de ida es de 24 minutos, con una desviación estándar de 3,8 minutos. Suponiendo que la distribución es normal, obtener:

   (a) La probabilidad de que un viaje tome al menos media hora.

   (b) El porcentaje de las veces que llega tarde al trabajo, si la empresa abre a las 9:00 h y él sale diariamente de su casa a las 8:45 h.

   (c) La probabilidad de que pierda el café si sale de su casa a las 8:35 h y el café se sirve en la oficina de 8:50 h a 9:00 h.

(d) La longitud de tiempo por arriba de la cual encontramos el 15 % de los viajes más lentos.

Solución: Definiendo la variable aleatoria *X: tiempo de viaje (que tarda el ingeniero desde su casa hasta su oficina) en minutos.*

(a) A partir del enunciado se tiene que $X \sim N(\mu = 24; \sigma^2 = 3{,}8^2)$. La probabilidad de que un viaje tome al menos media hora, $\mathbb{P}(X \geqslant 30)$, queda determinada por:

$$\begin{aligned}\mathbb{P}(X \geqslant 30) &= 1 - \mathbb{P}(X < 30) = 1 - \mathbb{P}\left(\frac{X-\mu}{\sigma} < \frac{30-24}{3{,}8}\right)\\ &= 1 - \mathbb{P}(Z < 1{,}58) = 1 - \Phi(1{,}58)\\ &= 1 - 0{,}9429 = 0{,}0571.\end{aligned}$$

(b) El ingeniero llegará tarde al trabajo si se demora más de 15 minutos, luego la probabilidad pedida es:

$$\begin{aligned}\mathbb{P}(X > 15) &= 1 - \mathbb{P}(X \leqslant 15) = 1 - \mathbb{P}\left(\frac{X-\mu}{\sigma} \leqslant \frac{15-24}{3{,}8}\right)\\ &= 1 - \mathbb{P}(Z \leqslant -2{,}37) = 1 - \Phi(-2{,}37)\\ &= 1 - 0{,}0089 = 0{,}9911.\end{aligned}$$

(c) Para que pierda el café, el ingeniero debería demorar más de 25 minutos, esto es:

$$\begin{aligned}\mathbb{P}(X > 25) &= 1 - \mathbb{P}(X \leqslant 25) = 1 - \mathbb{P}\left(\frac{X-\mu}{\sigma} \leqslant \frac{25-24}{3{,}8}\right)\\ &= 1 - \mathbb{P}(Z \leqslant 0{,}26) = 1 - \Phi(-2{,}37)\\ &= 1 - 0{,}6026 = 0{,}3974.\end{aligned}$$

(d) Esto se puede interpretar como determinar para qué valor de $X = t$ la probabilidad de que el ingeniero demore más de t minutos es igual a 15 %.

$$\begin{aligned}\mathbb{P}(X > t) = 0{,}15 \Rightarrow &\mathbb{P}(X \leqslant t) = 1 - 0{,}15 = 0{,}85\\ \mathbb{P}\left(\frac{X-\mu}{\sigma} \leqslant \frac{t-24}{3{,}8}\right) = &0{,}85,\end{aligned}$$

como $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$, luego:

$$\Phi\left(\frac{t-24}{3{,}8}\right) = 0{,}85 \Rightarrow \frac{t-24}{3{,}8} = \Phi^{-1}(0{,}85) = 1{,}04.$$

Con lo cual se obtiene que $t = 27{,}95$. Es decir, el 15 % de los viajes de mayor duración estarán por sobre los 28 minutos.

5. En una ciudad el consumo diario de agua (en millones de litros) sigue una distribución gamma con parámetros $\alpha = 2$ y $\beta = 3$. Si la capacidad diaria de dicha ciudad es de 9 millones de litros de agua, ¿cuál es la probabilidad de que en cualquier día dado el suministro de agua sea inadecuado? Además, se pide encontrar la cantidad media y la desviación estándar del consumo diario de agua.

   Solución: Se sabe que $X \sim \text{Gamma}(\alpha = 2, \beta = 3) \Rightarrow f(x) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-\frac{x}{\beta}}$, $x > 0$, entonces se conoce el valor esperado y la varianza de X, que son dados por $E[X] = \alpha\beta = 6$ y $Var[X] = \alpha\beta^2 = 18$. También se sabe que el suministro es inadecuado si no alcanza para cubrir la demanda diaria de agua; luego la probabilidad de que en un día dado el suministro de agua sea inadecuado se obtiene mediante:

$$\mathbb{P}(X > 9) = \int_9^{\infty} \frac{1}{\Gamma(2)3^2} x^1 e^{-\frac{x}{3}}\, dx = \int_9^{\infty} \frac{1}{9} x e^{-\frac{x}{3}}\, dx = 4e^{-3}.$$

6. La longitud de tiempo para que un individuo sea atendido en una cafetería es una variable aleatoria que sigue una distribución exponencial con media de 4 minutos. ¿Cuál es la probabilidad de que una persona sea atendida en menos de 3 minutos en al menos 4 de los siguientes 6 días?

   Solución: Se sabe que $X \sim \text{Exp}(\lambda = 4)$ y se llama al evento de interés A (*que una persona sea atendida en menos de 3 minutos en al menos 4 de los siguientes 6 días*). Además, el evento A establece una repetición de experimentos (esperar por la atención en una cafetería) en 6 ocasiones que consideraremos como independientes; estos experimentos siguen una distribución exponencial. Luego, se debe calcular la probabilidad de que, de esas ocasiones, en al menos

4 sea atendida antes de 3 minutos, es decir, puede ser atendida antes de 3 minutos en 4, 5 o 6 ocasiones, los cuales se consideran eventos independientes y se denominarán A_i, $i = 4, 5, 6$, respectivamente.

Así:

$$\mathbb{P}(A) = \sum_{i=4}^{6} \mathbb{P}(A_i),$$

donde:

$$\begin{aligned}\mathbb{P}(A_i) =& \binom{6}{i} (\mathbb{P}(X) < 3)^i (\mathbb{P}(X \geqslant 3))^{6-i} \\ =& \binom{6}{i} (F_X(3))^i (1 - F_X(3))^{6-i}, \quad i = 4, 5, 6.\end{aligned}$$

Además, el valor $\binom{6}{i}$ es producto de la cantidad de maneras en que puede ocurrir el evento A_i. Luego:

$$\mathbb{P}(A) = \sum_{i=4}^{6} \binom{6}{i} (F_X(3))^i (1 - F_X(3))^{6-i}. \tag{3.3}$$

La función de distribución acumulada de X se obtiene de una distribución exponencial con parámetro $\lambda = 4$. En efecto:

$$F_X(x) = \int_0^x \lambda e^{-\lambda t}\, dt = 1 - e^{-\lambda x}.$$

Finalmente, evaluando ($x = 3$ y $\lambda = 4$) y reemplazando en la ecuación 3.3, se tiene que:

$$\mathbb{P}(A) = \sum_{i=4}^{6} \binom{6}{i} (1 - e^{-12})^i (e^{-12})^{6-i}.$$

7. Una distribución que se aplica para el estudio de sobrevida de aparatos electrónicos es la distribución de Weibull. Sea $X \sim \text{Weibull}(\alpha, \theta)$, entonces la función de densidad asociada está dada por la expresión:

$$f(x) = \alpha \cdot \theta^{-\alpha} x^{\alpha-1} \exp\left\{-\left(\frac{x}{\theta}\right)^{\alpha}\right\}, \quad x > 0.$$

Al respecto, se sabe que la vida de servicio, en años, de la batería de un aparato para sordos es una variable aleatoria que tiene una distribución de Weibull con parámetros $\alpha = 2$ y $\theta = \sqrt{2}$. Se desea determinar:

(a) ¿Cuánto tiempo se puede esperar que dure la batería?

(b) ¿Cuál es la probabilidad de que la batería esté en operación después de dos años?

Solución: Se sabe que $X \sim \text{Weibull}(\alpha = 2, \theta = \sqrt{2})$, entonces:

$$f(x) = x \cdot \exp\left\{-\frac{x^2}{2}\right\}, \quad x > 0.$$

(a) El tiempo esperado de duración corresponde a la esperanza de X, el cual está dado por $E[X] = \theta\,\Gamma(1 + \frac{1}{\alpha}) = \frac{\sqrt{2\pi}}{2}$.

(b) Por su parte, la probabilidad de que la batería esté en operación después de dos años está dada por:

$$\mathbb{P}(X > 2) = \int_2^\infty x \exp\left\{-\frac{x^2}{2}\right\} dx = e^{-2}.$$

8. Si la proporción de televisores de una cierta marca que requiere servicio técnico durante el primer año de operación es una variable aleatoria que sigue una distribución beta con $\alpha = 3$ y $\beta = 2$, ¿cuál es la probabilidad de que al menos el 80 % de los nuevos modelos de la marca que se vendieron este año requieran servicio durante su primer año de operación?

   Solución: Se sabe que $X \sim \text{Beta}(\alpha = 3, \beta = 2) \Rightarrow f(x) = 2 \cdot x^2\,(1 - x)$, $0 < x < 1$. La probabilidad de que al menos el 80 % de los nuevos modelos de la marca que se vendieron este año requieran servicio durante su primer año de operación está dada por la expresión:

$$\mathbb{P}(X > 0{,}8) = \int_{0{,}8}^1 2x^2(1 - x)\,dx = 1 - 1{,}6 \cdot 0{,}8^3.$$

9. Sea $X \sim N(\mu, \sigma^2)$, se pide:

   (a) Calcular el valor esperado de la función $g(X) = a + bX + cX^2$, donde a, b y c son constantes reales.

   (b) Demostrar que la distribución de la nueva variable aleatoria $T = (\frac{X-\mu}{\sigma})^2$ es chi-cuadrado con 1 grado de libertad.

Solución: (a) $E[g(X)] = E[a + bX + cX^2] = a + b\,E[X] + c\,E[X^2]$, como $X \sim N(\mu, \sigma^2)$, entonces $E[X] = \mu$ y $Var[X] = \sigma^2$. Luego:

$$E[g(X)] = a + b\mu + c\,\left(Var[X] + (E[X])^2\right) = a + b\mu + c\,(\sigma^2 + \mu^2).$$

(b) Nótese que la variable aleatoria $T = g(X)$, entonces su función de densidad se puede obtener mediante la función de distribución. Además, como $X \sim N(\mu, \sigma^2)$, entonces es conocida su función de densidad. En efecto, utilizando este último resultado se tiene que:

$$\begin{aligned}
F_T(t) =& \mathbb{P}(T \leqslant t) = \mathbb{P}\left(\left[\frac{X-\mu}{\sigma}\right]^2 \leqslant t\right) \\
=& \mathbb{P}\left(\frac{|X-\mu|}{\sigma} \leqslant \sqrt{t}\right) = \mathbb{P}\left(|X-\mu| \leqslant \sigma\sqrt{t}\right) \\
=& \mathbb{P}\left(-\sigma\sqrt{t} \leqslant X-\mu \leqslant \sigma\sqrt{t}\right) \\
=& \mathbb{P}\left(\mu - \sigma\sqrt{t} \leqslant X \leqslant \mu + \sigma\sqrt{t}\right) \\
=& F_X(\mu + \sigma\sqrt{t}) - F_X(\mu - \sigma\sqrt{t}).
\end{aligned}$$

Aplicando la derivada parcial respecto a x a $F_T(t)$, se obtiene la función de densidad de la variable aleatoria T:

$$\begin{aligned}
f_T(t) =& f_X(\mu + \sigma\sqrt{t})\frac{\sigma}{2\sqrt{t}} + f_X(\mu - \sigma\sqrt{t})\frac{\sigma}{2\sqrt{t}} \\
=& \frac{1}{\sqrt{2\pi}\cancel{\sigma}}\left[\exp\left\{-\frac{1}{2\sigma^2}(\cancel{\mu} + \sigma\sqrt{t} - \cancel{\mu})^2\right\} + \right. \\
& \left. + \exp\left\{-\frac{1}{2\sigma^2}(\cancel{\mu} - \sigma\sqrt{t} - \cancel{\mu})^2\right\}\right]\frac{\cancel{\sigma}}{2\sqrt{t}} \\
=& \frac{1}{\sqrt{2\pi}}t^{-\frac{1}{2}}\exp\left\{-\frac{t}{2}\right\},\ t > 0.
\end{aligned}$$

Esta función de densidad corresponde a la función de densidad de una distribución chi-cuadrado con 1 grado de libertad, correspondiente a la distribución de T.

10. Si la variable aleatoria X tiene función de densidad $f(x) = \theta x^{\theta-1}$, $0 < x < 1$, $\theta > 0$, se puede probar que la distribución de la variable aleatoria $Y = -\log(X)$ es exponencial de parámetro θ ($E[Y] = \theta^{-1}$).

Solución: Encontraremos la función de densidad de la variable aleatoria Y. En efecto:

$$F_Y(y) = \mathbb{P}(Y \leqslant y) = \mathbb{P}(-\log X \leqslant y) = \mathbb{P}(X > e^{-y}) = 1 - F_X(e^{-y}).$$

Luego, derivando se obtiene:

$$f_Y(y) = f_X(e^{-y})e^{-y} = \theta(e^{-y})^{\theta-1}e^{-y} = \theta\, e^{-\theta y},\ y > 0,$$

que corresponde a la función de densidad de una distribución exponencial de parámetro θ.

11. Sea la variable aleatoria X que mide la concentración de cierto contaminante, en partes por millón, la cual se sabe que sigue una distribución logarítmica normal:

$$Y = \log(X) \sim N(\mu, \sigma^2).$$

Con parámetros $\mu = 3{,}2$ y $\sigma^2 = 1{,}0$, determinar la probabilidad de que la concentración exceda las ocho partes por millón.

Solución: La probabilidad pedida es $\mathbb{P}(X > 8)$; el cálculo consiste en encontrar la forma de la distribución normal estándar Z. En efecto:

$$\begin{aligned}\mathbb{P}(X > 8) =& 1 - \mathbb{P}(X \leqslant 8) = 1 - \mathbb{P}(\log(X) \leqslant \log(8)) \\ =& 1 - \mathbb{P}\left(Z = \frac{\log(X) - 3{,}2}{1{,}0} \leqslant \frac{\log(8) - 3{,}2}{1{,}0}\right) \\ =& 1 - \Phi(-1{,}12) = 0{,}8686.\end{aligned}$$

3.6. Ejercicios propuestos

Distribuciones de variables aleatorias discretas

1. Mostrar que $X \sim \text{Bernoulli}(p) \Rightarrow Y = X^2 \sim \text{Bernoulli}(p)$.

2. Mostrar que para una variable aleatoria X con distribución de Poisson se cumple la propiedad de pérdida de memoria dada por la expresión:

$$\mathbb{P}(X > a + b | X > a) = \mathbb{P}(X > b),\ \text{con } a, b \in \mathbb{Z}^+.$$

3. Un pediatra desea reclutar a 5 parejas que esperan su primer hijo para participar en un régimen de parto natural. Por estudios realizados con anterioridad, el médico sabe que la probabilidad de que una pareja elegida al azar acceda a participar es de 0,2, entonces desea calcular la probabilidad de que se requiera preguntar a 15 parejas antes de encontrar a 5 que estén de acuerdo en participar.

4. En un sistema de transportes, los expertos creen que la probabilidad de que un tren sufra un desperfecto mecánico es de 0,001 por viaje (entendido como un trayecto completo) y que este hecho no influye en un futuro desperfecto. Sea la variable aleatoria X: *número de desperfectos que se presentan en los próximos 100 viajes*:

   (a) Especificar la distribución de probabilidades de X.

   (b) Calcular la probabilidad de que no se produzca ningún desperfecto.

   (c) ¿Cuál es la probabilidad de que se produzca exactamente un desperfecto?

   (d) Para (c), realizar la aproximación mediante la distribución de Poisson para calcular el resultado obtenido en (c).

5. Una muestra de tamaño 3 es seleccionada sin reemplazo desde una urna que contiene 6 fichas, de las cuales 4 son azules y el resto, verdes. Determinar la probabilidad de que:

   (a) La primera y la segunda ficha sean azules y la tercera verde.

   (b) La muestra contenga exactamente dos fichas azules.

6. Un jugador de tiro con arco dispara al blanco y sabe que la probabilidad de acertar es 1 de 100. Determinar:

   (a) La distribución de la variable aleatoria *número de aciertos en n disparos*.

   (b) La cantidad de disparos que tendrá que realizar para que la probabilidad de dar en el blanco por lo menos una vez sea de 0,9.

7. Considerar la variable aleatoria $X \sim \text{Binomial}(n, p)$ y las funciones de esta variable aleatoria T_1 y T_2, dadas por:

$$T_1 = \frac{X}{n} \text{ y } T_2 = \frac{X+1}{n+2}.$$

Además, se define el *error cuadrático medio* (ECM) por:

$$ECM(T_i) = Var[T_i] + (E[T_i] - p)^2, \quad i = 1, 2.$$

Se pide calcular el ECM de T_1 y de T_2.

Distribuciones de variables aleatorias continuas

8. Demostrar que:

(a) $Z \sim N(0,1) \Rightarrow Z^2 \sim \chi^2(1)$.

(b) $X \sim \text{Gamma}(a, b) \Rightarrow T = 2Xb^{-1} \sim \chi^2(2a)$.

9. Sea $X \sim N(0,1)$, mostrar que:

(a) $E[X^{2k+1}] = 0$.

(b) $E[X^{2k}] = \frac{(2k)!}{2^k k!}$.

10. Se desea capturar 14 gacelas jóvenes con destino a una reserva animal. Para la expedición se dispone de 3 vehículos que pueden transportar una carga de 500, 700 y 1000 kg, respectivamente. El jefe de la expedición no sabe cómo elegir un vehículo y pide a un profesional con conocimientos de estadística que le ayude a decidir cuál es el vehículo más idóneo para la expedición, convencido de que este elegirá el vehículo más ligero capaz de transportar a las 14 gacelas. El jefe de la expedición le da la siguiente información: se sabe que el peso de una gacela normalmente tiene una media de 50 kg y una desviación típica observada de 6 kg.

11. El tiempo, en horas, que toma reparar una bomba de calor es una variable aleatoria X que tiene una distribución chi-cuadrado con 1 grado de libertad. ¿Cuál es la probabilidad de que la siguiente llamada de servicios requiera: (a) a lo más una hora para reparar la bomba de calor, (b) por lo menos (o al menos) dos horas para reparar la bomba de calor?

12. El consumo diario de energía eléctrica, en millones de kWh, es una variable aleatoria X que tiene una distribución gamma con media (o valor esperado) 6 y varianza 12. Se pide encontrar:

    (a) Los valores de los parámetros α y β.

    (b) La probabilidad de que en cualquier día dado el consumo de energía diario exceda 12 millones de kWh.

13. Se estudia el uso promedio de potencia (dB por hora) para una compañía particular y se sabe que tiene una distribución log-normal de parámetros $\mu = 4$ y $\sigma = 2$. Se requiere saber:

    (a) ¿Cuál es la probabilidad de que la compañía utilice más de 270 dB durante una hora particular?

    (b) ¿Cuál es el uso de potencia media (dB promedio por hora) y cuál es la varianza?

14. La distribución exponencial tiene la siguiente propiedad, llamada de pérdida de memoria:

$$\mathbb{P}(X > s + t \mid X > s) = \mathbb{P}(X > t).$$

    Probar esta propiedad. Además, utilizar este resultado para determinar $\mathbb{P}(3 < X < 8)$ y $\mathbb{P}(Y \geqslant 100 \mid Y > 50)$.

15. El cuantil de orden p-ésimo de la distribución de una variable aleatoria X, denotado con x_p, es definido como $\mathbb{P}(X < x_p) = p$, donde $0 < p < 1$. Cuando $p = 0{,}5$, el cuantil es llamado mediana de la variable aleatoria X. Determinar la mediana de las distribuciones normal, exponencial y log-normal.

16. X sigue una distribución de Cauchy ($X \sim \text{Cauchy}(a, b)$), si la función de densidad es $f(x) = (b\pi(1 + (\frac{x-a}{b})^2))^{-1}$, $x \in \mathbb{R}$.

    (a) Mostrar que el valor esperado se indetermina; usar $a = 0$ y $b = 1$.

    (b) Determinar la función de distribución para $X \sim \text{Cauchy}(a, b)$.

    (c) Calcular la mediana y la moda para $X \sim \text{Cauchy}(a, b)$.

17. Las distribuciones gamma, exponencial y de Weibull son aplicadas a problemas de confiabilidad y de prueba de vida, como los tiempos de falla o la duración de la vida de un componente medida en algún tiempo específico hasta la falla. La *confiabilidad de un componente o producto* se define como la probabilidad de que funcione apropiadamente por al menos un tiempo específico. Si $R(t)$ define la *confiabilidad de un componente o producto* en el tiempo t, podemos escribir $R(t) = \mathbb{P}(T > t)$. Por otro lado, la tasa de falla de un componente es definida por $Z(t) = f(t) \cdot (R(t))^{-1}$. Para las distribuciones exponencial y de Weibull, determinar la función $R(t)$ y la tasa de falla.

Capítulo 4

Vectores aleatorios

4.1. Vectores aleatorios bivariados

En el capítulo anterior, se estudiaron las características de una variable aleatoria desde un enfoque unidimensional. Ahora se estudiarán desde un enfoque bidimensional, donde se organizan las variables en arreglos matemáticos formando vectores de naturaleza aleatoria y, posteriormente, como una extensión al caso multivariado. El punto de partida es la definición de una variable aleatoria bidimensional o, si se lo considera como un ente matemático, vector aleatorio. Posteriormente, se analizarán las propiedades conjuntas y univariadas de los componentes del vector.

Definición **4.1** La expresión $(X_1, X_2)^T$ es llamada vector aleatorio si y solo si cada uno de sus componentes es una variable aleatoria:

$$\boldsymbol{X} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} : \Omega \to \mathbb{R}^2$$

Observación **4.1** Las variables aleatorias bidimensionales se clasifican en: discretas (si X_1 y X_2 son variables aleatorias discretas), continuas (si X_1 y X_2 son variables aleatorias continuas) y mixtas (si una variable aleatoria es discreta y la otra continua).

Variables aleatorias bidimensionales continuas

Definición **4.2** Sea $(X_1, X_2)^T$ una variable aleatoria bidimensional continua. La función de densidad conjunta asociada a esta variable aleatoria

bidimensional, denotada con $f(x_1, x_2)$, es una función real valuada que cumple con las siguientes condiciones:

(a) $f(x_1, x_2) \geqslant 0 \ \forall \ (x_1, x_2)^T \in Rec_{(X_1,X_2)^T} \subseteq \mathbb{R}^2$.

(b) $\displaystyle\iint_{Rec_{X_1,X_2}} f(x_1, x_2)\, dx_1 dx_2 = 1.$

Para calcular probabilidades, se debe tener en cuenta que el espacio ahora es $\mathbb{R}^2$, por lo que corresponde el cálculo mediante integrales dobles (caso continuo) o doble sumatoria (caso discreto). Si bien la definición es aplicable para el caso discreto y hay problemas de aplicación que se pueden enfrentar mediante distribuciones bivariadas discretas, este texto utiliza principalmente ejemplos para el caso continuo.

Desde el punto de vista gráfico, en este caso se trabaja con volúmenes (3 dimensiones). Para calcular probabilidades se trabajará con curvas de nivel en el plano (acotadas por la relación entre las dos variables aleatorias), el que estará acotado en la tercera dimensión por $f(x, y)$. Para obtener ese volumen es necesario utilizar integrales dobles.

Ejemplo **4.1** Considerar $(X_1, X_2)^T$ una variable aleatoria bidimensional continua con $f(x_1, x_2) = \exp\{-x_1 - x_2\}$, $x_1 > 0$, $x_2 > 0$.

(a) Verificar que $f(x_1, x_2)$ es función de densidad.

(b) Calcular (b.1) $\mathbb{P}(X_1 > 2, X_2 > 5)$, (b.2) $\mathbb{P}(X_1 > X_2)$ y (b.3) $\mathbb{P}(X_1 + X_2 > 1)$.

(c) Graficar (b.1), (b.2) y (b.3).

Solución: (a) Se tiene que $f(x_1, x_2) = e^{-x_1-x_2} \geqslant 0$. Además, se cumple que:

$$\begin{aligned}\int_0^\infty \int_0^\infty e^{-x_1-x_2}\, dx_1 dx_2 &= \int_0^\infty e^{-x_1} \left[\int_0^\infty e^{-x_2}\, dx_2\right] dx_1 \\ &= \Gamma(1) \int_0^\infty e^{-x_1}\, dx_1 = \Gamma(1) = 1.\end{aligned}$$

(b.1) La probabilidad pedida se obtiene como:

$$\begin{aligned}\mathbb{P}(X_1 > 2, X_2 > 5) &= \int_2^\infty \int_5^\infty e^{-x_1-x_2}\, dx_1 dx_2 \\ &= \int_2^\infty e^{-x_1} \left[\int_5^\infty e^{-x_2}\, dx_2\right] dx_1 \\ &= e^{-5} \int_2^\infty e^{-x_1}\, dx_1 = e^{-5}e^{-2} = e^{-7}.\end{aligned}$$

(b.2) La probabilidad está dada por:

$$\begin{aligned}\mathbb{P}(X_1 > X_2) &= \int_0^\infty \int_0^{x_1} e^{-x_1-x_2}\, dx_1 dx_2 = \int_0^\infty e^{-x_1} \left[\int_0^{x_1} e^{-x_2}\, dx_2\right] dx_1 \\ &= \int_0^\infty e^{-x_1}(1 - e^{-x_1})\, dx_1 = 1 - 0{,}5 = 0{,}5.\end{aligned}$$

(b.3) Finalmente, se puede calcular:

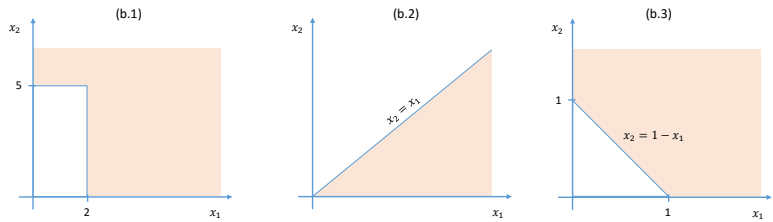
$$\mathbb{P}(X_1 + X_2 > 1) = \int_0^1 \int_{1-x_1}^\infty e^{-x_1-x_2}\, dx_1 dx_2 + \int_1^\infty \int_0^\infty e^{-x_1-x_2}\, dx_1 dx_2.$$

Una forma alternativa de obtener este último resultado es mediante:

$$\mathbb{P}(X_1 + X_2 > 1) = 1 - \int_0^1 \int_1^{1-x_2} e^{-x_1-x_2}\, dx_1 dx_2 = 2\, e^{-1}.$$

(c) Los gráficos que representan cada situación se pueden ver en la figura 4.1 y se obtienen directamente de la definición de probabilidad de variables aleatorias bivariadas. □

Figura 4.1: Región de integración dada por la probabilidad a calcular



Fuente: elaboración propia.

Función de densidad marginal

Definición **4.3** Sea $(X_1, X_2)^T$ una variable aleatoria bivariada con función de probabilidad conjunta conocida $f(x_1, x_2)$. La función de densidad marginal de la variable aleatoria X_1, está definida por:

$$f_{X_1}(x_1) = \int_{Rec_{X_2}} f(x_1, x_2)\, dx_2$$
$$f_{X_2}(x_2) = \int_{Rec_{X_1}} f(x_1, x_2)\, dx_1.$$

Ejemplo **4.2** Sea $(X_1, X_2)^T$ una variable aleatoria bivariada con función de probabilidad:

$$f(x_1, x_2) = x_1^2 + \frac{x_1 x_2}{3},\ 0 \leqslant x_1 \leqslant 1,\ 0 \leqslant x_2 \leqslant 2.$$

Se pide determinar la función de densidad marginal de X_1 y X_2.

Solución: (a) Densidad marginal de X_1:

$$f_{X_1}(x_1) = \int_0^2 (x_1^2 + \frac{x_1 x_2}{3})\, dx_2 = \left(x_1^2 x_2 + \frac{x_1 x_2^2}{3} \right) \Bigg|_{x_2=0}^{x_2=2}$$
$$= 2x_1^2 + \frac{2}{3} x_1 \ \ , 0 \leqslant x_1 \leqslant 1.$$

$f(x_1)$ es efectivamente función de densidad de X_1. En efecto:

$$\int_0^1 f_{X_1}(x_1)\, dx_1 = \int_0^1 2x_1^2 + \frac{2}{3} x_1\, dx_1$$
$$= \left(\frac{2}{3} x_1^3 + \frac{1}{3} x_1^2 \right) \Bigg|_0^1 = 1.$$

(b) Densidad marginal de X_2:

$$f_{X_2}(x_2) = \int_0^1 (x_1^2 + \frac{x_1 x_2}{3})\, dx_1$$
$$= \left(\frac{x_1^3}{3} + \frac{x_1^2 x_2}{6} \right) \Bigg|_{x_1=0}^{x_1=1}$$
$$= \frac{1}{3} + \frac{1}{6} x_2 \ \ , 0 \leqslant x_2 \leqslant 2.$$

$f(x_2)$ es efectivamente función de densidad de X_2, pues:

$$\int_0^2 f_{X_2}(x_2)\,dx_2 = \int_0^2 \frac{1}{3} + \frac{1}{6}x_2\,dx_2 = \left(\frac{1}{3}x_2 + \frac{1}{12}x_2^2\right)\Big|_0^2 = 1.$$

□

Ejemplo **4.3** Encontrar las funciones de densidad marginal a partir de la función de densidad bivariada:

(a) $f(x_1, x_2) = e^{-x_1-x_2}$, $x_1 > 0$, $x_2 > 0$.

(b) $f(x_1, x_2) = e^{-x_1}$, $0 < x_2 < x_1$.

Solución: (a) Las funciones de densidad marginal de X_1 y X_2 son, respectivamente:

$$f_{X_1}(x_1) = \int_0^\infty e^{-x_1-x_2}\,dx_2 = e^{-x_1}\int_0^\infty e^{-x_2}\,dx_2 = e^{-x_1}, \quad x_1 > 0.$$
$$f_{X_2}(x_2) = \int_0^\infty e^{-x_1-x_2}\,dx_1 = e^{-x_2}\int_0^\infty e^{-x_1}\,dx_1 = e^{-x_2}, \quad x_2 > 0.$$

(b) Las funciones de densidad marginal de X_1 y X_2 son, respectivamente:

$$f_{X_1}(x_1) = \int_0^{x_1} e^{-x_1}\,dx_2 = e^{-x_1}x_2\Big|_0^{x_1} = x_1e^{-x_1}, \quad x_1 > 0.$$
$$f_{X_2}(x_2) = \int_{x_2}^\infty e^{-x_1}\,dx_1 = -e^{-x_1}\Big|_{x_2}^\infty = e^{-x_2}, \quad x_2 > 0.$$

□

Ejercicio propuesto **4.1** Las variables aleatorias X e Y tienen función de densidad conjunta $f(x, y) = \frac{1}{4\pi}$, $x^2 + y^2 \leqslant 4$.

(a) Determinar $\mathbb{P}(Y > 2X)$.

(b) Determinar la función de densidad marginal de X.

(c) Calcular $\mathbb{P}(|X| < 1)$.

Otro de los conceptos importantes que surgen a partir del análisis bivariado es el de independencia entre dos variables aleatorias. Lo que interesa determinar es si los cambios que se producen en una variable

aleatoria X afectan o son consecuencia de los cambios en otra variable aleatoria Y. Si es así, se habla de que existe relación o correlación entre las variables X e Y; en caso contrario, se dice que las variables aleatorias son independientes o que no hay correlación entre ellas (o que esta es marginal).

Definición **4.4** Las variables aleatorias X_1 y X_2 son independientes si y solo si:

$$f_{X_1,X_2}(x_1, x_2) = f_{X_1}(x_1) \cdot f_{X_2}(x_2).$$

Función de densidad condicional

Definición **4.5** Sea $(X_1, X_2)^T$ una variable aleatoria bivariada con función de densidad conjunta $f(x_1, x_2)$. La función de densidad condicional de la variable aleatoria X_1, dado un valor fijo de X_2, digamos $X_2 = x_2$, está definida por:

$$f_{X_1|X_2=x_2}(x_1) = \frac{f(x_1, x_2)}{f_{X_2}(x_2)}.$$

Análogamente se define la función de densidad condicional de la variable aleatoria X_2, dado un valor fijo de $X_1 = x_1$:

$$f_{X_2|X_1=x_1}(x_2) = \frac{f(x_1, x_2)}{f_{X_1}(x_1)}.$$

Observación **4.2** Las funciones de densidad condicional son funciones de probabilidad; por tanto, son positivas y la integral de la función en el recorrido es igual a 1.

Ejemplo **4.4** Sea $f(x_1, x_2) = e^{-x_1}$, $0 < x_2 < x_1$. Además, se sabe que $f_{X_1}(x_1) = x_1 e^{-x_1}$ $x_1 > 0$ y $f_{X_2}(x_2) = e^{-x_2}$, $x_2 > 0$. Determinar las funciones de densidad condicional de $X_1 \mid X_2 = x_2$ y $X_2 \mid X_1 = x_1$.

Solución: Según la definición 4.5:

(a) $f_{X_1|X_2=x_2}(x_1) = \frac{e^{-x_1}}{e^{-x_2}} = e^{-x_1+x_2}$, $x_1 > x_2$.

(b) $f_{X_2|X_1=x_1}(x_2) = \frac{e^{-x_1}}{x_1 e^{-x_1}} = \frac{1}{x_1}$, $x_2 < x_1$.

Ambas funciones son densidades. En efecto:

$$\int f_{X_1|X_2=x_2}(x_1)\,dx_1 = \int_{x_2}^{\infty} e^{-x_1+x_2}\,dx_1 = e^{x_2}\frac{e^{-x_1}}{-1}\Big|_{x_2}^{\infty} = 1.$$

$$\int f_{X_2|X_1=x_1}(x_2)\,dx_2 = \int_0^{x_1} \frac{1}{x_1}\,dx_2 = \frac{x_2}{x_1}\Big|_0^{x_1} = 1.\square$$

Observación **4.3** Desde la definición 4.4 y la definición 4.5 es fácil probar que si X_1 y X_2 son variables aleatorias independientes, entonces:

(a) $f_{X_1|X_2}(x_1) = f_{X_1}(x_1)$.

(b) $f_{X_2|X_1}(x_2) = f_{X_2}(x_2)$.

Valor esperado de un vector aleatorio

Definición **4.6** Sea $(X_1, X_2)^T$ una variable aleatoria bivariada con una función de densidad conjunta $f(x_1, x_2)$. Sea G una función real valuada definida sobre $(X_1, X_2)^T$ $(G : \mathbb{R}^2 \to \mathbb{R})$. El valor esperado de $G(X_1, X_2)$ está dado por la siguiente expresión:

$$E[G(X_1, X_2)] = \iint_{\mathbb{R}} G(X_1, X_2) f(x_1, x_2)\,dx_2\,dx_1.$$

Ejemplo **4.5** Sea la función de densidad bivariada $f(x_1, x_2) = 6(1 - x_1 - x_2)$, $0 < x_2 < 1 - x_1$, $x_1 > 0$ y sea $G(X_1, X_2) = X_1^2 X_2$. Obtener una expresión para $E[G(X_1, X_2)]$.

Solución: $E[G(X_1, X_2)] = \int_0^1 \int_0^{1-x_1} x_1^2 x_2 6(1-x_1-x_2)\,dx_2\,dx_1 = \frac{1}{60}$. $\square$

Ejemplo **4.6** Sea $(X_1, X_2)^T$ una variable aleatoria bivariada con función de probabilidad conjunta $f(x_1, x_2) = \exp\left\{-\theta x_1 - \theta^{-1} x_2\right\}$, $x_1 > 0$, $x_2 > 0$ y θ constante positiva. Calcular el valor esperado de:

(a) $G_1(X_1, X_2) = X_1 \cdot X_2$.

(b) $G_2(X_1, X_2) = X_1$.

(c) $G_3(X_1, X_2) = \exp\{t\,(X_1 + X_2)\}$.

Solución: (a) Valor esperado de X_1X_2:

$$\begin{aligned} E[X_1X_2] &= \int_0^\infty \int_0^\infty x_1x_2 \exp\left\{-\theta x_1 - \frac{x_2}{\theta}\right\} dx_2\, dx_1 \\ &= \int_0^\infty x_1 e^{-\theta x_1} \left[\int_0^\infty x_2 \exp\left\{-\frac{x_2}{\theta}\right\} dx_2\right] dx_1 \\ &= \int_0^\infty x_1 e^{-\theta x_1} \theta \left[\int_0^\infty \frac{x_2}{\theta} \exp\left\{-\frac{x_2}{\theta}\right\} dx_2\right] dx_1, \end{aligned}$$

donde $\displaystyle\int_0^\infty x_2\theta^{-1} \exp\left\{-\theta^{-1}x_2\right\} dx_2 = E[X_2]$ y $X_2 \sim \text{Exp}(\theta)$. Así $E[X_2] = \theta$, luego:

$$= \theta \int_0^\infty \theta x_1 e^{-\theta x_1} dx_1,$$

donde $\displaystyle\int_0^\infty x_1\theta e^{-\theta x_1} dx_1 = E[X_1]$ y $X_1 \sim \text{Exp}(\theta^{-1})$. Así $E[X_1] = \theta^{-1}$, luego:

$$E[X_1X_2] = \theta \cdot \theta^{-1} = 1.$$

(b) Valor esperado de X_1:

$$\begin{aligned} E[X_1] &= \iint x_1 f(x_1, x_2)\, dx_1\, dx_2 \\ &= \int_0^\infty \int_0^\infty x_1 \exp\left\{-\theta x_1 - x_2\theta^{-1}\right\} dx_2\, dx_1 \\ &= \int_0^\infty x_1 e^{-\theta x_1} \theta \left[\int_0^\infty \frac{1}{\theta} e^{-x_2\theta^{-1}}\, dx_2\right] dx_1, \end{aligned}$$

donde $\displaystyle\int_0^\infty \frac{1}{\theta} \exp\left\{-x_2\theta^{-1}\right\} dx_2 = 1$, pues es la integral en su recorrido de $X_2 \sim \text{Exp}(\theta)$. Luego:

$$= \int_0^\infty x_1\theta e^{-\theta x_1}\, dx_1.$$

Este último resultado corresponde al valor esperado de una variable aleatoria con distribución exponencial de parámetro θ^{-1}, cuyo valor es θ^{-1}. Finalmente:

$$E[X_1] = \theta^{-1}.$$

(c) Valor esperado de $\exp\{t\,(X_1+X_2)\}$:

$$\begin{aligned}
E\left[\exp\{t\,(X_1+X_2)\}\right] &= \iint \exp\{t\,(X_1+X_2)\}\,f(x_1,x_2)\,dx_1\,dx_2 \\
&= \int_0^\infty \int_0^\infty \exp\{t\,(X_1+X_2)\}\exp\left\{-\theta x_1-\theta^{-1}x_2\right\}\,dx_2\,dx_1 \\
&= \int_0^\infty e^{tx_1}e^{-\theta x_1}\left[\int_0^\infty e^{tx_2}e^{-\theta^{-1}x_2}\,dx_2\right]\,dx_1 \\
&= \int_0^\infty e^{tx_1}e^{-\theta x_1}\theta\left[\int_0^\infty e^{tx_2}\frac{1}{\theta}e^{-x_2\theta^{-1}}\,dx_2\right]\,dx_1,
\end{aligned}$$

donde:

$$\int_0^\infty \theta^{-1}\exp\{tx_2\}\exp\{-\theta^{-1}x_2\}\,dx_2 = E\big[\exp\{tx_2\}\big] = M_{X_2}(t)$$

y $X_2 \sim \text{Exp}(\theta) \Rightarrow M_{X_2}(t) = (1-\theta t)^{-1}$. Luego:

$$= (1-\theta t)^{-1}\int_0^\infty \exp\{tx_1\}\,x_1\theta e^{-\theta x_1}\,dx_1,$$

donde:

$$\int_0^\infty \theta\exp\{tx_1\}x_1\exp\{-x_1\theta\}\,dx_1 = E\big[\exp\{tx_1\}\big] = M_{X_1}(t)$$

y $X_1 \sim \text{Exp}(\theta^{-1}) \Rightarrow M_{X_1}(t) = \left(1-\frac{t}{\theta}\right)^{-1}$.

Finalmente:

$$E\left[\exp\{t\,(X_1+X_2)\}\right] = \frac{1}{(1-\theta t)(1-\frac{t}{\theta})}.\square$$

Ejercicio propuesto **4.2** Sea (X,Y) un vector aleatorio con función de densidad conjunta dada por $f(x,y) = x\exp\{-x(y+1)\}$, con $x>0, y>0$.

(a) Determinar las funciones de densidad marginal correspondientes y discutir sobre la independencia de X e Y.

(b) Calcular el valor esperado y la varianza de X.

(c) Determinar F_Y y encontrar la mediana para la variable Y.

(d) Calcular $\mathbb{P}(X>1|Y>1)$. $\square$

Propiedad **4.1** Sean las variables aleatorias X_1 y X_2, se define la covarianza entre X_1 y X_2 (denotada con $Cov(X_1, X_2)$) por:

$$Cov(X_1, X_2) = E[X_1X_2] - E[X_1] \cdot E[X_2].$$

Ejemplo **4.7** Considerando el ejemplo 4.6 para calcular la covarianza entre las variables X_1 y X_2.

Solución: En el ejemplo 4.6 se obtuvo que: $E[X_1X_2] = 1$ y $E[X_1] = \theta^{-1}$, además, calculando, se tendrá que $E[X_2] = \theta$. Finalmente, por la propiedad 4.1, la covarianza entre X_1 y X_2 es: $Cov(X_1, X_2) = 1 - \theta^{-1} \cdot \theta = 0$. □

Coeficiente de correlación

En términos prácticos, y al igual que la covarianza, el coeficiente de correlación es una medida de asociación entre dos variables aleatorias, que mide la asociación de tipo lineal entre las variables (aumentos constantes en la proporción que define la relación, en cualquier unidad de cambio de una de las variables).

Definición **4.7** Sea el vector aleatorio $(X_1, X_2)^T$, se define el coeficiente de correlación lineal entre las variables aleatorias X_1, X_2, denotado con $\rho(X_1, X_2)$, también llamado coeficiente de correlación de Pearson, por la expresión siguiente:

$$\rho(X_1, X_2) = \frac{Cov(X_1, X_2)}{\sqrt{Var(X_1)Var(X_2)}}, \quad -1 \leqslant \rho \leqslant 1. \tag{4.1}$$

Propiedad **4.2** X_1 y X_2 son variables aleatorias independientes, entonces:

$$E[X_1 \cdot X_2] = E[X_1] \cdot E[X_2].$$

Demostración: Considerando el valor esperado de $X_1 \cdot X_2$, se tiene:

$$E[X_1X_2] = \iint_{\mathbb{R}} x_1x_2f(x_1, x_2)\, dx_2\, dx_1.$$

Como X_1 y X_2 son variables aleatorias independientes, entonces $f(x_1, x_2) = f(x_1) \cdot f(x_2)$. Así:

$$\begin{aligned} E[X_1X_2] &= \iint_{\mathbb{R}} x_1x_2f(x_1)f(x_2)\,dx_2\,dx_1 \\ &= \int x_1f(x_1)\left[\int x_2f(x_2)\,dx_2\right]dx_1, \end{aligned}$$

donde $\int x_2f(x_2)\,dx_2 = E[X_2]$. Luego:

$$\begin{aligned} E[X_1X_2] &= E[X_2]\int x_1f(x_1)\,dx_1 \\ &= E[X_1] \cdot E[X_2]. \end{aligned}$$

Observación **4.4** A continuación, se presentan algunos resultados para el valor esperado y la covarianza entre variables aleatorias independientes:

1. El resultado anterior se puede generalizar para el caso en que se desee determinar el valor esperado de $G_1(X_1) \cdot G_2(X_2)$, donde las variables aleatorias X_1 y X_2 son independientes. En tal caso:

$$E\left[G_1(X_1)\,G_2(X_2)\right] = E\left[G_1(X_1)\right]\,E\left[G_2(X_2)\right].$$

2. Como la $Cov(X_1, X_2) = E[X_1X_2] - E[X_1]E[X_2]$ y siendo X_1 y X_2 variables aleatorias independientes:

$$E[X_1X_2] = E[X_1]\,E[X_2] \Rightarrow Cov(X_1, X_2) = 0.$$

Propiedad **4.3** Sean X_1 y X_2 variables aleatorias, entonces se cumple que:

$$Var(X_1 \pm X_2) = Var(X_1) + Var(X_2) \pm 2 \cdot Cov(X_1, X_2).$$

Demostración: Para la varianza de la diferencia de variables aleatorias

(análogo para la suma):

$$\begin{aligned}V(X_1 - X_2) &= E\left[(X_1 - X_2)^2\right] - (E[X_1 - X_2])^2 \\ &= E[X_1^2 + X_2^2 - 2X_1X_2] - \left((E[X_1])^2 + (E[X_2])^2 - 2E[X_1]E[X_2]\right) \\ &= E[X_1^2] + E[X_2^2] - 2\,E[X_1X_2] - (E[X_1])^2 - (E[X_2])^2 + 2E[X_1]E[X_2] \\ &= \left(E[X_1^2] - (E[X_1])^2\right) + \left(E[X_2^2] - (E[X_2])^2\right) - 2\left(E[X_1X_2] + E[X_1]E[X_2]\right),\end{aligned}$$

donde $E[X_i^2] - (E[X_i])^2 = Var(X_i)$ y $E[X_1X_2] + E[X_1]E[X_2] = Cov(X_1, X_2)$:

$$= Var(X_1) + Var(X_2) - 2 \cdot Cov(X_1, X_2).$$

Esperanza y varianza de distribuciones condicionales

Definición **4.8** Sean las variables aleatorias X_1, X_2, $X_1 \mid X_2 = x_2$ y $X_2 \mid X_1 = x_1$, a partir de las cuales se definen expresiones para el valor esperado y la varianza condicional:

$$\begin{aligned}(1)\quad & E[X_1 \mid X_2 = x_2] = \int x_1 f_{X_1|X_2=x_2}(x_1)\,dx_1 \\ & E[X_2 \mid X_1 = x_1] = \int x_2 f_{X_2|X_1=x_1}(x_2)\,dx_2. \\ (2)\quad & Var(X_1 \mid X_2 = x_2) = E[X_1^2 \mid X_2 = x_2] - (E[X_1 \mid X_2 = x_2])^2 \\ & Var(X_2 \mid X_1 = x_1) = E[X_2^2 \mid X_1 = x_1] - (E[X_2 \mid X_1 = x_1])^2.\end{aligned}$$

Ejemplo **4.8** Sea $f(x_1, x_2) = e^{-x_1}$, $0 < x_2 < x_1 < \infty$ la función de densidad conjunta de (X_1, X_2) y sea $f_{X_1|X_2=x_2}(x_1) = \exp\{x_2 - x_1\}$, $x_2 < x_1 < \infty$, con x_2 un valor fijo. Determinar la esperanza y la varianza de $X_1 \mid X_2 = x_2$.

Solución: Aplicando la definición 4.8:

$$\begin{aligned}E[X_1 \mid X_2 = x_2] &= \int_{x_2}^{\infty} x_1 f_{X_1|X_2=x_2}(x_1)\,dx_1 = \int_{x_2}^{\infty} x_1 e^{x_2 - x_1}\,dx_1 \\ &= e^{x_2}\int_{x_2}^{\infty} x_1 e^{-x_1}\,dx_1 = 1 + x_2.\end{aligned}$$

En el caso de la varianza es necesario obtener:

$$E[X_1^2 \mid X_2 = x_2] = \int_{x_2}^{\infty} x_1^2 f_{X_1|X_2=x_2}(x_1)\, dx_1 = \int_{x_2}^{\infty} x_1^2 e^{x_2-x_1}\, dx_1$$
$$= e^{x_2} \int_{x_2}^{\infty} x_1^2 e^{-x_1}\, dx_1 = 2 + 2x_2 + x_2^2.$$

Finalmente, la varianza es:

$$Var(X_1 \mid X_2 = x_2) = 2 + 2x_2 + x_2^2 - (1 + x_2)^2$$
$$= 2 + 2x_2 + x_2^2 - 1 - 2x_2 - x_2^2 = 1.$$

□

Propiedad **4.4** Sean X_1 y X_2 variables aleatorias y sea $X_1 \mid X_2 = x_2$ la variable aleatoria condicional de X_1 dado un valor fijo x_2 de la variable aleatoria X_2, se tienen los siguientes resultados:

(a) $E\left[E[X_1 \mid X_2]\right] = E[X_1]$.

(b) $Var(X_1) = E[Var(X_1 \mid X_2)] + Var(E[X_1 \mid X_2])$.

Demostración: (a) Valor esperado del valor esperado condicional:

$$E[E[X_1 \mid X_2]] = \int E[X_1 \mid X_2 = x_2] f_{X_2}(x_2)\, dx_2$$
$$= \int \left(\int x_1 f_{X_1|X_2}(x_1) dx_1 \right) f_{X_2}(x_2)\, dx_2$$
$$= \iint x_1 f(x_1, x_2)\, dx_1\, dx_2 = E[X_1].$$

(b) Varianza marginal como función de distribuciones condicionales:

$$E[Var(X_1 \mid X_2)] = E[E[X_1^2 \mid X_2] - (E[X_1 \mid X_2])^2]$$
$$= E[E[X_1^2 \mid X_2]] - E[(E[X_1 \mid X_2])^2]$$
$$= E[X_1^2] - E[(E[X_1 \mid X_2])^2].$$
$$Var(E[X_1 \mid X_2]) = E[(E[X_1 \mid X_2])^2] - (E[E[X_1 \mid X_2]])^2$$
$$= E[(E[X_1 \mid X_2])^2] - (E[X_1])^2.$$

Finalmente, se obtiene: $E[Var(X_1 \mid X_2)] - Var(E[X_1 \mid X_2]) = E[X_1^2] - (E[X_1])^2 = Var(X_1)$.

Observación **4.5** Sea $G(X_i)$ función de la variable aleatoria X_i, $i = 1, 2$ y sea $X_1 \mid X_2 = x_2$ la variable aleatoria condicional de X_1, dado un valor fijo x_2 de la variable aleatoria X_2, se tiene que:

$$E\left[E[G(X_1 \mid X_2)]\right] = E[G(X_1)].$$

Observación **4.6** Nótese que para los resultados expresados en la propiedad 4.4, se obtienen resultados análogos para $X_2 \mid X_1$.

4.2. Transformaciones de variables aleatorias bivariadas

En ocasiones el tratamiento analítico de distribuciones de variables aleatorias bivariadas genera complicaciones. En estos casos se puede recurrir a transformaciones que permitan resolver el problema, las que implican cambios en el espacio donde está definida la función de densidad por medio de la función que define la transformación. Así, es necesario recurrir a herramientas de álgebra lineal donde se definen las condiciones y el algoritmo de resolución.

Definición **4.9** Sea $(X_1, X_2)^T$ una variable aleatoria bivariada con función de densidad $f(x_1, x_2)$ y sea, además, la nueva variable aleatoria bivariada $(Y_1, Y_2)^T$, tal que:

$$(Y_1, Y_2)^T = (G_1(X_1, X_2), G_2(X_1, X_2))^T.$$

Es decir: $Y_1 = G_1(X_1, X_2)$ e $Y_2 = G_2(X_1, X_2)$. Donde G_1 y G_2 son transformaciones continuas que admiten derivadas parciales. Entonces, la función de densidad conjunta de $(Y_1, Y_2)^T$ está dada por:

$$f_{Y_1,Y_2}(y_1, y_2) = f_{X_1,X_2}(G_1^{-1}(y_1, y_2), G_2^{-1}(y_1, y_2))\,|J|\,, \tag{4.2}$$

donde $|J|$ es el valor absoluto del jacobiano, que es el determinante de una matriz construida con las derivadas parciales de la transformación definida por:

$$J = \begin{bmatrix} \frac{\partial}{\partial y_1} G_1^{-1}(y_1, y_2) & \frac{\partial}{\partial y_2} G_1^{-1}(y_1, y_2) \\ \frac{\partial}{\partial y_1} G_2^{-1}(y_1, y_2) & \frac{\partial}{\partial y_2} G_2^{-1}(y_1, y_2) \end{bmatrix}$$

Observación **4.7** Nótese que si X_1 y X_2 son variables aleatorias independientes, entonces la ecuación 4.2 se puede escribir como:

$$f_{Y_1,Y_2}(y_1,y_2) = f_{X_1}(G_1^{-1}(y_1,y_2))f_{X_2}(G_2^{-1}(y_1,y_2))\,|J|\,.$$

Ejemplo **4.9** Sea $f(x_1,x_2) = \frac{1}{2\pi\sqrt{x_2}}\exp\left\{-\frac{x_1^2+x_2}{2}\right\}$, $x_1 \in \mathbb{R}$, $x_2 > 0$, considerar la transformación dada por:

$$Y_1 = \frac{X_1}{\sqrt{X_2}} \text{ y } Y_2 = X_2.$$

Determinar la función de densidad conjunta de la variable aleatoria bivariada $(Y_1,Y_2)^T$.

Solución: Desde la transformación indicada se obtiene $x_1(y_1,y_2)$ y $x_2(y_1,y_2)$ y, posteriormente, el jacobiano de la transformación:

$$y_1 = \frac{x_1}{\sqrt{x_2}} \Rightarrow x_1(y_1,y_2) = y_1\sqrt{y_2} \qquad y_2 = x_2 \Rightarrow x_2(y_1,y_2) = y_2.$$

$$J = \begin{bmatrix} \sqrt{y_2} & \frac{y_1}{2\sqrt{y_2}} \\ 0 & 1 \end{bmatrix} \Rightarrow |J| = \sqrt{y_2}.$$

La función de densidad conjunta de $(Y_1,Y_2)^T$ está dada por:

$$\begin{aligned} f_{Y_1,Y_2}(y_1,y_2) &= f_{X_1,X_2}(y_1\sqrt{y_2},y_2)\cdot\sqrt{y_2} \\ &= \frac{1}{2\pi\sqrt{y_2}}\exp\left\{-\frac{y_1^2y_2+y_2}{2}\right\}\sqrt{y_2} \\ &= \frac{1}{2\pi}\exp\left\{-\frac{y_1^2y_2+y_2}{2}\right\},\ y_1 \in \mathbb{R},\ y_2 > 0. \end{aligned}$$

□

Ejemplo **4.10** Sean $X_i \sim \chi_1^2$, $i = 1,2$ variables aleatorias independientes. Encontrar la función de probabilidad de la nueva variable aleatoria bivariada $(Y_1,Y_2)^T$ dada por la transformación:

$$(Y_1,Y_2)^T = \left(\frac{X_1}{X_1+X_2}, X_2\right)^T$$

.

Solución: Si $X_i \sim \chi^2_1$, entonces $f_{X_i}(x_i) = \left(\sqrt{2\pi\, x_i}\right)^{-1} \exp\left\{-\frac{x_i}{2}\right\}$, $x_i > 0$, $i = 1, 2$ y por independencia se tiene que la función de densidad conjunta de $(X_1, X_2)^T$ está dada por:

$$f_{X_1,X_2}(x_1, x_2) = \frac{1}{2\pi}(x_1 x_2)^{-\frac{1}{2}} \exp\left\{-\frac{x_1 + x_2}{2}\right\}, \quad x_1 > 0, \quad x_2 > 0.$$

Luego, $x_1(y_1, y_2)$, $x_2(y_1, y_2)$ y el jacobiano de la transformación están dados por:

$$y_1 = \frac{x_1}{x_1 + x_2} \Rightarrow x_1(y_1, y_2) = \frac{y_1 y_2}{1 - y_1} \quad \text{y} \quad y_2 = x_2 \Rightarrow x_2(y_1, y_2) = y_2.$$

$$J = \begin{bmatrix} \frac{y_2}{(1-y_1)^2} & \frac{y_1}{1-y_1} \\ 0 & 1 \end{bmatrix} \Rightarrow |J| = \frac{y_2}{(1 - y_1)^2}.$$

La función de densidad conjunta de $(Y_1, Y_2)^T$ está dada por:

$$\begin{aligned} f_{Y_1,Y_2}(y_1, y_2) &= f_{X_1,X_2}\left(\frac{y_1 y_2}{1 - y_1}, y_2\right) \cdot \frac{y_2}{(1 - y_1)^2} \\ &= \frac{1}{2\pi}\left(\frac{y_1 y_2^2}{1 - y_1}\right)^{-\frac{1}{2}} \exp\left\{-\left(\frac{y_1 y_2}{2(1 - y_1)} + \frac{y_2}{2}\right)\right\} \cdot \frac{y_2}{(1 - y_1)^2} \\ &= \frac{1}{2\pi} \frac{1}{\sqrt{y_1(1 - y_1)^3}} \exp\left\{-\left(\frac{y_1 y_2}{2(1 - y_1)} + \frac{y_2}{2}\right)\right\}, \\ & y_2 > 0, \quad 0 < y_1 < 1. \end{aligned}$$

□

Ejercicio propuesto **4.3** Sea (X, Y) una variable aleatoria bidimensional con función de densidad conjunta:

$$f(x, y) = \frac{1}{\pi}, \quad 0 < x^2 + y^2 \leqslant 1.$$

Demostrar que las variables aleatorias $U = \sqrt{X^2 + Y^2}$ y $V = \arctan\left\{\frac{Y}{X}\right\}$ son independientes.

Algunas distribuciones de probabilidad construidas a través de transformaciones bivariadas

En esta sección se presentan algunas distribuciones de probabilidad generadas a partir de transformaciones bivariadas. Es decir, se obtiene una

nueva variable aleatoria desde la definición de una operación entre variables aleatorias. Para obtener la distribución de la nueva variable aleatoria se puede utilizar la fgm (en casos aditivos), pero en casos multiplicativos se puede aplicar la técnica de transformación de variables aleatorias considerando un vector bivariado y luego calculando la función de densidad marginal para la variable de interés.

Distribución chi-cuadrado

Propiedad **4.5** Sea una variable aleatoria $Z \sim N(0,1)$, entonces la nueva variable aleatoria $X = Z^2$ sigue una distribución chi-cuadrado con 1 grado de libertad, denotado por $X \sim \chi^2(1)$.

Además, si se tiene X_1 y X_2, variables aleatorias que siguen una distribución chi-cuadrado independientes con n_1 y n_2 grados de libertad, respectivamente, entonces la variable aleatoria $W = X_1 + X_2$ distribuye chi-cuadrado con $n_1 + n_2$ grados de libertad, lo que se denota con:

$$W = X_1 + X_2 \sim \chi^2(n_1 + n_2).$$

Demostración: Para mostrar que $X \sim \chi^2(1)$, se puede utilizar la función de distribución, pues $F'(x) = f(x)$:

$$F_X(t) = \mathbb{P}(X \leqslant t) = \mathbb{P}(Z^2 \leqslant t) = F_Z\left(\sqrt{t}\right) - F_Z\left(-\sqrt{t}\right).$$

Aplicando la derivada parcial respecto a t se tiene que:

$$f_X(t) = \frac{1}{\sqrt{t}} f_Z\left(\sqrt{t}\right) = \frac{1}{\sqrt{2\pi t}} \exp\left\{-\frac{1}{2}t\right\}, \quad t > 0.$$

Por su parte, para demostrar que $W \sim \chi^2(n_1 + n_2)$, se puede utilizar la fgm de la distribución chi-cuadrado. Es decir:

$$\begin{aligned} M_W(t) =& E\left[\exp\left\{t(X_1 + X_2)\right\}\right] = E\left[\exp\left\{t \cdot X_1\right\}\right] \cdot E\left[\exp\left\{t \cdot X_2\right\}\right] \\ =& M_{X_1}(t) \cdot M_{X_2}(t), \end{aligned}$$

donde que X_1 y X_2 son independientes, se tiene que:

$$= \left[M_{X_1}(t)\right]^2 = (1 - 2t)^{-\frac{2}{2}}.$$

La última expresión corresponde a la fgm de una variable aleatoria con distribución chi-cuadrado con 2 grados de libertad.

Distribución t-Student

Propiedad **4.6** Considerando las variables aleatorias independientes $Z \sim N(0,1)$ y $Y \sim \chi^2(n)$, la nueva variable aleatoria T, dada por la expresión:

$$T = \frac{Z}{\sqrt{Y \cdot n^{-1}}},$$

se dice que tiene una distribución *t-Student* con n grados de libertad.

Demostración: Para determinar la función de densidad de la distribución t-Student, se utilizará una transformación bivariada de $(T,Z)^T \rightarrow (T_1,T_2)^T$ definida convenientemente por:

$$T_1 = \frac{Z}{\sqrt{Y \cdot n^{-1}}} \text{ y } T_2 = Y,$$

luego se obtiene la función de densidad de $(T_1,T_2)^T$ y la función de densidad marginal de interés, que en este caso es la asociada a T_1. La función de densidad de $(T_1,T_2)^T$ es:

$$t_1 = \frac{z}{\sqrt{y \cdot n^{-1}}} \Rightarrow z = t_1\sqrt{t_2 \cdot n^{-1}} \text{ y } t_2 = y$$

$$J = \begin{bmatrix} \sqrt{t_2 \cdot n^{-1}} & h(t,y) \\ 0 & 1 \end{bmatrix} \Rightarrow |J| = \sqrt{t_2 \cdot n^{-1}}.$$

Por otra parte, la función de densidad conjunta de $(T,Y)^T$, por ser variables aleatorias independientes, es:

$$f_{Z,Y}(z,y) = f_Z(z)f_Y(y) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{z^2}{2}\right\} \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} \exp\left\{-\frac{y}{2}\right\}.$$

Finalmente, la función de densidad conjunta de $(T_1,T_2)^T$ es:

$$\begin{aligned} f_{T_1,T_2}(t_1,t_2) &= f_{X_1,X_2}(t_1\sqrt{t_2 \cdot n^{-1}}, t_2) \cdot \sqrt{t_2 \cdot n^{-1}} \\ &= \frac{1}{\sqrt{2\pi}2^{\frac{n}{2}}\Gamma(\frac{n}{2})} \exp\left\{-\frac{t_2}{2}\left(\frac{t_1^2}{n}+1\right)\right\}, \quad t_2 > 0, \ \ t_1 > 0. \end{aligned}$$

Luego, la densidad marginal de T_1 es:

$$\begin{aligned} f_{T_1}(t_1) &= \int_0^\infty f_{T_1,T_2}(t_1,t_2)\,dt_2 \\ &= \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n}{2})} \left(\frac{t_1^2}{n}+1\right)^{-\frac{n+1}{2}}, -\infty < t_1 < +\infty. \end{aligned}$$

Donde la última expresión corresponde a la función de densidad de una variable aleatoria con distribución t-Student con n grados de libertad.

Ejercicio propuesto **4.4** Demostrar que las expresiones para el valor esperado y la varianza para una variable aleatoria con distribución t-Student están dadas por: 0 y $\frac{n}{n-2}$, respectivamente.

Distribución F de Fisher

Propiedad **4.7** Considerar las variables aleatorias independientes $X_i \sim \chi^2_{n_i}$, $i = 1, 2$. Se define la variable aleatoria F por la expresión:

$$F = \frac{X_1 \cdot n_1^{-1}}{X_2 \cdot n_2^{-1}},$$

se dice que sigue una distribución *F de Fisher* con n_1 y n_2 grados de libertad (correspondientes a los grados de libertad del numerador y denominador, respectivamente).

Demostración: La función de densidad de la distribución F de Fisher se puede obtener definiendo una transformación bivariada y luego la distribución marginal adecuada. En efecto, sean:

$$y_1 = \frac{x_1 \cdot n_1^{-1}}{x_2 \cdot n_2^{-1}} \Rightarrow x_1 = y_1 \cdot y_2 \cdot \frac{n_1}{n_2} \quad \text{y} \quad y_2 = x_2.$$

$$J = \begin{bmatrix} \frac{n_1}{n_2} y_2 & h(t,y) \\ 0 & 1 \end{bmatrix} \Rightarrow |J| = y_2 \frac{n_1}{n_2}.$$

Por su parte, la función de densidad conjunta de $(Y_1, Y_2)^T$, por ser variables aleatorias independientes, es el producto de las funciones de densidad marginal de ambas variables aleatorias. Donde:

$$f_{X_i}(x_i) = \frac{1}{2^{\frac{n_i}{2}} \Gamma\left(\frac{n_i}{2}\right)} x_i^{\frac{n_i}{2}-2} \exp\left\{-\frac{x_i}{2}\right\} \quad x_i > 0.$$

Luego, la función de densidad conjunta de $(Y_1, Y_2)^T$ está dada por:

$$\begin{aligned} f_{Y_1,Y_2}(y_1, y_2) &= f_{X_1}\left(y_1\, y_2 \frac{n_1}{n_2}\right) \cdot f_{X_2}(y_2) \cdot y_2 \frac{n_1}{n_2} \\ &= \frac{1}{\Gamma\left(\frac{n_1}{2}\right) \Gamma\left(\frac{n_2}{2}\right) 2^{\frac{n_1+n_2}{2}}} y_1^{\frac{n_1}{2}-1} y_2^{\frac{n_1+n_2}{2}-1} \cdot \\ &\cdot \exp\left\{-\frac{1}{2}\left(y_1 \frac{n_1}{n_2} + 1\right) y_2\right\}, \quad y_1 > 0, \quad y_2 > 0. \end{aligned}$$

Finalmente, la función de densidad asociada a la distribución F de Fisher corresponde a la función de densidad marginal de Y_1. En efecto:

$$f_{Y_1}(y_1) = \frac{1}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)2^{\frac{n_1+n_2}{2}}} y_1^{\frac{n_1}{2}-1} \int_0^{\infty} y_2^{\frac{n_1+n_2}{2}-1} \exp\left\{-\frac{1}{2}\left(y_1\frac{n_1}{n_2}+1\right)y_2\right\} dy_2.$$

Luego, llevando la integral a una del tipo gamma se puede obtener lo siguiente:

$$= \frac{\Gamma\left(\frac{n_1+n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)} \left(y_1\frac{n_1}{n_2}+1\right) y_1^{\frac{n_1}{2}-1}, \quad y_1 > 0.$$

Ejercicio propuesto **4.5** Como Y_1 sigue una distribución F de Fisher, utilizar la función de densidad para demostrar que la esperanza y la varianza están dadas respectivamente por:

$$\frac{n_2}{n_2-2}, \quad n_2 > 2 \quad \text{y} \quad \frac{2\,n_2^2\,(n_1+n_2-2)}{n_1(n_2-2)^2(n_2-4)}, \quad n_2 > 4.$$

4.3. Vectores aleatorios multivariados

Esta sección es una extensión de vectores aleatorios bivariados considerando ahora un vector de dimensión $(n \times 1)$, cuyas componentes son variables aleatorias. La sección comienza con algunas definiciones acerca del espacio en que se construyen la medida de probabilidad y la función de densidad multivariada (y los elementos que pueden obtenerse a partir de esta), para luego abordar el concepto de muestra aleatoria, que es el que permite dar el paso a las distribuciones muestrales y, posteriormente, a la inferencia estadística.

Definición **4.10** Sea Ω un espacio muestral asociado a un experimento aleatorio. Se llamará vector aleatorio $\boldsymbol{X}$ aquel vector cuyas componentes solo son variables aleatorias unidimensionales, es decir, X_i, $i = 1, 2, \cdots, n$,

es una variable aleatoria en $\mathbb{R}$ y se denota con:

$$\boldsymbol{X} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} = (X_1, X_2, \cdots, X_n)^T$$

Función de densidad conjunta

Definición **4.11** Sea $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)^T$ un vector aleatorio. La función $f : \mathbb{R}^n \to \mathbb{R}$ es una función de densidad conjunta del vector aleatorio $\boldsymbol{X}$ si y solo si cumple con las siguientes condiciones:

(1) $f(x_1, x_2, \cdots, x_n) \geqslant 0 \ \forall (x_1, x_2, \cdots, x_n)^T \in \mathbb{R}^n$.

(2) $\iint \cdots \int f(x_1, x_2, \cdots, x_n)\, dx_n\, dx_{n-1} \cdots dx_1 = 1$.

Ejemplo **4.11** Probar que la función asociada al vector aleatorio $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)^T$ es función de densidad y está dada por:

$$f(x_1, x_2, \cdots, x_n) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2\right\},$$
$$x_i \in \mathbb{R}, i = 1, 2, \cdots, n,$$

donde $\mu \in \mathbb{R}$ y $\sigma \in \mathbb{R}^+$ son constantes.

Solución: Para que $f(\boldsymbol{x})$ sea función de densidad debe cumplir con las dos condiciones dadas por la definición 4.11. En efecto:

(1) $f(\boldsymbol{x}) > 0$ y (2) la integral multivariada es:

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}(2\pi\sigma^2)^{-\frac{1}{2}}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right\} =$$
$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}(2\pi\sigma^2)^{-\frac{n}{2}}\cdot(2\pi\sigma^2)^{-\frac{n}{2}}\cdots(2\pi\sigma^2)^{-\frac{n}{2}}\cdot$$
$$\exp\left\{-\frac{1}{2\sigma^2}(x_1-\mu)^2\right\}\cdots\exp\left\{-\frac{1}{2\sigma^2}(x_n-\mu)^2\right\}dx_n\cdots dx_1.$$

En el resultado anterior, cada una de las integrales que se generan tiene la siguiente forma:

$$\int_{-\infty}^{+\infty}(2\pi\sigma^2)^{-\frac{n}{2}}\exp\left\{-\frac{1}{2\sigma^2}(x_i-\mu)^2\right\}dx_i = 1,\ i = 1, 2, \cdots, n.$$

Nótese que la función en cada una de las n integrales corresponde a la función de densidad de la distribución normal, por lo que el resultado es: $1^n = 1$. Luego, $f(\boldsymbol{x})$ es función de densidad de la variable aleatoria multivariada $\boldsymbol{X}$. □

Función de densidad marginal multivariada

Se denotará $f_{X_j}(x_j)$, $j = 1, \cdots, n$ la función de densidad de la variable aleatoria X_j. La obtención de la función de densidad marginal desde un vector aleatorio multivariado es una extensión análoga al caso bivariado; dado que ahora se cuenta con n variables aleatorias, la función de densidad de la j-ésima variable aleatoria corresponde a la integral multivariada con respecto a las restantes $n-1$ variables.

Definición **4.12** Sea $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)^T$ un vector aleatorio. Las n variables aleatorias que componen este vector son variables aleatorias independientes si y solo si la función de densidad conjunta puede ser escrita como:

$$f(x_1, x_2, \cdots, x_n) = \prod_{j=1}^{n} f_{X_j}(x_j),$$

donde $f_{X_j}(x_j)$ es la función de densidad marginal de la variable aleatoria X_j, con $j = 1, 2, \cdots, n$ valor fijo.

Ejemplo **4.12** Desde el ejemplo 4.11, se probará que las variables aleatorias $X_1, X_2, \cdots, X_n$ son variables aleatorias independientes.

Solución: Desde la definición 4.12 se establece que el producto de las n funciones de densidad marginal debe ser igual a la función de densidad conjunta. En efecto, sea X_j la j-ésima variable aleatoria que compone el vector aleatorio $\boldsymbol{X}$, es decir, $j = 1, 2, \cdots, n$ es un valor fijo. Así:

$$\begin{aligned} f_{X_j}(x_j) =& (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma^2}(x_j-\mu)^2\right\} \cdot \\ & \cdot \iint \cdots \int (2\pi\sigma^2)^{-\frac{n-1}{2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i\neq j}^{n}(x_i-\mu)\right\} dx_n \cdots dx_1. \end{aligned}$$

Cada una de las $n-1$ integrales del resultado anterior es de la función de densidad de una variable aleatoria con distribución normal, por lo que

son iguales a 1. Finalmente, se tiene que la función de densidad marginal de X_j es:

$$f_{X_j}(x_j) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma^2}(x_j-\mu)^2\right\}, \quad -\infty < x_j < +\infty.$$

Dado que la expresión anterior es válida para $j = 1, 2, \cdots, n$, se puede probar que el producto de las n funciones de densidad es igual a la función de densidad de la variable aleatoria multivariada $\boldsymbol{X}$, con lo que se prueba que las variables aleatorias $X_1, X_2, \cdots, X_n$ son independientes. Es decir:

$$\begin{aligned}
\prod_{j=1}^{n} f_{X_j}(x_j) &= \prod_{j=1}^{n} (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma^2}(x_j-\mu)^2\right\} \\
&= \left((2\pi\sigma^2)^{-\frac{1}{2}}\right)^n \exp\left\{-\frac{1}{2\sigma^2}\sum_{j=1}^{n}(x_j-\mu)^2\right\} \\
&= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{j=1}^{n}(x_j-\mu)^2\right\} \\
&= f(x_1, x_2, \cdots, x_n).
\end{aligned}$$

□

Definición **4.13** Sea $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)^T$ un vector aleatorio. Las variables aleatorias que componen el vector aleatorio $\boldsymbol{X}$: $X_1, X_2, \cdots, X_n$ son llamadas idénticamente distribuidas (*id*) si y solo si cada una de estas tiene la misma distribución con los mismos parámetros (igual valor). Además, si las variables aleatorias idénticamente distribuidas son independientes, se habla de variables aleatorias *iid*.

Ejemplo **4.13** Sean las variables aleatorias $X_1, X_2, \cdots, X_n$ independientes idénticamente distribuidas y $X_i \sim \text{Exp}(\beta)$. Se desea obtener la función de densidad conjunta de las n variables.

Solución: Como $X_i \sim \text{Exp}(\beta)$ (con $i = 1, 2, \cdots, n$), entonces $f(x_i) = \frac{1}{\beta}\exp\left\{-\frac{1}{\beta}x_i\right\}$, $x_i > 0$. Por lo tanto, la función de densidad conjunta está

dada por:

$$f(x_1, x_2, \ldots, x_n) = \prod_{j=1}^{n} f_{X_j}(x_j) = \prod_{j=1}^{n} \frac{1}{\beta} \exp\left\{-\frac{1}{\beta} x_j\right\}$$
$$= \left(\frac{1}{\beta}\right)^n \exp\left\{-\frac{1}{\beta} \sum_{j=1}^{n} x_j\right\}.$$

□

Ejemplo **4.14** Las variables aleatorias $X_1, X_2, \cdots, X_n$ son independientes y $X_i \sim N(\mu, \sigma_i^2)$, para $i = 1, 2, \cdots, n$. Probar que las variables no son idénticamente distribuidas.

Solución: Las variables aleatorias no son idénticamente distribuidas pues el parámetro de la distribución $\boldsymbol{\theta} = (\mu, \sigma_i^2)^T$ cambia para cada función marginal. □

Definición **4.14** Los componentes de un vector aleatorio definido por $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)^T$ constituyen una muestra aleatoria si y solo si las n variables son:

(1) Independientes.

(2) Idénticamente distribuidas.

Ejemplo **4.15** Las variables aleatorias $X_1, X_2, \cdots, X_n$ constituyen una muestra aleatoria, donde $X_i \sim Binomial(1, p)$. Determinar la función de densidad conjunta.

Solución: La función de densidad conjunta es:

$$f(x_1, x_2, \ldots, x_n) = \prod_{j=1}^{n} f_{X_j}(x_j) = \prod_{j=1}^{n} p^{x_j}(1-p)^{1-x_j}$$
$$= p^{\sum_{j=1}^{n} x_j}(1-p)^{\sum_{j=1}^{n}(1-x_j)}$$
$$= p^{\sum_{j=1}^{n} x_j}(1-p)^{n-\sum_{j=1}^{n} x_j}.$$

□

Propiedad **4.8** Si $X_1, X_2, \cdots, X_n$ son variables aleatorias independientes y si $Y_1 = G_1(X_1), Y_2 = G_2(X_2), \cdots, Y_n = G_n(X_n)$ son nuevas variables aleatorias (funciones de las variables aleatorias iniciales), entonces:

$$E[Y_1 \cdot Y_2 \cdots Y_n] = E[Y_1] \cdot E[Y_2] \cdots E[Y_n].$$

Demostración: Dado que las variables aleatorias X_i $(i = 1, 2, \cdots, n)$ son independientes, entonces $f(\boldsymbol{x}) = f(x_1) \cdots f(x_n)$ y además $G(\boldsymbol{X}) = G_1(X_1) \cdot G_2(X_2) \cdots G_n(X_n)$, luego:

$$\begin{aligned}
E[Y_1 \cdot Y_2 \cdots Y_n] =& E\left[G(\boldsymbol{X})\right] = \iint \cdots \int G(\boldsymbol{x}) \cdot f(\boldsymbol{x})\, d\boldsymbol{x} \\
=& \iint \cdots \int G_1(X_1) \cdots G_n(X_n) \cdot f(x_1) \cdots f(x_n)\, dx_1 \cdots dx_n \\
=& \iint \cdots \int Y_1 \cdots G_n(X_n) \cdot f(x_1) \cdots f(x_n)\, dx_1 \cdots dx_n \\
=& \left(\int Y_1 \cdot G_1(X_1)\, dx_1 \right) \cdots (G_n(X_n) \cdot f(x_n)\, dx_n) \\
=& E[G_1(X_1)] \cdot E[G_1(X_2)] \cdots E[G_n(X_n)] \\
=& E[Y_1] \cdot E[Y_2] \cdots E[Y_n].
\end{aligned}$$

Propiedad **4.9** Si $X_1, X_2, \cdots, X_n$ son variables aleatorias, entonces la función generadora de momentos de la nueva variable aleatoria $Y = \sum_{i=1}^{n} X_i$ está dada por:

(1) $M_Y(t) = \prod_{i=1}^{n} M_{X_i}(t)$, si $X_1, X_2, \cdots, X_n$ son variables aleatorias independientes.

(2) $M_Y(t) = [M_{X_i}(t)]^n$, si $X_1, X_2, \cdots, X_n$ son variables aleatorias iid.

Demostración: (1) Si se asume que $G_i(X_i) = E\left[e^{t X_i}\right]$, $i = 1, 2, \cdots, n$, entonces por la propiedad 4.8 es inmediato el resultado. (2) Las variables $X_1, X_2, \cdots, X_n$ son iid, entonces:

$$M_Y(t) = M_{X_1}(t) \cdot M_{X_2}(t) \cdots M_{X_n}(t) = [M_{X_i}(t)]^n .$$

Ejemplo **4.16** Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde $X_i \sim N(0, 1)$. Determinar la distribución de las nuevas variables aleatorias:

(a) $Y = \sum_{i=1}^{n} X_i$.

(b) $Y = \sum_{i=1}^{n} X_i^2$.

Solución: (a) $X_1 \sim N(0,1) \Rightarrow M_{X_1}(t) = \exp\left\{\frac{1}{2}\,t^2\right\}$, además aplicando la propiedad 4.8:

$$\begin{aligned} M_Y(t) &= \left[\exp\left\{\frac{1}{2}\cdot t^2\right\}\right]^n = \exp\left\{\frac{1}{2}\cdot t^2\cdot n\right\} \\ &= \exp\left\{t\cdot 0 + \frac{1}{2}\cdot t^2\cdot n\right\}. \end{aligned}$$

Esta última expresión corresponde a la fgm de una variable que sigue una distribución normal con valor esperado $\mu = 0$ y varianza $\sigma^2 = n$. Entonces:

$$Y = \sum_{i=1}^{n} X_i \sim N(0,n).$$

(b) En el ejercicio resuelto 9(b) del capítulo 3, se comprobó que $Z \sim N(0,1) \Rightarrow T = Z^2 \sim \chi^2(1)$ (chi-cuadrado con un grado de libertad).

A partir de la fgm se obtiene:

$$M_Y(t) = [M_T(t)]^n = \left[(1-2t)^{-\frac{1}{2}}\right]^n = (1-2t)^{-\frac{n}{2}}.$$

La última expresión corresponde a la fgm de una variable aleatoria con distribución chi-cuadrado con n grados de libertad. □

Ejemplo **4.17** Para realizar una obra de construcción se utiliza arena y concreto. Quienes planifican necesitan conocer el costo de estos materiales. Así, considerando las variables X: *cantidad de arena* (en toneladas) e Y: *cantidad de concreto* (cientos de libras) y, siendo $X \sim N(\mu_1; 0{,}25)$, $Y \sim N(\mu_2; 0{,}04)$, donde se sabe que la arena tiene un costo de US 7 por tonelada y, el concreto, de US 3 por 100 libras y, además, se agregan US 100 por otro tipo de costos. El costo total está dado por $C(X,Y) = 100 + 7X + 3Y$, encontrar la distribución de la variable aleatoria C.

Solución: Se obtiene la distribución de C usando la función generadora de momentos:

$$\begin{aligned} M_C(t) =& E[e^{tX}] = E[\exp\{t(100+7X+3Y)\}] \\ =& E[\exp\{100t + 7Xt + 3Yt\}] \\ =& E[\exp\{100t\}\cdot\exp\{7Xt\}\cdot\exp\{3Yt\}] \\ =& \exp\{100t\}\,E[\exp\{7Xt\}]\cdot E[\exp\{3Yt\}] \\ =& \exp\{100t\}\cdot M_X(7t)\cdot M_Y(3t). \end{aligned}$$

Como $X \sim N(\mu_1; 0{,}25)$, entonces $M_X(t) = \exp\left\{7t\mu_1 + \frac{1}{2}(7t)^2 \cdot 0{,}25\right\}$; de forma análoga se obtiene $M_Y(t)$, luego:

$$\begin{aligned} M_C(t) &= \exp\{100t\} \cdot \exp\left\{7t\mu_1 + \frac{(7t)^2}{2} \cdot 0{,}25\right\} \cdot \exp\left\{3t\mu_2 + \frac{(3t)^2}{2} \cdot 0{,}04\right\} \\ &= \exp\left\{t(100 + 7\mu_1 + 3\mu_2) + \frac{1}{2}t^2(12{,}25 + 0{,}36)\right\}. \end{aligned}$$

Finalmente, a partir de este último resultado se tiene que $C \sim \mathrm{N}(100 + 7\mu_1 + 3\mu_2; 12{,}61)$. □

Distribución del promedio y la varianza muestral

La distribución del promedio y de la varianza muestral es un elemento previo al capítulo sobre estimación de parámetros. En este sentido, si bien tanto el promedio como la varianza muestral son elementos puntuales de una distribución, se pueden entender como variables aleatorias toda vez que el experimento de obtener una muestra aleatoria puede realizarse en las mismas condiciones repetidas veces, por lo que es posible obtener (teóricamente) un conjunto de valores puntuales que constituyen las variables aleatorias.

Definición **4.15** El promedio de una muestra aleatoria está definido como la media aritmética de un conjunto de variables aleatorias y se fundamenta en el hecho de que un experimento aleatorio puede reproducirse bajo las mismas características. El promedio, denotado con $\bar{X}_n$ para X_i, $i = 1, 2, \cdots, n$ o, simplemente, con $\bar{X}$, está dado por la expresión:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

La idea es obtener la distribución de una nueva variable aleatoria que representa al promedio. En efecto, a continuación se presenta una propiedad mediante la cual se puede asociar la distribución de una muestra aleatoria, basada en una distribución normal, con los parámetros de la distribución del promedio.

Propiedad **4.10** Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde $X_i \sim N(\mu, \sigma^2)$. La distribución de la nueva variable aleatoria $\bar{X}$ conocida como

el promedio de la muestra aleatoria sigue una distribución normal dada por:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \sim N\left(\mu, \frac{\sigma^2}{n}\right).$$

Demostración: Por la propiedad 4.14 y la propiedad 4.9(b), se tiene que:

$$\begin{aligned} M_T(t) =& E\left[e^{t\cdot T}\right] = E\left[\exp\left\{t\frac{1}{n}\sum_{i=1}^{n} X_i\right\}\right] = \left(E\left[\exp\left\{\frac{t}{n}X_1\right\}\right]\right)^n \\ &= \left(M_{X_1}\left(\frac{t}{n}\right)\right)^n = \left(\exp\left\{\frac{t}{n}\mu + \frac{1}{2}\left(\frac{t}{n}\right)^2\sigma^2\right\}\right)^n \\ &= \exp\left\{t\mu + \frac{1}{2}t^2\frac{\sigma^2}{n}\right\}. \end{aligned}$$

Esta última expresión corresponde a la distribución de una variable aleatoria normal con los parámetros pedidos.

Además, se puede estandarizar la distribución del promedio y calcular probabilidades desde una distribución normal estándar. Si llamamos $\bar{X}$ al promedio de una muestra aleatoria $X_1, X_2, \cdots, X_n$, donde $X_i \sim N(\mu, \sigma^2)$, la distribución Z estandarizada se puede construir mediante:

$$Z = \frac{\bar{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}} \sim N(0, 1).$$

Definición **4.16** Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, se define la varianza muestral por:

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2.$$

Propiedad **4.11** Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria donde $X_i \sim N(\mu, \sigma^2)$, entonces la variable aleatoria Y, que contiene la varianza muestral S^2, tiene una distribución:

$$Y = \frac{n-1}{\sigma^2}S^2 \Rightarrow Y \sim \chi^2(n-1).$$

Demostración: Para mostrar este resultado, hay que escribir una expresión equivalente a Y, de tal manera que aparezcan μ y σ^2. En efecto:

$$\begin{aligned} Y =& \frac{n-1}{\sigma^2}S^2 = \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \bar{X})^2 \\ =& \frac{1}{\sigma^2}\sum_{i=1}^{n}\left[(X_i - \mu) - (\bar{X} - \mu)\right]^2 \\ =& \frac{1}{\sigma^2}\sum_{i=1}^{n}\left[(X_i - \mu)^2 + (\bar{X} - \mu)^2 - 2(X_i - \mu)(\bar{X} - \mu)\right] \\ =& \frac{1}{\sigma^2}\left[\sum_{i=1}^{n}(X_i - \mu)^2 + \sum_{i=1}^{n}(\bar{X} - \mu)^2 - 2(\bar{X} - \mu)\sum_{i=1}^{n}(X_i - \mu)\right]. \end{aligned}$$

Luego, se puede reexpresar la última expresión considerando que $\sum_{i=1}^{n}(X_i - \mu) = n\bar{X} - n\mu = n(\bar{X} - \mu)$, con lo cual:

$$Y = \frac{1}{\sigma^2}\left[\sum_{i=1}^{n}(X_i - \mu)^2 - n(\bar{X} - \mu)^2\right]. = \sum_{i=1}^{n}\left(\frac{X_i - \mu}{\sigma}\right)^2 - \left(\frac{\bar{X} - \mu}{\sigma \cdot n^{-\frac{1}{2}}}\right)^2.$$

Por otra parte, analizando la distribuciones de la última expresión:

$$\begin{aligned} X_i \sim N(\mu, \sigma^2) \Rightarrow & \frac{X_i - \mu}{\sigma} \sim N(0,1) \Rightarrow \left(\frac{X_i - \mu}{\sigma}\right)^2 \sim \chi^2(1) \\ & \Rightarrow \sum_{i=1}^{n}\left(\frac{X_i - \mu}{\sigma}\right)^2 \sim \chi^2(n). \end{aligned}$$

Efectuando un análisis similar al anterior:

$$\left(\frac{\bar{X} - \mu}{\sigma \cdot n^{-\frac{1}{2}}}\right)^2 \sim \chi^2(1) \Rightarrow Y \sim \chi^2(n-1).$$

Distribución del mínimo y del máximo

Sean $X_1, X_2, \cdots, X_n$ variables aleatorias independientes, donde X_i tiene una función de densidad $f_{X_i}(x_i)$, se conoce como el *mínimo* a la variable aleatoria más pequeña (en el sentido matemático de la expresión) y al *máximo* como la variable aleatoria de mayor valor. El mínimo y el máximo se denotan respectivamente con:

(a) $M = \text{mín}\{X_1, X_2, \cdots, X_n\}$.

(b) $T = \text{máx}\{X_1, X_2, \cdots, X_n\}$.

Luego, conociendo la función de densidad de las variables aleatorias X_i, es de interés conocer la distribución de la variables aleatorias M y T definidas anteriormente. A continuación, se presenta una propiedad que permite solucionar esta inquietud.

Propiedad **4.12** Las funciones de densidad de las variables aleatorias M y T, conocidas como el mínimo y el máximo de las variables aleatorias $X_1, X_2, \cdots, X_n$, están dadas por:

$$f_T(t) = n \cdot \left[F_{X_1}(t)\right]^{n-1} \cdot f_{X_1}(t). \tag{4.3}$$

$$f_M(m) = n \cdot \left[1 - F_{X_1}(m)\right]^{n-1} \cdot f_{X_1}(m). \tag{4.4}$$

Demostración: (4.3) En este caso:

$$\begin{aligned} F_T(t) =& \mathbb{P}(T \leqslant t) = \mathbb{P}(\text{máx}\{X_1, X_2, \cdots, X_n\} \leqslant t) \\ =& \mathbb{P}(\{X_1 \leqslant t\} \cap \{X_2 \leqslant t\} \cap \cdots \cap \{X_n \leqslant t\}) \\ =& \mathbb{P}(X_1 \leqslant t) \cdot \mathbb{P}(X_2 \leqslant t) \cdots \mathbb{P}(X_n \leqslant t), \text{ donde } X_i \text{ es } id \\ =& F_{X_1}(t) \cdot F_{X_2}(t) \cdots F_{X_n}(t) \\ =& \left[F_{X_1}(t)\right]^n. \end{aligned}$$

Realizando la derivada parcial respecto a t:

$$f_T(t) = n \cdot \left[F_{X_1}(t)\right]^{n-1} \cdot f_{X_1}(t).$$

(4.4) Al igual que el caso anterior, la función de densidad del mínimo se obtiene mediante:

$$\begin{aligned} F_M(m) =& \mathbb{P}(M \leqslant m) = \mathbb{P}(\text{mín}\{X_1, X_2, \cdots, X_n\} \leqslant m) \\ =& 1 - \mathbb{P}(\text{mín}\{X_1, X_2, \cdots, X_n\} > m) \\ =& 1 - \mathbb{P}(\{X_1 > m\} \cap \{X_2 > m\} \cap \cdots \cap \{X_n > m\}) \\ =& 1 - \mathbb{P}(X_1 > m) \cdot \mathbb{P}(X_2 > m) \cdots \mathbb{P}(X_n > m) \\ =& 1 - (1 - F_{X_1}(m)) \cdot (1 - F_{X_2}(m)) \cdots (1 - F_{X_n}(m)) \\ =& 1 - \left[1 - F_{X_1}(m)\right]^n. \end{aligned}$$

Realizando la derivada parcial respecto a m:

$$f_M(m) = n \cdot \left[1 - F_{X_1}(m)\right]^{n-1} \cdot f_{X_1}(m).$$

Ejemplo **4.18** Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde $X_i \sim \text{Exp}(\beta)$. Determinar la función de densidad del mínimo y del máximo.

Solución: Para obtener las funciones de densidad para las variables aleatorias se utiliza la propiedad 4.12. En efecto, dado que $X_i \sim \text{Exp}(\beta)$, se tiene que:

$$f_{X_i}(x) = \frac{1}{\beta}\exp\left\{-\frac{1}{\beta}x\right\}, \quad x > 0.$$
$$F_{X_i}(x) = \int_0^x \frac{1}{\beta}\exp\left\{-\frac{1}{\beta}t\right\}\, dt = 1 - \exp\left\{-\frac{1}{\beta}x\right\}, \quad x > 0.$$

Luego, la función de densidad del máximo es:

$$f_T(t) = n \cdot \left[1 - \exp\left\{-\frac{1}{\beta}t\right\}\right]^{n-1} \cdot \frac{1}{\beta}\exp\left\{-\frac{1}{\beta}t\right\}, \ \ t > 0.$$

Por su parte, la función de densidad del mínimo es:

$$\begin{aligned} f_M(m) =& n \cdot \left[1 - 1 + \exp\left\{-\frac{1}{\beta}m\right\}\right]^{n-1} \cdot \frac{1}{\beta}\exp\left\{-\frac{1}{\beta}m\right\} \\ =& \frac{n}{\beta} \cdot \exp\left\{-\frac{n}{\beta}m\right\}, \ \ m > 0. \end{aligned}$$

De los resultados obtenidos es inmediato que $M \sim \text{Exp}(\frac{\beta}{n})$. □

4.4. Ejercicios resueltos

1. Sean A y B en $(\Omega, \mathcal{F}, P)$, se definen las siguientes variables aleatorias:

$$X = \begin{cases} 1, & \text{si } A \text{ ocurre} \\ 0, & \text{si no} \end{cases} \qquad Y = \begin{cases} 1, & \text{si } B \text{ ocurre} \\ 0, & \text{si no.} \end{cases}$$

Obtener: (a) Ω, (b) Rec_{XY}, (c) $\mathbb{P}(X, Y)$, (d) $p_X(x)$ y $p_Y(y)$.

Solución: (a) $\Omega = (A^c \cap B^c) \cup (A^c \cap B) \cup (A \cap B^c) \cup (A \cap B)$.

(b) $Rec_{XY} = \{(0,0), (0,1), (1,0), (1,1)\} = \{0,1\} \times \{0,1\} = Rec_X \times Rec_Y$. (c)

$$P_{XY}(x,y) = \begin{cases} \mathbb{P}(A^c \cap B^c), & \text{si} \quad x = y = 0 \\ \mathbb{P}(A^c \cap B), & \text{si} \quad x = 0, y = 1 \\ \mathbb{P}(A \cap B^c), & \text{si} \quad x = 1, y = 0 \\ \mathbb{P}(A \cap B), & \text{si} \quad x = y = 1. \end{cases}$$

(d) Las distribuciones marginales de X y de Y se pueden deducir desde la relación de la tabla siguiente:

X \ Y	0	1	$\mathbb{P}(X=x)$
0	$\mathbb{P}(A^c \cap B^c)$	$\mathbb{P}(A^c \cap B)$	$\mathbb{P}(A^c)$
1	$\mathbb{P}(A \cap B^c)$	$\mathbb{P}(A \cap B)$	$\mathbb{P}(A)$
$\mathbb{P}(Y=y)$	$\mathbb{P}(B^c)$	$\mathbb{P}(B)$	1

Entonces, siendo $\mathbb{P}(X=x) = p_X(x)$ y $\mathbb{P}(Y=y) = p_Y(y)$:

$$p_X(x) = \sum_y p_{XY}(x,y) = \begin{cases} 1-\mathbb{P}(A), & \text{si } x=0 \\ \mathbb{P}(A), & \text{si } x=1. \end{cases}$$

$$p_Y(y) = \sum_x p_{XY}(x,y) = \begin{cases} 1-\mathbb{P}(B), & \text{si } y=0 \\ \mathbb{P}(B), & \text{si } y=1. \end{cases}$$

2. Sean $X_i \sim \chi^2(4)$, $i=1,2$ (chi-cuadrado con 4 grados de libertad) variables aleatorias independientes. Encontrar:

   (a) $E[Y-1]$, donde $Y = X_1 \cdot X_2^{-1}$.

   (b) $Var[X_1^2 - X_2^2]$.

   Solución: (a) Se sabe que $X_i \sim \chi^2(4) \Rightarrow f_X(x) = \frac{1}{4}\, x_i \, \exp\{-\frac{x_i}{2}\}$ y $E[X_i] = 4$, $i=1,2$. Así:

$$\begin{aligned} E[Y-1] =& E[X_1 \cdot X_2^{-1} - 1] = E[X_1] \cdot E[{X_2}^{-1}] - 1 \\ =& 4 \cdot E[{X_2}^{-1}] - 1, \end{aligned}$$

   donde:

$$\begin{aligned} E[{X_2}^{-1}] &= \int_0^\infty x_2^{-1}\, \frac{1}{4}\, x_2 \exp\left\{-\frac{x_2}{2}\right\} dx_2 \\ &= \frac{1}{4}\int_0^\infty \exp\left\{-\frac{x_2}{2}\right\} dx_2 \\ &= -\frac{2}{4}\exp\left\{-\frac{x_2}{2}\right\}\Big|_0^\infty = \frac{1}{2}. \end{aligned}$$

   Finalmente, $E[Y-1] = 2-1 = 1$.

(b) $Var[X_1^2 - X_2^2] = Var[X_1^2] + Var[X_2^2] - 2 \cdot Cov(X_1^2, X_2^2)$. Como las variables aleatorias X_i son independientes, entonces $Cov(X_1^2, X_2^2) = 0$. Por su parte, $Var[{X_i}^2] = E[{X_i}^4] - \left(E[{X_i}^2]\right)^2$.

Es posible obtener los momentos de orden 2 y 4 (en torno a 0) a través de la *fgm*. Recordando que la *fgm* de una chi-cuadrado con 4 grados de libertad es $M_{X_i}(t) = (1-2t)^{-2}$. En efecto:

$$\begin{aligned}
\frac{d}{dt} M_{X_i}(t) &= 4\,(1-2t)^{-3} \\
\frac{d^2}{dt^2} M_{X_i}(t) &= 24\,(1-2t)^{-4} \\
\frac{d^3}{dt^3} M_{X_i}(t) &= 192\,(1-2t)^{-5} \\
\frac{d^4}{dt^4} M_{X_i}(t) &= 1920\,(1-2t)^{-6}.
\end{aligned}$$

Luego, haciendo $t = 0$, se tiene que $E[X_i^2] = 24$ y que $E[X_i^4] = 1920$. Finalmente:

$$Var[{X_1}^2 + {X_2}^2] = 2 \cdot (1920 - 24^2) = 2688.$$

3. La distribución conjunta de las proporciones de dos enzimas en cierto sistema biológico es descrita adecuadamente por la función de densidad conjunta:

$$f(x,y) = 3\theta^{-3}(x+y), \quad 0 < x < \theta, \quad 0 < y < \theta, \quad 0 < x + y < \theta,$$

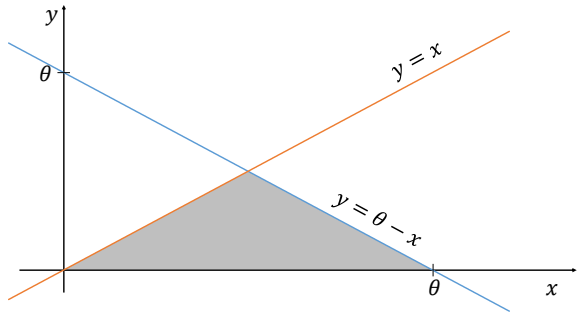
donde $0 < \theta < 1$ es un parámetro desconocido. Se pide:

(a) Calcular $\mathbb{P}(X > Y)$.

(b) Las funciones de densidad marginal de X y de Y.

Solución: (a) Aquí la región de integración está acotada por la recta $y = x$ y las restricciones iniciales de las variables (véase la figura 4.2). Luego, se tiene que:

Figura 4.2: Gráfico para $\mathbb{P}(X > Y)$



Fuente: elaboración propia.

$$\begin{aligned}
\mathbb{P}(X > Y) &= \int_0^{\frac{\theta}{2}} \int_y^{\theta-y} f(x,y)\,dx\,dy = 3\theta^{-3} \int_0^{\frac{\theta}{2}} \left(\frac{x^2}{2} - xy\right)\Big|_y^{\theta-y} dy \\
&= 3\theta^{-3} \int_0^{\frac{\theta}{2}} \left(\frac{(\theta-y)^2}{2} - (\theta-y)y - \frac{y^2}{2} - y^2\right) dy \\
&= 3\theta^{-3} \int_0^{\frac{\theta}{2}} \left(\frac{\theta^2}{2} - 2y^2\right) dy = 3\theta^{-3} \left(\frac{\theta^2}{2}y - \frac{2}{3}y^3\right)\Big|_0^{\frac{\theta}{2}} \\
&= 3\theta^{-3} \left(\frac{\theta^2}{2}\frac{\theta}{2} - \frac{2}{3}\left(\frac{\theta}{2}\right)^3\right) = \frac{1}{2}.
\end{aligned}$$

(b) Las funciones de densidad marginal están dadas por:

$$f_X(x) = \int_0^{\theta-x} f(x,y)\,dy = \frac{3}{2}\,\theta^{-3}(\theta^2 - x^2), \quad 0 < x < \theta.$$

$$f_Y(y) = \int_0^{\theta-y} f(x,y)\,dx = \frac{3}{2}\,\theta^{-3}(\theta^2 - y^2), \quad 0 < y < \theta.$$

4. La variable aleatoria X tiene función de densidad $f(x) = \alpha\, e^{-\alpha x}$, $x > 0$. Por su parte, la variable aleatoria continua Y, independiente de X, tiene función de densidad $f(y) = \beta\, e^{-\beta y}$, $y > 0$. Al respecto, se pide:

   (a) Determinar la función de densidad conjunta de la variable aleatoria $(U, V)^T$, definida por la transformación: $U = X - Y$ y

$$V = X + Y.$$

(b) Determinar la función de densidad de la nueva variable U.

Solución: La densidad conjunta de $(U, V)^T$ se obtiene mediante la transformación de variables aleatorias bidimensionales. En efecto, $u(x, y)$, $v(x, y)$ y el jacobiano de la transformación están dados por:

$$\begin{array}{l} u = x - y \\ v = x + y \end{array} \Rightarrow \begin{array}{l} x = \frac{u+v}{2} \\ y = \frac{v-u}{2} \end{array}$$

$$J = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \Rightarrow |J| = \frac{1}{2}.$$

La función de densidad conjunta de $(U, V)^T$ es:

$$\begin{aligned} f_{U,V}(u, v) &= \frac{1}{2} f_X\left(\frac{u+v}{2}\right) \cdot f_Y\left(\frac{v-u}{2}\right) \\ &= \frac{1}{2}\,\alpha \exp\left\{-\alpha\frac{u+v}{2}\right\} \beta \exp\left\{-\beta\frac{v-u}{2}\right\} \\ &= \frac{\alpha\beta}{2} \exp\left\{-\frac{u}{2}(\alpha-\beta)\right\} \exp\left\{-\frac{v}{2}(\alpha+\beta)\right\}, \\ &- v < u < v, \;\; v > 0. \end{aligned}$$

La función de densidad marginal de U es:

$$\begin{aligned} f_U(u) &= \frac{\alpha\beta}{2} \exp\left\{-\frac{u}{2}(\alpha-\beta)\right\} \left[\int_{-u}^{\infty} \exp\left\{-\frac{v}{2}(\alpha+\beta)\right\} dv + \right. \\ &\left. + \int_{u}^{\infty} \exp\left\{-\frac{v}{2}(\alpha+\beta)\right\} dv\right] \\ &= \frac{\alpha\beta}{2} \exp\left\{-\frac{u}{2}(\alpha-\beta)\right\} \left[\exp\left\{-\frac{v}{2}(\alpha+\beta)\right\} \frac{-2}{\alpha+\beta}\Big|_{-u}^{\infty} + \right. \\ &\left. + \exp\left\{-\frac{v}{2}(\alpha+\beta)\right\} \frac{-2}{\alpha+\beta}\Big|_{u}^{\infty}\right] \\ &= \frac{\alpha\beta}{\alpha+\beta}(e^{u\alpha} + e^{-u\beta}), \;\; -\infty < u < +\infty. \end{aligned}$$

5. Sean X_1, X_2 variables aleatorias independientes idénticamente distribuidas, donde $X_i \sim \text{Exp}(\lambda = 1)$. Obtener la distribución de $Y_1 = X_1 + X_2$.

Solución: Desde el enunciado se tiene que la función de densidad conjunta de $(X_1, X_2)^T$ es $f(x_1, x_2) = e^{-x_1}e^{-x_2}$, $x_1 > 0$, $x_2 > 0$. Luego, se define (convenientemente) la transformación bivariada siguiente y se obtiene el respectivo jacobiano:

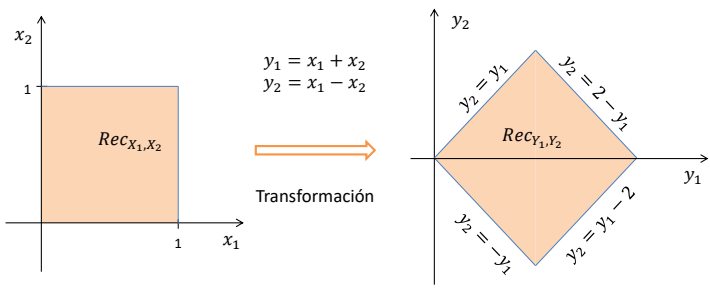
$$\begin{aligned} y_1 &= x_1 + x_2 \\ y_2 &= x_1 - x_2 \end{aligned} \quad \Rightarrow \quad J = -\tfrac{1}{2} \Rightarrow |J| = \tfrac{1}{2}.$$

Finalmente, obteniendo la densidad marginal para Y_1 se tiene que:

$$f_{Y_1}(y_1) = \frac{1}{2}e^{-y_1}, \text{ si } |y_2| \leqslant y_1, \quad y_1 \geqslant 0.$$

Gráficamente, la transformación modifica la región donde se define la función de densidad conjunta para la curva de nivel dada en el plano $z = 0$ (véase figura 4.3).

Figura 4.3: Gráfico de la transformación de $(X_1, X_2)^T$ a $(Y_1, Y_2)^T$



Fuente: elaboración propia.

6. Probar que si $Y_1 \sim \text{Gamma}(\alpha, 1)$ y $Y_2 \sim \text{Gamma}(\beta, 1)$ (independientes), entonces se tiene que:

$$X = \frac{Y_1}{Y_1 + Y_2} \sim \text{Beta}(\alpha, \beta).$$

Solución: Se define (convenientemente) la transformación:

$$\begin{aligned} x_1 &= \tfrac{y_1}{y_1+y_2} \\ x_2 &= y_2 \end{aligned} \quad \Rightarrow \quad \begin{aligned} y_1 &= \tfrac{x_1 x_2}{1-x_1} \\ y_2 &= x_2. \end{aligned}$$

Como $y_1 > 0$ y $y_2 > 0$, se tiene que $x_2 > 0$ y $0 < x_1 < 1$.

Luego, el jacobiano de la transformación está dado por:

$$J = \begin{bmatrix} \frac{x_2}{(1-x_1)^2} & \frac{x_1}{1-x_1} \\ 0 & 1 \end{bmatrix} \Rightarrow |J| = \left| \frac{x_2}{(1-x_1)^2} \right| = \frac{x_2}{(1-x_1)^2}.$$

Con esto, y el hecho de que las variables Y_1 e Y_2 son independientes, la función de densidad conjunta de estas es:

$$f_{Y_1,Y_2}(y_1, y_2) = \frac{1}{\Gamma(\alpha)\Gamma(\beta)} y_1^{\alpha-1} y_2^{\beta-1} e^{-y_1-y_2} \quad y_1 > 0 \;\; y_2 > 0.$$

La función de densidad conjunta del vector $(X_1, X_2)^T$ es:

$$f_{X_1,X_2}(x_1, x_2) = \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \cdot \left(\frac{x_1 x_2}{1-x_1} \right)^{\alpha-1} \cdot x_2^{\beta-1} \cdot \frac{x_2 \cdot e^{-\left(\frac{x_1 x_2}{1-x_1}\right) - x_2}}{(1-x_1)^2},$$

con $x_2 > 0$ y $0 < x_1 < 1$.

La función de densidad de la variable aleatoria X corresponde a la densidad de la variable aleatoria X_1. La densidad marginal de X_1 se puede obtener resolviendo la integral de la función de densidad conjunta, respecto a la variable aleatoria X_2.

Siendo $c = \Gamma(\alpha)\Gamma(\beta)$, se tiene que:

$$\begin{aligned} f_{X_1}(x_1) &= \int_0^\infty f_{X_1X_2}(x_1 x_2)\, dx_2 \\ &= \int_0^\infty \frac{1}{c} \left(\frac{x_1 x_2}{1-x_1} \right)^{\alpha-1} \frac{x_2^{\beta-1} x_2}{(1-x_1)^2} \exp\left\{ -\left(\frac{x_1 x_2}{1-x_1} \right) - x_2 \right\} dx_2 \\ &= \frac{1}{c} \frac{x_1^{\alpha-1}}{(1-x_1)^{\alpha+1}} \int_0^\infty x_2^{\alpha+\beta-1} \exp\left\{ -\frac{x_2}{1-x_1} \right\} dx_2. \end{aligned}$$

Haciendo el cambio de variable $u = \frac{1}{1-x_1} x_2 \rightarrow x_2 = u(1-x_1)$ y $dx_2 = (1-x_1)du$, la integral toma la forma:

$$\begin{aligned} f_{X_1}(x_1) &= \frac{x_1^{\alpha-1}}{c\,(1-x_1)^{\alpha+1}} \cdot \int_0^\infty (1-x_1)^{\alpha+\beta} u^{\alpha+\beta-1} e^{-u} du \\ &= \frac{1}{c} \cdot \frac{x_1^{\alpha-1}}{(1-x_1)^{\alpha+1}} \cdot (1-x_1)^{\alpha+\beta} \cdot \int_0^\infty u^{\alpha+\beta-1} e^{-u} du \\ &= \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \cdot x_1^{\alpha-1} \cdot (1-x_1)^{\alpha-1} \cdot \Gamma(\alpha+\beta), 0 < x_1 < 1. \end{aligned}$$

Como $X_1 = X = \dfrac{Y_1}{Y_1 + Y_2}$, se tiene que la función de densidad es:

$$f_X(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1},\ 0 < x < 1,$$

que corresponde a la función de densidad de una variable aleatoria con distribución beta con parámetros α y β.

7. Sean X_1 y X_2 variables aleatorias y sea $X_1|X_2 = x_2$ la variable aleatoria X_1 dado un valor fijo x_2 de X_2. Mostrar que:

$$E\left[E\left[X_1X_2|X_2\right]\right] = E\left[X_2\,E\left[X_1|X_2\right]\right] = E\left[X_1X_2\right]$$

Solución: Dado que se tienen tres igualdades, se procederá a mostrar la igualdad de a pares:

$$\begin{aligned}
E\left[E\left[X_1X_2|X_2\right]\right] &= \iint E\left[x_1x_2|X_2\right] f(x_1,x_2)dx_1dx_2 \\
&= \int \left(\int x_1x_2 \cdot f(x_1|x_2)dx_1\right) f(x_1,x_2)dx_1dx_2 \\
&= \int x_2 \left(\int x_1 \cdot f(x_1|x_2)dx_1\right) f(x_1,x_2)dx_1dx_2 \\
&= \iint x_2E\left[x_1|X_2\right] f(x_1,x_2)\,dx_1dx_2 \\
&= E\left[x_2 \cdot E\left[x_1|X_2\right]\right].
\end{aligned}$$

$$\begin{aligned}
E\left[X_2E\left[X_1|X_2\right]\right] &= \int x_2E\left[X_1|X_2\right] f(x_2)\,dx_2 \\
&= \int x_2 \left(\int x_1 f(x_1|X_2)\,dx_1\right) f(x_2)\,dx_2 \\
&= \iint x_2x_1 f(x_1|x_2)f(x_2)\,dx_1dx_2 \\
&= \iint x_1x_2 f(x_1,x_2)dx_1dx_2 \\
&= E[X_1X_2].
\end{aligned}$$

8. Sea $(X_1, X_2)^T$ una variable aleatoria continua bivariada, para la cual la función de densidad marginal de X_1 es $f_{X_1}(x_1) = e^{-x_1}$,

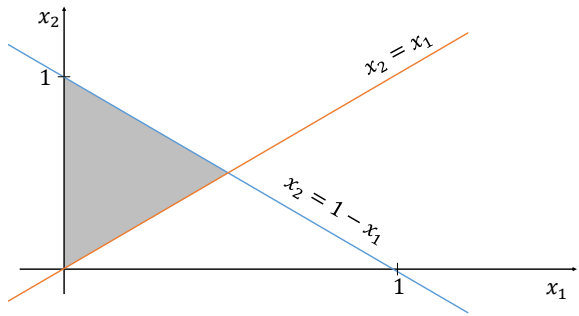
$x_1 > 0$ y la función de densidad condicional de $X_2 \mid X_1 = x_1$ es $f_{X_2|X_1}(x_2) = \exp\{-x_2 + x_1\}$, $0 < x_1 < x_2 < \infty$. Se pide:

(a) Calcular $\mathbb{P}(X_1 + X_2 < 1)$.

(b) Determinar la varianza de la variable aleatoria $Y = X_1 + X_2$.

Solución: (a) Para obtener $\mathbb{P}(X_1 + X_2 < 1)$, el área de integración está acotada por la probabilidad pedida y por las restricciones iniciales de las variables (véase la figura 4.4), luego se plantea y resuelve:

Figura 4.4: Gráfico para $\mathbb{P}(X_1 + X_2 < 1)$



Fuente: elaboración propia.

$$\mathbb{P}(X_1 + X_2 < 1) = \int_0^{0,5} \int_{x_1}^{1-x_1} f(x_1, x_2)\, dx_2\, dx_1.$$

Se sabe que $f(x_1, x_2) = f_{X_2|X_1=x_1}(x_2) \cdot f_{X_1}(x_1)$, luego:

$$\begin{aligned}
&= \int_0^{0,5} \int_{x_1}^{1-x_1} f_{X_2|X_1=x_1}(x_2) \cdot f_{X_1}(x_1)\, dx_2\, dx_1 \\
&= \int_0^{0,5} \left(-e^{-x_2}\right)\Big|_{x_1}^{1-x_1}\, dx_1 \\
&= \int_0^{0,5} (e^{x_1} - e^{-1+x_1})\, dx_1 \\
&= (1 - e^{x_1}) \int_0^{0,5} e^{-x_1}\, dx_1 \\
&= (1 - e^{x_1})(-e^{-x_1})\Big|_0^{0,5} \\
&= (1 - e^{x_1})(1 - e^{-0,5}) = 0{,}249.
\end{aligned}$$

(b) La varianza de la variable aleatoria $Y = X_1 + X_2$ se calcula mediante la siguiente expresión:

$$Var[Y] = Var[X_1 + X_2] = Var[X_1] + Var[X_2] + 2\,Cov(X_1, X_2).$$

Cada uno de los componentes de la expresión anterior se pueden obtener por:

$$X_1 \sim \text{Exp}(1) \Rightarrow E[X_1^2] = Var[X_1] + (E[X_1])^2 = 2.$$

Por otra parte, $E[X_2] = E[E[X_2 \mid X_1]]$ y $E[X_2^2] = E[E[X_2^2 \mid X_1]]$, entonces

$$Var[X_2] = E[E[X_2^2 \mid X_1]] - (E[E[X_2 \mid X_1]])^2,$$

donde los valores se calculan mediante:

$$\begin{aligned}
E[X_2 \mid X_1] &= \int_{x_1}^{\infty} x_2 f_{X_2|X_1}(x_2) dx_2 = \int_{x_1}^{\infty} x_2 e^{-x_2}\, dx_2 \\
&= -x_2 e^{-x_2} - e^{-x_2}\big|_{x_1}^{\infty} = x_1 + 1. \\
E[E[X_2 \mid X_1]] &= E[X_1 + 1] = E[X_1] + 1 = 2. \\
E[X_2^2 \mid X_1] &= \int_{x_1}^{\infty} x_2^2 f_{X_2|X_1}(x_2) dx_2 \\
&= e^{x_1}\left(-x_2^2 e^{-x_2} - 2x_2 e^{-x_2} - 2e^{-x_2}\right)\big|_{x_1}^{\infty} \\
&= e^{x_1}(x_1^2 e^{-x_1} + 2x_1 e^{-x_1} + 2e^{-x_1}) \\
&= x_1^2 + 2x_1 + 2. \\
E[E[X_2^2 \mid X_1]] &= E[X_1^2 + 2X_1 + 2] = E[X_1^2] + 2\,E[X_1] + 2 = 6.
\end{aligned}$$

Por lo tanto, $Var[X_2] = 6 - 2^2 = 2$. Y para la última expresión, el resultado se obtiene por:

$Cov(X_1, X_2) = E[X_1 X_2] - \cancel{E[X_1]}^{1} \cdot \cancel{E[X_2]}^{2}$, donde:

$$\begin{aligned}
E[X_1 X_2] &= \int_0^{\infty} \int_{x_1}^{\infty} x_1 x_2 e^{-x_2}\, dx_2\, dx_1 \\
&= \int_0^{\infty} x_1 \left(-x_2 e^{-x_2} - e^{-x_2}\right)\big|_{x_1}^{\infty}\, dx_1 \\
&= \int_0^{\infty} (x_1^2 e^{-x_1} - x_1 e^{-x_1})\, dx_1 \\
&= \int_0^{\infty} x_1^2 e^{-x_1}\, dx_1 - \int_0^{\infty} x_1 e^{-x_1}\, dx_1 \\
&= E[X_1^2] - (E[X_1])^2 = Var[X_1] = 1.
\end{aligned}$$

Finalmente, $Var[Y] = 1 + 2 + 2 \cdot (1 - 1 \cdot 2) = 1$.

9. Sea $(X_1, X_2)^T$ una variable aleatoria bivariada. Determinar la correlación entre X_1 y X_2 sabiendo que la función de densidad condicional de $X_2|X_1 = x_1$ y la función de densidad marginal de X_1 están dadas, respectivamente, por:

$$\begin{aligned}
fX_2|X_1(x_2) &= (\alpha \cdot x_1)^{-1} \exp\left\{-(\alpha \cdot x_1)^{-1} x_2\right\}, x_2 > 0, \alpha > 0 \\
f_{X_1}(x_1) &= \beta \exp\{-\beta x_1\}, x_1 > 0, \beta > 0.
\end{aligned}$$

Solución: A partir de las funciones de densidad, se obtiene lo siguiente:

$$\begin{aligned} X_1 \sim \text{Exp}(\beta^{-1}) \Rightarrow & E[X_1] = \beta^{-1} \text{ y } Var[X_1] = \beta^{-2} \\ X_2|X_1 \sim \text{Exp}(\alpha\, x_1) \Rightarrow & E[X_2|X_1] = \alpha\, x_1 \text{ y } Var[X_2|X_1] = (\alpha\, x_1)^2. \end{aligned}$$

Para calcular la correlación entre X_1 y X_2, se requiere obtener los siguientes resultados:

$$\begin{aligned} E[X_1X_2] =& E[X_1E[X_2|X_1]] = E[X_1 \cdot \alpha \cdot X_1] = \alpha E[X_1^2] \\ =& \alpha \left(Var[X_1] + (E[X_1])^2\right) = 3\alpha\beta^{-2} \\ E[X_2] =& E[E[X_2|X_1]] = E[\alpha X_1] = \alpha\beta^{-1} \\ Var[X_2] =& E[V[X_2|X_1]] + V[E[X_2|X_1]] = E[(\alpha X_1)^2] + V[\alpha X_1] \\ =& 3\alpha^2\beta^{-2}. \end{aligned}$$

Con estos resultados, la correlación entre X_1 y X_2 es:

$$Corr(X_1, X_2) = \frac{2\alpha\beta^{-2} - \beta^{-1}\alpha\beta^{-1}}{\sqrt{\beta^{-2}} \cdot \sqrt{3\alpha^2\beta^{-2}}} = \frac{1}{\sqrt{3}}.$$

10. Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde $X_i \sim N(0,1)$. Determinar la función de densidad de las siguientes variables aleatorias:

    (a) $Y_1 = \exp\left\{9\bar{X} - 5\right\}$.

    (b) $Y_2 = \sum_{i=1}^{4} X_i^2 \cdot \left(\sum_{i=5}^{n} X_i^2\right)^{-1}$.

    Solución: (a) Dado que $X_i \sim N(0,1)$, entonces $\bar{X} \sim N(0, n^{-1})$. De esta manera, es conocida la función de densidad de la variable aleatoria $\bar{X}$, la cual está dada por:

$$f_{\bar{X}}(x) = \frac{1}{2\pi\, n^{-1}} \exp\left\{-\frac{1}{2n^{-1}}x^2\right\} \qquad -\infty < x < \infty.$$

Para determinar la función de densidad de Y_1, se utilizará la función de probabilidad acumulada, en efecto:

$$\begin{aligned} F_{Y_1}(y) =& \mathbb{P}\left(Y_1 \leqslant y\right) = \mathbb{P}\left(\exp\left\{9\bar{X} - 5\right\} \leqslant y\right) \\ =& \mathbb{P}\left(9\bar{X} - 5 \leqslant \log y\right) = \mathbb{P}\left(\bar{X} \leqslant \frac{1}{9}\log y - \frac{5}{9}\right) \\ =& F_{\bar{X}}\left(\bar{X} \leqslant \frac{1}{9}\log y - \frac{5}{9}\right). \end{aligned}$$

Aplicando la derivada parcial respecto a y, se tiene que:

$$\begin{aligned} f_{Y_1}(y) =& f_{\bar{X}}\left(\bar{X} \leqslant \frac{1}{9}\log y - \frac{5}{9}\right) \cdot \frac{1}{9y}, \quad y > 0 \\ =& \frac{1}{\sqrt{2\pi(9y)^2}} \exp\left\{-\frac{1}{162}(\log y + 5)^2\right\}, \quad y > 0. \end{aligned}$$

(b) Para obtener la función de densidad de Y_2, se puede reescribir la expresión, tal que $Y_2 = \frac{T_1}{T_2}$, donde:

$$T_1 = \sum_{i=1}^{4} X_i^2 \quad \text{y} \quad T_2 = \sum_{i=5}^{n} X_i^2, \text{ donde } T_1 \sim \chi^2(4) \text{ y } T_2 \sim \chi^2(n-4).$$

Estos resultados llevan a una distribución F de Fisher:

$$\frac{n-4}{4} \cdot Y_2 = \frac{T_1 \cdot 4^{-1}}{T_2 \cdot (n-4)^{-1}} \sim F(4, n-4).$$

Finalmente, la función de densidad de Y_2, está dada por:

$$f_{Y_2}(y) = \frac{4}{n-4} f_{F(4,n-4)}(y), \; y > 0.$$

11. Sean $X_1, X_2, \cdots, X_n$ variables aleatorias independientes, donde $X_i \sim N(\mu_i, \sigma_i^2)$. Determinar la distribución de:

$$Y = \sum_{i=1}^{n} X_i$$

Solución: Por la propiedad 4.9(a) se tiene que:

$$\begin{aligned} M_Y(t) &= \prod_{i=1}^{n} M_{X_i}(t) = \prod_{i=1}^{n} \exp\left\{t\mu_i + \frac{1}{2}t^2\sigma_i^2\right\} \\ &= \exp\left\{t\sum_{i=1}^{n}\mu_i + \frac{1}{2}t^2\sum_{i=1}^{n}\sigma_i^2\right\} \end{aligned}$$

De la última expresión se tiene la fgm de una distribución normal, identificando los parámetros se tiene:

$$Y \sim N\left(\sum_{i=1}^{n}\mu_i, \sum_{i=1}^{n}\sigma_i^2\right).$$

4.5. Ejercicios propuestos

1. Sea $f(x,y) = c(|x| + |y|)$, $|x| + |y| \leqslant 1$.
   (a) Verificar que $c = \frac{3}{4}$.
   (b) Probar que $\mathbb{P}(X+Y \leqslant 1, X \geqslant 0, Y \geqslant 0) = \frac{1}{4}$. Graficar la región de integración.
   (c) Sean $f_X(x)$ y $f_Y(y)$ las funciones de densidad marginal de X e Y respectivamente. Probar que:
   $$f_X(x) = \frac{3}{4}(1-x^2),\ |x| \leqslant 1 \quad f_Y(y) = \frac{3}{4}(1-y^2),\ |y| \leqslant 1.$$
   (d) Mostrar que X, Y no son independientes.

2. Sea $f(x,y) = k(x^2+y^2)$, $0 < x < y < 1$ función de densidad conjunta del vector aleatorio $(X,Y)^T$. Se pide:
   (a) Encontrar el valor de k para que $f(x,y)$ sea función de densidad.
   (b) Determinar la función de densidad marginal de X e Y.
   (c) Determinar la función de densidad condicional de $Y \mid X = x$.
   (d) Calcular $Var[X+Y]$.
   (e) Calcular el coeficiente de correlación.
   (f) Determinar una expresión para $\mathbb{P}(0 < X < 0{,}2 \mid 0{,}1 < Y < 0{,}5)$, no se requiere resolver las integrales.

3. Considerar la función $f(x_1,x_2) = \exp\{-x_1\}$, $x_1 > x_2 > 0$, asociada a un vector aleatorio bivariado $(X_1,X_2)^T$.
   (a) Verificar que la función definida es efectivamente función de densidad del vector aleatorio bivariado $(X_1,X_2)^T$.
   (b) Calcular $\mathbb{P}(X_1 + X_2 < 1)$.

4. Sean $X_1, X_2 \sim$ Uniforme$(0,1)$ variables aleatorias iid, sea la transformación bivariada $Y_1 = X_1 + X_2$ y $Y_2 = X_1 - X_2$.

   (a) Demostrar que $f(y_1, y_2) = 0{,}5$; $-y_2 < y_1 < 2 - y_2$, $-2 + y_1 < y_2 < y_1$.

   (b) Graficar la transformación.

5. Sean $X_1, X_2 \sim \text{Exp}(\lambda)$ variables aleatorias iid. Probar que $Y = X_1 + X_2 \sim \text{Gamma}(2, \lambda)$.

6. Sean $X_1 \sim \text{Gamma}(\alpha_1, \lambda)$ y $X_2 \sim \text{Gamma}(\alpha_2, \lambda)$ variables aleatorias independientes. Probar que:

   (a) $Y_1 = \frac{X_1}{X_1 + X_2} \sim \text{Beta}(\alpha_1, \alpha_2)$.

   (b) $Y_2 = X_1 + X_2 \sim \text{Gamma}(\alpha_1 + \alpha_2, \lambda)$.

   (c) Y_1, Y_2 son variables aleatorias independientes.

7. La covarianza entre dos variables aleatorias X_1 y X_2 es:

$$Cov(X_1, X_2) = E\left[(X_1 - E[X_1])(X_2 - E[X_2])\right].$$

   Probar los siguientes resultados:

   (a) $Cov(X_1, X_2) = E[X_1 X_2] - E[X_1]E[X_2]$.

   (b) $Cov(a_1 X_1 + b_1, a_2 X_2 + b_2) = a_1 a_2\, Cov(X_1, X_2)$.

   (c) $Cov(X_1, X_2 + X_3) = Cov(X_1, X_2) + Cov(X_1, X_3)$.

   (d) $Var[U] = \sum_{i=1}^{n} a_i^2 Var[Y_i] + 2 \sum \sum_{i<j} a_{ij}\, Cov(Y_i, Y_j)$, donde la doble suma está dada para los pares (i, j), con $i < j$, siendo $U = \sum_{i=1}^{n} a_i Y_i$, tal que $E[Y_i] = \mu_i$ y a_i constantes, ambos para $i = 1, 2, \cdots, n$.

8. Determinar el coeficiente de correlación entre X_1 y X_2, sabiendo que la función de densidad de la distribución condicional de $X_1 | X_2 = x_2$, asociada a la variable aleatoria bivariada $(X_1, X_2)^T$, es:

$$f_{X_1|X_2=x_2}(x_1) = \frac{1}{\sqrt{2\pi x_2^2}} \exp\left\{-\frac{1}{2x_2^2}(x_1 - x_2)^2\right\},\ -\infty < x_1 < +\infty,$$

   y la función de densidad marginal de X_2: $f_{X_2}(x_2) = 1$, $0 < x_2 < 1$.

9. Sean $X_1 \sim N(\mu_1, \sigma_1^2)$ y $X_2 \sim N(\mu_2, \sigma_2^2)$, donde X_1 y X_2 son variables aleatorias independientes. Calcular:

   (a) $E[X_1^2 + X_2]$, nótese que $E[G_1(X_1) + G_2(X_2)] = E[G_1(X_1)] + E[G_2(X_2)]$.

   (b) $E\left[(X_1 - \mu_1) \cdot \sigma_1^{-1} \exp\{(X_2 - \mu_2) \cdot \sigma_2^{-1}\}\right]$.

   (c) La fgm de $X_1 + X_2$. Luego, a partir de este resultado señalar la distribución de $X_1 + X_2$.

10. Sea $f_{X,Y}(x, y) = 2$, $x, y > 0$, $x + y < 1$ la función de densidad conjunta del vector aleatorio $(X, Y)^T$:

    (a) Graficar el recorrido de $(X, Y)^T$.

    (b) Probar que $f_X(x) = 2(1 - x)$, si $0 < x < 1$.

    (c) Probar que $Y|X = x \sim \text{Uniforme}(0, 1 - x)$ para cada $x \in (0, 1)$. Análogamente, $X|Y = y \sim \text{Uniforme}(0, 1 - y)$, $y \in (0, 1)$.

11. Probar que si $X \sim \text{Poisson}(\lambda)$ y $Y|X = x \sim \text{Binomial}(x, p)$, entonces $Y \sim \text{Poisson}(\lambda p)$.

12. Sean X_1 y X_2 variables aleatorias y sea $X_1 \mid X_2 = x_2$ la variable aleatoria condicional de X_1 para un valor fijo x_2 de la variable aleatoria X_2. Mostrar que:

$$E[E[X_1 X_2 \mid X_2]] = E[X_2\, E[X_1 \mid X_2]] = E[X_1 X_2].$$

13. Mostrar que $E\left[(\alpha X_1 + \beta X_2) \mid Y\right] = \alpha\, E[X_1 \mid Y] + \beta\, E[X_2 \mid Y]$.

14. Dadas las distribuciones condicionales siguientes, probar que:

    (a) $Y|X = x \sim \text{Binomial}(x, p) \Rightarrow Var(Y|X = x) = xp(1 - p)$. Además, probar que $Var(Y) = p(1 - p)E[X] + p^2 Var(x)$.

    (b) $Y|X = x \sim \text{Uniforme}(0, 1 - x) \Rightarrow Var(Y|X = x) = \dfrac{(1 - x)^2}{12}$.

15. Sea la distribución condicional de $Y \mid X = x \sim N(\rho x, 1 - \rho^2)$ y la distribución de la variable aleatoria $X \sim N(0, 1)$. Determinar el coeficiente de correlación entre X e Y.

16. Sean Y_1 e Y_2 variables aleatorias independientes, con $Y_i \sim \chi^2(2)$. Determinar la función de densidad de la variable aleatoria V:

$$V = \frac{Y_1}{Y_1 + Y_2}.$$

17. Sean $Y_1 \sim \chi^2(n_1)$ e $Y_2 \sim \chi^2(n_2)$ y, además, Y_1 e Y_2 son variables aleatorias independientes. Determinar:

    (a) $E\left[\frac{Y_1 \cdot n_1^{-1}}{Y_2 \cdot n_2^{-1}}\right]$.

    (b) $Var\left(\frac{Y_1 \cdot n_1^{-1}}{Y_2 \cdot n_2^{-1}}\right)$.

18. Sea $(X_1, X_2)^T$ un vector aleatorio bivariado continuo tal que las variables aleatorias $X_i \sim \chi^2(n_i)$, $i = 1, 2$ son independientes. Determinar una expresión para la función de probabilidad de la nueva variable aleatoria bivariada $(Y_1, Y_2)^T$, dada por la transformación $(Y_1, Y_2)^T = (X_1 \cdot X_2^{-1}, X_2)^T$ y, además, la función de densidad marginal de la variable aleatoria Y_1.

19. Sea $f_{XY}(x, y) = 2(x + y - 2xy)$, $0 \leqslant x, y \leqslant 1$, la función de densidad conjunta del vector $(X, Y)^T$. Probar que:

    (a) $X \sim$ Uniforme$(0, 1)$. Análogamente, $Y \sim$ Uniforme$(0, 1)$.

    (b) $Cov(X, Y) = -\frac{1}{36}$. Donde $E[X] = E[Y] = \frac{1}{2}$ y $E[XY] = \frac{2}{9}$.

20. Sean las variables aleatorias independientes idénticamente distribuidas X_1, X_2, donde $X_i \sim N(\mu, \sigma^2)$, $i = 1, 2$. Se define $Y_1 = X_1 + X_2$ e $Y_2 = X_1 - X_2$, determinar si Y_1 e Y_2 son variables aleatorias independientes.

21. Sea el experimento aleatorio de sacar una ficha desde una urna (están numeradas del 1 al N) y sea X: *número de la ficha*. Luego, con la ficha seleccionada se lanza una moneda honesta la cantidad de veces que indica la ficha y se define la variable aleatoria Y: *número de caras obtenidas*. Sea $X = x$ (la ficha seleccionada tiene número x), luego se lanza la moneda x veces, entonces se pide mostrar que:

    (a) $Y|X = x \sim$ Binomial$(x, \frac{1}{2})$ y que $\mathbb{P}(X = x) = \frac{1}{N}$.

(b) $\mathbb{P}(X = x, Y = y) = \binom{x}{y}\left(\frac{1}{2}\right)^y\left(1-\frac{1}{2}\right)^{x-y}\frac{1}{N}$, $y = 0, \cdots, x$ $x = 1, \cdots, N$.

(c) $Cov(X, Y) = \frac{(N+1)(N-1)}{24}$.

22. Sea $X_1, X_2, \cdots, X_n$, donde $X_i \sim$ Bernoulli(p). Probar que:

$$\sum_{i=1}^{n} X_i \sim \text{Binomial}(n, p).$$

23. Sea $X_1, X_2, \cdots, X_n$, donde $X_i \sim N(\mu, \sigma^2)$. Asumir que las variables aleatorias $\bar{X}$ y S^2 son independientes y están dadas por:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{y} \quad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2.$$

Se requiere determinar:

(a) La función de densidad de la variable aleatoria S^2.

(b) La función de densidad conjunta de $\bar{X}$ y S^2.

(c) El valor esperado de la variable aleatoria $Q = \bar{X}^2 \exp\left\{(n-1)S^2\right\}$.

24. Sean $X_1, X_2, \cdots, X_k$ variables aleatorias independientes:

(a) Si $X_i \sim \chi^2(1)$, demostrar que:

$$\sum_{i=1}^{k} X_i \sim \chi^2(k).$$

(b) Si $X_i \sim \chi^2(n_i)$, $i = 1, \cdots, k$, demostrar que:

$$\sum_{i=1}^{k} X_i \sim \chi^2\left(\sum_{i=1}^{k} n_i\right).$$

25. $X_1, X_2, \cdots, X_n$ y $Y_1, Y_2, \cdots, Y_n$ son muestras aleatorias independientes, donde $X_i \sim N(\mu_1, \sigma_1^2)$ y $Y_i \sim N(\mu_2, \sigma_2^2)$.

(a) Demostrar que:

$$\frac{1}{n}\sum_{i=1}^{n} X_i \sim N(\mu_1, \frac{\sigma_1^2}{n}).$$

(b) Encontrar el valor esperado y la varianza de la nueva variable aleatoria T:

$$T = \frac{1}{n}\sum_{i=1}^{n} Y_i \exp\left\{\frac{1}{n}\sum_{i=1}^{n} X_i\right\}.$$

26. Las variables aleatorias $X_1, X_2, \cdots, X_n$ constituyen una muestra aleatoria. Determinar la función de densidad conjunta de las n variables, si:

(a) $X_i \sim \text{Beta}(\theta, 2)$.

(b) $X_i \sim \text{Gamma}(\alpha, \beta)$.

27. Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde $X_i \sim \text{Poisson}(\lambda)$. Determinar el valor esperado de las nuevas variables T_1 y T_2, dadas por las expresiones:

$$T_1 = \bar{X} \quad T_2 = \frac{2}{n(n+1)}\sum_{i=1}^{n} i\, X_i.$$

28. Las variables aleatorias $X_1, X_2, \cdots, X_n$ son independientes, $X_i \sim \text{Binomial}(1, p)$. Determinar el valor esperado de las funciones T_1 y T_2 siguientes:

$$T_1 = \bar{X} \quad T_2 = \frac{1}{3}(X_1 + X_2 + X_3).$$

29. Sean las variables aleatorias $X_1, X_2, \cdots, X_n$ independientes, donde $X_i \sim \text{Uniforme}(0, \theta)$. Determinar el valor esperado de las variables:

(a) $U_n = 2\bar{X}$.

(b) $T = \text{máx}\{X_1, x_2, \cdots, X_n\}$.

30. Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde la función de densidad de X_i es $f(x_i) = 2\theta^{-2} x_i \exp\left\{-\theta^{-2} x_i^2\right\}$, $x_i > 0$, $\theta > 0$. Determinar el valor esperado de la función:

$$T = \frac{2}{\sqrt{\pi} n}\sum_{i=1}^{n} X_i.$$

31. Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, con $X_i \sim \text{Binomial}(2, \theta)$. Determinar el valor esperado de T, con:

$$T = \frac{1}{2n(2n-1)} U(U-1). \text{ Con } U = \sum_{i=1}^{n} X_i.$$

32. Sean $X_1, X_2, \cdots, X_{n_1}$ una muestra aleatoria de tamaño n_1, donde $X_i \sim N(\mu_1, \sigma_1^2)$, y sean, además, $Y_1, Y_2, \cdots, Y_{n_2}$ una muestra aleatoria de tamaño n_2 (independiente de la anterior), donde $Y_i \sim N(\mu_2, \sigma_2^2)$. A través de la función generadora de momentos, obtener la distribución de la variable aleatoria T:

$$T = \frac{1}{n_1} \sum_{i=1}^{n} X_i - \frac{1}{n_2} \sum_{i=1}^{n} Y_i.$$

33. Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde $X_i \sim N(\mu, \sigma^2)$. Determinar el valor esperado de las siguientes variables aleatorias:

    (a) $Y_1 = \dfrac{1}{n(n+1)} \left((n-2) \sum_{i=1}^{n} X_i^2 + (n\bar{X})^2 \right)$.

    (b) $Y_2 = \dfrac{1}{n} \sum_{i=1}^{n} X_i^2$.

    (c) $Y_3 = \dfrac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2$.

34. Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde $X_i \sim N(0, 1)$. Se definen las siguientes variables aleatorias:

$$Y_1 = \exp\left\{9\bar{X} - 5\right\} \text{ y } Y_2 = \frac{\sum_{i=1}^{4} X_i^2}{\sum_{i=5}^{n} X_i^2}.$$

    (a) Determinar, por separado, la función de densidad de Y_1 e Y_2.

    (b) Determinar, por separado, el valor esperado de Y_1 e Y_2.

35. Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde la función de densidad de X_i está dada por $f(x_i) = \theta \exp\{-\theta x_i\}$, $x_i > 0$. Determinar:

    (a) La función de densidad conjunta de $X_1, X_2, \cdots, X_n$.

(b) La función de densidad de las nuevas variables aleatorias: $M = \min\{X_1, X_2, \cdots, X_n\}$ y $T = \max\{X_1, X_2, \cdots, X_n\}$.

(c) $\mathbb{P}(T \leqslant 5)$.

(d) La distribución de la variable aleatoria $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

36. Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde la función de densidad de la variable aleatoria X_i es $f(x_i) = \theta x_i^{\theta-1}$, $0 < x < 1$, $\theta > 0$:

(a) Mostrar que $Y_1 = \log\left(X_1^{-1}\right) \sim \text{Exp}(\theta^{-1})$.

(b) Mostrar que $T = \frac{1}{n}\sum_{i=1}^{n} \log\left(X_i^{-1}\right) \sim \text{Gamma}(n, (n\theta)^{-1})$.

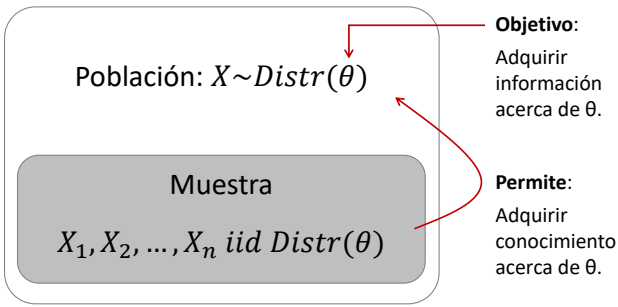
Capítulo 5

Estimación de parámetros

5.1. Inferencia estadística

La inferencia estadística está constituida por un conjunto de fundamentos, métodos y técnicas que permiten describir y, en general, conocer las características de la población a través de la información de una muestra (aleatoria). Considerando que la población está definida por parámetros, en la muestra los elementos respectivos se llaman estadísticos, que son los utilizados para obtener aproximaciones (bajo un nivel de confianza aceptable) de los parámetros y/o para poner a prueba creencias acerca de los mismos parámetros (esta relación se presenta en el esquema de la figura 5.1).

Figura 5.1: Esquema del proceso de inferencia estadística



Fuente: elaboración propia.

Ejemplo **5.1** Sea la variable aleatoria X: *tiempo de falla de un componente electrónico*, la cual (en la población) es $X \sim f(x|\theta) = \frac{1}{\theta}\exp\left\{-\frac{1}{\theta}x\right\}$, $x > 0$. Se desea:

(a) Tener conocimiento acerca del tiempo medio de falla.

(b) Contar con un intervalo que cubra el tiempo medio de falla con un 95 % de confianza.

(c) Verificar que el tiempo promedio de falla es mayor o igual a 10 000 horas.

Solución: El tiempo medio de falla en la población corresponde al valor esperado, el cual es $E[X] = \theta$. Si se cuenta con una muestra aleatoria, cuyas componentes son $X_1, X_2, \cdots, X_n$ variables aleatorias independientes e idénticamente distribuidas que siguen la misma distribución que la variable en la población, entonces:

(a) Una estimación de θ (denotada por $\widehat{\theta}$) está dada por el tiempo de falla medio observado en la muestra:

$$\widehat{\theta} = \frac{\sum_{i=1}^{n} X_i}{n} = \bar{X}_n.$$

(b) Determinar un intervalo que cubra el tiempo promedio de falla con un 95 % de confianza. Se puede obtener resolviendo la ecuación de la probabilidad de que θ se encuentre entre dos valores, a_1 y a_2, sea de 0,95, en la cual las incógnitas son a_1 y a_2 y se expresa como:

$$\mathbb{P}\left(a_1 < \theta < a_2\right) = 0{,}95.$$

(c) Aquí se puede enfrentar este problema obteniendo la probabilidad de que $\theta \geqslant 10\,000$, para lo cual se debe asumir algún supuesto adicional para especificar la incógnita. □

Los tres problemas que aparecen en el ejemplo 5.1 son los que resuelve la inferencia estadística a través de fundamentos teóricos y métodos o técnicas de análisis de información. Específicamente, (a) y (b) son problemas de estimación de parámetros, mientras que (c) es un problema de pruebas de hipótesis.

En este capítulo se presentan fundamentos y métodos (técnicas) para obtener estimaciones puntuales e intervalos en que se encuentre un parámetro. Además, se estudia el comportamiento del parámetro estimado bajo supuestos de muestreo aleatorio, referido a consistencia y eficiencia en el contexto estadístico.

5.2. Conceptos de inferencia estadística

Población

La población se define como el conjunto del universo en estudio (individuos u objetos).

Un elemento observable llamado variable aleatoria, X, es lo relevante para el estudio. La variable aleatoria posee una densidad (cuantía), que depende de un vector de parámetros ($\boldsymbol{\theta}$) desconocido; es de naturaleza no-observable y es una característica de la población. Esto se denota con:

$$X \sim f(x|\boldsymbol{\theta}),\ \boldsymbol{\theta} \in \Theta,$$

donde Θ es el espacio paramétrico, el cual es un subconjunto de $\mathbb{R}^p$.

En la literatura, el vector de parámetros $\boldsymbol{\theta}$ es conocido como el *estado de la naturaleza*.

Ejemplo **5.2** Algunos ejemplos de población asociada a un experimento aleatorio son:

(1) Sea la variable aleatoria X definida como:

$$X = \begin{cases} 1, & \text{presencia del atributo} \\ 0, & \text{ausencia del atributo,} \end{cases}$$

donde $X \sim$ Bernoulli(p); $p \in (0,1)$ y $f(x|p) = p^x(1-p)^{1-x}$; $E[X] = p$, $Var[X] = p \cdot (1-p)$. p: proporción de individuos u objetos que poseen el atributo.

(2) La variable aleatoria X es *número de clientes que llegan a la fila de un banco en un cierto periodo*. Entonces $X \sim$ Poisson(λ), $f(x|\lambda) = \frac{\lambda^x e^{-\lambda}}{x!}$, $x = 0, 1, 2, \cdots, \infty$; $\lambda > 0$; $\Theta = (0, +\infty)$; $E[X] = \lambda$, $Var[X] = \lambda$ y λ: tasa media de llegada.

(3) La variable aleatoria X es *tiempo de falla de un cierto componente de un sistema*. Entonces $X \sim \text{Exp}(\lambda, \theta)$ y $f(x|\lambda, \theta) = \frac{1}{\lambda} \exp\left\{-\frac{1}{\lambda}(x-\theta)\right\}$; $x \geqslant \theta$, $\lambda > 0$, $\theta > 0$ (λ: parámetro de escala, θ: parámetro de posición).

(4) Modelo normal: $X \sim N(\mu, \sigma^2)$, donde X: *puntaje obtenido en un test* o *tensión a la que es sometida una viga*. □

Muestra aleatoria

Una muestra aleatoria es obtenida por medio de la realización de un experimento aleatorio desde una población $X \sim f(x|\theta)$, la cual es un vector aleatorio $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)^T$, donde las variables aleatorias $X_1, X_2, \cdots, X_n$ son independientes y poseen densidad $f(x|\theta)$. Entonces $\boldsymbol{X} \in \Omega$ (Ω: espacio muestral).

Luego $(\Omega, \mathcal{F}, \{f(\cdot|\theta) : \theta \in \Theta\})$ es un modelo estadístico, es decir, una familia de espacios de probabilidad indexada por el parámetro.

El concepto de muestra aleatoria fue presentado con anterioridad en la definición 4.14 (del capítulo 4), donde además se trabaja de forma operativa como concepto resumido de variables aleatorias *iid*. Aquí se vuelve a presentar, dada la importancia que tiene en la estimación de parámetros.

Función de verosimilitud

Definición **5.1** La función de verosimilitud de θ de una muestra aleatoria $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)^T$, proveniente de la población X con función de densidad $f(x|\theta)$, viene dada por:

$$L(\boldsymbol{\theta}) = L(\boldsymbol{\theta}|x_i) = \prod_{i=1}^{n} f(x_i, \boldsymbol{\theta}), \quad \theta \in \Theta.$$

El principio de verosimilitud dice que la información que entrega la muestra aleatoria $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)^T$ acerca del parámetro θ está representada en la función de verosimilitud.

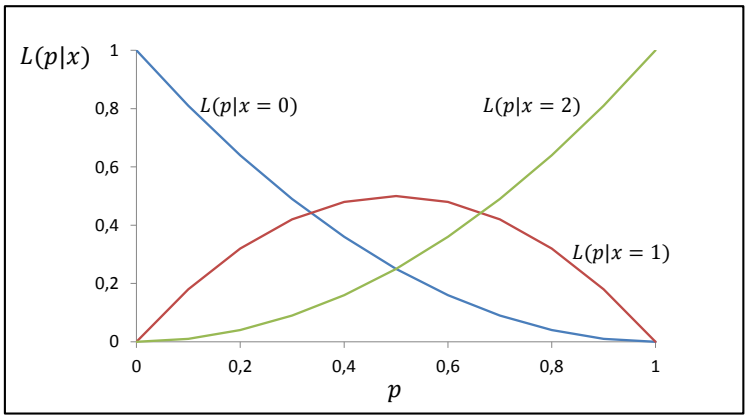
Ejemplo **5.3** Algunos ejemplos de función de verosimilitud son:

(1) Sea $X \sim \text{Binomial}(n = 2, p)$. Luego la función de verosimilitud y

el gráfico para p (figura 5.2) están dados por:

$$L(p) = \binom{2}{x} p^x (1-p)^{2-x}, \quad x = 0, 1, 2.$$

Figura 5.2: Gráfico de $L(p|x)$ para $X \sim \text{Binomial}(n = 2, p)$



Fuente: elaboración propia.

(2) Sea la muestra aleatoria $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)^T$ $X_i \sim \text{Bernoulli}(p)$, la función de densidad es $p(x|p) = p^x(1-p)^{1-x}$. Así, la función de verosimilitud está dada por:

$$L(p) = p^{\sum x_i}(1-p)^{n-\sum x_i}.$$

Nótese que $\sum X_i$ es una función de la muestra, que representa la cantidad de éxitos en la muestra. □

Estadístico (estadígrafo)

Definición **5.2** Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, una estadística T puede ser definida como:

(1) T es una función de la muestra aleatoria, $T = T(\boldsymbol{X})$, luego también es una variable aleatoria que asume valores en algún espacio Y.

(2) T es una función definida sobre el espacio muestral y asumiendo

valores en Y.

$$T : \mathbb{R}^n \to \mathbb{R}$$
$$\boldsymbol{x} \mapsto y = T(\boldsymbol{X})$$

Ejemplo **5.4** Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde $X \sim$ Bernoulli(p). El siguiente es un ejemplo de estadística:

$T = T(\boldsymbol{X}) = \sum X_i \sim$ Binomial(n, p) y asume valores en $\{0, 1, \cdots, n\}$.

$$\begin{aligned} T : \{(0,1), \cdots, (0,1)\} &\to \{0, 1, 2, \cdots, n\} \\ \boldsymbol{x} = (X_1, X_2, \cdots, X_n) &\mapsto y = T(X_1, X_2, \cdots, X_n) = \sum X_i \end{aligned}$$

□

5.3. Estimación puntual

Para realizar la estimación se contará con la información de una variable aleatoria a través de una muestra aleatoria tomada desde una población de interés. Es decir, sea la muestra aleatoria $X_1, X_2, \cdots, X_n$, tomada desde una población con distribución conocida, los parámetros asociados a dicha distribución, y que se pueden anotar como un vector de parámetros $\boldsymbol{\theta}$, son los que se estimarán. El proceso de asignar un valor (obtenido como función de la muestra aleatoria únicamente) a un parámetro es llamado *estimación puntual* y el valor obtenido, *estimador puntual*, entendiendo como parámetro un vector cuyas componentes son los parámetros de la distribución subyacente.

A continuación, se estudiarán algunos métodos para conseguir estimadores de dichos parámetros (estimaciones puntuales).

Método de máxima verosimilitud

Este método se basa en la maximización de la función de verosimilitud respecto de los parámetros, procedimiento que determina funciones de la muestra como estimadores de los parámetros respectivos, que no dependen de los parámetros. Estos estimadores son llamados estimadores de máxima verosimilitud.

Definición **5.3** Se define un método para estimar los parámetros, conocido como método de máxima verosimilitud, que consiste en obtener los valores de los parámetros que maximizan la función de verosimilitud.

Observación **5.1** Para simplificar el proceso de maximización de la función de verosimilitud se utiliza el logaritmo de dicha función, llamada función de log-verosimilitud y denotada por $\ell(\boldsymbol{\theta}|x_i)$. Luego, se tiene que:

$$\ell(\boldsymbol{\theta}|x_i) = \log\left(L(\boldsymbol{\theta}|x_i)\right) = \sum_{i=1}^{n} f(x_i, \boldsymbol{\theta}).$$

Definición **5.4** Sea $\boldsymbol{\theta}$ un vector de parámetros asociado a una cierta función de densidad para una muestra aleatoria $X_1, X_2, \cdots, X_n$. El vector $\widehat{\boldsymbol{\theta}}_{MV}$ es el vector que contiene los valores estimados para el parámetro $\boldsymbol{\theta}$ mediante el método de máxima verosimilitud.

Definición **5.5** Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde la variable sigue una distribución que es conocida, el estimador de máxima verosimilitud (EMV) del parámetro $\boldsymbol{\theta}$ se define por:

$$\widehat{\boldsymbol{\theta}}_{MV} = \underset{\boldsymbol{\theta}}{\text{máx}}\{L(\boldsymbol{\theta}|x_i)\} = \underset{\boldsymbol{\theta}}{\text{máx}}\{\ell(\boldsymbol{\theta}|x_i)\}.$$

Ejemplo **5.5** Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde $X_i \sim \text{Exp}(\theta)$. Obtener el estimador de máxima verosimilitud del parámetro θ.

Solución: Según la definición 5.5, para obtener el EMV se requiere la función de verosimilitud y de log-verosimilitud. Así, se tiene:

$$\begin{aligned}
\ell(\theta|x_i) &= \sum_{i=1}^{n} \log\left(\frac{1}{\theta} e^{-\frac{1}{\theta}x_i}\right) \\
&= \sum_{i=1}^{n} \left[-\log(\theta) - \frac{1}{\theta}x_i\right] \\
&= -\, n\log(\theta) - \frac{1}{\theta}\sum_{i=1}^{n} x_i.
\end{aligned}$$

Igualando la primera derivada respecto a cero de la función anterior, arroja un valor candidato a valor extremo. Entonces:

$$\begin{aligned}
\frac{\partial}{\partial\theta}\ell(\theta|x_i) &= -\,\frac{n}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n} x_i \\
\frac{\partial}{\partial\theta}\ell(\theta|x_i) &= 0 \Rightarrow \widehat{\theta}_{MV} = \frac{1}{n}\sum_{i=1}^{n} X_i.
\end{aligned}$$

Además, se puede comprobar que el valor anteriormente obtenido como candidato a máximo lo es realmente mediante la segunda derivada evaluada en dicho valor:

$$\frac{\partial^2}{\partial\theta^2}\ell(\theta|x_i) = \frac{n}{\theta^2} - \frac{2}{\theta^3}\sum_{i=1}^{n} x_i,$$
$$\left.\frac{\partial^2}{\partial\theta^2}\ell(\theta|x_i)\right|_{\theta=\widehat{\theta}_{MV}} = -\frac{n^3}{(\sum_{i=1}^{n} x_i)^2}.$$

Este último es un valor negativo, luego $\widehat{\theta}_{MV}$ es un máximo de $L(\theta|x_i)$ y, por lo tanto, el estimador de máxima verosimilitud del parámetro θ. □

Ejemplo **5.6** Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria de una variable $X \sim N(\mu, \sigma^2)$. Se desea obtener el estimador de máxima verosimilitud del parámetro $\boldsymbol{\theta} = (\mu, \sigma^2)^T$.

Solución: Se obtiene la función de log-verosimilitud:

$$\begin{aligned}\ell(\boldsymbol{\theta}|x_i) &= \sum_{i=1}^{n} \log\left(\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{1}{2\sigma^2}(x_i-\mu)^2\right\}\right)\\ &= \sum_{i=1}^{n}\left[-\log\sqrt{2\pi} - \log\sqrt{\sigma^2} - \frac{1}{2\sigma^2}(x_i-\mu)^2\right]\\ &= -\frac{n}{2}\log 2 - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2.\end{aligned}$$

La primera derivada de la función de log-verosimilitud, respecto al vector que se quiere maximizar igual a cero, arroja un vector candidato a valor extremo. Como se trata de un vector, se deben hacer las derivadas parciales respecto de cada uno de los componentes del vector. Entonces:

$$\frac{\partial}{\partial\mu}\ell(\boldsymbol{\theta}|x_i) = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu) = \frac{1}{\sigma^2}\left(\sum_{i=1}^{n} x_i - n\mu\right) \tag{5.1}$$

$$\frac{\partial}{\partial\sigma^2}\ell(\boldsymbol{\theta}|x_i) = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(x_i-\mu)^2. \tag{5.2}$$

Luego, los máximos para ambos parámetros se obtienen resolviendo el

sistema de ecuaciones formado al igualar a cero las ecuaciones 5.1 y 5.2:

$$\widehat{\mu}_{MV} = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}. \tag{5.3}$$

$$\widehat{\sigma^2}_{MV} = \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2. \tag{5.4}$$

Finalmente, para verificar que los valores dados por la expresión 5.3 y la expresión 5.4 son, efectivamente, máximos de la función de verosimilitud se debe tener que la matriz hessiana evaluada en $\widehat{\boldsymbol{\theta}}_{MV}$ es definida negativa (queda de ejercicio para el lector).

Ejemplo **5.7** Sean $X_{i1}, X_{i2}, \cdots, X_{in}$ con $i = 1, 2, \cdots, k$ un conjunto de k muestras aleatorias independientes, donde en las poblaciones se tiene $X_i \sim N(\mu_i, \sigma^2)$. Se desea obtener el estimador de máxima verosimilitud para $\boldsymbol{\theta} = (\mu_1, \mu_2, \cdots, \mu_k, \sigma^2)^T$.

Solución: Se comienza por la función de log-verosimilitud:

$$\begin{aligned}
L\left(\boldsymbol{\theta}|x_1, x_2, \cdots, x_k\right) &= L\left(\mu_1, \sigma^2|x_1\right) \cdot L\left(\mu_2, \sigma^2|x_2\right) \cdot \cdots \cdot L\left(\mu_k, \sigma^2|x_k\right) \\
&= \prod_{i=1}^{k} L\left(\mu_i, \sigma^2|x_i\right). \\
\ell\left(\boldsymbol{\theta}|x_1, x_2, \cdots, x_k\right) &= \sum_{i=1}^{k} \ell\left(\mu_i, \sigma^2\right) \\
&= \sum_{i=1}^{k}\sum_{j=1}^{n} \log\left[\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(x_{ij}-\mu)^2\right\}\right] \\
&= -\frac{nk}{2}\log(2\pi) - \frac{nk}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{k}\sum_{j=1}^{n}\left(x_{ij}-\mu_i\right)^2.
\end{aligned}$$

Las derivadas respecto a los $k+1$ parámetros son las siguientes:

$$\frac{\partial}{\partial \mu_i}\ell\left(\boldsymbol{\theta}|x_{ij}\right) = \frac{1}{\sigma^2}\sum_{j=1}^{n}(x_{ij}-\mu_i) = \frac{1}{\sigma^2}\left(\sum_{j=1}^{n} x_{ij} - n\mu_i\right) \tag{5.5}$$

$$\frac{\partial}{\partial \sigma^2}\ell\left(\boldsymbol{\theta}|x_{ij}\right) = -\frac{nk}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{k}\sum_{j=1}^{n}(x_{ij}-\mu_i)^2. \tag{5.6}$$

En la ecuación 5.5 el resultado es válido para $i = 1, 2, \cdots, k$. Igualando a cero las derivadas y resolviendo el sistema de ecuaciones, se

obtienen las componentes del vector de parámetros estimados $\widehat{\boldsymbol{\theta}}_{MV} = (\widehat{\mu}_1, \widehat{\mu}_2, \cdots, \widehat{\mu}_k, \widehat{\sigma}^2)^T$ dados por:

$$\widehat{\mu}_i = \frac{1}{n}\sum_{j=1}^{n} X_{ij} = \bar{X}_i, \quad i = 1, 2, \cdots, k.$$

$$\widehat{\sigma}^2 = \frac{1}{nk}\sum_{i=1}^{k}\sum_{j=1}^{n}(x_{ij} - \widehat{\mu}_i)^2 = \frac{1}{nk}\sum_{i=1}^{k}\sum_{j=1}^{n}(X_{ij} - \bar{X}_i)^2.$$

□

Una propiedad del EMV de mucha utilidad es que, si se cuenta con el EMV de un parámetro, entonces el EMV de una función de este parámetro es el EMV evaluado en dicha función; esto se cumple solo si la función es 1 a 1. A continuación, se presenta este resultado de invarianza del EMV.

Propiedad **5.1** Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria con $X \sim f(x|\theta)$ y sea $\widehat{\theta}_{MV}$ el EMV de θ, entonces $g(\widehat{\theta}_{MV})$ es el EMV de $g(\theta)$, cuando $g(\cdot)$ es una función 1 a 1. Este resultado es conocido como *propiedad de invarianza* del EMV.

Demostración: Siendo $g(\cdot)$ una función 1 a 1, tal que $\theta = g^{-1}(g(\theta))$, la cual se utiliza para obtener la función de verosimilitud:

$$L(\theta; x) = L\left(g^{-1}(g(\theta)); x\right).$$

Luego, $\widehat{\theta}$ maximiza ambos lados de la igualdad, así:

$$\widehat{\theta} = g^{-1}\left(\widehat{g(\theta)}\right).$$

Desde este último resultado se obtiene que $g(\widehat{\theta}) = \widehat{g(\theta)}$. O sea, el EMV de $g(\theta)$ es $g(\widehat{\theta})$.

Ejemplo **5.8** Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde la distribución de la variable aleatoria es $X \sim$ Bernoulli(θ). Se desea obtener el estimador de máxima verosimilitud para $\delta = \theta - \theta^2$.

Solución: Nótese que el parámetro δ se puede expresar como función de θ, siendo $\delta = g(\theta) = \theta - \theta^2$. Luego, por la propiedad de invarianza se tiene que $\widehat{\delta} = \widehat{g(\theta)} = g(\widehat{\theta})$.

En efecto, el análisis anterior implica obtener el estimador de máxima verosimilitud de θ:

$$\ell(\theta|x_i) = \log\theta \cdot \sum_{i=1}^{n} x_i + \log(1-\theta) \cdot (n - \sum_{i=1}^{n} x_i)$$

$$\frac{\partial}{\partial\theta}\ell(\theta|x_i) = \frac{1}{\theta} \cdot \sum_{i=1}^{n} x_i - \frac{1}{1-\theta} \cdot (n - \sum_{i=1}^{n} x_i)$$

$$\widehat{\theta}_{MV} = \bar{X}.$$

Finalmente, $\widehat{\delta} = g(\bar{X}) = \bar{X} - \bar{X}^2$.

En este caso, la propiedad 5.1 de invarianza del estimador de máxima verosimilitud simplifica el cálculo para obtener el estimador. □

Ejemplo **5.9** Desde el ejemplo 4.17, considerar el parámetro θ igual al costo real promedio, es decir, $\theta = E(C)$. Obtener un estimador para θ.

Solución: Como $\theta = E(C) = E(100 + 7X + 3Y) = 100 + 7\mu_1 + 3\mu_2$, entonces, usando la propiedad 5.1 de invarianza:

$$\widehat{\theta} = 100 + 7\widehat{\mu_1} + 3\widehat{\mu_2}.$$

Luego, usando el método de máxima verosimilitud (véase ejercicio 5.6), se obtiene que $\widehat{\mu_1} = \bar{X}$ y $\widehat{\mu_2} = \bar{Y}$. Finalmente:

$$\widehat{\theta} = 100 + 7\bar{X} + 3\bar{Y}.\square$$

Ejercicio propuesto **5.1** Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria con $X_i \sim \text{Exp}(\beta)$. Encontrar los estimadores de máxima verosimilitud para los parámetros:

(a) $\delta_1 = E[X]$.

(b) $\delta_2 = \mathbb{P}(X \geqslant 1)$.

Método de los momentos

Este método, a diferencia del método de máxima verosimilitud, no requiere conocer la distribución de las variables aleatorias que componen la muestra aleatoria, sino solo el valor esperado de X_i^r, $r = 1, 2, \cdots, q$, donde el vector de parámetros $\boldsymbol{\theta}$ es de dimensión $q \times 1$. Es decir, si el parámetro tiene solo un componente, es de 1×1, entonces se requiere conocer $E[X_i]$.

Definición **5.6** Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde se conoce el valor esperado de X_i^r. El método de los momentos para estimar el parámetro $\boldsymbol{\theta}$ (de dimensión $q \times 1$), denotado por $\widehat{\boldsymbol{\theta}}_{MM}$, se obtiene resolviendo el sistema de ecuaciones siguiente:

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n} x_i &= E[X] \\
\frac{1}{n}\sum_{i=1}^{n} x_i^2 &= E[X^2] \\
&\vdots \\
\frac{1}{n}\sum_{i=1}^{n} x_i^q &= E[X^q].
\end{aligned}$$

Ejemplo **5.10** Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde $E[X] = \mu$ y la varianza $Var[X] = \sigma^2$. Se desea encontrar un estimador para el vector $\boldsymbol{\theta} = (\mu, \sigma^2)^T$, que, como tiene dimensión 2, esa será también la cantidad de ecuaciones que compongan el sistema. Entonces:

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n} x_i &= E[X] = \mu. \\
\frac{1}{n}\sum_{i=1}^{n} x_i^2 &= E[X^2] = Var[X] + (E[X])^2 = \sigma^2 + \mu^2.
\end{aligned}$$

De estas ecuaciones se tiene que el estimador de momentos es:

$$\begin{aligned}
\widehat{\mu}_{MM} &= \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}. \\
\widehat{\sigma^2}_{MM} &= \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \widehat{\mu}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2.
\end{aligned}$$

□

En caso de que no se disponga de una muestra aleatoria y, en su lugar, se disponga de variables aleatorias independientes, entonces el siguiente resultado es una posibilidad para obtener el estimador de momentos.

Observación **5.2** Sean $X_1, X_2, \cdots, X_n$ variables aleatorias independientes (no necesariamente identicamente distribuidas), donde X_i tiene el mismo

supuesto que la variable aleatoria X medida en la población. Entonces el estimador de momentos para el parámetro $\boldsymbol{\theta}$ de dimensión q se obtiene resolviendo:

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n} x_i &= E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] \\
\frac{1}{n}\sum_{i=1}^{n} x_i^2 &= E\left[\frac{1}{n}\sum_{i=1}^{n} X_i^2\right] \\
&\vdots \\
\frac{1}{n}\sum_{i=1}^{n} x_i^q &= E\left[\frac{1}{n}\sum_{i=1}^{n} X_i^q\right].
\end{aligned}$$

Ejemplo **5.11** Las variables aleatorias $Y_1, Y_2, \cdots, Y_n$ son independientes, con función de densidad $f(y_i) = (\beta x_i)^{-1} \exp\left\{-y_i(\beta x_i)^{-1}\right\}$, $y_i > 0$, donde los elementos x_i son constantes positivas y $\beta > 0$. Se requiere obtener el estimador de momentos para el parámetro β.

Solución: El estimador de momentos se obtiene considerando que la esperanza de X_i está dada por $E[Y_i] = \beta x_i$. Entonces:

$$\frac{1}{n}\sum_{j=1}^{n} Y_j = E\left[\frac{1}{n}\sum_{j=1}^{n} Y_j\right] = \frac{1}{n}\sum_{j=1}^{n} E[Y_j] = \frac{1}{n}\beta \cdot \left(\sum_{j=1}^{n} x_j\right).$$

Finalmente, evaluando $\beta = \widehat{\beta}_{MM}$, se tiene el estimador:

$$\widehat{\beta}_{MM} = \left(\frac{1}{n}\sum_{j=1}^{n} X_j\right)^{-1} \frac{1}{n}\sum_{j=1}^{n} Y_j = \left(\bar{X}\right)^{-1} \bar{Y}.$$

□

Método de mínimos cuadrados

A diferencia de los métodos anteriores, este método no requiere conocer la distribución de la muestra aleatoria, ni el momento de orden r. El requerimiento es solo conocer el valor esperado de la variable aleatoria.

Definición **5.7** Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde se conoce el valor esperado de la variable X, $E[X]$, el que depende de q parámetros $\theta_1, \theta_2, \cdots, \theta_q$. El método de mínimos cuadrados es definido como el

mínimo con respecto a los q parámetros de la expresión:

$$Q(\boldsymbol{\theta}|\boldsymbol{x_i}) = \sum_{i=1}^{n} \left(X_i - E[X]\right)^2 .$$

Ejemplo **5.12** Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde $X \sim Exp(\beta)$. Luego, se sabe que $E[X] = \beta$. Entonces:

$$Q = \sum_{i=1}^{n} \left[x_i - \beta\right]^2 \Rightarrow \widehat{\beta}_{MC} = \min_{\beta}\{\sum_{i=1}^{n} \left[x_i - \beta\right]^2\}.$$

Finalmente, $\dfrac{\partial}{\partial\beta}Q = -2\sum_{i=1}^{n}(x_i - \beta) \Rightarrow \widehat{\beta}_{MC} = \dfrac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}.$ □

Ejemplo **5.13** Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, tal que $E[X] = \alpha + \beta A_i$, donde A_i son constantes positivas. Obtener el estimador de mínimos cuadrados del vector de parámetros $\boldsymbol{\theta} = (\alpha, \beta)^T$.

Solución: La ecuación $Q(\boldsymbol{\theta}|\boldsymbol{x_i})$ en este caso es:

$$Q = \sum_{i=1}^{n} \left[x_i - (\alpha + \beta\, A_i)\right]^2 .$$

Las derivadas respecto de $\boldsymbol{\theta}$ son:

$$\begin{aligned}
\frac{\partial}{\partial\alpha}Q &= -\,2\sum_{i=1}^{n}\left[x_i - (\alpha + \beta\, A_i)\right] \\
\frac{\partial}{\partial\beta}Q &= -\,2\sum_{i=1}^{n}\left[(x_i - (\alpha + \beta\, A_i))\cdot A_i\right].
\end{aligned}$$

Igualando a cero y resolviendo, los estimadores son:

$$\begin{aligned}
\widehat{\beta}_{MC} &= \frac{1}{\sum_{i=1}^{n}(A_i - \bar{A})^2}\left(\sum_{i=1}^{n} A_i x_i - \bar{x}\sum_{i=1}^{n} A_i\right) \\
\widehat{\alpha}_{MC} &= \bar{X} - \widehat{\beta}_{MC}\bar{A}. \square
\end{aligned}$$

5.4. Propiedades de los estimadores

Hasta el momento se han presentado métodos para obtener estimadores para un parámetro $\boldsymbol{\theta}$. Ahora, si se tienen dos estimadores distintos, digamos $\widehat{\theta}_1$ y $\widehat{\theta}_2$ para el parámetro θ de interés, surge la interrogante de

cómo seleccionar entre estos estimadores el mejor para el parámetro en cuestión.

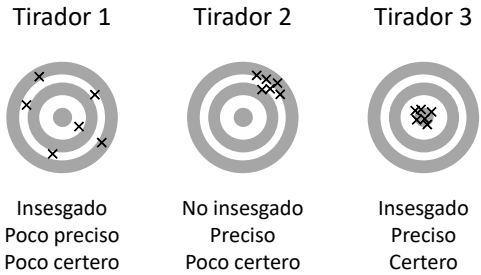
Puede hacerse la analogía con jugadores de tiro al blanco: suponer que se tienen los resultados de los lanzamientos de tres tiradores (véase la figura 5.3), se establece que el mejor tirador es el tercero pues sus lanzamientos cumplen con las siguientes características:

(1) En general, los lanzamientos están cerca del blanco.

(2) La distancia que hay entre los tiros realizados es pequeña.

(3) Los tiros están alrededor del blanco.

Figura 5.3: Analogía de los tiradores con propiedades de los estimadores



Fuente: adaptación de Sharon L. Lohr. *Sampling: Design and Analysis*. Arizona State University, 2009.

Asumiendo que el blanco es el valor del parámetro y los tiros son la estimación, se tiene que un estimador debe cumplir con las tres características anteriores (o al menos informarse acerca de estas), las cuales son:

Insesgamiento: $E[\widehat{\theta}] = \theta$.

Precisión: $Var[\widehat{\theta}]$ es pequeño.

Certero: (eficiente): $E\left[(\widehat{\theta} - \theta)^2\right]$ es pequeño.

Luego, se avanza a lo que se llamará consistencia del estimador, estudiando el comportamiento del estimador en muestras grandes ($n \to \infty$).

Esta propiedad puede ser complementaria con la anterior, pues si el estimador depende de n (tamaño de la muestra) se debería mejorar la estimación en la medida que se cuente con mayor información acerca de la población.

Ejemplo **5.14** Se tiene una muestra aleatoria $X_1, X_2, \cdots, X_n$, donde $X \sim N(\mu, \sigma^2)$. Son estimadores de los parámetros μ y σ^2, respectivamente:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \text{ y } S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2.$$

Entonces, se quiere conocer qué tan buenos son estos estimadores para los parámetros respectivos, para lo cual se definen propiedades asociadas a la distribución del estimador del parámetro (como variable aleatoria); el valor esperado debería ser el parámetro y la varianza pequeña. □

Insesgamiento

Definición **5.8** Sean $\widehat{\theta} = T(X_1, X_2, \cdots, X_n)$ un estimador del parámetro θ. El estimador $\widehat{\theta}$ es un estimador insesgado si y solo si:

$$E[\widehat{\theta}] = \theta.$$

Este es el primer requisito que se le pide a un estimador: ser insesgado. Esta propiedad se fundamenta en que, dada una muestra aleatoria, esta se puede repetir en las mismas condiciones y en cada una de ellas obtener un nuevo estimador, por lo que se tiene que este estimador tiene una distribución asociada. Lo que se espera es que la esperanza de la distribución del parámetro estimado sea el parámetro.

Ejemplo **5.15** Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde $X \sim N(\mu, \sigma^2)$. Verificar si los siguientes estimadores para los parámetros σ^2 y e^{μ}, respectivamente, son insesgados:

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 \quad \text{y} \quad e^{\bar{X}}.$$

Solución: Para obtener $E[S^2]$ recordar que:

$$X \sim N(\mu, \sigma^2) \Rightarrow \frac{n-1}{\sigma^2} \cdot S^2 \sim \chi^2\,(n-1).$$

Con este resultado, $E[S^2] = \sigma^2$, lo cual verifica que S^2 es un estimador insesgado para σ^2.

Siendo $\bar{X} \sim N(\mu, \sigma^2 \cdot n^{-1})$, se obtiene $E[e^{\bar{X}}]$ usando la fgm de $\bar{X}$:

$$E\left[e^{\bar{X}}\right] = M_{\bar{X}}(t=1) = \exp\left\{\mu + \frac{1}{2}\frac{\sigma^2}{n}\right\}.$$

Por lo tanto, $e^{\bar{X}}$ no es un estimador insesgado de e^{μ}. □

Observación **5.3** En el ejemplo anterior, se puede observar que si $n \to \infty$, entonces $e^{\bar{X}}$ es un estimador insesgado de e^{μ}. Este resultado lleva al concepto de *estimador asintóticamente insesgado*; entonces, un estimador $\widehat{\theta}$ es un estimador asintóticamente insesgado para el parámetro θ si:

$$\lim_{n\to\infty} E[\widehat{\theta}] = \theta.$$

La observación anterior dice que, si no se dispone de un estimador insesgado, una alternativa es que el valor esperado del estimador converja al parámetro en la medida que n es grande. Por lo tanto, en estos casos se requiere de una muestra aleatoria de un n suficientemente grande.

Observación **5.4** En ciertos casos, cuando no se dispone de un estimador insesgado, es posible construir estimadores insesgados. Por ejemplo, en el caso de que se tenga una muestra aleatoria $X_1, X_2, \cdots, X_n$ con $X \sim N(\mu, \sigma^2)$, el EVM de σ^2 y el valor esperado de este son:

$$\widehat{\sigma^2}_{MV} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{n-1}{n}S^2 \text{ y } E[\widehat{\sigma}^2_{MV}] = \frac{n-1}{n}\sigma^2.$$

Entonces, $\frac{n}{n-1} \cdot \widehat{\sigma}^2_{MV}$ es un estimador insesgado para el parámetro σ^2.

Precisión de la estimación (eficiencia)

Otra de las propiedades de un estimador se refiere a la variabilidad. Suponer que se dispone de dos parámetros, ambos insesgados, aunque uno presenta mayor variabilidad que el otro, entonces se dice que el segundo estimador es más eficiente que el primero.

Propiedad **5.2** Sea $\widehat{\theta}$ un estimador insesgado del parámetro θ y Sean $\widehat{\theta}^*$ otro estimador insesgado del mismo parámetro θ. El estimador $\widehat{\theta}$ es de mínima varianza si y solo si:

$$Var[\widehat{\theta}] \leqslant Var[\widehat{\theta}^*].$$

Para medir la eficiencia de un estimador θ se utiliza un indicador para comparar la variabilidad del estimador, llamado cota de Cramer-Rao.

Propiedad **5.3** Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde X tiene una distribución conocida (y luego también es conocida la función de densidad). Además, sea $\widehat{\theta}$ un estimador insesgado del parámetro θ. Entonces:

$$Var[\widehat{\theta}] \geqslant CCR,$$

donde CCR es la cota de Cramer-Rao y se define por:

$$CCR = \frac{1}{n \cdot E\left[\left(\frac{\partial}{\partial\theta}\log f(x_i)\right)^2\right]} = \frac{-1}{n \cdot E\left[\frac{\partial^2}{\partial\theta^2}\log f(x_i)\right]}.$$

Ejemplo **5.16** Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde $X \sim N(\mu, \sigma_0^2)$, con σ_0^2 conocido. Analizar la eficiencia del EMV, EMM y EMC para μ es $\bar{X}$ (que son insesgados).

Solución: Nótese que $Var[\widehat{\mu}] = Var[\bar{X}] = \frac{\sigma_0^2}{n}$. Además, para calcular la cota de Cramer-Rao se tiene que:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left\{-\frac{1}{2\sigma_0^2}(x-\mu)^2\right\}$$

$$\log f(x) = -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log\sigma_0^2 - \frac{1}{2\sigma_0^2}(x-\mu)^2$$

$$\frac{\partial}{\partial\mu}\log f(x) = \frac{1}{\sigma_0^2}(x-\mu)$$

$$\frac{\partial^2}{\partial\mu^2}\log f(x) = -\frac{1}{\sigma_0^2}.$$

Luego, $CCR = \frac{\sigma_0^2}{n} = Var[\widehat{\mu}]$, se dice que el estimador es eficiente. □

Observación **5.5** En la propiedad 5.3 la expresión de la cota de Cramer-Rao es la primera expresión (basada en la primera derivada de la función

de log-verosimilitud); la segunda igualdad se cumple bajo ciertas condiciones de regularidad. Para los casos en estudio de este texto se han seleccionado aquellos en que se cumplen las condiciones.

Definición **5.9** En caso de que la varianza del estimador sea mayor que la cota de Cramer-Rao, es conveniente definir un coeficiente de eficiencia. Se define la eficiencia relativa de un estimador (ER) como:

$$ER = \frac{CCR}{Var[\widehat{\theta}]}.$$

Observación **5.6** A partir de la propiedad 5.3, se tiene que $ER \leqslant 1$.

Ejemplo **5.17** Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde $X \sim N(\mu, \sigma^2)$. En el ejemplo 5.15 se probó que S^2 es un estimador insesgado para el parámetro σ^2; se verificará la eficiencia de este estimador.

Solución: Dado que $\frac{n-1}{\sigma^2}S^2 \sim \chi^2(n-1)$, entonces $Var[S^2] = 2(\sigma^2)^2 \cdot (n-1)^{-1}$. Por su parte, la cota de Cramer-Rao se obtiene mediante:

$$\begin{aligned}
f(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\} \\
\log f(x) &= -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log\sigma^2 - \frac{1}{2\sigma^2}(x-\mu)^2 \\
\frac{\partial}{\partial\sigma^2}\log f(x) &= -\frac{1}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}(x-\mu)^2 \\
\frac{\partial^2}{\partial(\sigma^2)^2}\log f(x) &= \frac{1}{2(\sigma^2)^2} - \frac{1}{(\sigma^2)^3}(x-\mu)^2.
\end{aligned}$$

Luego, la cota de Cramer-Rao es:

$$\text{CCR} = -\frac{1}{n}\left[\frac{1}{2(\sigma^2)^2} - \frac{1}{(\sigma^2)^3}E\left[(x-\mu)^2\right]\right]^{-1} = \frac{2(\sigma^2)^2}{n}.$$

De este resultado se tiene que la varianza del estimador no alcanza la CCR. Luego, se puede obtener la eficiencia relativa del estimador, la que está dada por:

$$ER = \frac{n-1}{n}. \qquad \square$$

Error cuadrático medio

Sea $\widehat{\theta}$ un estimador cualquiera (posiblemente no insesgado) del parámetro θ. El error cuadrático medio (ECM) del estimador $\widehat{\theta}$ está dado por:

$$ECM(\widehat{\theta}) = E\left[\left(\widehat{\theta} - \theta\right)^2\right].$$

Propiedad **5.4** Una expresión equivalente para el $ECM(\widehat{\theta})$ es:

$$ECM(\widehat{\theta}) = Var[\widehat{\theta}] + \left(E[\widehat{\theta}] - \theta\right)^2.$$

Demostración: El resultado se obtiene a partir de la definición 5.4:

$$\begin{aligned} ECM(\widehat{\theta}) &= E\left[\left(\widehat{\theta} - \theta\right)^2\right] = E\left[\left(\widehat{\theta} - E[\widehat{\theta}] + E[\widehat{\theta}] - \theta\right)^2\right] \\ &= E\left[\left(\widehat{\theta} - E[\widehat{\theta}]\right)^2 + \left(E[\widehat{\theta}] - \theta\right)^2 + 2(\widehat{\theta} - E[\widehat{\theta}])(E[\widehat{\theta}] - \theta)\right] \\ &= E\left[\left(\widehat{\theta} - E[\widehat{\theta}]\right)^2\right] + E\left[\left(E[\widehat{\theta}] - \theta\right)^2\right] + 2E\left[(\widehat{\theta} - E[\widehat{\theta}])(E[\widehat{\theta}] - \theta)\right] \\ &= Var[\widehat{\theta}] + \left(E[\widehat{\theta}] - \theta\right)^2. \end{aligned}$$

Ejemplo **5.18** Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria con $X \sim N(\mu, \sigma^2)$. Se propone como estimador del parámetro $\theta = e^{\mu}$ a $\widehat{\theta} = e^{\bar{X}}$. Obtener una expresión para el ECM de este estimador.

Solución: Usando la propiedad 5.4:

$$E[\widehat{\theta}] = E[e^{\bar{X}}] = M_{\bar{X}}(t=1) = \exp\left\{\mu + \frac{1}{2n}\sigma^2\right\} = e^{\mu} \cdot \exp\left\{\frac{1}{2n}\sigma^2\right\}.$$

Nótese del resultado anterior que $\widehat{\theta}$ no es insesgado.

$$Var[\widehat{\theta}] = E[\widehat{\theta}^2] - \left(E[\widehat{\theta}]\right)^2,$$

donde $E[\widehat{\theta}^2] = E[e^{2\bar{X}}] = M_{\bar{X}}(t=2) = \exp\left\{2\mu + 2\frac{\sigma^2}{n}\right\}$. Luego:

$$\begin{aligned} Var[\widehat{\theta}] &= \exp\left\{2\mu + 2\frac{\sigma^2}{n}\right\} - \exp\left\{2\mu + \frac{\sigma^2}{n}\right\} \\ \Rightarrow ECM(\widehat{\theta}) &= \exp\{2\mu\} \cdot \left(1 + \exp\left\{\frac{2\sigma^2}{n}\right\} - 2\exp\left\{\frac{\sigma^2}{2n}\right\}\right). \square \end{aligned}$$

Ejemplo **5.19** Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde $X \sim N(\mu, \sigma^2)$. Encontrar una expresión para el *EMC* del estimador de máxima verosimilitud de σ^2:

$$\widehat{\sigma^2}_{MV} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2.$$

Solución: El $ECM(\hat{\sigma}^2)$ se obtiene resolviendo:

$$E[\hat{\sigma}^2] = \frac{n-1}{n}\sigma^2.$$

$$\frac{n-1}{\sigma^2}S^2 \sim \chi^2(n-1) \Rightarrow Var[\hat{\sigma}^2_{MV}] = \frac{2(n-1)}{n^2}\sigma^4.$$

Con estos resultados el *ECM* es:

$$ECM(\hat{\sigma}^2_{MV}) = \frac{2(n-1)}{n^2}\sigma^4 + \left[\frac{n-1}{n}\sigma^2 - \sigma^2\right]^2.\square$$

Estimadores consistentes

Esta propiedad controla tanto el valor esperado como la variabilidad del estimador y se fundamenta en que, si la muestra es tan grande como la población, el comportamiento del estimador debería tender al parámetro en el valor esperado y la varianza disminuir tendiendo a cero.

Definición **5.10** El estimador $\widehat{\theta}$ del parámetro θ es llamado un estimador consistente si:

$$\lim_{n\to\infty} E\left[\widehat{\theta}\right] = \theta \quad \text{y} \quad \lim_{n\to\infty} Var\left[\widehat{\theta}\right] = 0.$$

Es decir, $\widehat{\theta}$ es un estimador consistente para el parámetro θ si es asintóticamente insesgado y la variabilidad asintótica tiende a cero.

Ejemplo **5.20** Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde $X \sim N(\mu, \sigma^2)$. Se quiere probar que $\hat{\sigma}^2_{MV}$ es un estimador consistente del parámetro σ^2.

Se sabe que $\hat{\sigma}^2_{MV} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$. Entonces:

$$E\left[\hat{\sigma}^2_{MV}\right] = \frac{n-1}{n}\sigma^2$$

$$Var\left[\hat{\sigma}^2_{MV}\right] = \frac{2(n-1)}{n^2}(\sigma^2)^2.$$

Aplicando límites de $n \to \infty$ a las expresiones anteriores:

$$\lim_{n\to\infty} E\left[\widehat{\sigma}^2_{MV}\right] = \sigma^2$$
$$\lim_{n\to\infty} Var\left[\widehat{\sigma}^2_{MV}\right] = 0.$$

Por lo tanto, $\widehat{\sigma}^2_{MV}$ es un estimador consistente del parámetro σ^2. □

Ejercicio propuesto **5.2** (1) Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde la función de densidad de las variables aleatorias X_i es:

$$f(x_i) = \frac{1}{\theta} x_i^{-(\theta^{-1}+1)},\ x_i \geqslant 1,\ \theta > 0.$$

(a) Estudiar la consistencia del estimador de máxima verosimilitud de θ.

(b) Determinar si el estimador de máxima verosimilitud de θ es el mejor estimador dentro de todos los estimadores insesgados.

(2) Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde $X_i \sim N(0, \sigma^2)$.

(a) Determinar cuál de los siguientes estimadores es consistente:

$$T_1 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 \quad T_2 = \left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^2 \quad T_3 = \frac{1}{n}\sum_{i=1}^{n} (X_i - \bar{X})^2.$$

(b) Si tuviese que elegir un estimador de los anteriores, ¿cuál sería la elección? Fundamentar.

5.5. Estimación por intervalos

En la sección anterior se estudiaron métodos para obtener estimaciones puntuales para un parámetro de una variable aleatoria por medio de una muestra (además de las propiedades que se esperan para dichos estimadores). Ahora se estudiarán métodos para obtener un intervalo en que se encuentre el valor del parámetro, con una cierta probabilidad (o confianza) de acertar; es decir, aquí la estimación no se hace mediante un valor puntual, sino mediante un conjunto de valores y, por ello, estos métodos se llaman estimación por intervalos de confianza.

Por ejemplo, Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde $X \sim N(\mu, \sigma_0^2)$, donde σ_0^2 es un valor conocido. Se sabe que $\bar{X}$ es un estimador del parámetro μ, además:

$$\bar{X} \sim N(\mu, \frac{\sigma_0^2}{n}) \Rightarrow Z = \frac{\bar{X} - \mu}{\sqrt{\frac{\sigma_0^2}{n}}} \sim N(0,1). \tag{5.7}$$

Por otra parte, como $Z \sim N(0,1)$, entonces $\mathbb{P}(-1{,}96 \leqslant Z \leqslant 1{,}96) = 0{,}95$ y, reemplazando Z por la equivalencia dada en la expresión 5.7, entonces la probabilidad siguiente es igual a 0,95:

$$\mathbb{P}\left(-1{,}96 \leqslant \frac{\bar{X} - \mu}{\sqrt{\frac{\sigma_0^2}{n}}} \leqslant 1{,}96\right) = \mathbb{P}\left(\bar{X} - 1{,}96 \cdot \sqrt{\frac{\sigma_0^2}{n}} \leqslant \mu \leqslant \bar{X} + 1{,}96 \cdot \sqrt{\frac{\sigma_0^2}{n}}\right).$$

Este resultado se interpreta como que la probabilidad de que el parámetro μ (la media) esté entre los valores $\bar{X} - 1{,}96 \cdot \sqrt{\frac{\sigma_0^2}{n}}$ y $\bar{X} + 1{,}96 \cdot \sqrt{\frac{\sigma_0^2}{n}}$ es de 0,95. Es decir, se ha encontrado un intervalo en el cual se hallará el parámetro bajo cierta probabilidad. A esto se le llamará *estimación por intervalos de confianza*.

Si bien el resultado anterior fue obtenido para un valor de probabilidad particular, se puede generalizar a un valor cualquiera. En efecto, considérese una probabilidad de $1 - \alpha$ y una variable aleatoria $Z \sim N(0,1)$:

$$\mathbb{P}(-z_{1-\frac{\alpha}{2}} \leqslant Z \leqslant z_{1-\frac{\alpha}{2}}) = 1 - \alpha.$$

El gráfico de la figura 5.4 muestra la región donde se calcula la probabilidad, la cual está definida entre $-z_{1-\frac{\alpha}{2}}$ y $z_{1-\frac{\alpha}{2}}$ en el recorrido de la variable aleatoria $Z \sim N(0,1)$.

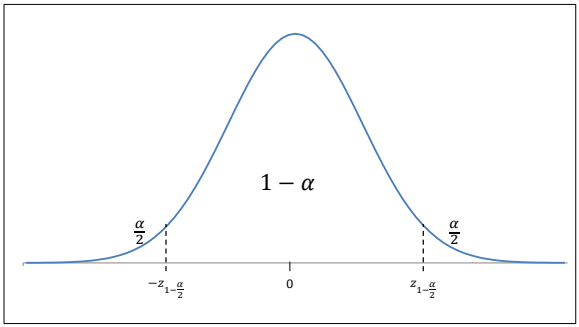
Reemplazando Z según la expresión 5.7 y despejando μ:

$$\mathbb{P}\left(\bar{X} - z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\sigma_0^2}{n}} \leqslant \mu \leqslant \bar{X} + z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\sigma_0^2}{n}}\right) = 1 - \alpha.$$

Luego, el intervalo de confianza del $1 - \alpha$ para el parámetro μ es:

$$\left[\bar{X} - z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\sigma_0^2}{n}}, \bar{X} + z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\sigma_0^2}{n}}\right].$$

Figura 5.4: Gráfico de $Z \sim N(0,1)$ para la construcción de una región crítica



Fuente: elaboración propia.

Cantidad pivotal

Una de las claves para resolver la situación anterior fue considerar la distribución de la variable aleatoria Z según la ecuación 5.7. Nótese que en dicha ecuación están μ y el estimador de máxima verosimilitud de μ, $\bar{X}$, y la distribución asociada no depende de ningún parámetro (o valor desconocido). Estos tres elementos son los que se deben considerar para construir intervalos de confianza para un parámetro. A continuación, se presenta el concepto de *pivote*, una función que contiene el parámetro y el estimador de máxima verosimilitud del parámetro y sigue una distribución que no depende de ningún parámetro.

Definición **5.11** Sea $\widehat{\theta}$ un estimador del parámetro θ. Una función f del estimador $\widehat{\theta}$ y del parámetro θ, es decir, $f(\widehat{\theta}, \theta)$, se conoce como una cantidad pivotal o pivote si y solo si la distribución de $f(\widehat{\theta}, \theta)$ no depende del parámetro θ.

Ejemplo **5.21** Las siguientes funciones son ejemplos de pivotes:

(1) Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde $X_i \sim N(\mu, \sigma_0^2)$, donde σ_0^2 es un valor conocido. Se sabe que $\bar{X}$ es un buen estimador del parámetro μ, además $\bar{X} \sim N(\mu, \sigma_0^2 \cdot n^{-1})$. Sea:

$$f(\bar{X}, \mu) = \frac{\bar{X} - \mu}{\sqrt{\frac{\sigma_0^2}{n}}} \sim N(0,1).$$

Entonces, $f(\bar{X}, \mu)$ satisface las condiciones para ser función pivote.

(2) Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde $X_i \sim N(\mu, \sigma^2)$. Se sabe de las siguientes variables aleatorias independientes que:

$$\frac{\bar{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}} \sim N(0,1) \text{ y } \frac{n-1}{\sigma^2} S^2 \sim \chi^2(n-1). \tag{5.8}$$

Por los resultados obtenidos anteriormente, la siguiente función es pivote para el parámetro μ:

$$f(\bar{X}, \mu) = \frac{\bar{X} - \mu}{S \cdot \sqrt{n^{-1}}} \sim TS(n-1).$$

(3) Obsérvese que la variable aleatoria W de la expresión 5.8 del punto anterior es una cantidad pivotal para el parámetro σ^2:

$$f(\widehat{\sigma^2}, \sigma^2) = \frac{n-1}{\sigma^2} S^2 \sim \chi^2(n-1).$$

(4) Sean $X_1, X_2, \cdots, X_{n_1}$ y $Y_1, Y_2, \cdots, Y_{n_2}$ muestras aleatorias independientes, donde $X \sim N(\mu_1, \sigma_1^2)$ y $Y \sim N(\mu_2, \sigma_2^2)$, siendo σ_1^2 y σ_2^2 valores conocidos. Se desea construir una cantidad pivotal para el parámetro $\theta = \mu_2 - \mu_1$.

Aquí se puede obtener que el EMV de θ es $\widehat{\theta} = \bar{Y} - \bar{X}$. Por otra parte:

$$\bar{X} \sim \mathrm{N}(\mu_1, \frac{\sigma_1^2}{n_1}) \text{ y } \bar{Y} \sim \mathrm{N}(\mu_2, \frac{\sigma_2^2}{n_2}).$$

Luego, $\bar{Y} - \bar{X} \sim N\left(\mu_2 - \mu_1, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$. Finalmente, recordando que σ_1^2 y σ_2^2 son valores conocidos, el pivote para θ es la función:

$$f(\widehat{\theta}, \theta) = \frac{\bar{Y} - \bar{X} - (\mu_2 - \mu_1)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1). \ \square$$

Método del pivote

Los pivotes pueden ser utilizados para construir intervalos de confianza para el parámetro θ, del cual depende la función pivotal. Para ello, se pueden considerar los siguientes pasos:

Paso 1: disponer de un pivote, $f(\widehat{\theta}, \theta)$, para el parámetro θ (de interés).

Paso 2: fijar los valores de U_I y de U_S de modo que el pivote se encuentre entre estos valores, donde U_I es el percentil de orden $\frac{\alpha}{2}$ y U_S, el percentil de orden $1-\frac{\alpha}{2}$ de la distribución de $f(\widehat{\theta},\theta)$. Entonces, se puede determinar la probabilidad de que el parámetro se encuentre entre U_I y U_S, la cual es $1-\alpha$, es decir:

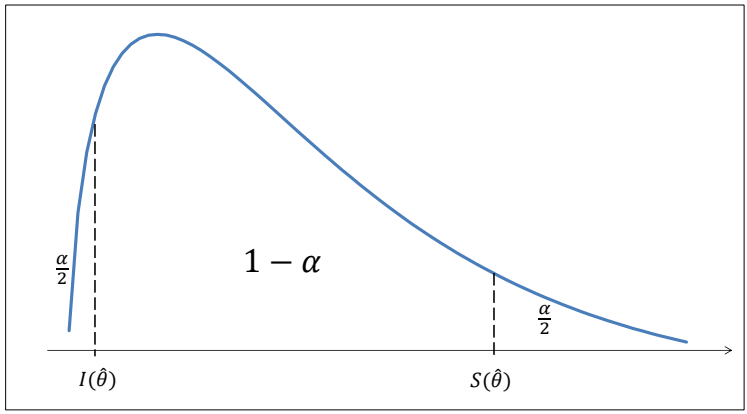
$$\mathbb{P}\left(U_I \leqslant f(\widehat{\theta},\theta) \leqslant U_S\right) = 1-\alpha.$$

Paso 3: despejar el parámetro θ a partir de la cantidad pivotal, de tal manera que la expresión anterior se transforme en:

$$\mathbb{P}\left(I(\widehat{\theta}) \leqslant \theta \leqslant S(\widehat{\theta})\right) = 1-\alpha.$$

Finalmente, el intervalo de confianza para el parámetro θ está dado por $\left[I(\widehat{\theta}), S(\widehat{\theta})\right]$, con una probabilidad o confianza de $1-\alpha$ (véase figura 5.5).

Figura 5.5: Intervalo de confianza para el parámetro θ



Fuente: elaboración propia.

Observación **5.7** Para distribuciones simétricas en torno a cero, se tiene que $U_S = -U_I$.

Ejemplo **5.22** Los siguientes intervalos de confianza se obtienen para los parámetros y pivotes obtenidos en el ejemplo 5.21, respectivamente.

(1) Como se dispone del pivote y este tiene asociada una distribución simétrica en torno a cero, en el paso 2 de la construcción de un intervalo de confianza, la siguiente expresión es igual a $1-\alpha$:

$$\mathbb{P}\left(-z_{1-\frac{\alpha}{2}} \leqslant \frac{\bar{X}-\mu}{\sqrt{\frac{\sigma_0^2}{n}}} \leqslant z_{1-\frac{\alpha}{2}}\right) = \mathbb{P}\left(\bar{X} - z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_0^2}{n}} \leqslant \mu \leqslant \bar{X} + z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_0^2}{n}}\right).$$

Luego, el intervalo de confianza para μ con una confianza de $1-\alpha$ es:

$$\left[\bar{X} - z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\sigma_0^2}{n}}\ ,\ \bar{X} + z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\sigma_0^2}{n}}\right].$$

(2) El pivote sigue una distribución simétrica en torno a cero, entonces la siguiente expresión es igual a $1-\alpha$:

$$\mathbb{P}\left(-t_{1-\frac{\alpha}{2}} \leqslant \frac{\bar{X}-\mu}{\sqrt{\frac{S^2}{n}}} \leqslant t_{1-\frac{\alpha}{2}}\right) = \mathbb{P}\left(\bar{X} - t_{1-\frac{\alpha}{2}}\sqrt{\frac{S^2}{n}} \leqslant \mu \leqslant \bar{X} + t_{1-\frac{\alpha}{2}}\sqrt{\frac{S^2}{n}}\right).$$

Luego, el intervalo de confianza para μ, con una confianza de $1-\alpha$, es:

$$\left[\bar{X} - t_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{S^2}{n}}, \bar{X} + t_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{S^2}{n}}\right].$$

(3) El pivote para σ^2 sigue una distribución que no es simétrica; la construcción del intervalo de confianza $1-\alpha$ es la siguiente:

$$\mathbb{P}\left(\chi^2_{\frac{\alpha}{2}} \leqslant \frac{n-1}{\sigma^2}S^2 \leqslant \chi^2_{1-\frac{\alpha}{2}}\right) = \mathbb{P}\left(\frac{(n-1)\cdot S^2}{\chi^2_{1-\frac{\alpha}{2}}} \leqslant \sigma^2 \leqslant \frac{(n-1)\cdot S^2}{\chi^2_{\frac{\alpha}{2}}}\right).$$

Finalmente, con una confianza de $1-\alpha$, el intervalo para σ^2 es:

$$\left[\frac{(n-1)\cdot S^2}{\chi^2_{1-\frac{\alpha}{2}}}\ ,\ \frac{(n-1)\cdot S^2}{\chi^2_{\frac{\alpha}{2}}}\right].$$

(4) La distribución de la función del pivote es simétrica en torno a cero, entonces:

$$\mathbb{P}\left(-z_{1-\frac{\alpha}{2}} \leqslant \frac{\bar{Y}-\bar{X}-(\mu_2-\mu_1)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \leqslant z_{1-\frac{\alpha}{2}}\right) = 1-\alpha.$$

Despejando el parámetro de interés, $\mu_2 - \mu_1$:

$$\mathbb{P}\left(\bar{Y} - \bar{X} - z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \leqslant \mu \leqslant \bar{Y} - \bar{X} + z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right) = 1 - \alpha.$$

Luego, el intervalo de confianza para $\mu_2 - \mu_1$, con una confianza de $1 - \alpha$, es:

$$\left[\bar{Y} - \bar{X} - z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\ ,\ \bar{Y} - \bar{X} + z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right]. \quad \square$$

Ejemplo **5.23** Con los intervalos de confianza obtenidos en el ejemplo 5.22 y la información empírica que se entrega a continuación, se cuantificarán algunos resultados para los parámetros de interés, suponiendo en todos los casos que el nivel de confianza es de 95 % ($\alpha = 0{,}05$).

(1) Se tiene una muestra aleatoria para la variable X, donde $n = 35$ y $\bar{X} = 12{,}1$. Además, se sabe por estudios anteriores que la varianza de la población es de 1,35 ($\sigma^2 = 1{,}35$), con lo cual se tiene que:

$$\mathbb{P}\left(12{,}1 - 1{,}96 \cdot \sqrt{\frac{1{,}35}{35}} \leqslant \mu \leqslant 12{,}1 + 1{,}96 \cdot \sqrt{\frac{1{,}35}{35}}\right) = 1 - \alpha.$$

Este resultado se tiene del hecho de que $z_{1-\frac{\alpha}{2}} = z_{0{,}975} = 1{,}96$. Luego, el intervalo de confianza para μ está dado por: $[11{,}72\ ;\ 12{,}49]$, con una confianza del 95 %.

(2) Considerando los siguientes datos: $n = 20$, $\bar{x} = 10{,}3$, $s^2 = 0{,}79$ y $\alpha = 0{,}05$, y sabiendo que $t_{0{,}975}(20) = 2{,}093$, se tiene que:

$$\mathbb{P}\left(10{,}3 - 2{,}093 \cdot \sqrt{\frac{0{,}79}{20}} \leqslant \mu \leqslant 10{,}3 + 2{,}093 \cdot \sqrt{\frac{0{,}79}{20}}\right) = 0{,}95.$$

Finalmente, con una confianza del 95 %, $\mu \in [9{,}88\ ;\ 10{,}72]$.

(3) Considerar dos muestras aleatorias, tal que $X \sim N(\mu_1, 100)$ e $Y \sim N(\mu_2, 160)$ independientes, donde $n_1 = 32$ y $n_2 = 40$. Las medias de las muestras son $\bar{Y} = 250$ y $\bar{X} = 190$, luego:

$$\left[250 - 190 - 1{,}96 \cdot \sqrt{\frac{100}{32} + \frac{160}{40}}\ ;\ 250 - 190 + 1{,}96 \cdot \sqrt{\frac{100}{32} + \frac{160}{40}}\right]$$

Finalmente, el intervalo de confianza es: $[54{,}77;\ 65{,}23]$. $\square$

Ejemplo **5.24** A partir del ejemplo 5.9, (a) construir un intervalo de confianza de $1-\alpha$ para el costo real promedio, en función de n_1 y n_2 (cantidad de muestras de arena y concreto, respectivamente), y (b) si la cantidad de arena y concreto que se necesita es de 30 toneladas y 40 libras, respectivamente, interpretar el intervalo de confianza del 95 % para el promedio real en términos del problema.

Solución: (a) Como $\widehat{\theta} = 100 + 7\bar{X} + 3\bar{Y}$, además $\bar{X} \sim \mathrm{N}(\mu_1, \frac{12{,}25}{n_1})$ y $\bar{Y} \sim \mathrm{N}(\mu_2, \frac{0{,}36}{n_2})$. Entonces el pivote es:

$$Z = \frac{100 + 7\bar{X} + 3\bar{Y} - \theta}{\sqrt{\frac{12{,}25}{n_1} + \frac{0{,}36}{n_2}}} \sim \mathrm{N}(0,1).$$

$$\mathbb{P}\left(-z_{1-\frac{\alpha}{2}} \leqslant \frac{100 + 7\bar{X} + 3\bar{Y} - \theta}{\sqrt{\frac{12{,}25}{n_1} + \frac{0{,}36}{n_2}}} \leqslant z_{1-\frac{\alpha}{2}}\right) = 1 - \alpha.$$

$$\mathbb{P}\left(I(\widehat{\theta}) \leqslant \theta \leqslant S(\widehat{\theta})\right) = 1 - \alpha,$$

donde:

$$I(\widehat{\theta}) = 100 + 7\bar{X} + 3\bar{Y} - z_{1-\frac{\alpha}{2}}\sqrt{\frac{12{,}25}{n_1} + \frac{0{,}36}{n_2}}.$$
$$S(\widehat{\theta}) = 100 + 7\bar{X} + 3\bar{Y} + z_{1-\frac{\alpha}{2}}\sqrt{\frac{12{,}25}{n_1} + \frac{0{,}36}{n_2}}.$$

(b) $\sum X_i = 30$, $\sum Y_i = 40$ y $z_{1-\frac{\alpha}{2}} = z_{0{,}975} = 1{,}96$, luego:

$$I(\widehat{\theta}) = 100 + \frac{210}{n_1} + \frac{120}{n_2} - 1{,}96\sqrt{\frac{12{,}25}{n_1} + \frac{0{,}36}{n_2}}.$$
$$S(\widehat{\theta}) = 100 + \frac{210}{n_1} + \frac{120}{n_2} + 1{,}96\sqrt{\frac{12{,}25}{n_1} + \frac{0{,}36}{n_2}}.$$

Finalmente, $\theta \in \left[I(\widehat{\theta}); S(\widehat{\theta})\right]$. □

Intervalos de confianza asintóticos

En caso de que se disponga de una muestra suficientemente grande, es posible construir intervalos de confianza asintóticos para el parámetro de interés. Estos intervalos de confianza se obtienen mediante la construcción de una cantidad pivotal que se aproxima a la distribución normal.

Esta propiedad es bastante útil en casos en que sea complejo obtener una cantidad pivotal y en que se disponga de una muestra aleatoria con un n grande.

Propiedad **5.5** Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde la variable aleatoria X tiene alguna distribución (conocida) que depende de un parámetro θ. Sea $\widehat{\theta}_{MV}$ el estimador de máxima verosimilitud del parámetro θ. Entonces:

$$f(\widehat{\theta}, \theta) = \frac{g(\widehat{\theta}_{MV}) - g(\theta)}{\sqrt{\left.\frac{\left[\frac{\partial}{\partial\theta} g(\theta)\right]^2}{-n\, E\left[\frac{\partial^2}{\partial\theta^2} \log f(x_i)\right]}\right|_{\theta=\widehat{\theta}_{MV}}}} \dot{\sim} N(0,1), \tag{5.9}$$

donde $g(\cdot)$ es una función real valuada y $\dot{\sim}$ es un símbolo que representa una *distribución aproximada*. Es decir, la expresión 5.9 corresponde a una distribución asintótica, por ende, es válida solo si n es grande ($n \to \infty$).

Así, la función $f(\widehat{\theta}, \theta)$ es pivote para θ.

Ejemplo **5.25** Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde la variable aleatoria $X \sim \text{Bernoulli}(\theta)$. Se quieren construir intervalos de confianza para los parámetros θ y $\delta = (1-\theta)\,\theta$.

Solución: Se sabe que la función de cuantía de la variable aleatoria es $f(x_i) = \theta^{x_i} \cdot (1-\theta)^{1-x_i}$, $i = 1, \cdots, n$, a partir de la cual se puede obtener la función de log-verosimilitud y luego el estimador de máxima verosimilitud. En efecto:

$$\ell(\theta|x_i) = \log\theta \cdot \sum_{i=1}^{n} x_i + \log(1-\theta) \cdot (n - \sum_{i=1}^{n} x_i).$$

$$\frac{\partial}{\partial\theta}\ell(\theta|x_i) = \frac{1}{\theta} \cdot \sum_{i=1}^{n} x_i - \frac{1}{1-\theta} \cdot (n - \sum_{i=1}^{n} x_i).$$

$$\widehat{\theta}_{MV} = \bar{X}.$$

(1) Para construir el intervalo de confianza para θ, se define la función $g(\theta) = \theta$. Para obtener el valor esperado que aparece en el denominador de la expresión 5.9, se necesita la segunda derivada del logaritmo de la función de densidad, realizando el cálculo:

$$\frac{\partial^2}{\partial\theta^2}\ell(\theta|x) = -\frac{x}{\theta^2} - \frac{1-x}{(1-\theta)^2}.$$

Considerando que $E[X] = \theta$ y evaluando la expresión en $\theta = \widehat{\theta}_{MV}$:

$$-n \cdot E\left[\frac{\partial^2}{\partial\theta^2}\ell(\theta|x)\right] = \frac{n}{\theta \cdot (1-\theta)}$$
$$-n \cdot E\left[\frac{\partial^2}{\partial\theta^2}\ell(\theta|x)\right]\Big|_{\theta=\widehat{\theta}_{MV}} = \frac{n}{\bar{X} \cdot (1-\bar{X})}.$$

Luego, reemplazando en la expresión 5.9, se obtiene $f(\widehat{\theta}, \theta)$, una función pivotal para θ:

$$f(\widehat{\theta}, \theta) = \frac{\bar{X} - \theta}{\sqrt{\frac{\bar{X} \cdot (1-\bar{X})}{n}}} \dot{\sim} N(0,1).$$

Utilizando $f(\widehat{\theta}, \theta)$, se construye el intervalo de confianza:

$$\mathbb{P}\left(-z_{1-\frac{\alpha}{2}} \leqslant \frac{\bar{X} - \theta}{\sqrt{\frac{\bar{X} \cdot (1-\bar{X})}{n}}} \leqslant z_{1-\frac{\alpha}{2}}\right) \dot{\approx} 1 - \alpha.$$

$$\mathbb{P}\left(\bar{X} - z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\bar{X} \cdot (1-\bar{X})}{n}} \leqslant \theta \leqslant \bar{X} + z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\bar{X} \cdot (1-\bar{X})}{n}}\right) \dot{\approx} 1-\alpha.$$

Entonces, el intervalo en que se encuentra el parámetro θ, con una confianza aproximada de $100 \cdot (1-\alpha)\,\%$, es:

$$\left[\bar{X} - z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\bar{X} \cdot (1-\bar{X})}{n}}\ ,\ \bar{X} + z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\bar{X} \cdot (1-\bar{X})}{n}}\right].$$

(2) Para construir el intervalo de confianza para el parámetro δ, se considera que $g(\theta) = \delta(\theta) = \theta \cdot (1-\theta)$. El estimador de máxima verosimilitud de δ, $\widehat{\delta}_{MV}$ se obtiene considerando la propiedad 5.1 de invarianza de estimadores obtenidos con este método:

$$\widehat{\delta}_{MV} = g(\widehat{\theta}_{MV}) = \widehat{\theta}_{MV} \cdot (1 - \widehat{\theta}_{MV}) = \bar{X} \cdot (1 - \bar{X}).$$

El pivote para el parámetro δ se puede obtener desde la expresión 5.9, para lo cual se necesita obtener una expresión para:

$$\frac{\partial}{\partial\theta} g(\theta) = 1 - 2\,\theta \Rightarrow \frac{\partial}{\partial\theta} g(\theta)\Big|_{\theta=\widehat{\theta}_{MV}} = 1 - 2\bar{X}.$$

Con este resultado, más los obtenidos en (1), se tiene que la función pivotal para δ: $f(\widehat{\delta}, \delta)$ está dada por:

$$f(\widehat{\delta}, \delta) = \frac{\bar{X}(1-\bar{X}) - \delta}{\sqrt{\frac{(1-2\bar{X})^2 \bar{X}(1-\bar{X})}{n}}} \dot{\sim} N(0,1).$$

Definiendo $h(\bar{X}) = \sqrt{\frac{(1-2\bar{X})^2 \bar{X}(1-\bar{X})}{n}}$. Finalmente, se tiene:

$$\mathbb{P}\left[\bar{X}(1-\bar{X}) - z_{1-\frac{\alpha}{2}} \cdot h(\bar{X}) \leqslant \delta \leqslant \bar{X}(1-\bar{X}) + z_{1-\frac{\alpha}{2}} \cdot h(\bar{X})\right] \dot{\approx} 1-\alpha.$$

Entonces, el intervalo en que se encuentra el parámetro δ, con una confianza aproximada de $100 \cdot (1-\alpha)\,\%$, es $\theta \in \left[I(\widehat{\theta}); S(\widehat{\theta})\right]$, donde:

$$\begin{aligned} I(\widehat{\theta}) =& \bar{X}(1-\bar{X}) - z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{(1-2\bar{X})^2 \bar{X} \cdot (1-\bar{X})}{n}} \\ S(\widehat{\theta}) =& \bar{X}(1-\bar{X}) + z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{(1-\bar{X})^2 \bar{X} \cdot (1-\bar{X})}{n}}. \quad \square \end{aligned}$$

Ejercicio propuesto **5.3** Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria con $X \sim \text{Exp}(\alpha)$, considerar los estimadores de máxima verosimilitud de los parámetros $\delta_1 = E[X]$ y $\delta_2 = F(x, \alpha) = \mathbb{P}(X \leqslant x)$. Obtener intervalos de confianza aproximados para estos dos parámetros.

Teorema central del límite

El teorema central del límite es de gran importancia para el desarrollo de la estadística inferencial, pues, a partir de una muestra aleatoria con cualquier distribución, permite obtener una nueva variable aleatoria mediante una transformación lineal que siga una distribución normal estándar. Ya hemos visto las ventajas que presenta la distribución normal estándar para el cálculo de probabilidades y para la obtención de intervalos de confianza.

Propiedad **5.6** Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria desde una población con media $E[X]$ y varianza $Var[X]$. El teorema central del límite (TCL) muestra que Z_n tiene distribución asintótica normal con media 0 y varianza 1. Es decir, $Z_n \dot{\sim} N(0,1)$ cuando $n \to \infty$, donde:

$$Z_n = \frac{\bar{X} - E[X]}{\sqrt{\frac{Var[X]}{n}}} \sim \text{N}(0,1).$$

Una forma alternativa de expresar Z_n es:

$$Z_n = \frac{\sum_{i=1}^{n} X_i - E\left[\sum_{i=1}^{n} X_i\right]}{\sqrt{Var\left[\sum_{i=1}^{n} X_i\right]}}.$$

Ejemplo **5.26** Sean $Y_1, Y_2, \cdots, Y_{10\,000}$ una muestra aleatoria donde $Y_i \sim$ Bernoulli$(p = 0{,}20)$. Se desea calcular la probabilidad:

$$\mathbb{P}\left(1980 \leqslant \sum_{i=1}^{10\,000} X_i \leqslant 2050\right).$$

Solución: Si se define $Y = \sum_{i=1}^{10\,000} X_i$, entonces $Y \sim$ Binomial$(n = 10\,000; p = 0{,}20)$. Desde este último resultado se conoce el valor esperado y la varianza de Y: $E[Y] = 10\,000 \cdot 0{,}20 = 2000$ y $Var[Y] = 10\,000 \cdot 0{,}20 \cdot 0{,}80 = 1600$.

Utilizando el teorema central del límite:

$$\mathbb{P}\left(\frac{1980 - 2000}{\sqrt{1600}} \leqslant Z_n \leqslant \frac{2050 - 2000}{\sqrt{1600}}\right) = \mathbb{P}\left(-0{,}50 \leqslant Z_n \leqslant 1{,}25\right).$$

Finalmente, utilizando el hecho de que $Z_n \dot{\sim} N(0,1)$:

$$\mathbb{P}\left(-0{,}50 \leqslant Z_n \leqslant 1{,}25\right) = \Phi(1{,}25) - \Phi(-0{,}50) = 0{,}5859. \quad \square$$

Observación **5.8** Nótese que para aplicar el TCL es suficiente conocer el valor esperado y la varianza. En el ejemplo 5.26, si bien se conoce la distribución de la variable aleatoria y con esto se puede obtener directamente la probabilidad pedida mediante $F_X(x)$, para efectos de ejemplificar la propiedad 5.6, se resuelve mediante la aproximación asintótica a la distribución normal.

Ejemplo **5.27** Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde $X_i \sim$ Bernoulli(θ) y el tamaño de la muestra es suficientemente grande para que se cumplan las condiciones del TCL, se sabe, a partir del ejemplo 5.25, que un estimador para θ es $\widehat{\theta}$.

(a) Asumiendo el valor máximo para la varianza, obtener un intervalo de confianza para θ.

(b) Calcular el tamaño de muestra mínimo para que el ancho del intervalo no sea superior a δ y con un nivel de confianza de $1 - \alpha$.

Solución: (a) Según los datos entregados, $X_i \sim \text{Bernoulli}(\theta) \Rightarrow E[X] = \theta$ y $Var[X] = \theta(1-\theta)$. El valor máximo para la varianza se obtiene por:

$$Var[X] = h(\theta) = \theta(1-\theta) \Rightarrow \frac{d}{d\theta}h(\theta) = 1 - 2\theta = 0 \Rightarrow \theta = 0{,}5.$$

Con lo cual, usando el TCL, se obtiene un pivote para θ:

$$\bar{X} \sim N\left(\theta, \frac{0{,}5^2}{n}\right) \Rightarrow f(\widehat{\theta}, \theta) = \frac{\bar{X} - \theta}{\sqrt{\frac{0{,}5^2}{n}}} \sim N(0,1).$$

Con lo cual el intervalo de confianza es:

$$\mathbb{P}\left(-z_{1-\frac{\alpha}{2}} \leqslant f(\widehat{\theta}, \theta) \leqslant z_{1-\frac{\alpha}{2}}\right) = 1 - \alpha.$$

$$\theta \in \left[-z_{1-\frac{\alpha}{2}}\frac{0{,}5}{\sqrt{n}} + \bar{X}, z_{1-\frac{\alpha}{2}}\frac{0{,}5}{\sqrt{n}} + \bar{X}\right].$$

(b) Desde la parte (a), aplicando las condiciones dadas para el intervalo de confianza, se tiene:

$$z_{1-\frac{\alpha}{2}}\frac{0{,}5}{\sqrt{n}} + \cancel{\bar{X}} + z_{1-\frac{\alpha}{2}}\frac{0{,}5}{\sqrt{n}} - \cancel{\bar{X}} \leqslant \delta \Rightarrow 2 \cdot z_{1-\frac{\alpha}{2}}\frac{0{,}5}{\sqrt{n}} \leqslant \delta.$$

Por lo tanto, $n \geqslant \delta^{-1} \cdot z_{1-\frac{\alpha}{2}}$. □

5.6. Ejercicios resueltos

1. La muestra aleatoria $X_1, X_2, \cdots, X_n$ fue obtenida de una población con $X_i \sim \text{Gamma}(\alpha, \beta)$. Bajo el supuesto de que α es conocido, encontrar el estimador de máxima verosimilitud del parámetro β.

   Solución: Para encontrar el estimador mediante el método de máxima verosimilitud, se obtiene la función de verosimilitud (L) y la de

log-verosimilitud (ℓ):

$$\begin{aligned} L &= \prod_{i=1}^{n} \frac{1}{\beta^{\alpha}\Gamma(\alpha)} \cdot x_i^{\alpha-1} \cdot \exp\left\{\frac{-x_i}{\beta}\right\} \\ &= \left(\frac{1}{\beta^{\alpha}\Gamma(\alpha)}\right)^{n} \left(\prod_{i=1}^{n} x_i\right)^{\alpha-1} \exp\left\{-\frac{1}{\beta}\sum_{i=1}^{n} x_i\right\}. \\ \ell &= -n\log\left(\beta^{\alpha}\Gamma(\alpha)\right) - \frac{1}{\beta}\sum x_i + (\alpha-1)\log\left(\prod_{i=1}^{n} x_i\right) \\ &= -n\alpha\log\beta - n\log\Gamma(\alpha) - \frac{1}{\beta}\sum x_i + (\alpha-1)\log\left(\prod_{i=1}^{n} x_i\right). \end{aligned}$$

Realizando $\frac{\partial}{\partial\beta}\ell$ e igualándola a 0, se obtiene el estimador de máxima verosimilitud de β:

$$\widehat{\beta}_{MV} = \frac{1}{\alpha}\cdot\frac{1}{n}\cdot\sum X_i = \frac{1}{\alpha}\overline{X}.$$

2. El tiempo de falla (en horas) de un instrumento electrónico es una variable aleatoria T con función de densidad $f(t) = \theta\exp\{-\theta(t-3)\}$, $t > 3$. Se prueban n artículos seleccionados al azar y se anotan sus tiempos de falla $T_1, T_2, \cdots, T_n$. Encontrar el estimador de máxima verosimilitud del parámetro desconocido θ utilizando la muestra aleatoria.

   Solución: Para realizar la estimación solicitada se requiere de la función de verosimilitud, para luego maximizarla y obtener los candidatos a valor extremo (en este caso máximo):

$$\begin{aligned} L(\theta) =& \prod_{j=1}^{n} \theta\exp\{-\theta(t_j-3)\} \\ =& \theta^n \exp\left\{-\sum_{j=1}^{n}\theta(t_j-3)\right\} \\ =& \theta^n \exp\left\{-\theta\left(\sum_{j=1}^{n} t_j - 3n\right)\right\}. \\ \ell(\theta) =& n\log(\theta) - \theta\left(\sum_{j=1}^{n} t_j - 3n\right). \end{aligned}$$

Luego, la derivada parcial de $\ell(\theta)$ respecto a θ es:

$$\frac{\partial}{\partial\theta}\ell(\theta) = \frac{n}{\theta} - \left(\sum_{j=1}^{n} t_j - 3n\right).$$

Finalmente, igualando la derivada parcial a cero y evaluando en la segunda derivada de la función de log-verosimilitud respecto al parámetro:

$$\widehat{\theta}_{MV} = n \cdot \left(\sum_{j=1}^{n} t_j - 3n\right)^{-1}.$$

3. Sean $X_1, X_2, \cdots, X_{n_1}$ una muestra aleatoria, donde $X_i \sim N(\mu, k\sigma^2)$ y k es una constante positiva. Además, Sean $Y_1, Y_2, \cdots, Y_{n_2}$ una muestra aleatoria (ambas independientes), donde $Y_i \sim N(\mu, \sigma^2)$. Determinar el estimador de máxima verosimilitud para μ y σ^2 utilizando las $n_1 + n_2$ observaciones.

   Solución: Si se utilizan las $n_1 + n_2$ observaciones, la función de verosimilitud es igual al producto de la función de verosimilitud de las n_1 y las n_2 observaciones. Entonces:

$$\begin{aligned}
L(\mu,\sigma^2) =& \prod_{j=1}^{n} f_{X_j}(x_j) \cdot \prod_{j=1}^{n} f_{Y_j}(y_j) \\
=& \sqrt{2\pi k\sigma^2}^{-n_1} \exp\left\{-\frac{1}{2k\sigma^2}\sum_{j=1}^{n_1}(x_j-\mu)^2\right\} \cdot \sqrt{2\pi\sigma^2}^{-n_2} \\
& \cdot \exp\left\{-\frac{1}{2\sigma^2}\sum_{j=1}^{n_2}(y_j-\mu)^2\right\} \\
\ell(\mu,\sigma^2) =& -\frac{n_1}{2}\log(2\pi k\sigma^2) - \frac{1}{2k\sigma^2}\sum_{j=1}^{n_1}(x_j-\mu)^2 - \\
& -\frac{-n_2}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{j=1}^{n_2}(y_j-\mu)^2.
\end{aligned}$$

Luego, las derivadas parciales de la función de log-verosimilitud respecto de cada parámetro son:

$$\frac{d}{d\mu}\ell = \frac{1}{k\sigma^2}\sum_{j=1}^{n_1}(x_j-\mu) + \frac{1}{\sigma^2}\sum_{j=1}^{n_2}(y_j-\mu).$$
$$\frac{d}{d\sigma^2}\ell = \frac{1}{2k(\sigma^2)^2}\sum_{j=1}^{n_1}(x_j-\mu)^2 + \frac{1}{2(\sigma^2)^2}\sum_{j=1}^{n_2}(y_j-\mu)^2 - (n_1+n_2)\frac{1}{2\sigma^2}.$$

Finalmente, igualando las derivadas parciales a cero y resolviendo el sistema de dos ecuaciones, se obtienen los estimadores de máxima verosimilitud para los parámetros pedidos. Los cuales son, respectivamente:

$$\widehat{\mu}_{MV} = \left(\frac{n_1}{k} - n_2\right)^{-1}\left(\frac{1}{k}\sum_{j=1}^{n_1}x_j - \sum_{j=1}^{n_2}y_j\right).$$
$$\widehat{\sigma}^2_{MV} = (n_1+n_2)^{-1}\sum_{j=1}^{n_1}(x_j-\widehat{\mu}_{MV})^2(k^{-1}-k^{-2}).$$

4. Se obtienen 200 observaciones utilizando una muestra aleatoria. El resumen de estas observaciones es $\sum_{i=1}^{200} X_i = 300$ y $\sum_{i=1}^{200} X_i^2 = 3754$. Utilizando estos valores, obtener valores numéricos de estimadores de momentos de la media y la varianza poblacional.

   Solución: A través de la definición del método para obtener estimadores de momentos:

$$E[X_1] = \frac{1}{200}\sum_{i=1}^{200}X_i \Rightarrow \widehat{\mu}_{MV} = E[X_1] = \frac{300}{200} = \frac{3}{2}$$
$$E[X_1^2] = \frac{1}{200}\sum_{i=1}^{200}X_i^2 = \frac{3754}{200} = 18{,}77.$$

   Finalmente, $\widehat{\sigma^2}_{MV} = VarX_1 = E[X_1^2] - (E[X_1])^2 = 16{,}52$.

5. Sean $X_1, X_2, \cdots, X_n$ e $Y_1, Y_2, \cdots, Y_n$ muestras aleatorias independientes, con $X_i \sim$ Poisson(λ) e $Y_i \sim$ Poisson($\lambda(1-\theta)$). Se pide:

   (a) Escribir la función de verosimilitud para $X_1, X_2, \cdots, X_n$ y para $Y_1, Y_2, \cdots, Y_n$ en forma conjunta (para las $2n$ observaciones).

   (b) Encontrar los estimadores de máxima verosimilitud para los parámetros λ y θ.

Solución: (a) Las funciones de verosimilitud y log-verosimilitud en forma conjunta para las $2n$ observaciones son:

$$\begin{aligned} L(\lambda,\theta;\boldsymbol{x},\boldsymbol{y}) =& \prod_{i=1}^{n} f(x_i|\lambda)\cdot\prod_{i=1}^{n} f(y_i|\lambda,\theta) \\ =& e^{-n\lambda(2-\theta)}\lambda^{\sum_{i=1}^{n}x_i}(\lambda(1-\theta))^{\sum_{i=1}^{n}y_i}\left(\prod_{i=1}^{n}x_i!y_i!\right)^{-1} \\ \ell(\lambda,\theta) =& -n\lambda(2-\theta)+\sum_{i=1}^{n}x_i\log\lambda+\log(\lambda(1-\theta))\sum_{i=1}^{n}y_i-c. \end{aligned}$$

(b) Los estimadores de máxima verosimilitud se obtienen de las derivadas parciales de la función de log-verosimilitud:

$$\begin{aligned} \frac{\partial}{\partial\lambda}\ell(\lambda,\theta) =& n(\theta-2)+\frac{1}{\lambda}\left(\sum_{i=1}^{n}x_i+\sum_{i=1}^{n}y_i\right) \\ \frac{\partial}{\partial\theta}\ell(\lambda,\theta) =& n\lambda-\frac{1}{1-\theta}\sum_{i=1}^{n}y_i. \end{aligned}$$

Igualando a cero ambas derivadas parciales, se tiene un sistema de ecuaciones del que es posible obtener los estimadores pedidos. En efecto, los estimadores son:

$$\begin{aligned} \widehat{\lambda}_{MV} =& \frac{1}{n}\sum_{i=1}^{n}x_i=\bar{X}. \\ \widehat{\theta}_{MV} =& 1-\sum_{i=1}^{n}y_i\cdot\left(\sum_{i=1}^{n}x_i\right)^{-1}. \end{aligned}$$

6. Las variables aleatorias $Y_1, Y_2, \cdots, Y_n$ son independientes, con función de densidad:

$$f(y_i)=\frac{1}{\beta x_i}\exp\left\{-\frac{1}{\beta x_i}y_i\right\},\ y_i>0,\ (x_i>0 \text{ constantes}, \beta>0).$$

   (a) Obtener una expresión para el estimador de máxima verosimilitud para el parámetro β.

   (b) Determinar el valor esperado y la varianza del estimador obtenido en (a).

Solución: (a) La función de verosimilitud (y de log-verosimilitud) y su posterior maximización está dada por:

$$\begin{aligned} L(\beta) =& \prod_{i=1}^{n} \left((\beta\, x_i)^{-1} \exp\left\{ -(\beta\, x_i)^{-1} y_i \right\} \right) \\ =& \beta^{-n} \prod_{i=1}^{n} (x_i)^{-1} \exp\left\{ -\beta^{-1} \sum_{i=1}^{n} (x_i)^{-1} y_i \right\}. \\ \ell(\beta) =& -n \log\beta + \log\left(\prod_{i=1}^{n} (x_i)^{-1} \right) - \beta^{-1} \sum_{i=1}^{n} (x_i)^{-1} y_i. \\ \frac{\partial}{\partial\beta}\ell(\beta) =& -n\beta^{-1} + \beta^{-2} \sum_{i=1}^{n} (x_i)^{-1} y_i. \end{aligned}$$

Igualando la última expresión a cero, se tiene el estimador pedido:

$$\widehat{\beta}_{MV} = \frac{1}{n} \sum_{i=1}^{n} \frac{y_i}{x_i}.$$

(b) El valor esperado y la varianza del estimador son:

$$E\left[\widehat{\beta}_{MV}\right] = E\left[\frac{1}{n} \sum_{i=1}^{n} x_i^{-1} Y_i \right] = \frac{1}{n} \sum_{i=1}^{n} x_i^{-1} E\left[Y_i\right].$$

Nótese que $Y_i \sim \text{Exp}(\beta x_i)$, entonces $E\left[Y_i\right] = \beta x_i$, luego:

$$\begin{aligned} =& \frac{1}{n} \sum_{i=1}^{n} \cancel{x_i^{-1}} \beta \cancel{x_i} = \beta. \\ Var[\widehat{\beta}_{MV}] =& Var\left[\frac{1}{n} \sum_{i=1}^{n} x_i^{-1} Y_i \right] = \frac{1}{n^2} \sum_{i=1}^{n} (x_i^{-1})^2 Var\left[Y_i\right] \end{aligned}$$

Como $Y_i \sim \text{Exp}(\beta x_i)$, entonces $Var\left[Y_i\right] = \beta^2 x_i^2$, luego:

$$= \frac{1}{n^2} \sum_{i=1}^{n} \cancel{(x_i^{-1})^2} \beta^2 \cancel{x_i^2} = \frac{1}{n} \beta^2.$$

7. Sean cuatro poblaciones ($i = 1, 2, 3, 4$), donde en cada una de ellas se obtienen muestras aleatorias independientes $X_{i1}, X_{i2}, \cdots, X_{in_i}$ de tamaño n_i. En la i-ésima población, se asume que $X_{ij} \sim N(\mu_i, \sigma^2)$, $i = 1, 2, 3, 4$, $j = 1, 2, \cdots, n_i$. Además, se asume que las poblaciones X_{ij} constituyen un conjunto de $4n$ variables independientes. Se pide:

(a) Considerando solo la población 2, encontrar el estimador de máxima verosimilitud para μ_2.

(b) Encontrar el estimador de máxima verosimilitud para el parámetro $\theta = \mu_1 - \mu_2 + \mu_3$.

Solución: (a) Considerando $i = 2$, se tiene la muestra aleatoria $X_{21}, X_{22}, \cdots, X_{2n_2}$, con $X_{2j} \sim N(\mu_2, \sigma^2)$. Desde esta definición, los estimadores de máxima verosimilitud han sido obtenidos en el ejercicio 5.6 y son:

$$\widehat{\boldsymbol{\theta}}_{MV} = (\widehat{\mu_2}, \widehat{\sigma}_2^2)^T = \left(\bar{x}_2, \frac{1}{n_2} \sum_{i=1}^{n_2} (x_{2i} - \bar{x}_2)^2 \right)^T .$$

(b) Sea $W = X_{1j} - X_{2j} + X_{3j}$, dada la distribución de X_{ij}, se tiene que:

$$W \sim N(\mu_1 - \mu_2 + \mu_3, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} + \frac{\sigma_3^2}{n_3}).$$

Luego, el estimador pedido es $\widehat{\theta}_{MV} = \bar{x}_1 - \bar{x}_2 + \bar{x}_3$.

8. Sean $X_1, X_2, ..., X_n$ una muestra aleatoria, donde la función de densidad de X_i es:

$$f(x_i) = 2\theta^{-2} x_i \exp\left\{ -\frac{x_i^2}{\theta^2} \right\}, \; x_i > 0, \theta > 0$$

(a) Encontrar el estimador de máxima verosimilitud para θ^2.

(b) Determinar si el estimador encontrado en (a) es insesgado y consistente.

(c) Encontrar un intervalo de confianza aproximado para el parámetro θ^2.

Solución: (a) El estimador se puede obtener maximizando la fun-

ción de verosimilitud o de log-verosimilitud:

$$\begin{aligned} L\left(\theta^2\right) =& \prod 2\theta^{-2} x_i \exp\left\{-X_i^2\theta^{-2}\right\} \\ =& \left(2\theta^{-2}\right)^n \exp\left\{-\theta^{-2}\sum X_i^2\right\} \prod X_i. \\ \ell(\theta^2) =& n\log\left(2\theta^{-2}\right) - \theta^{-2}\sum X_i^2 + \log\left(\prod X_i\right) \\ =& n\log 2 - n\log\left(\theta^2\right) - \theta^{-2}\cdot\sum X_i^2 + \log\left(\prod X_i\right). \end{aligned}$$

Realizando $\dfrac{\partial \ell}{\partial \theta^2} = 0\Big|_{\theta^2=\widehat{\theta}^2_{MV}}$:

$$-\frac{n}{\widehat{\theta}^2_{MV}} + \frac{1}{\left(\widehat{\theta}^2_{MV}\right)^2}\cdot\sum X_i^2 = 0$$

$$\frac{1}{\left(\widehat{\theta}^2_{MV}\right)^2}\cdot\sum X_i^2 = \frac{n}{\widehat{\theta}^2_{MV}}$$

$$\widehat{\theta}^2_{MV} = \frac{1}{n}\sum X_i^2.$$

(b) Para estudiar si el estimador es insesgado, se debe tener que la esperanza del estimador de máxima verosimilitud es igual al parámetro, esto es:

$$E\left[\widehat{\theta}^2_{MV}\right] = E\left[\frac{1}{n}\sum X_i^2\right] = E\left[X^2\right].$$

Para calcular este último resultado, se puede hacer uso de la definición de esperanza de una función de variable aleatoria:

$$\begin{aligned} E\left[X^2\right] =& \int_0^\infty x_i^2 \cdot 2\theta^{-2} x_i \exp\left\{-x_i^2\theta^{-2}\right\}\, dx_i \\ =& \int_0^\infty 2\theta^{-2} x_i^3 \exp\left\{-x_i^2\theta^{-2}\right\}\, dx_i \\ =& \theta^2 \int_0^\infty u e^{-u} du = \theta^2, \text{ con } u = x_i^2\theta^{-2}. \end{aligned}$$

De esta manera, se tiene que el estimador es insesgado.

Por su parte, para analizar la consistencia, debemos calcular la varianza del estimador y analizar su límite para un número muy grande de n, es decir, $\lim_{n\to\infty} Var\left[\theta^2_{MV}\right]$. Si este límite es cero, entonces el estimador es consistente.

$$Var\left[\widehat{\theta}^2_{MV}\right] = \frac{1}{n^2} Var\left[\sum X_i^2\right] = \frac{\not{n}}{n^{\not{2}}} Var[X_i^2] = \frac{1}{n} Var[X_i^2] \quad (5.10)$$

$$= \frac{1}{n}\left(E\left[X_i^4\right] - (E[X_i^2])^2\right). \quad (5.11)$$

Para calcular $E\left[X_i^4\right]$, nuevamente se aplica la definición y se hace uso de integrales.

$$\begin{aligned} E\left[X_i^4\right] &= \int_0^\infty x_i^4 \cdot 2\theta^{-2} x_i \exp\left\{-\frac{x_i^2}{\theta^2}\right\} dx_i \\ &= \int_0^\infty 2\theta^{-2} x_i^5 \exp\left\{-\frac{x_i^2}{\theta^2}\right\} dx_i \\ &= \int_0^\infty u_i^2 \theta^4 e^{-u_i} du_i, \ \text{ con } u_i = \frac{x_i^2}{\theta^2} \\ &= \theta^4 \Gamma(3) = 2\theta^4. \end{aligned}$$

Reemplazando en la ecuación 5.11:

$$\begin{aligned} Var\left[\widehat{\theta}^2_{MV}\right] &= \frac{1}{n}\left(E\left[X_i^4\right] - (E\left[X_i^2\right])^2\right) \\ &= \frac{1}{n}\left(2\theta^4 - \theta^4\right) \\ &= \frac{1}{n}\theta^4. \end{aligned}$$

Finalmente:

$$\lim_{n\to\infty} Var\left[\widehat{\boldsymbol{\theta}}^2_{MV}\right] = \lim_{n\to\infty} \frac{\theta^4}{n} = 0.$$

Esto prueba que el estimador es consistente.

(c) Para la distribución aproximada, $g(\theta^2) = \theta^2$ y la segunda derivada de $g(\theta)$ respecto a θ^2 es $4\theta^2$. Así, la función pivotal es:

$$f(\widehat{\theta}^2_{MV}, \theta^2) = \frac{\widehat{\theta}^2_{MV} - \theta^2}{\sqrt{\frac{(\widehat{\theta}^2_{MV})^2}{n}}} = \frac{\widehat{\theta}^2_{MV} - \theta^2}{\frac{(\widehat{\theta}_{MV})^2}{\sqrt{n}}} \sim N(0,1)$$

Luego, como la distribución del pivote es simétrica, se tiene:

$$\mathbb{P}\left(-z_{1-\frac{\alpha}{2}} \leqslant f(\widehat{\theta}^2_{MV}, \theta^2) \leqslant z_{1-\frac{\alpha}{2}}\right) = 1-\alpha.$$

Finalmente, despejando el parámetro θ^2, se obtiene que el intervalo de confianza es:

$$\theta^2 \in \left[\widehat{\theta}^2_{MV} - \frac{\widehat{\theta}^2_{MV}}{\sqrt{n}} z_{1-\frac{\alpha}{2}}; \widehat{\theta}^2_{MV} + \frac{\widehat{\theta}^2_{MV}}{\sqrt{n}} z_{1-\frac{\alpha}{2}}\right].$$

9. Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria. Encontrar el intervalo de confianza aproximado de $(1-\alpha)\cdot 100\,\%$ para el parámetro θ, cuando la función de densidad de X en la población es:

   (a) $f(x) = \theta^x(1-\theta)^{1-x}$, $x = 0, 1$.

   (b) $f(x) = \theta c^\theta x^{-(\theta+1)}$, $x > c$, donde c es una constante positiva conocida.

   (c) $f(x) = x\theta^{-2}\exp\left\{-0,5\theta^{-2}x^2\right\}$, $x > 0$.

   (d) $f(x) = \theta c x^{c-1}\exp\{-\theta x^c\}$, $x > c$, donde c es una constante positiva conocida.

   Solución: Para obtener un intervalo aproximado para el parámetro θ se necesita el estimador de máxima verosimilitud en cada caso. Si se maximiza la función de verosimilitud respecto de θ, se tiene que los estimadores para (a), (b), (c) y (d) son, respectivamente:

$$\bar{X};\ n\left(n\log(c) - \sum \log(x_i)\right)^{-1};\ \sqrt{\frac{1}{2}\sum x_i^2};\ n\left(\sum x_i^c\right)^{-1}$$

   Luego, la función pivotal $f(\widehat{\theta}, \theta)$ requiere de la función $g(\theta)$, que en los cuatro casos es $g(\theta) = \theta$. Obteniendo los resultados que requiere la propiedad 5.5 para obtener el pivote y aplicar la igualdad:

$$\mathbb{P}(-z_{1-\frac{\alpha}{2}} \leqslant f(\widehat{\theta}, \theta) \leqslant z_{1-\frac{\alpha}{2}}) = 1-\alpha,$$

   con lo cual los intervalos de confianza aproximados para θ son:

   (a) $\left[\bar{x} - z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{x}(1-\bar{x})}{n}}; \bar{x} + z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{x}(1-\bar{x})}{n}}\right]$.

(b) $\left[\widehat{\theta}_b - z_{1-\frac{\alpha}{2}} R_b; \widehat{\theta}_b + z_{1-\frac{\alpha}{2}} R_b\right]$,
donde $R_b = \sqrt{n^{-1}\left(\widehat{\theta}_b^{-2} - \log(c\widehat{\theta}_b(\widehat{\theta}_b - 1)^{-1})\right)}$.

(c) $\left[\widehat{\theta}_c - z_{1-\frac{\alpha}{2}} R_c; \widehat{\theta}_c + z_{1-\frac{\alpha}{2}} R_c\right]$, donde $R_c = \sqrt{\widehat{\theta}_c \left(2n(\pi - \theta^{-1})\right)^{-1}}$.

(d) $\left[\widehat{\theta}_d \left(1 - \sqrt{n^{-1}} z_{1-\frac{\alpha}{2}}\right); \widehat{\theta}_d \left(1 + \sqrt{n^{-1}} z_{1-\frac{\alpha}{2}}\right)\right]$, donde $\widehat{\theta}_j$ $(j = a, b, c, d)$ es el estimador de máxima verosimilitud de θ obtenido anteriormente para (a), (b), (c) y (d).

10. Sea la muestra aleatoria $X_1, X_2, \cdots, X_n$, donde la función de densidad de X está dada por:

$$f(x) = \frac{1}{\beta} x \exp\left\{-\frac{1}{2\beta} x^2\right\},\ x > 0,\ (\beta > 0).$$

(a) Mostrar que el estimador de máxima verosimilitud de β es $\widehat{\beta}_{MV} = \frac{1}{2n}\sum X_i^2$. Verificar la consistencia de $\widehat{\beta}_{MV}$.

(b) Mostrar que $X^2 \sim \text{Exp}(2\beta)$ y, a partir de este resultado, mostrar además que $\sum X_i^2 \sim \text{Gamma}(n, 2\beta)$.

(c) Utilizando (b), mostrar que $\beta^{-1}\sum X_i^2 \sim \chi^2(2n)$.

(d) Construir una función pivotal para β.

(e) Hallar un intervalo de confianza para β con un nivel de confianza de $1 - \alpha$.

Solución: (a) Estimador de máxima verosimilitud de β:

$$\ell(\beta, x) = -n\log\beta - \frac{1}{2\beta}\sum x_i^2 + \log\left(\prod x_i\right)$$
$$\frac{\partial}{\partial}\ell = -\frac{n}{\beta} + \frac{1}{2\beta^2}\sum x_i^2 = 0.$$

A partir de este último resultado se tiene que el estimador es:

$$\widehat{\beta}_{MV} = \frac{1}{2n}\sum x_i^2.$$

Para verificar la consistencia de $\widehat{\beta}_{MV}$, utilizar el resultado de (c) (queda de ejercicio para el lector).

(b) Para mostrar que $X^2 \sim \text{Exp}(2\beta)$, se puede utilizar la función de distribución, entonces:

$$\begin{aligned} F_{X^2}(t) =& \mathbb{P}\left(X^2 \leqslant t\right) = \mathbb{P}\left(-\sqrt{t} \leqslant X \leqslant \sqrt{t}\right) \\ =& F_X\left(\sqrt{t}\right) - F_X\left(-\sqrt{t}\right). \end{aligned}$$

Derivando la expresión respecto a t, se obtiene la función de densidad:

$$\begin{aligned} f_{X^2}(t) =& \frac{1}{2\sqrt{t}} f_X\left(\sqrt{t}\right) + \frac{1}{2\sqrt{t}} f_X\left(-\sqrt{t}\right) \\ =& \frac{1}{2\sqrt{t}} \frac{1}{\beta} \sqrt{t} \exp\left\{-\frac{1}{2\beta} t\right\} \\ =& \frac{1}{2\beta} \exp\left\{-\frac{1}{2\beta} t\right\}, \quad t > 0. \end{aligned}$$

La última expresión corresponde a la función de densidad de una distribución exponencial de parámetro 2β. Usando la función generadora de momentos y que las variables X_i tienen todas la misma distribución, entonces:

$$M_{\sum X_i^2}(t) = (1 - 2\beta t)^{-n},$$

con lo cual, efectivamente, $\sum X_i^2 \sim \text{Gamma}(n, 2\beta)$.

(c) Siendo $T = \beta^{-1} \sum X_i^2$, luego utilizando la fgm:

$$M_T(t) = M_{\sum X_i^2}(\beta^{-1} t) = (1 - 2t)^{-n} \Rightarrow T \sim \chi^2(2n).$$

(d) El pivote se obtiene directamente de (c):

$$f(\beta, \widehat{\beta}) = \beta^{-1} \sum X_i^2 \sim \chi^2(2n).$$

(e) Usando el pivote anterior, el intervalo de confianza para β con un nivel de confianza de $1 - \alpha$ es:

$$\mathbb{P}\left(\chi^2_{\frac{\alpha}{2}} \leqslant \beta^{-1} \sum X_i^2 \leqslant \chi^2_{1-\frac{\alpha}{2}}\right) = \mathbb{P}\left(\frac{\sum X_i^2}{\chi^2_{1-\frac{\alpha}{2}}} \leqslant \beta \leqslant \frac{\sum X_i^2}{\chi^2_{\frac{\alpha}{2}}}\right) = 1 - \alpha.$$

11. En un taller automotriz se ha tomado una muestra aleatoria de 100 vehículos, para los cuales se midió el tiempo promedio de reparación, que resultó de 4 días. La cuasi desviación típica (desviación estándar de la muestra) es de 4,5.

    (a) Estimar un intervalo de confianza del 95 % para la media.

    (b) Para una probabilidad del 95 %, determinar un tamaño n para estimar el tiempo medio, de forma que el error sea a lo más de un día.

    Solución: (a) Desde el enunciado del problema, los datos son:

    - Muestra aleatoria de 100 vehículos: $n = 100$.
    - Tiempo promedio de reparación es de 4 días: $\overline{x} = 4$.
    - La desviación típica de la muestra es de 4,5 días: $s = 4{,}5$.
    - Tanto para (a) como para (b), se pide utilizar una confianza del 95 %, entonces $1 - \alpha = 0{,}95 \Rightarrow z_{1-\frac{\alpha}{2}} = z_{0{,}975} = 1{,}96$.

    Conocidos $\overline{x}$ y s, el intervalo de confianza para la media es:

$$\begin{aligned} \overline{x} \pm z_{1-\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}} &= 4 \pm 1{,}96 \cdot \frac{4{,}5}{\sqrt{100}} \\ &= 4 \pm 0{,}88. \end{aligned}$$

    Es decir, con el 95 % de confianza, se estima que el tiempo medio que tarda en ser reparado un vehículo en el taller es de entre 3,12 y 4,88 días (entre 3 y 5 días).

    (b) Se pide que $\mu \in (\overline{x} \pm \delta)$ (con $\delta = 1$), como $\delta = z_{1-\frac{\alpha}{2}} \cdot \sqrt{Var\left[\widehat{\theta}_{MV}\right]}$, entonces:

$$\begin{aligned} \delta = z_{1-\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}} &= 1{,}96 \cdot \frac{4{,}5}{\sqrt{n}} = 1 \\ \sqrt{n} = 8{,}82 \Rightarrow n &= 77{,}79 \approx 78. \end{aligned}$$

    Con un tamaño de muestra de 78 vehículos y una probabilidad del 95 %, se puede estimar el tiempo medio que tarda en ser reparado un vehículo con un error de a lo más un día.

12. En una prueba de aptitud se sabe que el número de aciertos es de 1000 con una desviación estándar de 125, sobre la base de la experiencia. Si se aplica la prueba a 100 personas seleccionadas al azar, calcular las siguientes probabilidades que involucran al promedio $\bar{X}$:

    (a) $\mathbb{P}\left(985 \leqslant \bar{X} \leqslant 1015\right)$.

    (b) $\mathbb{P}\left(\bar{X} > 1020\right)$.

    (c) $\mathbb{P}\left(960 < \bar{X} < 1040\right)$.

    Solución: A partir del enunciado del problema, se conocen los datos de valor esperado, desviación estándar y tamaño de muestra, que son, respectivamente, $E[X] = 1000$, $\sqrt{Var[X]} = 125$ y $n = 100$. Utilizando el teorema central del límite:

$$\begin{aligned}(a)\quad \mathbb{P}\left(985 \leqslant \bar{X} \leqslant 1015\right) &= \mathbb{P}\left(\frac{985-1000}{12{,}5} \leqslant Z_n \leqslant \frac{1015-1000}{12{,}5}\right)\\ &= \mathbb{P}(-1{,}2 \leqslant Z_n \leqslant 1{,}2) = 0{,}7698.\end{aligned}$$

    (b) En este caso, se tiene que:

$$\begin{aligned}\mathbb{P}\left(\bar{X} > 1020\right) &= 1 - \mathbb{P}\left(\bar{X} \leqslant 1020\right) = 1 - \mathbb{P}\left(Z_n \leqslant 1{,}6\right)\\ &= 0{,}05848.\end{aligned}$$

    (c) Este resultado es análogo a (a): $\mathbb{P}\left(960 < \bar{X} < 1040\right) = \mathbb{P}(-3{,}2 < Z_n < 3{,}2) = 0{,}9986$.

5.7. Ejercicios propuestos

1. Sean $Y_i \sim N(\alpha + \beta x_i, \sigma^2)$, $i = 1, 2, \cdots, n$ variables aleatorias independientes y los elementos x_i, constantes conocidas. Utilizar la función:

$$g(y_i|\alpha, \beta) = \sum_{i=1}^{n} \left(Y_i - (\alpha + \beta x_i)\right)^2,$$

    para encontrar los estimadores de máxima verosimilitud de α y β y la función de verosimilitud para obtener el estimador de σ^2.

2. Sean $X_1, X_2, \cdots, X_{n_1}$ e $Y_1, Y_2, \cdots, Y_{n_2}$ muestras aleatorias independientes. Encontrar los estimadores de máxima verosimilitud para θ_1 y θ_2 $(\theta_1, \theta_2 > 0)$, siendo las funciones de densidad:

$$f(x) = \frac{x}{\theta_1} \exp\left\{-\frac{x^2}{2\theta_1}\right\},\ x > 0, \quad f(y) = \frac{y}{\theta_2} \exp\left\{-\frac{y^2}{2\theta_2}\right\}\ y > 0.$$

3. Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria. Determinar los estimadores de máxima verosimilitud para los parámetros λ y μ, donde la función de densidad es:

$$f(x) = \sqrt{\frac{\lambda}{2\pi}} x^{-\frac{3}{2}} \exp\left\{-\frac{\lambda}{2\mu^2}\frac{(x-\mu)^2}{x}\right\}, \quad x > 0.$$

4. Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde $X_i \sim \text{Beta}(\theta, 2)$.

   (a) Encontrar el estimador de máxima verosimilitud para θ.

   (b) Evaluar el estimador si los valores en la muestra son:

0,30	0,80	0,27	0,35	0,62	0,55	0,60	0,35	0,45
0,51	0,63	0,59	0,68	0,82	0,56	0,39	0,40	0,69
0,44	0,55	0,64	0,34	0,67	0,92	0,34		

5. Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde la función de densidad de X_i está dada por $f(x) = \theta x^{\theta-1}$, $0 < x < 1$, $\theta > 0$.

   (a) Mostrar que el estimador de máxima verosimilitud de θ es función de $\bar{X}_g$, donde $\bar{X}_g = \left(\prod_{i=1}^n X_i\right)^{\frac{1}{n}}$.

   (b) Estudiar el insesgamiento del estimador encontrado en (a).

6. Las variables aleatorias $Y_1, Y_2, \cdots, Y_n$ son independientes, con función de densidad $f(y_i) = (\beta a_i)^{-1} \exp\left\{-y_i(\beta a_i)^{-1}\right\}$, $y_i > 0$ y a_i son constantes conocidas y positivas, $\beta > 0$.

   (a) Obtener expresiones para los estimadores de máxima verosimilitud, momentos y mínimos cuadrados para el parámetro β.

   (b) Estudiar el insesgamiento de los estimadores obtenidos en (a). En aquellos que son insesgados, determinar la varianza.

   (c) Si los $a_i = 1$ para $i = 1, 2, \cdots, n$, determinar un intervalo de confianza aproximado para el parámetro β.

7. Sean Y_1, Y_2, Y_3 una muestra aleatoria con $f_Y(y) = \theta^{-1} \exp\left\{-\theta^{-1} y\right\}$, $y > 0$. Considerar los siguientes estimadores del parámetro θ: $\widehat{\theta}_1 = Y_1$, $\widehat{\theta}_2 = 0,5(Y_1 + Y_2)$, $\widehat{\theta}_3 = \frac{1}{3}(Y_1 + 2Y_2)$, $\widehat{\theta}_4 = \text{mín}(Y_1, Y_2, Y_3)$ y $\widehat{\theta}_5 = \bar{Y}$. Para cada uno de estos estimadores, determinar (a) el insesgamiento y (b) el estimador de menor varianza.

8. Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde $X_i \sim \text{Beta}(\theta, 1)$. Determinar:

   (a) El estimador de máxima verosimilitud para θ.

   (b) El valor esperado y la varianza del estimador encontrado en (a).

   (c) El insesgamiento y la consistencia del estimador obtenido en (a).

9. Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria con $X_i \sim \text{Poisson}(\lambda)$. Estudiar el insesgamiento y la consistencia de estos dos estimadores de λ:

$$T_1 = \bar{X};\ T_2 = \frac{2}{n(n+1)} \sum_{i=1}^{n} i\ X_i.$$

   (a) Determinar el valor esperado y la varianza de T_1 y T_2.

   (b) Estudiar el insesgamiento y la consistencia de T_1 y T_2.

10. Sean $Y_1, Y_2, \cdots, Y_n$ una muestra aleatoria, donde la variable Y_i tiene una función de densidad dada por $f(y) = \theta^{-1}$, $0 < y < \theta$. Dos estimadores son propuestos para estimar el parámetro θ:

$$T_1 = k_1 \bar{Y},\ T_2 = k_2 M,\ M = \text{máx}\left\{Y_1, Y_2, \cdots, Y_n\right\}.$$

   (a) Encontrar los valores k_1 y k_2 para que T_1 y T_2 sean estimadores insesgados del parámetro θ.

   (b) Determinar si los estimadores de (a) son consistentes para el parámetro θ.

   (c) Proponer un estimador insesgado para el parámetro θ^r, $r \in \mathbb{N}$ y calcular su varianza.

11. Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria con $X_i \sim \text{Uniforme}(0, \theta)$, se define la estadística $T = \text{mín}(X_1, X_2, \cdots, X_n)$.

(a) Mostrar que $f_T(t) = n(1-\theta^{-1}t)^{n-1}\theta^{-1}$ es la función de densidad de T.

(b) Mostrar que $T_1 = \frac{1}{n-1}T$ es un estimador insesgado para θ.

12. Sean $X_1, \cdots, X_n$ una muestra aleatoria con $f_X(x) = 3\beta^3 y^{-4}$, $y > \beta$; se define la estadística $T = \sqrt{X_1 \cdot X_2}$.

(a) Mostrar que $|E[T] - \beta| = \frac{1}{3}\beta$.

(b) Mostrar que $ECM(T) = \frac{31}{36}\beta^2$.

13. Sean $X_1, \cdots, X_n$ una muestra aleatoria, donde la función de densidad de X_i está dada por: $f(x) = \theta^{-1} \exp\{-\theta^{-1}x_i\}$, $x_i > 0$.

(a) Obtener el estimador de mínimos cuadrados para θ.

(b) Determinar si el estimador de mínimos cuadrados es consistente.

(c) Encontrar un intervalo del 95 % para el parámetro θ utilizando el estimador de mínimos cuadrados.

14. Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde $X_i \sim N(\mu, k\sigma^2)$ y k es una constante positiva. Obtener, con $100 \cdot (1-\alpha)\,\%$, un intervalo de confianza exacto para el parámetro μ.

15. Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde $X_i \sim N(100, \sigma^2)$. Obtener, con $100 \cdot (1-\alpha)\,\%$, el intervalo de confianza:

(a) Exacto para el parámetro σ^2,

(b) Aproximado para el parámetro $\theta = \sigma^4$.

16. Sean $X_1, X_2, \cdots, X_n$ y $Y_1, Y_2, \cdots, Y_n$ muestras aleatorias independientes, donde se tiene que $X_i \sim N(\mu_1, \sigma_1^2)$ e $Y_i \sim N(\mu_2, \sigma^2)$, y sean $\bar{X}$ e $\bar{Y}$ los estimadores de máxima verosimilitud de los parámetros μ_1 y μ_2, respectivamente.

(a) Asumiendo que σ_1^2 y σ_2^2 son parámetros desconocidos pero iguales, es decir, $\sigma_1^2 = \sigma_2^2 = \sigma^2$, entonces construir un intervalo de confianza exacto del $(1-\alpha) \cdot 100\,\%$ para el parámetro $\mu_1 - \mu_2$.

(b) Construir un intervalo de confianza exacto del $(1-\alpha) \cdot 100\,\%$ para el parámetro $\theta = \sigma_1^2 \cdot (\sigma_2^2)^{-1}$.

17. Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria donde $X_i \sim \text{Bernoulli}(p)$. Construir un intervalo de confianza aproximado para:

    (a) El parámetro p.

    (b) El parámetro p^2.

18. Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde $X_i \sim \text{Gamma}(\alpha, \beta)$. Bajo el supuesto de que α es conocido, encontrar un intervalo de confianza aproximado de $(1 - p_0) \cdot 100\,\%$ para el parámetro β.

19. Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde $X_i \sim N(\mu, \sigma_0^2)$, con σ_0^2 conocido. Determinar un intervalo de confianza aproximado de $(1 - \alpha) \cdot 100\,\%$ para el parámetro $\theta = \exp\{\mu\}$.

20. La función de densidad de la variable aleatoria bivariada $(X, Y)^T$ es:

$$f(x, y) = \exp\left\{-(\theta x + \theta^{-1} y)\right\},\ x > 0,\ y > 0 \text{ y } \theta > 0.$$

    Se obtiene una muestra aleatoria $U_1, U_2, \cdots, U_n$, donde $U_i = \left(\frac{X}{Y}\right)^{-\frac{1}{2}}$.

    (a) Obtener un estimador insesgado para θ como una función de $\frac{1}{n}\sum_{i=1}^{n} U_i$.

    (b) Calcular la varianza del estimador insesgado propuesto en (a).

21. En una muestra aleatoria de tamaño $n = 4$, se tiene:

$$x_1 = 3{,}3 \qquad x_2 = -0{,}3$$
$$x_3 = -0{,}6 \qquad x_4 = -0{,}9$$

    Con una desviación estándar de $\sigma = 3$, encontrar un intervalo de confianza del $90\,\%$ para μ.

22. En un taller de construcción se preparan tablas de variados anchos. Se sabe que el ancho es una variable aleatoria que sigue una distribución normal con varianza desconocida.

    Para una muestra de tamaño $n = 10$ se tiene:

$$x_1 = 20{,}0 \qquad x_2 = 19{,}7 \qquad x_3 = 20{,}1 \qquad x_4 = 19{,}9$$
$$x_5 = 20{,}2 \qquad x_6 = 19{,}5 \qquad x_7 = 20{,}3 \qquad x_8 = 20{,}4$$
$$x_9 = 19{,}6 \qquad x_{10} = 20{,}0$$

Determinar un intervalo de confianza del 95 % para σ^2.

23. El director de los medios de comunicación de una agencia publicitaria publicó un anuncio de un banco en dos revistas. En cada una de ellas se colocaron anuncios semejantes.

    Un mes más tarde, una compañía de investigación encontró que 226 de los 473 lectores de la primera revista sabían de los servicios bancarios, al igual que 165 de los 493 lectores de la segunda revista.

    (a) ¿Son adecuados los tamaños muestrales para utilizar la aproximación normal?

    (b) Encontrar un intervalo de confianza del 95 % para la diferencia de proporciones de los lectores que tienen conocimiento de los servicios.

24. Para $i = 1, 2, ..., k$ y $j = 1, 2, ..., n$, Sean $X_{ij} \sim N\left(\mu_i, \sigma^2\right)$ (k poblaciones de tamaño n cada una). Se asume que las variables X_{ij} constituyen un conjunto de variables aleatorias independientes. Sabiendo que para $i = 1, 2, ..., k$:

$$\overline{X}_i = \frac{1}{n}\sum_{j=1}^{n} X_{ij} \sim N(\mu_i, \frac{\sigma^2}{n})$$

$$S^2{}_i = \frac{1}{n-1}\sum_{j=1}^{n}\left(X_{ij} - \overline{X}_i\right) \Rightarrow \frac{(n-1)}{\sigma^2}S_i^2 \sim \chi^2(n-1).$$

    (a) Siendo $\theta = \sum_{i=1}^{k} c_i\mu_i$, mostrar que el estimador de máxima verosimilitud es:

$$\widehat{\theta} = \sum_{i=1}^{k} c_i\bar{X}_i.$$

    (b) Mostrar que:

$$\widehat{\theta} \sim N\left(\theta, \frac{\sigma^2}{n}\sum_{i=1}^{k} c_i^2\right).$$

    (c) Probar que con $(1-\alpha)\cdot 100\,\%$, se tiene que:

$$\theta \in \left(\widehat{\theta} \pm \sqrt{\frac{1}{kn}\sum_{i=1}^{k} c_i \sum_{i=1}^{k} S_i^2}\right).$$

(d) Probar que con $(1-\alpha)\cdot 100\,\%$, se tiene que:

$$\sigma^2 \in \left(\frac{k(n-1)}{\chi^2_{1-\frac{\alpha}{2}}(k(n-1))} \sum_{i=1}^{k} S_i^2;\ \frac{k(n-1)}{\chi^2_{\frac{\alpha}{2}}(k(n-1))} \sum_{i=1}^{k} S_i^2 \right).$$

Capítulo 6

Pruebas de hipótesis

El conjunto de secciones secuenciales que abarca el área de la inferencia estadística comienza con la estimación de parámetros y continúa con pruebas de hipótesis. Según lo expuesto en el capítulo anterior, se cuenta con una técnica que permite conseguir estimaciones puntuales e intervalos en que se encuentra un parámetro con una cierta probabilidad. Este capítulo entrega técnicas para enfrentar diseños experimentales en los que se requiere decidir acerca de la veracidad de una aseveración sobre una población o de la comparación de características de interés entre dos o más poblaciones.

En concreto, se estudiarán técnicas para comprobar:

(a) La validez de afirmaciones acerca de parámetros poblacionales:

$\mu = \mu_0$: la media puede asumir el valor μ_0.

(b) Comparaciones entre poblaciones (a través del promedio):

$\mu_1 = \mu_2$: la producción de una empresa mediante la implementación de dos sistemas de turnos es igual.

(c) Comparaciones entre mediciones en diferentes instantes de tiempo:

$\mu_A = \mu_B$: la producción antes del entrenamiento es igual a la que se observa después, medidas en los tiempos: A (antes del entrenamiento) y B (después del entrenamiento).

Es decir, estas técnicas permitirán abordar hipótesis y comprobar objetivos de una investigación cuantitativa.

Este capítulo y el anterior constituyen la base de la inferencia estadística, que luego continúa con la modelación estadística de fenómenos reales por medio de la utilización de la axiomática de probabilidad.

6.1. Elementos iniciales de pruebas de hipótesis

A continuación, se presentan conceptos para comprender y luego aplicar las técnicas de pruebas de hipótesis.

Definición **6.1** Una prueba de hipótesis estadística es el procedimiento que permite estudiar una conjetura acerca de los parámetros de la distribución de una variable aleatoria.

Ejemplo **6.1** Sea la variable aleatoria X: *tiempo de falla de una máquina*, donde $X \sim \text{Exp}(\lambda)$, con $E[X] = \lambda$. Se desea responder si el tiempo medio de falla de una máquina es superior a 500 horas.

Solución: La conjetura se debe presentar para los parámetros de la distribución, en este caso, para el tiempo medio que corresponde al parámetro λ. Es decir, la hipótesis es si $\lambda > 500$. □

Las pruebas estadísticas se basan en el contraste de dos hipótesis complementarias, las que se llamarán y denotarán con:

- Hipótesis nula: H_0.
- Hipótesis alternativa (a investigar): H_1.

Definición **6.2** Se define el espacio paramétrico Θ como el conjunto de todos los valores posibles que puede tomar un parámetro $\boldsymbol{\theta}$ y se tendrá que $\boldsymbol{\theta} \in \Theta$. Además, la hipótesis nula y alternativa definen, respectivamente, Θ_0 y Θ_1 ($\Theta = \Theta_0 \cup \Theta_1$) de la siguiente manera:

- $H_0 : \theta \in \Theta_0$.
- $H_1 : \theta \in \Theta_1$.

Ejemplo **6.2** Considerar la variable aleatoria $X \sim \text{Exp}(\beta)$. Se desea probar la siguiente hipótesis $H_0 : \beta \leqslant 1$ versus $H_1 : \beta > 1$.

Solución: Dado que la variable aleatoria $X \sim \text{Exp}(\beta)$, se sabe que $\beta > 0$. Entonces, el espacio paramétrico está dado por $\Theta = \{\beta | \beta > 0\} = \mathbb{R}^+$. Luego, el espacio paramétrico para ambas hipótesis está dado por:

- $H_0 : \beta \leqslant 1 \Rightarrow \Theta_0 = \{\beta | 0 < \beta \leqslant 1\}$.
- $H_1 : \beta > 1 \Rightarrow \Theta_1 = \{\beta | \beta > 1\}$. □

Si se considera el valor del parámetro dado en H_0, entonces queda especificado el valor del parámetro asociado a la distribución. En aquellos casos en que esto no ocurre, como, por ejemplo, distribuciones que tienen más de un parámetro, es necesario definir alguna regla que permita evaluar la hipótesis y distinguirlos. En este sentido, la siguiente definición ayuda a clasificar la hipótesis.

Definición **6.3** Si la hipótesis estadística especifica completamente la distribución de la variable aleatoria (es decir, dada la hipótesis no hay parámetros desconocidos en la distribución), se denomina hipótesis simple. De lo contrario, la hipótesis se llama compuesta.

Ejemplo **6.3** (1) Sean la variable aleatoria $X \sim \text{Exp}(\beta)$ y la hipótesis:

$$H_0 : \beta = 1 \text{ versus } H_1 : \beta \neq 1.$$

Entonces, $\Theta = \{\beta | \beta > 0\}$ y $\Theta_0 = \{\beta | \beta = 1\}$. Por lo tanto, es una hipótesis simple.

(2) Sea la variable aleatoria $X \sim N(\mu, \sigma^2)$, considerar la hipótesis:

$$H_0 : \mu = 1 \text{ versus } H_1 : \mu \neq 1.$$

Así, $\Theta = \left\{(\mu, \sigma^2) | -\infty < \mu < +\infty, \sigma^2 > 0\right\}$ y $\Theta_0 = \left\{\sigma^2 | \sigma^2 > 0\right\}$. Esta es una hipótesis compuesta, debido a que σ^2 no queda especificado en H_0 (es un parámetro desconocido). □

Luego, para aplicar las técnicas que permitan decidir entre H_0 y H_1, se requiere definir para las hipótesis compuestas supuestos referidos al

parámetro desconocido, como, por ejemplo, $\sigma^2 = \sigma_0^2$, es decir, asumir que la varianza es conocida.

El paso siguiente, ya definidos H_0 y H_1, es, dada la información de una muestra aleatoria, seleccionar una técnica cuya aplicación conduzca a la decisión en favor de H_0 o en favor de H_1.

Definición **6.4** Un test estadístico de una hipótesis es una regla, basada en una muestra aleatoria, que cuando es aplicado conduce a la decisión en favor de H_0 o en favor de H_1.

Esta definición implica que la decisión acerca de la hipótesis planteada se realiza con la muestra y se extiende a la población, proceso que genera la incertidumbre sobre que lo que ocurre en la muestra sea lo mismo que en la población. Entonces, se define la probabilidad de equivocarse al decidir acerca de una hipótesis basándose en el resultado de un test estadístico.

Definición **6.5** Al decidir acerca de una hipótesis mediante una prueba estadística basada en una muestra aleatoria, se definen dos tipos de errores que se podrían cometer en el proceso:

- Error tipo I: Rechazar H_0 en la muestra, dado que H_0 es verdadera en la población. Luego, se define α como la probabilidad de cometer un *error tipo I*, conocida como nivel de significación. El error tipo I es igual a α y, análogamente, se define $1 - \alpha$ como el nivel de confianza (ya utilizado anteriormente en el capítulo sobre intervalos de confianza).

- Error tipo II: No rechazar H_0 en la muestra, dado que H_0 es falsa en la población. Luego, a la probabilidad de cometer un *error tipo II* se la llama β. A partir del error tipo II se puede definir la llamada potencia del test, que se obtiene por $1-\beta$, es decir, es la probabilidad de rechazar H_0 sobre la base de una muestra aleatoria, dado que H_0 es falsa en la población. Teniendo en cuenta esta definición, se puede notar que la potencia depende del valor de $\theta \in \Theta$, por lo que se habla de la función potencia.

El siguiente esquema resume la definición 6.5, presentando los errores tipo I y tipo II como productos de la decisión asumida con la evidencia

que entrega la muestra y lo que realmente ocurre, que está dado por la población.

Decisión	Población: H_0 es falsa	Población: H_0 es verdadera
Rechazar H_0 (H_0 es falsa).	Acierto $1-\beta$	Error tipo I α
No rechazar H_0 (H_0 es verdadera).	Error tipo II β	Acierto $1-\alpha$

Se utiliza un test estadístico T, que es una función real valuada que depende solamente de los elementos $X_1, X_2, \cdots, X_n$, que constituyen una muestra aleatoria $T : \mathbb{R}^+ \to \mathbb{R}$. Por medio de este test estadístico se podrá tomar la decisión en favor de H_0 o en favor de H_1.

Observación **6.1** Habitualmente, se habla de *rechazar* H_0 o de *aceptar* H_0, cuando se decide en favor de H_1 o en favor de H_0, respectivamente. Esta forma de expresar la decisión de una hipótesis se debe a que H_1 es la hipótesis de investigación (que plantea el investigador y que busca evidencia que la respalde) y decir que se rechaza H_0 (que es falsa) implica aceptar H_1.

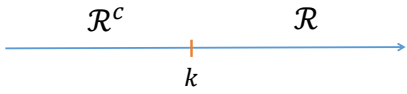
6.2. Región de rechazo

El criterio de la región de rechazo permitirá decidir en favor de H_0 o de H_1, bajo un cierto nivel de confianza.

Definición **6.6** Sea $\Re \subset \mathbb{R}$ llamada *región de rechazo*, en la que se rechazará H_0 si un estadístico T pertenece a ella. En caso contrario, se dice que no hay evidencia suficiente para rechazar H_0, sobre la base de la muestra aleatoria.

Desde la definición 6.6, surge la necesidad de determinar un valor $k \in \mathbb{R}$ (dependiente de α), tal que para valores mayores o menores a k (según corresponda) se rechazará H_0, véase la figura 6.1.

Figura 6.1: Región de rechazo



Fuente: elaboración propia.

Lema de Neyman-Pearson

Este lema entrega un método que permitirá calcular un valor c que defina la mejor región que podría obtenerse.

Propiedad **6.1** Se desea contrastar la hipótesis $H_0 : \theta = \theta_0$ versus $H_1 : \theta = \theta_1$, donde θ_0 y θ_1 definen completamente el espacio paramétrico. Entonces $\Re$ es la mejor región crítica de nivel α (es decir, si se fija *a priori* el nivel α, es esta región la de mayor potencia), y siendo c (valor que depende de α, por lo que en algunos textos es llamado c_α) el valor que define el punto desde el cual se define $\Re^c$ y el restante es $\Re$. Así, se puede determinar el valor de c desde la expresión:

$$\mathbb{P}(\Re | H_0 : \theta = \theta_0) = \alpha.$$

En términos prácticos, para determinar $\Re$ de la propiedad 6.1, se considera rechazar H_0 sobre la base de la información que entrega la muestra (a través del estimador de θ) en comparación con un valor crítico c que definirá la región crítica $\Re$, tal que lleve a rechazar H_0. Desde esta inecuación, se construye la función pivotal $f(\theta, \widehat{\theta})$, con la cual es posible determinar el valor de c en función de α.

A continuación, se presentan ejemplos de construcción de test estadísticos basados en fijar el error tipo I.

Ejemplo **6.4** Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria con $X \sim N(\mu, \sigma_0^2)$. Considerar la hipótesis $H_0 : \mu \leqslant \mu_0$ y $H_1 : \mu > \mu_0$; construir un test de nivel α para decidir acerca de la hipótesis planteada.

Solución: Considerando que $\widehat{\mu} = \bar{X}$ tiene distribución conocida, la expresión para determinar la región crítica, dado que se rechazará H_0

para valores grandes del parámetro, es:

$$\Re = \left\{ \bar{X} | \bar{X} > c \right\}.$$

Luego, fijando el error tipo I en α y notando que para decidir en favor de H_0 es suficiente que $\mu = \mu_0$, entonces, usando el lema de Neyman-Pearson:

$$\mathbb{P}(\bar{X} > c | H_0 : \mu = \mu_0) = \alpha.$$

Luego, como σ_0^2 es conocido, entonces el pivote es $Z \sim \mathrm{N}(0,1)$:

$$\mathbb{P}\left(\frac{\bar{X} - \mu}{\sqrt{\frac{\sigma_0^2}{n}}} > \frac{c - \mu}{\sqrt{\frac{\sigma_0^2}{n}}} \middle| H_0 : \mu = \mu_0 \right) = \mathbb{P}\left(Z > \frac{c - \mu_0}{\sqrt{\frac{\sigma_0^2}{n}}} \right) = \alpha$$

$$\Rightarrow \Phi^{-1}(1-\alpha) = z_{1-\alpha} = \frac{c - \mu_0}{\sqrt{\frac{\sigma_0^2}{n}}} \Rightarrow c = \mu_0 + z_{1-\alpha} \cdot \sqrt{\frac{\sigma_0^2}{n}}.$$

Entonces, $\Re = \left\{ \bar{X} | \bar{X} > \mu_0 + z_{1-\alpha} \cdot \sqrt{\frac{\sigma_0^2}{n}} \right\}$. □

Observación **6.2** Gráfico de α y β: El valor de α (error tipo I) se fija inicialmente y β depende del valor de θ_0 (el parámetro bajo H_1). En este sentido, se puede obtener el gráfico para α y β considerando una hipótesis de interés. En efecto, si se utiliza el ejemplo 6.4 se tendrá que el error tipo I y el error tipo II están dados, respectivamente, por:

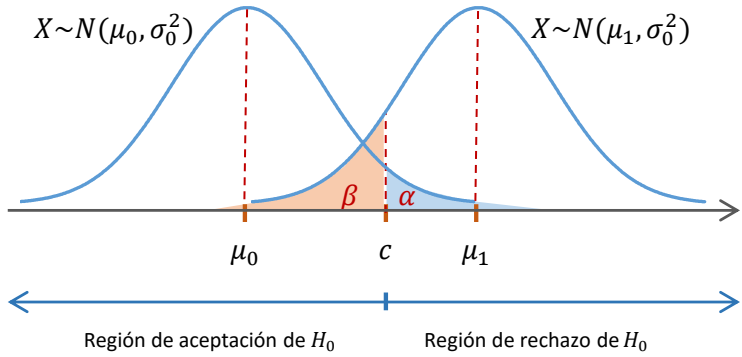
$$\alpha = \mathbb{P}(\bar{X} > c | H_0 : \mu = \mu_0).$$
$$\beta = \mathbb{P}(\bar{X} < c | H_1 : \mu > \mu_0).$$

Luego, la figura 6.2 tiene los gráficos de las ecuaciones anteriores: la primera para la media bajo H_0, donde se tendría que $X \sim N(\mu_0, \sigma_0^2)$, y la segunda para la media bajo H_1, donde se tendría que $X \sim N(\mu_1, \sigma_0^2)$, con $\mu_1 > \mu_0$. Además, considera el valor de c que define la región de rechazo.

Ejemplo **6.5** Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria con $X_i \sim N(\mu; 25)$. Se necesita desarrollar un test para la hipótesis $H_0 : \mu = 1$ versus $H_1 : \mu > 1$.

Solución: Primero, nótese que la hipótesis propuesta es simple. Luego, el test se construye fijando la probabilidad de error tipo I, con una

Figura 6.2: Región de rechazo para un caso particular



Fuente: elaboración propia.

cierta magnitud α. Dado que la hipótesis está formulada para la media poblacional μ, es decir, la región crítica se define para $\bar{X}$, entonces la región crítica está dada por $\Re = \left\{\bar{X} | \bar{X} > c\right\}$. Luego, c se puede calcular controlando el error tipo I y estandarizando para $\bar{X} \sim N(\mu, 25 \cdot n^{-1})$ de la siguiente forma:

$$\begin{aligned}
\alpha &= \mathbb{P}\left(\bar{X} > c | H_0 : \mu = 1\right) \\
&= \mathbb{P}\left(\frac{\bar{X} - 1}{5 \cdot \sqrt{n^{-1}}} > \frac{c - 1}{5 \cdot \sqrt{n^{-1}}}\right) \\
&= \mathbb{P}\left(Z > \frac{c - 1}{5 \cdot \sqrt{n^{-1}}}\right).
\end{aligned}$$

Como $Z \sim N(0, 1)$, entonces:

$$1 - \alpha = \mathbb{P}\left(Z < \frac{(c-1)\sqrt{n}}{5}\right) = \Phi\left((c-1)\frac{\sqrt{n}}{5}\right).$$

Dado que α fue fijado inicialmente y n es conocido, aplicando la función inversa para despejar c, entonces:

$$\Phi^{-1}(1-\alpha) = z_{1-\alpha} = (c-1)\frac{\sqrt{n}}{5} \Rightarrow c = 1 + \frac{5}{\sqrt{n}}\, z_{1-\alpha}.$$

Finalmente, la región crítica queda definida por:

$$\Re = \left\{\bar{X} | \bar{X} > 1 + \frac{5}{\sqrt{n}} z_{1-\alpha}\right\}.$$

Es decir, si $\bar{X} \in \Re$, entonces hay evidencia estadística para rechazar la hipótesis $H_0 : \mu = 1$, con un nivel de significación α.

La función potencia para este test se calcula mediante:

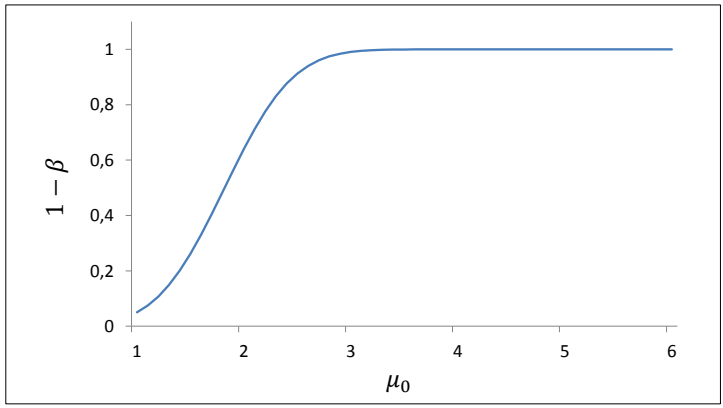
$$\begin{aligned} 1-\beta &= \mathbb{P}\left(\bar{X} > 1 + \frac{5}{\sqrt{n}} z_{1-\alpha} \middle| H_1 : \mu > 1 \right) \\ &= \mathbb{P}\left(\frac{\bar{X} - \mu_0}{5 \cdot \sqrt{n^{-1}}} > \left[1 + \frac{5}{\sqrt{n}} z_{1-\alpha} - \mu_0 \right] \cdot \left(\frac{5}{\sqrt{n}} \right)^{-1} \right). \\ &= \mathbb{P}\left(Z > \frac{1-\mu_0}{5} \sqrt{n} + z_{1-\alpha} \right), \quad \mu_0 \in (1, +\infty). \end{aligned}$$

Si $\alpha = 0{,}05$ y $n = 100$, entonces $c = 1{,}82$. Es decir, se rechaza H_0 si en la muestra $\bar{X} > 1{,}82$. Por otra parte, la potencia del test es:

$$1 - \beta = 1 - \mathbb{P}\left(Z \leqslant 2 \cdot (1 - \mu_0) + 1{,}64 \right), \mu_0 \in (1, +\infty).$$

La figura 6.3 muestra el comportamiento de la potencia del test para distintos valores de μ_0 ($\mu_0 \in (1, +\infty)$). □

Figura 6.3: Potencia del test como función de μ_0



Fuente: elaboración propia.

Ejemplo **6.6** Sean $X_1, X_2, \cdots, X_{n_1}$ una muestra aleatoria con $X_i \sim N(\mu_1, \sigma_1^2)$ y sean $Y_1, Y_2, \cdots, Y_{n_2}$ una muestra aleatoria donde $Y_i \sim N(\mu_2, \sigma_2^2)$. Se desea probar la siguiente hipótesis:

$$H_0 : \mu_1 - \mu_2 = 0 \text{ versus } H_1 : \mu_1 - \mu_2 > 0.$$

Dado que se trata de una hipótesis compuesta, para construir el test se deben asumir ciertos supuestos relacionados con el parámetro de escala (varianzas de las poblaciones). Entonces, se puede construir el test bajo las siguientes consideraciones:

(a) Asumiendo que σ_1^2 y σ_2^2 son conocidos.

(b) Asumiendo que σ_1^2 y σ_2^2 son desconocidos pero iguales, es decir, $\sigma_1^2 = \sigma_2^2 = \sigma^2$.

Solución: (a) La región de rechazo es:

$$\Re = \left\{ \bar{X} - \bar{Y} | \bar{X} - \bar{Y} > c \right\}.$$

Luego, el valor de c se obtiene desde la probabilidad de error tipo I:

$$\begin{aligned} \alpha &= \mathbb{P}\left(\bar{X} - \bar{Y} > c | H_0 : \mu_1 - \mu_2 = 0 \right) \\ &= \mathbb{P}\left(\frac{\bar{X} - \bar{Y} - 0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} > \frac{c - 0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \right) \\ &= \mathbb{P}\left(Z > \frac{c - 0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \right). \end{aligned}$$

Aplicando la función de probabilidad inversa de $Z \sim \mathrm{N}(0,1)$ para despejar c:

$$z_{1-\alpha} = \frac{c}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \Rightarrow c = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \cdot z_{1-\alpha}.$$

Con lo cual queda definida la región de rechazo por:

$$\Re = \left\{ \bar{X} - \bar{Y} | \bar{X} - \bar{Y} > \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \cdot z_{1-\alpha} \right\}.$$

(b) Se asume que σ_1^2 y σ_2^2 son desconocidos pero iguales, es decir, $\sigma_1^2 = \sigma_2^2 = \sigma^2$ (homogeneidad de varianzas). La región de rechazo, al igual que en el caso anterior, está dada por:

$$\Re = \left\{ \bar{X} - \bar{Y} | \bar{X} - \bar{Y} > c \right\}.$$

Luego, el valor de c se obtiene desde la probabilidad de error tipo I.

$$\alpha = \mathbb{P}\left(\bar{X} - \bar{Y} > c | H_0 : \mu_1 - \mu_2 = 0\right).$$

Como $\bar{X} - \bar{Y} \sim N\left(\mu_1 - \mu_2, \sigma^2(\frac{1}{n_1} + \frac{1}{n_1})\right)$ y σ^2 es desconocido, el pivote corresponde a una variable aleatoria con distribución t-Student:

$$\alpha = \mathbb{P}\left(\frac{\bar{X} - \bar{Y} - 0}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} > \frac{c - 0}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}\right) = \mathbb{P}\left(T > \frac{c}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}\right),$$

donde $S_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}$. Como $T \sim \text{TS}(n_1 + n_2 - 2)$, entonces:

$$t_{1-\alpha} = \frac{c}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \Rightarrow c = S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \cdot t_{1-\alpha}.$$

Luego, se dice que hay evidencia estadística para rechazar H_0, con un nivel de confianza de $1 - \alpha$, si se cumple que:

$$\bar{X} - \bar{Y} > S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \cdot t_{1-\alpha}. \square$$

Ejercicio propuesto **6.1** A partir del ejemplo 6.6(a), sea $n_1 = n_2 = n$, obtener una expresión para calcular el tamaño muestral n, cuando se utiliza una probabilidad de error tipo I de α, una probabilidad de error tipo II de β y una diferencia de medias $\delta = \mu_1 - \mu_2$ (asumiendo $\delta > 0$).

Ejemplo **6.7** A partir del ejemplo 6.6(2), suponer que se tiene un grupo 1 y un grupo 2 de estudiantes, a los cuales se les aplican tipos de enseñanza diferentes con el objetivo de mejorar sus calificaciones. Finalizada la implementación, se registran las siguientes calificaciones:

Grupo 1:	4,1	4,8	7,0	6,1	5,6	4,0	3,5
Grupo 2:	3,6	4,1	4,2	5,2	5,3	3,2	3,5

Solución: Los grupos definen dos poblaciones independientes, X (grupo 1) e Y (grupo 2), luego, con los datos entregados, $\bar{X} = 5{,}0$, $\bar{Y} = 4{,}2$, $S_1 = 1{,}4$ y $S_2 = 0{,}82$.

Si $\alpha = 0{,}05$, entonces $c = 0{,}94$ y $\Re = \left\{\bar{X} - \bar{Y} | \bar{X} - \bar{Y} > 0{,}94\right\}$ (región de rechazo), desde la muestra se obtiene que $\bar{X} - \bar{Y} = 0{,}8$, como $\bar{X} - \bar{Y} \notin \Re$,

se concluye que no hay evidencia estadística suficiente para rechazar la hipótesis nula $H_0 : \mu_1 = \mu_2$. Es decir, no hay evidencia suficiente para afirmar que los métodos producen diferencias en las calificaciones logradas por los estudiantes. □

Ejemplo **6.8** A partir del ejemplo 5.24, construir un test para probar la hipótesis de que el costo promedio es superior a US 170.

Solución: El parámetro es para el costo promedio o esperado, luego la hipótesis es $H_0 : \theta = 170$ versus $H_1 : \theta > 170$. Entonces:

$$\Re = \left\{\widehat{\theta}|\widehat{\theta} > c\right\} = \left\{100 + 7\bar{X} + 3\bar{Y}|100 + 7\bar{X} + 3\bar{Y} > c\right\}.$$

Luego, usando la propiedad 6.1:

$$\begin{aligned}
\alpha =& \mathbb{P}\left(100 + 7\bar{X} + 3\bar{Y} > c|\theta = 170\right) \\
=& \mathbb{P}\left(Z > \frac{c-\theta}{\sqrt{\frac{12{,}25}{n_1} + \frac{0{,}36}{n_2}}}|\theta = 170\right) = \mathbb{P}\left(Z > \frac{c-170}{\sqrt{\frac{12{,}25}{n_1} + \frac{0{,}36}{n_2}}}\right) \\
\Rightarrow & c = 170 + z_{1-\alpha}\sqrt{\frac{12{,}25}{n_1} + \frac{0{,}36}{n_2}}.\square
\end{aligned}$$

Observación **6.3** Las pruebas de hipótesis que se plantean pueden clasificarse en *unilateral* o de *una cola*, o *bilateral* o de *dos colas*, de acuerdo a cómo reparten la probabilidad de error tipo I en la región de rechazo:

- Unilateral: $H_0 : \theta = \theta_0$ y $H_1 : \theta > \theta_0$.
- Unilateral: $H_0 : \theta = \theta_0$ y $H_1 : \theta < \theta_0$.
- Bilateral: $H_0 : \theta = \theta_0$ y $H_1 : \theta \neq \theta_0$.

6.3. Estadístico de prueba

Un estadístico de prueba es una función T definida como:

$$\begin{aligned}
&T : \mathbb{R}^n \to \mathbb{R} \\
&\boldsymbol{x} \in \mathbb{R}^n \mapsto T = T(X_1, X_2, \cdots, X_n) \in \mathbb{R}.
\end{aligned}$$

Donde T tiene una distribución que queda completamente especificada bajo H_0 (no hay parámetros desconocidos). Luego, el estadístico de

prueba se puede calcular desde la región de rechazo conociendo c (según la propiedad 6.1) u obteniendo el pivote para el parámetro condicionado en H_0. Por su parte, el criterio de rechazo se define según H_1 (de igual manera como se define la región de rechazo).

Ejemplo **6.9** A partir del ejemplo 6.4, la hipótesis $H_0 : \mu = \mu_0$ versus $H_1 : \mu > \mu_0$. En este caso, el pivote es $Z \sim \mathrm{N}(0,1)$ y el estadístico de prueba es $Z_0 = Z|H_0$, donde:

$$Z = \frac{\bar{X} - \mu}{\sqrt{\frac{\sigma_0^2}{n}}} \Rightarrow Z_0 = \frac{\bar{X} - \mu_0}{\sqrt{\frac{\sigma_0^2}{n}}}.$$

Finalmente, se rechazará H_0 si $Z_0 > z_{1-\alpha}$.

Análogamente, se puede probar que si la hipótesis es $H_0 : \mu = \mu_0$ versus $H_1 : \mu < \mu_0$, entonces se rechazará H_0 si:

$$Z_0 = \frac{\bar{X} - \mu_0}{\sqrt{\frac{\sigma_0^2}{n}}} < z_\alpha . \square$$

Ejemplo **6.10** Considerando los datos del ejemplo 6.9, encontrar el estadístico de prueba si ahora la hipótesis es $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$.

Solución: Al igual que en el ejemplo 6.9, el estadístico de prueba es:

$$Z_0 = \frac{\bar{X} - \mu_0}{\sqrt{\frac{\sigma_0^2}{n}}}.$$

Finalmente, considerando que la prueba es bilateral y que la distribución de Z es simétrica respecto de cero, se rechazará H_0 si:

$$|Z_0| > z_{1-\frac{\alpha}{2}} . \square$$

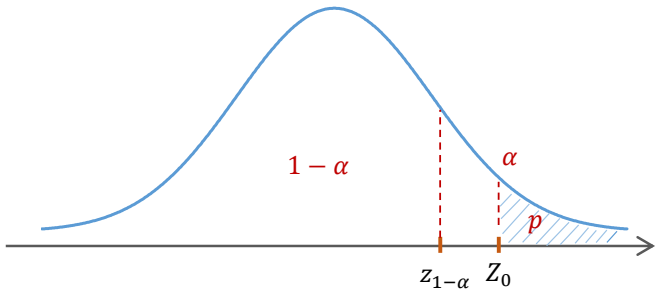
6.4. Valor-*p*

El método de la región de rechazo, anteriormente visto para decidir en una prueba de hipótesis, requiere fijar previamente el nivel de significación α (error tipo I), entonces dos investigadores que trabajan sobre el mismo problema (y con la misma información) pueden llegar a una decisión distinta por el solo hecho de utilizar un nivel de significación distinto. Ahora se presenta un método para decidir en una prueba de hipótesis, que no requiere fijar *a priori* el nivel de confianza.

Definición **6.7** El método consiste en obtener el menor nivel de significación para el cual se rechazaría H_0 sobre la base del estadístico de prueba. Por lo tanto, si el investigador dispone de un nivel crítico α (tamaño del error tipo I), rechazaría H_0 si el valor-p es menor que α.

Ejemplo **6.11** Para ilustrar la definición anterior, se utiliza el ejemplo 6.9, en el cual se concluyó que se rechazará H_0 para $Z_0 > z_{1-\alpha}$. De la figura 6.4 se tiene que valor-$p = \mathbb{P}(Z > Z_0)$, por lo tanto, se rechaza H_0 si valor-$p \leqslant \alpha$. □

Figura 6.4: Gráfico del estadístico de prueba y el valor-p



Fuente: elaboración propia.

Ejemplo **6.12** Considerar una población con distribución $N(\mu, \sigma^2 = 36)$, desde la que se obtiene una muestra aleatoria de tamaño 25, en la que se encontró que $\bar{X} = 14$. Se quiere decidir acerca de las siguientes hipótesis $H_0 : \mu \geqslant 17$ versus $H_1 : \mu < 17$.

Solución: Para obtener el valor-p se requiere obtener antes Z_0 (estadístico de prueba bajo H_0):

$$Z_0 = \frac{\bar{X} - \mu}{\sigma \cdot \sqrt{n^{-1}}} = \frac{14 - 17}{6 \cdot 5^{-1}} = -2{,}5$$

$$\text{Valor-}p = \mathbb{P}(Z \leqslant -2{,}5) = \Phi(-2{,}5) = 0{,}0062.$$

Es decir, la probabilidad de que $\bar{X}$ sea menor o igual a 14 es de 0,62 %, por lo que se puede considerar altamente improbable que, basado en una

muestra de tamaño 25, se encuentre un promedio muestral de 14 o menos, cuando $\mu = 17$ (H_0 es verdadero). Es decir, cuando $\mu = 17$, solo en 62 de 10 000 muestras de tamaño 25 el valor del estadístico de prueba $\bar{X}$ será igual o menor que 14. Por lo tanto, se puede decir que hay una evidencia importante para rechazar $H_0 : \mu \geqslant 17$. □

Ahora, en el ejemplo 6.12 se puede notar además que, para cualquier nivel de significación mayor que 0,0062, se rechazará la hipótesis nula, puesto que, en este caso, el valor muestral $\bar{X} = 14$ caerá en la región crítica. Por el contrario, un valor de α menor que 0,0062 conduce a aceptar la hipótesis nula, pues el valor muestral $\bar{X} = 14$ caerá fuera de la región crítica.

En resumen, basándose en el valor-p, la decisión acerca de H_0 será:

- Rechazar H_0, si el valor-p es menor que α.
- Aceptar H_0, si el valor-p es mayor que α.

Valor-p según hipótesis unilateral y bilateral

Definición **6.8** Según las definiciones y ejercicios anteriores, calculando el valor-p a partir del estadístico de prueba, se obtiene, para una hipótesis unilateral (cola inferior o cola superior) o una bilateral, lo siguiente:

Tipo de prueba	Hipótesis	Valor-p
Unilateral	$H_0 : \theta \leqslant \theta_0$ vs $H_1 : \theta > \theta_0$	$\mathbb{P}(T(X) \geqslant T(\boldsymbol{x}))$
	$H_0 : \theta \geqslant \theta_0$ vs $H_1 : \theta < \theta_0$	$\mathbb{P}(T(X) \leqslant T(\boldsymbol{x}))$
Bilateral	$H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$	$2 \cdot \mathbb{P}(T(X) \geqslant \lvert T(\boldsymbol{x})\rvert)$

Ejemplo **6.13** En el ejemplo 6.12, si se consideran las hipótesis como $H_0 : \mu = 17$ versus $H_1 : \mu \neq 17$, entonces se trata de una hipótesis bilateral. Luego, el valor-p se obtiene por:

$$\text{Valor-}p = 2 \cdot \mathbb{P}(Z \geqslant |-2{,}5|) = 2 \cdot \mathbb{P}(Z \geqslant 2{,}5) = 2 \cdot 0{,}0062 = 0{,}0124$$

Ahora, se puede rechazar la hipótesis nula para un nivel de significación mayor o igual a 0,0124. □

Valor-p según cantidad de evidencia

Como se señaló en la presentación de esta técnica, para decidir acerca de una hipótesis estadística no se requiere del nivel de significación, pero sí al momento de hacer la comparación final y decidir sobre H_0. Sin embargo, se puede definir una regla empírica que relaciona el valor-p con la cantidad de evidencia en contra de H_0 que está contenida en la muestra; aunque no constituye una regla (pues los errores están relacionados con los problemas particulares), puede utilizarse como referencia. La escala es:

- Si el valor-$p > 0{,}10$, entonces la muestra no contiene evidencia en contra de H_0.
- Si $0{,}05 <$ valor-$p < 0{,}10$, entonces la muestra contiene evidencia débil en contra de H_0.
- Si $0{,}01 <$ valor-$p < 0{,}05$, entonces la muestra contiene evidencia fuerte en contra de H_0.
- Si el valor-$p < 0{,}01$, entonces la muestra contiene evidencia muy fuerte en contra de H_0.

Ejemplo **6.14** Se sabe que el 10 % de los huevos de una especie de pescado no madurarán. Se obtiene una muestra aleatoria de 20 huevos de esos peces, de los cuales 5 efectivamente no maduraron. Se quiere saber cuánta es la evidencia en contra de la hipótesis planteada.

Solución: En este caso, la hipótesis es $H_0 : p = 0{,}1$ versus $H_1 : p \neq 0{,}1$ y $\widehat{p} = 0{,}25$. Aquí se tiene que el estadístico de prueba es la proporción estimada, que bajo H_0 distribuye:

$$\widehat{P} \sim N\left(0{,}1; \frac{0{,}1 \cdot 0{,}9}{20}\right).$$

Entonces el valor-p (bilateral) está dado por:

$$\begin{aligned}\text{Valor} - p =& 2 \cdot \mathbb{P}\left(Z \geqslant \frac{0{,}25 - 0{,}1}{\sqrt{\frac{0{,}1 \cdot 0{,}9}{20}}}\right) \\ =& 2 \cdot \mathbb{P}\left(Z \geqslant \frac{0{,}15}{0{,}067}\right) = \Phi(2{,}24) = 0{,}0252.\end{aligned}$$

Finalmente, según la cantidad de evidencia que arroja el valor-p, se dice que la muestra contiene evidencia fuerte en contra de H_0. □

Observación **6.4** MS-Excel dispone de una función llamada `PRUEBA.T.N`, la cual permite realizar pruebas de comparación de muestras independientes y relacionadas (para la media) de una (unilateral) y dos colas (bilateral), devolviendo el valor-p según la tabla de la definición 6.8. La sintaxis de la función es:

```
PRUEBA.T.N(matriz 1; matriz 2; colas; tipo)
```

Donde `matriz 1` y `matriz 2` corresponden al primer y segundo conjunto de datos, respectivamente (muestra aleatoria 1 y 2 en caso de muestras independientes, y medición A y B en caso de muestras relacionadas). La opción `colas`, permite indicar si la prueba T es de una o dos colas (valores 1 y 2 respectivamente). El último argumento, `tipo`, es 1: muestras pareadas, 2: dos muestras independientes con varianzas iguales (homogeneidad u homoscedástica) y 3: dos muestras independientes con varianzas diferentes (heterogeneidad o heteroscedasticidad). Adaptación de la información extraída de la ayuda de la función $PRUEBA.T.N$ de MS-Excel.

6.5. Cálculo del tamaño muestral basado en α y β

El problema es obtener el menor tamaño muestral n que garantice un cierto nivel para el error tipo I y para el error tipo II, asociados a un contraste de hipótesis estadística. En efecto, si se consideran las hipótesis propuestas en el ejemplo 6.4 y utilizadas en la observación 6.2, se tendrán las ecuaciones:

$$\alpha = \mathbb{P}\left(Z > \frac{c-\mu_0}{\sqrt{\frac{\sigma_0^2}{n}}}\right) \Rightarrow z_{1-\alpha} = \frac{c-\mu_0}{\sqrt{\frac{\sigma_0^2}{n}}}$$

$$\beta = \mathbb{P}\left(Z < \frac{c-\mu_1}{\sqrt{\frac{\sigma_0^2}{n}}}\right) \Rightarrow z_{\beta} = \frac{c-\mu_1}{\sqrt{\frac{\sigma_0^2}{n}}}, \text{ con } \mu_1 > \mu_0.$$

Luego, despejando c e igualando las ecuaciones, se tiene que:

$$\mu_0 + z_{1-\alpha}\sqrt{\frac{\sigma_0^2}{n}} = \mu_1 + z_{\beta}\sqrt{\frac{\sigma_0^2}{n}}.$$

A partir de esta última ecuación se puede despejar el valor de n:

$$n = \left(\frac{z_{1-\alpha} - z_{\beta}}{\mu_1 - \mu_0} \right)^2 \sigma_0^2.$$

Ejemplo **6.15** Para el ejemplo 6.5 se cree que el verdadero valor de la media será de al menos 2, además se asume $\alpha = \beta = 0{,}05$ y $\sigma_0^2 = 25$. Se desea determinar un tamaño de muestra para esta hipótesis bajo estos niveles de error tipo I y error tipo II.

Solución: Como $\alpha = \beta = 0{,}05$, entonces $z_{1-\alpha} = z_{0{,}95} = 1{,}65$ y $z_{\beta} = z_{0{,}05} = -1{,}65$, luego se tiene que:

$$n = \frac{(1{,}65 + 1{,}65)^2}{(2-1)^2} \cdot 25 = 272{,}25.$$

Finalmente, se requiere de al menos 273 observaciones en la muestra para los niveles de error tipo I y tipo II indicados. □

6.6. Comparación de muestras independientes

En esta sección se construirán test estadísticos para diseños experimentales de comparación de muestras independientes, usando parámetros como media, varianza o proporción.

Comparación de medias

Sean $X_1, X_2, \cdots, X_{n_1}$ una muestra aleatoria donde $X_i \sim N(\mu_1, \sigma_1^2)$ y sean $Y_1, Y_2, \cdots, Y_{n_2}$ una muestra aleatoria donde $Y_i \sim N(\mu_2, \sigma_2^2)$, ambas independientes. La hipótesis que se desea contrastar es la comparación de medias, la que tiene como aplicación práctica comparar una característica cuantitativa entre dos grupos en que las mediciones son independientes.

Ejemplo **6.16** Fundamentar que la prueba de hipótesis para resolver el problema corresponde a una prueba T de muestras independientes y definir la hipótesis nula y alternativa en cada caso.

(1) Comparación de dos métodos de enseñanza, a través del rendimiento académico, utilizando dos grupos de estudiantes distintos.

(2) Se desea determinar cuál de los dos incentivos a la producción tiene mejor efecto, considerando a los operarios de una empresa divididos aleatoriamente en dos grupos.

Solución: (1) El rendimiento académico se puede obtener mediante la aplicación de una prueba estándar a ambos grupos; este proceso arroja dos muestras independientes. Nótese que la independencia entre las muestras se da porque se trata de estudiantes distintos. La hipótesis es: H_0: Los dos grupos no presentan diferencias significativas en el rendimiento académico (no hay efecto del método de enseñanza), y H_1: Hay diferencias significativas en el rendimiento de ambos grupos (hay efecto del método de enseñanza).

(2) Se tienen dos muestras independientes con las mediciones de la productividad de los operarios. La hipótesis es: H_0: No hay diferencia en la producción (el efecto que tienen los incentivos en la producción es marginal), y H_1: Hay diferencias en la producción (uno de los incentivos produce un efecto significativo en la productividad de la empresa). □

Para determinar si se observan diferencias entre los grupos se comparan las medias. Sin embargo, esta comparación está condicionada por la variabilidad de los grupos (puede darse que el promedio de uno de los grupos esté sostenido por pocas observaciones con muy alta o muy baja puntuación respecto de las restantes), por lo que construir una prueba estadística de comparación de grupos requiere asumir un supuesto acerca de la variabilidad de los grupos: grupos homogéneos o grupos heterogéneos.

Formalmente, $X_1, X_2, \cdots, X_{n_1}$ es una muestra aleatoria con $X \sim N(\mu_1, \sigma_2^2)$ e $Y_1, Y_2, \cdots, Y_{n_2}$ es una muestra aleatoria con $Y \sim N(\mu_2, \sigma_2^2)$, ambas independientes. Luego, la hipótesis estadística se plantea como:

$$H_0 : \mu_1 - \mu_2 = 0 \text{ versus } H_1 : \mu_1 - \mu_2 \neq 0.$$

Respecto de la variabilidad, se asumirá que $\sigma_1^2 = \sigma_2^2 = \sigma^2$ (homogeneidad de varianzas). Luego, la región de rechazo se definirá para:

$$\Re = \left\{ \bar{X} - \bar{Y} | \bar{X} - \bar{Y} < c_1 \vee \bar{X} - \bar{Y} > c_2 \right\}.$$

El estadístico de prueba bajo H_0 sigue una distribución t-Student y el percentil se obtiene considerando una hipótesis bilateral. En efecto, el

estadístico bajo H_0 está dado por:

$$T_0 = \frac{\bar{X} - \bar{Y}}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}, \text{ donde } S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}.$$

Finalmente, con un nivel de significación α, se rechaza H_0 si $|T_0| > t_{1-\frac{\alpha}{2}}$.

Comparación de varianzas

Esta prueba es para comprobar si dos poblaciones independientes tienen la misma varianza. La prueba de hipótesis se define a continuación.

Sean $X_1, X_2, \cdots, X_{n_1}$ una muestra aleatoria donde $X_i \sim N(\mu_1, \sigma_1^2)$ y sean $Y_1, Y_2, \cdots, Y_{n_2}$ una muestra aleatoria donde $Y_i \sim N(\mu_2, \sigma_2^2)$, ambas independientes, la hipótesis de comparación de varianzas es:

$$H_0 : \sigma_1^2 = \sigma_2^2 \text{ versus } H_1 : \sigma_1^2 \neq \sigma_2^2.$$

Desde la distribución de X e Y se tiene, respectivamente, que $F_1 = \frac{n_1-1}{\sigma_1^2}S_1^2$ y $F_2 = \frac{n_2-1}{\sigma_2^2}S_2^2$. Luego, la variable aleatoria F siguiente tiene distribución F de Fisher:

$$F = \frac{F_1}{F_2}\frac{n_2 - 1}{n_1 - 1} \sim F(n_1 - 1, n_2 - 1).$$

Luego, el estadístico de prueba, F_0, está dado por:

$$F_0 = \frac{S_1^2}{S_2^2}.$$

Como se trata de una prueba bilateral, se rechazará H_0 si:

$$F_0 < F_{\frac{\alpha}{2}}(n_1 - 1, n_2 - 1) \vee F_0 > F_{1-\frac{\alpha}{2}}(n_1 - 1, n_2 - 1).$$

Ejemplo **6.17** Para producir una cierta pieza, una compañía utiliza dos máquinas. La persona a cargo está interesada en conocer si la variabilidad entre las piezas producidas por ambas máquinas es similar. Para esto toma una muestra aleatoria para cada máquina de 10 y 20 piezas, obteniendo que la variabilidad es de 0,003 y 0,001 unidades cuadradas, respectivamente. Realizar el contraste utilizando un 95 % de confianza.

Solución: La hipótesis es de comparación de varianza para el caso bilateral (no se tiene sospecha de que una tenga mayor variabilidad por sobre la otra) y se expresa como:

$$H_0 : \sigma_1^2 = \sigma_2^2 \text{ versus } H_1 : \sigma_1^2 \neq \sigma_2^2.$$

Luego, el estadístico de prueba es:

$$F_0 = \frac{S_1^2}{S_2^2} = \frac{0{,}003}{0{,}001} = 3.$$

Finalmente, como $F_0 = 3 > F_{0,975}(9, 19) = 2{,}880$, se rechaza H_0 con un 95 % de nivel de confianza. □

Comparación de proporciones

Para este tipo de problemas es de interés comparar dos grupos independientes de observaciones en que la característica de interés es de tipo dicotómico, con una cierta probabilidad de optar por una de las opciones. Es decir, la variable aleatoria de interés sigue una distribución de Bernoulli, cuyo parámetro es la probabilidad de éxito al realizar el ensayo.

Entonces, el planteamiento de la prueba de comparación de proporciones de muestras independientes se realiza desde dos poblaciones independientes de $X \sim \text{Bernoulli}(p_1)$ e $Y \sim \text{Bernoulli}(p_2)$. La hipótesis de comparación de proporciones es:

$$H_0 : p_1 = p_2 \text{ versus } H_1 : p_1 \neq p_2.$$

Como el estimador de $p_1 - p_2$ es $\bar{X} - \bar{Y}$, para calcular la probabilidad se necesita una distribución conocida. En efecto, es posible construir dicha distribución utilizando el resultado obtenido en el ejemplo 5.8, de donde:

$$\bar{X} \sim N\left(p_1, \frac{\bar{X}(1-\bar{X})}{n_1}\right) \text{ y } \bar{Y} \sim N\left(p_2, \frac{\bar{Y}(1-\bar{Y})}{n_2}\right).$$

Entonces, se puede escribir:

$$\bar{X} - \bar{Y} \sim N\left(p_1 - p_2, \frac{\bar{X}(1-\bar{X})}{n_1} + \frac{\bar{Y}(1-\bar{Y})}{n_2}\right).$$

Luego, considerando que la hipótesis es bilateral, se tiene:

$$\begin{aligned}\frac{\alpha}{2} =& \mathbb{P}(\bar{X} - \bar{Y} > c | H_0 : p_1 - p_2 = 0) \\ =& \mathbb{P}\left(Z > \frac{c}{\sqrt{\frac{\bar{X}(1-\bar{X})}{n_1} + \frac{\bar{Y}(1-\bar{Y})}{n_2}}} \right).\end{aligned}$$

Finalmente, despejando c desde la última ecuación se tiene que se rechazará H_0 si:

$$|\bar{X} - \bar{Y}| > z_{1-\frac{\alpha}{2}} \sqrt{\frac{\bar{X}(1-\bar{X})}{n_1} + \frac{\bar{Y}(1-\bar{Y})}{n_2}}.$$

Además, del resultado anterior, se obtiene que mediante el estadístico de prueba se rechaza H_0 para:

$$|Z_0| > z_{1-\frac{\alpha}{2}}, \text{ donde } Z_0 = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\bar{X}(1-\bar{X})}{n_1} + \frac{\bar{Y}(1-\bar{Y})}{n_2}}}.$$

A continuación, se presenta una situación problemática en que la solución pasa por la construcción de un contraste de hipótesis.

Ejemplo **6.18** En un ensayo clínico para comparar dos tratamientos (nueva droga versus antigua droga) para una enfermedad cardiovascular, con la nueva droga 80 de 120 pacientes tuvieron mejora, mientras que con la antigua droga 32 de 80 pacientes presentaron mejora de la enfermedad. Aplicar una prueba de hipótesis para comparar las proporciones de pacientes que mejoraron con ambas drogas y definir el resultado acerca de la efectividad de la nueva droga frente a la antigua para mejorar la enfermedad .

Solución: Las hipótesis para la prueba que se propone es:

$$H_0 : p_1 = p_2 \text{ versus } H_1 : p_1 \neq p_2,$$

donde p_1 es la proporción de pacientes que mejoran con la nueva droga y p_2 es la proporción de pacientes que mejoran con la antigua droga. A partir de los datos se tiene que: $\bar{X} = 0{,}667$ e $\bar{Y} = 0{,}400$; además $X_i \sim$ Bernoulli(p_1) e $Y_i \sim$ Bernoulli(p_2).

Teniendo en cuenta estos antecedentes, luego de definir el error tipo I como α, se tiene que:

$$\alpha = \mathbb{P}(\bar{X} - \bar{Y} > c | p_1 - p_2 = 0) = \mathbb{P}\left(Z > \frac{c}{n^{-1}\left(\bar{X}(1-\bar{X}) + \bar{Y}(1-\bar{Y})\right)}\right).$$

Por lo tanto:

$$c = z_{1-\alpha}\sqrt{\frac{\bar{X}(1-\bar{X}) + \bar{Y}(1-\bar{Y})}{n}}.$$

Además, $\bar{X} - \bar{Y} = 0,267$ y, con un nivel de confianza de 95 %, $c = 0,079$. Con lo que se tiene que la diferencia de proporciones muestrales está en la región de rechazo de H_0. Es decir, se puede concluir que con la nueva droga se obtienen mejores resultados en comparación con la antigua droga, con un 95 % de confianza. □

6.7. Comparación de muestras dependientes

Aquellos diseños experimentales en que se quiere comparar dos mediciones que son obtenidas (sobre un mismo sujeto o unidad) antes y después de un tratamiento o intervención, se llamarán comparación de muestras dependientes.

Considerar las observaciones $(X_1, Y_1), (X_2, Y_2), \cdots, (X_n, Y_n)$, que constituyen una muestra aleatoria con las mediciones en dos momentos sobre las mismas unidades, con $(X_i, Y_i)^T \sim N_2((\mu_1, \mu_2)^T, \Sigma)$, donde Σ es una matriz de 2×2 que contiene la estructura de correlación entre las variables. Luego, se define la hipótesis de comparación de mediciones (o muestras relacionadas) como sigue:

$$H_0 : \mu_1 - \mu_2 = 0 \text{ versus } H_1 : \mu_1 - \mu_2 \neq 0.$$

Para construir el test estadístico es necesario definir una nueva variable aleatoria D, tal que $D_1, D_2, \cdots, D_n$ constituyen una muestra aleatoria, la cual se construye considerando $D_i = X_i - Y_i$, $i = 1, 2, \cdots, n$, $\bar{D} = \frac{1}{n}\sum D_i$ y $S_D^2 = \frac{1}{n-1}\sum (D_i - \bar{D})^2$. Entonces, la hipótesis de comparación de muestras relacionadas se puede expresar como:

$$H_0 : \mu_D = 0 \text{ versus } H_1 : \mu_D \neq 0.$$

Que corresponde a una prueba t-Student para una muestra (véase el ejemplo 6.10). Luego, el estadístico de prueba está dado por:

$$T_0 = \frac{\bar{D}}{S_D\sqrt{n^{-1}}} \sim \text{TS}(n-1).$$

En este caso, la hipótesis es bilateral, por lo que se rechaza H_0 para valores pequeños y grandes de T:

$$\Re = \left\{ |T| > t_{1-\frac{\alpha}{2}} \right\}.$$

Ejemplo **6.19** Se desea evaluar la eficacia de un método de intervención social. Este método está orientado a hogares que están bajo la línea de la pobreza y contiene un conjunto de acciones cuyo objetivo es que la superen al finalizar la intervención.

Por otra parte, para medir si un hogar está bajo la línea de la pobreza, se aplica un instrumento validado y probado, que arroja un puntaje (los puntajes menores indican mayor pobreza). Luego, considerando la complejidad del tema, para esta evaluación se espera aumentar el puntaje de los hogares (no se considerará la comparación de cantidad de hogares que han salido de la línea de la pobreza); la evaluación considera un puntaje inicial (antes de la intervención) y otro posterior (después de la intervención).

Solución: El problema implica la comparación de mediciones antes y después de un tratamiento sobre las mismas unidades de observación; la prueba de hipótesis para contrastar la eficacia de la intervención está dada por la media antes y después: $H_0 : \mu_D = 0$. No hay efecto de la intervención (los puntajes promedio de los hogares son iguales), y $H_1 : \mu_D \neq 0$. Hay efecto de la intervención (los puntajes promedio de los hogares presentan diferencias). □

6.8. Test de razón de verosimilitud

El método revisado anteriormente, basado en el lema de Neyman-Pearson, permite construir pruebas de máxima potencia para hipótesis simples (es decir, se conoce la distribución de las observaciones, excepto para un solo parámetro). Situaciones problemáticas en que sea más de uno el parámetro desconocido no son poco frecuentes y en aquellos casos es necesario

recurrir a otro método para obtener test estadísticos. En este sentido, el llamado test de razón de verosimilitud (TRV) es apropiado para este tipo de problemas, toda vez que funciona tanto para hipótesis simples como para hipótesis compuestas. A continuación, se define el método para construir un TRV basado en la información de una muestra aleatoria.

Sea la muestra aleatoria $X_1, X_2, \cdots, X_n$, donde la función de densidad es $f(x|\theta)$, recordando que $\Theta_0 \cup \Theta_1 = \Theta$ y $\Theta_0 \cap \Theta_1 = \phi$, y la hipótesis:

$$H_0 : \theta \in \Theta_0 \text{ versus } H_1 : \theta \in \Theta_1.$$

Se define la estadística de razón de verosimilitud por:

$$\lambda(\boldsymbol{x}) = \frac{L(\widehat{\theta})}{L(\widetilde{\theta})}.$$

En esta última expresión, $\widehat{\theta}$ es el estimador de máxima verosimilitud de θ considerando el espacio paramétrico y $\widetilde{\theta}$ es el estimador de máxima verosimilitud de θ considerando el espacio paramétrico dado por la hipótesis nula Θ_0 (también llamado bajo H_0).

Considerando un nivel de confianza $1 - \alpha$, se rechaza H_0 para los valores grandes de $\lambda(\boldsymbol{x})$, es decir:

$$\Re = \{\boldsymbol{x} | \lambda(\boldsymbol{x}) \geqslant \lambda_\alpha\}.$$

Y, λ_α se determina por:

$$\underset{\theta \in \Theta_0}{\text{máx}}\, \mathbb{P}(\Re|\theta) = \alpha.$$

Ejemplo **6.20** Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde $X \sim N(0, \sigma^2)$. Construir un test de razón de verosimilitud para la hipótesis bilateral siguiente:

$$H_0 : \sigma^2 = 1 \text{ versus } \; H_1 : \sigma^2 \neq 1.$$

Solución: Para desarrollar este tipo de test se requiere la función de verosimilitud, la cual está dada por:

$$L(\sigma^2; x) = (2\pi)^{-\frac{n}{2}} (\sigma^2)^{-\frac{n}{2}} \exp\left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^{n} x_i^2 \right\}.$$

Luego, la verosimilitud evaluada en $\sigma^2 = \widehat{\sigma^2} = s^2$ y en $\sigma^2 = \sigma_0^2 = 1$, corresponde a:

$$L(\sigma^2 = s^2) = (2\pi)^{-\frac{n}{2}} (s^2)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2s^2}\sum x_i^2\right\}.$$
$$L(\sigma^2 = 1) = (2\pi)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2}\sum x_i^2\right\}.$$

Con esto, el estadístico de razón de verosimilitud es:

$$\widehat{\lambda} = (s^2)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2}\left(1 - \frac{1}{s^2}\right)\sum x_i^2\right\}.$$

Finalmente, la región de rechazo, con un nivel de significación α, es: $\Re = \left\{\widehat{\lambda} | -2\log\widehat{\lambda} > \chi^2_{1-\alpha}(1)\right\}$. □

Test T (para una muestra)

Uno de los resultados ampliamente utilizados en la práctica es el llamado test T para una muestra (o *test T de Student*), que se puede construir utilizando un test de razón de verosimilud. En efecto, en esta subsección se presenta este resultado a partir de una muestra aleatoria extraída de una población con distribución normal, tal como sigue.

Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria con $X \sim N(\mu, \sigma^2)$, con μ y σ^2 desconocidos. La hipótesis siguiente es la que define un test T para una muestra:

$$H_0 : \mu = \mu_0 \text{ versus } H_1 : \mu \neq \mu_0.$$

Para construir el test es necesario identificar el espacio paramétrico, el cual está dado por:

$$\Theta = \left\{(\mu, \sigma^2) | -\infty < \mu < +\infty, \sigma^2 > 0\right\} \text{ y } \Theta_0 = \left\{\sigma^2 | \sigma^2 > 0\right\}.$$

Por su parte, los estimadores bajo el espacio paramétrico general, $\widehat{\mu}$ y $\widehat{\sigma}^2$, y los estimadores bajo la hipótesis nula, $\widetilde{\mu}$ y $\widetilde{\sigma}^2$, están dados respectivamente por:

$$(\widehat{\mu}, \widehat{\sigma}^2)^T = (\bar{X}, \frac{1}{n}\sum (X_i - \bar{X})^2)^T \Rightarrow n\widehat{\sigma}^2 = \sum (X_i - \bar{X})^2.$$
$$(\widetilde{\mu}, \widetilde{\sigma}^2)^T = (\mu_0, \frac{1}{n}\sum (X_i - \mu_0)^2)^T \Rightarrow n\widehat{\sigma}^2 = \sum (X_i - \mu_0)^2.$$

De esta manera, el estadístico de razón de verosimilitud, $\lambda(\boldsymbol{x})$, está dado por la expresión:

$$\begin{aligned}\lambda(\boldsymbol{x}) &= \frac{(2\pi\hat{\sigma}^2)^{-\frac{n}{2}}\exp\left\{-\frac{1}{2\hat{\sigma}^2}\sum(x_i-\bar{x})^2\right\}}{(2\pi\tilde{\sigma}^2)^{-\frac{n}{2}}\exp\left\{-\frac{1}{2\tilde{\sigma}^2}\sum(x_i-\mu_0)^2\right\}}\\ &= \left(\frac{\tilde{\sigma}^2}{\hat{\sigma}^2}\right)^{\frac{n}{2}}\\ &= \left(\frac{\sum(x_i-\mu_0)^2}{\sum(x_i-\bar{x})^2}\right)^{\frac{n}{2}}.\end{aligned}$$

Luego, la región de rechazo está dada por:

$$\Re = \left\{\boldsymbol{x}\,\middle|\,\frac{\sum(x_i-\mu_0)^2}{\sum(x_i-\bar{x})^2} \geqslant \lambda\right\}.$$

Luego, sabiendo que $\sum(x_i-\mu_0)^2 = \sum(x_i-\bar{x})^2 + n(\bar{x}-\mu_0)^2$:

$$\begin{aligned}\frac{\sum(x_i-\mu_0)^2}{\sum(x_i-\bar{x})^2} &= \frac{\sum(x_i-\bar{x})^2 + n(\bar{x}-\mu_0)^2}{\sum(x_i-\bar{x})^2}\\ &= 1 + \frac{n(\bar{x}-\mu_0)^2}{\frac{1}{n-1}\sum(x_i-\bar{x})^2}\cdot\frac{1}{n-1} \geqslant \lambda,\end{aligned}$$

donde $\frac{n(\bar{x}-\mu_0)^2}{\frac{1}{n-1}\sum(x_i-\bar{x})^2} = \left(\frac{\bar{x}-\mu_0}{s\sqrt{n^{-1}}}\right)^2 = T^2$, con $T \sim TS(n-1)$. Finalmente, se puede escribir:

$$= 1 + T^2 \geqslant \lambda^*.$$

Se rechaza H_0 si $T^2 \geqslant \lambda^{**} \Leftrightarrow |T| \geqslant \lambda_\alpha$. Como $T \sim \text{TS}(n-1)$, se rechaza H_0 si:

$$T \leqslant -t_{1-\frac{\alpha}{2}} \text{ o } T \geqslant t_{1-\frac{\alpha}{2}}.$$

Ejemplo **6.21** Se extrae de una población con distribución normal una muestra aleatoria de $n = 200$, la cual arrojó un promedio de 500 y una desviación estándar de 50. Probar la hipótesis:

$$H_0 : \mu = 510 \text{ y} \quad H_1 : \mu \neq 510.$$

Solución: Esta hipótesis está bajo las condiciones de un test T para una muestra. Luego, aquí el estadístico t es:

$$t = \frac{500-510}{50(\sqrt{200})^{-1}} = -2{,}828.$$

Finalmente, como $|t| = 2{,}828 > t_{0,975} = 1{,}972$, se puede concluir que la muestra arroja evidencia suficiente para rechazar la hipótesis nula, con un 95 % de confianza. □

Test T para muestras independientes

En esta subsección se desarrolla el llamado test T para muestras independientes o de comparación de medias de poblaciones independientes, utilizando el test de razón de verosimilitudes.

Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde $X \sim N(\mu, \sigma^2)$. Luego, la hipótesis estadística se plantea como:

$$H_0 : \mu_1 - \mu_2 = 0 \text{ versus } H_1 : \mu_1 - \mu_2 \neq 0.$$

Respecto de la variabilidad, se asumirá que $\sigma_1^2 = \sigma_2^2 = \sigma^2$ (homogeneidad de varianzas). Luego, se tiene que para la muestra completa, es decir, con $n_1 + n_2$ observaciones, el espacio paramétrico general y el restringido por la hipótesis nula es:

$$\Theta = \left\{(\mu_1, \mu_2, \sigma^2) | -\infty < \mu_1 < +\infty, -\infty < \mu_2 < +\infty, \sigma^2 > 0\right\}.$$
$$\Theta_0 = \left\{(\mu, \sigma^2) | \mu_1 = \mu_2 = \mu, -\infty < \mu < +\infty, \sigma^2 > 0\right\}.$$

Los estimadores, $(\widehat{\mu_1}, \widehat{\mu_2}, \widehat{\sigma})$ y $(\widetilde{\mu_1}, \widetilde{\mu_2}, \widetilde{\sigma})$, son:

$$(\widehat{\mu_1}, \widehat{\mu_2}, \widehat{\sigma}) = (\bar{x}, \bar{y}, \frac{1}{n_1 + n_2}\left(\sum (x_i - \bar{x})^2 + (y_i - \bar{y})^2\right)).$$
$$\widetilde{\mu} = \frac{1}{n_1 + n_2}\left(\sum x_i + \sum y_i\right) = \frac{1}{n_1 + n_2}\left(n_1\bar{x} + n_2\bar{y}\right).$$
$$\widetilde{\sigma^2} = \frac{1}{n_1 + n_2}\left(\sum (x_i - \widetilde{x})^2 + \sum (y_i - \widetilde{\mu})^2\right).$$

El estadístico de razón de verosimilitudes está dado por:

$$\begin{aligned}\lambda(\boldsymbol{x}) &= \frac{L(\boldsymbol{x}|\mu_1, \sigma^2) \cdot L(\boldsymbol{y}|\mu_2, \sigma^2)}{L(\boldsymbol{x}, \boldsymbol{y}|\mu, \sigma^2)} \\ &= \left(\frac{\widetilde{\sigma^2}}{\widehat{\sigma^2}}\right)^{-\frac{n_1+n_2}{2}} \\ &= \left(\frac{\sum (x_i - \widetilde{\mu})^2 + \sum (y_i - \widetilde{\mu})^2}{\sum (x_i - \bar{x})^2 + \sum (y_i - \bar{y})^2}\right)^{\frac{n_1+n_2}{2}},\end{aligned}$$

donde:

$$\sum (x_i - \tilde{\mu})^2 + \sum (y_i - \tilde{\mu})^2 = \sum (x_i - \bar{x})^2 + \sum (y_i - \bar{y})^2 + \frac{n_1 n_2 (\bar{x} + \bar{y})^2}{n_1 + n_2},$$

con lo cual se tiene que:

$$1 + \frac{n_1 n_2 (\bar{x} + \bar{y})^2 (n_1 + n_2)^{-1}}{\sum (x_i - \bar{x})^2 + \sum (y_i - \bar{y})^2} \geqslant \lambda^*,$$

donde, $T = \frac{\bar{x}+\bar{y}}{\sqrt{(n_1+n_2-2)^{-1}(\sum (x_i-\bar{x})^2+\sum (y_i-\bar{y})^2)}} \sim \text{TS}(n_1+n_2-2)$. Entonces:

$$1 + \frac{T^2}{n_1 + n_2 - 2} \geqslant \lambda^{**}.$$

Finalmente, para un error tipo I de α, se rechazará H_0 si $|T| > t_{1-\frac{\alpha}{2}}$.

Test de razón de verosimilitud asintótico

Cuando $n \to \infty$ (n es grande) se tiene que $\lambda(\boldsymbol{x})$ bajo H_0:

$$2 \log(\lambda(\boldsymbol{x})) \sim \chi^2(r).$$

El test de razón de verosimilitud asintótico dice que se rechaza H_0 para valores grandes de la estadística $\chi^2(r)$:

$$\Re = \left\{ \boldsymbol{x} | 2 \log(\lambda(\boldsymbol{x})) \geqslant \chi^2_{1-\alpha}(r) \right\}.$$

Donde r es el número de parámetros del espacio paramétrico general menos el número de parámetros fijos bajo el espacio paramétrico restringido por H_0.

Ejemplo **6.22** $Y_1, Y_2, \cdots, Y_n$ son variables aleatorias independientes, con:

$$f(y_i) = (\beta x_i)^{-1} \exp\left\{ -y_i (\beta x_i)^{-1} \right\},\ y_i > 0, x_i > 0 \text{ (constantes) y } \beta > 0.$$

Construir un test de razón de verosimilitud para probar la hipótesis $H_0 : \beta = 50$ versus $H_1 : \beta \neq 50$.

Solución: El estimador de β bajo el espacio paramétrico general es $\widehat{\beta}_{MV} = \frac{1}{n} \sum y_i x_i^{-1}$, mientras que bajo la hipótesis nula el único parámetro

queda especificado por $\beta = 50$. Luego, el estadístico de razón de verosimilitud está dado por:

$$\lambda(\boldsymbol{y}) = \frac{\widehat{\beta}^{-n} \exp\left\{-\sum y_i(\beta x_i)^{-1}\right\}}{50^{-n} \exp\left\{-\sum y_i(50x_i)^{-1}\right\}} = \left(\frac{\widehat{\beta}}{50}\right)^{-n} \exp\left\{\frac{1}{50}\sum \frac{y_i}{x_i} + n\right\}.$$

Luego, $2\log(\lambda(\boldsymbol{y})) = \frac{n}{2}\log(\frac{1}{50n}\sum y_i x_i^{-1}) - \frac{1}{25}\sum y_i x_i^{-1} + 2n$. Recordando que $2\log(\lambda(\boldsymbol{y})) \sim \chi^2(1)$, se tiene que la región crítica con un nivel de significación de α, es:

$$\Re = \left\{\lambda(\boldsymbol{y}) | 2\log(\lambda(\boldsymbol{y})) > \chi^2_{1-\alpha}(1)\right\}. \ \square$$

6.9. Ejercicios resueltos

1. Sean $X_1, X_2, \cdots, X_n$ e $Y_1, Y_2, \cdots, Y_n$ una muestra aleatoria, donde $X_i \sim \text{Exp}(\alpha^{-1})$ y $Y_i \sim \text{Exp}(\beta^{-1})$. Se pide:

   (a) Determinar un test de hipótesis para (con nivel de confianza del 95 %) $H_0 : \alpha = \beta$ versus $H_1 : \alpha > \beta$. Utilizar como test estadístico la diferencia de los estimadores de máxima verosimilitud de los parámetros α y β y el hecho de que estos pueden aproximarse a una distribución normal.

   (b) Sabiendo que $\bar{X} = 1{,}25$; $\bar{Y} = 0{,}75$ y $n = 100$, decidir acerca de la hipótesis propuesta en (a).

   (c) Determinar la potencia del test propuesto en (a), si $\alpha - \beta = 3$.

   Solución: (a) Para construir el test de hipótesis pedido es necesario tener un pivote. Para esto se requiere identificar la forma del parámetro bajo las hipótesis: en este caso, la forma es $\theta = \alpha - \beta$; así, es necesario obtener $\widehat{\theta}_{MV}$. En efecto, por el ejemplo 5.5 y la propiedad 5.1 se tiene que:

   $$\widehat{\theta}_{MV} = \bar{X}^{-1} - \bar{Y}^{-1}.$$

   Luego, por la propiedad 5.5 se puede obtener una aproximación a la distribución normal para $\widehat{\theta}_{MV}$. Para esto se deben calcular algunos

elementos, estos son:

$$\begin{aligned}\log f(x_i) = \log\alpha - \alpha x_i \Rightarrow \frac{\partial}{\partial\alpha}\log f(x_i) &= \alpha^{-1} - x_i \\ \Rightarrow \frac{\partial^2}{\partial\alpha^2}\log f(x_i) &= -\alpha^{-2} \\ \Rightarrow \frac{\partial^2}{\partial\beta^2}\log f(x_i) &= -\beta^{-2},\end{aligned}$$

con lo cual:

$$-nE\left(\frac{\partial^2}{\partial\alpha^2}\log f(x_i) + \frac{\partial^2}{\partial\beta^2}\log f(x_i)\right) = n\left(\alpha^{-2} + \beta^{-2}\right).$$

Luego, evaluando la última expresión en $\alpha = \widehat{\alpha}_{MV}$ y $\beta = \widehat{\beta}_{MV}$ y reemplazando en la expresión de interés:

$$Z = \frac{\bar{X}^{-1} - \bar{Y}^{-1} - (\alpha - \beta)}{\sqrt{n^{-1}(\bar{X}^{-2} - \bar{Y}^{-2})}} \dot{\sim} N(0,1).$$

Luego, la región de rechazo se construye para el $\widehat{\theta}_{MV}$ y el error tipo I, de la siguiente manera:

$$\begin{aligned}\Re &= \left\{(\bar{X}^{-1}, \bar{Y}^{-1}) : \bar{X}^{-1} - \bar{Y}^{-1} > c\right\}. \\ \mathbb{P}(\text{Error tipo I}) =& \mathbb{P}\left(Z > \frac{c}{\sqrt{n^{-1}(\bar{X}^{-2} - \bar{Y}^{-2})}}\right) \\ \Rightarrow c =& z_{1-0,05} \cdot \sqrt{n^{-1}(\bar{X}^{-2} - \bar{Y}^{-2})} \\ c =& 1{,}69 \cdot \sqrt{n^{-1}(\bar{X}^{-2} - \bar{Y}^{-2})}.\end{aligned}$$

(b) Aquí $c = 0{,}256$ y $\bar{X}^{-1} - \bar{Y}^{-1} = -0{,}533$. Por lo tanto, no se puede rechazar H_0.

(c) La potencia está dada por:

$$\begin{aligned}\text{Potencia} =& \mathbb{P}\left(Z > 0{,}256 \middle| H_0 : \alpha - \beta = 3\right) \\ =& 1 - \mathbb{P}\left(Z \leqslant -17{,}567\right) = 0{,}99.\end{aligned}$$

2. La tensión de salida de un circuito eléctrico debe ser de $130V$. Una muestra de 40 mediciones independientes de este circuito dio un

promedio de $128{,}6V$ y una desviación estándar de $2{,}1V$. Probar, usando $\alpha = 0{,}05$, si la tensión de salida es menor a la que debe ser.

Solución: De los datos entregados, se tiene $\alpha = 0{,}05$, $\sigma = 2{,}1$, $n = 40$ y $\overline{X} = 128{,}6$. La hipótesis se expresa por:

$$H_0 : \mu = 130 \text{ versus } H_1 : \mu < 130.$$

Luego, la región de rechazo es $\Re = \left\{\overline{X} | \overline{X} < c\right\}$ y c se obtiene:

$$\mathbb{P}\left(\overline{X} < c | \mu_0 = 130\right) = \alpha.$$

Como $\overline{X} \sim \mathrm{N}(\mu, \frac{sigma^2}{n} = \frac{2{,}1^2}{40})$, entonces $Z = \frac{\overline{X}-\mu}{2{,}1(\sqrt{40})^{-1}} \sim \mathrm{N}(0, 1)$:

$$\mathbb{P}\left(Z < \frac{c - \mu_0}{\sigma(\sqrt{n})^{-1}}\right) = 0{,}05.$$
$$\mathbb{P}\left(Z < \frac{c - 130}{2{,}1(\sqrt{40})^{-1}}\right) = 0{,}05.$$

Aplicando la función inversa y como $\Phi^{-1}(0{,}05) = -1{,}65$, se obtiene que $c = 130 - 1{,}65\frac{2{,}1}{\sqrt{40}} = 129{,}45$.

Con este último resultado, $\Re = \left\{\bar{X} | \bar{X} < c = 129{,}45\right\}$ y como $\bar{X} = 128{,}6 \in \Re$, se rechaza $H_0 : \mu = 130$. Es decir, hay evidencia para decir que la tensión de salida del circuito electrónico es inferior a $130V$, con un 95 % de confianza.

3. Una muestra aleatoria de $n = 200$ estudiantes que rindieron la PSU de matemática en todo el país el año 2012 tuvo un puntaje promedio de 500 puntos con una desviación estándar de 50 puntos. Asumiendo que los puntajes siguen una distribución normal, determinar:

   (a) Un intervalo de confianza para la media.

   (b) Un intervalo de confianza para la varianza.

   (c) Si hay evidencia estadística para rechazar H_0, para $H_0 : \mu = 510$ versus $H_1 : \mu \neq 510$.

   Solución: (a) De los datos $n = 200$, $\overline{X} = 500$ y $S = 50$; entonces:

$$T = \frac{\overline{X} - \mu}{S(\sqrt{n})^{-1}} \sim TS(n - 1).$$

Luego, el intervalo de confianza del 95 % para μ (verdadero valor de la media) se obtiene:

$$\mathbb{P}\left(t_{0,025} < \frac{\overline{X}-\mu}{S(\sqrt{n})^{-1}} < t_{0,975}\right) = 0{,}05.$$

Despejando μ desde la ecuación anterior, se obtiene:

$$\mathbb{P}\left(\overline{X} - t_{0,025}\frac{s}{\sqrt{n}} < \mu < \overline{X} + t_{0,975}\frac{s}{\sqrt{n}}\right) = 0{,}05$$
$$\mathbb{P}\left(500 + t_{0,025}\frac{5}{2}\sqrt{2} < \mu < 500 + t_{0,975}\frac{5}{2}\sqrt{2}\right) = 0{,}05.$$

Como $t_{0,025}(199) = -1{,}972$ y $t_{0,975}(199) = 1{,}972$; entonces $\mu \in (493{,}03;\ 506{,}97)$.

(b) Recordando que $Y = \dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$. El intervalo para σ^2 con un 95 % de confianza es:

$$\mathbb{P}\left(\chi^2_{0,025} < \frac{(n-1)S^2}{\sigma^2} < \chi^2_{0,975}\right) = \mathbb{P}\left(\frac{199\cdot 2500}{\chi^2_{0,975}} < \sigma^2 < \frac{199\cdot 2500}{\chi^2_{0,025}}\right)$$

Como $\chi^2_{0,025} = 161{,}83$ y $\chi^2_{0,975} = 239{,}96$: $\sigma^2 \in (2073{,}3;\ 3074{,}3)$.

(c) El contraste pedido es el test T (bilateral), con $\mu_0 = 510$; en este caso, se tiene que el estadístico t bajo H_0 está dado por:

$$t = \frac{500-510}{50(\sqrt{200})^{-1}} = -2{,}828.$$

Luego, como $|t| = 2{,}828 > t_{0,975} = 1{,}972$ se rechaza H_0 con un 95 % de confianza.

4. Sea la muestra aleatoria $X_1, X_2, \cdots, X_n$, donde:

$$f_{X_i}(x|\theta) = \theta_1 \exp\{-\theta_1 x\}, x > 0,$$

y sean $Y_1, Y_2, \cdots, Y_m$ con $f_{Y_i}(y) = \theta_2 \exp\{-\theta_2 y\}$, $y > 0$. Construir un test para probar:

$$H_0 : \theta_1 = \theta_2 \text{ versus } H_1 : \theta_1 \neq \theta_2.$$

Solución: Se construirá un test de razón de verosimilitudes. Así, los estimadores de máxima verosimilitud bajo el espacio paramétrico general, $\widehat{\theta_1}$ y $\widehat{\theta_1}$, y bajo el espacio paramétrico restringido por la hipótesis nula $\theta = \theta_1 = \theta_2$, $\widetilde{\theta}$:

$$\widehat{\theta_1} = (\bar{x})^{-1} \text{ y } \widehat{\theta_2} = (\bar{y})^{-1}.$$

$$\widetilde{\theta} = (m+n)(\sum x_i + \sum y_i)^{-1}.$$

$$\begin{aligned}\lambda(\boldsymbol{x},\boldsymbol{y}) &= \frac{\widehat{\theta_1}^{\,n} e^{-n} \widehat{\theta_2}^{\,m} e^{-m}}{\widetilde{\theta}^{m+n} e^{-(m+n)}} = \frac{\widehat{\theta_1}^{\,n} \widehat{\theta_2}^{\,m}}{\widetilde{\theta}^{m+n}} \\ &= \frac{(n\bar{x}+m\bar{y})^{n+m}}{(\bar{x})^n (\bar{y})^m (n+m)^{m+n}} \\ &= \frac{\left(n\frac{\bar{x}}{\bar{y}}+m\right)^{n+m}}{\left(\frac{\bar{x}}{\bar{y}}\right)^n (n+m)^{n+m}} = \varphi\left(\frac{\bar{x}}{\bar{y}}\right).\end{aligned}$$

Bajo el espacio paramétrico restringido por $H_0 : \theta = \theta_1 = \theta_2$:

$$\sum X_i \sim \text{Gamma}(n,\theta) \Rightarrow 2\theta \sum X_i \sim \chi^2(n)$$
$$\sum Y_i \sim \text{Gamma}(m,\theta) \Rightarrow 2\theta \sum Y_i \sim \chi^2(m).$$

Entonces:

$$F = \frac{2\theta \sum X_i (2n)^{-1}}{2\theta \sum Y_i (2m)^{-1}} = \frac{\bar{X}}{\bar{Y}} \sim \text{F}(2n, 2m).$$

Se rechaza H_0 para valores pequeños y grandes de $F \sim \text{F}(2n, 2m)$, con lo cual la región de rechazo es:

$$\Re = \left\{ \frac{\bar{x}}{\bar{y}} \leqslant \text{F}_{\frac{\alpha}{2}} \vee \frac{\bar{x}}{\bar{y}} \geqslant \text{F}_{\frac{1-\alpha}{2}} \right\}.$$

6.10. Ejercicios propuestos

1. Considerar la siguiente hipótesis $H_0 : \mu = 3$ versus $H_1 : \mu < 3$. La desviación estándar poblacional es $\sigma = 0{,}18$. Sobre la base de una muestra aleatoria de $n = 36$ se obtuvo la siguiente región crítica: $\Re = \{\bar{X} | \bar{X} < 2{,}92\}$. Asumir una distribución normal de la variable en la población.

(a) Calcular la probabilidad de error tipo I.

(b) Calcular la probabilidad de error tipo II.

(c) Calcular la potencia del test cuando $\mu = 2{,}85$.

2. Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde $X \sim \text{Exp}(\beta)$. Considerar la hipótesis de interés:

$$H_0 : \beta \leqslant \beta_0 \text{ versus } H_1 : \beta > \beta_0 \ \ (\beta_0 > 0).$$

(a) Considerando una probabilidad de error tipo I α, construir un test para probar la hipótesis indicada.

(b) Calcular la potencia del test para un valor determinado de β_0.

(c) Decidir acerca de $H_0 : \beta \leqslant 1$ versus $H_1 : \beta > 1$, con un nivel de significación de 5 %, si se tiene una muestra aleatoria $X_1, X_2, \cdots, X_{36}$, con $\sum X_i = X_0$.

3. Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde $X \sim \text{Exp}(\alpha)$ y sean $Y_1, Y_2, \cdots, Y_k$ una muestra aleatoria, con $Y \sim \text{Exp}(\beta)$, ambas independientes. Construir un test para probar la hipótesis: $H_0 : \alpha = \beta$ versus $H_1 : \alpha < \beta$, con un nivel de confianza de $1 - p_0$.

4. Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, donde $X_i \sim N(\mu, \sigma^2)$.

(a) Utilizando un nivel de significación $\alpha = 0{,}05$, desarrollar la hipótesis $H_0 : \mu = 0$ versus $H_1 : \mu > 0$.

(b) Determinar un valor mínimo de n necesario para asegurar una potencia de al menos 0,80 cuando $\mu = 1$.

5. Sean $X_1, X_2, \cdots, X_n$ una muestra aleatoria, con $X_i \sim \text{Bernoulli}(\theta_1)$, y sean $Y_1, Y_2, \cdots, Y_n$ una muestra aleatoria, con $Y_i \sim \text{Bernoulli}(\theta_2)$.

(a) Determinar un test estadístico apropiado y la región crítica asociada a la siguiente hipótesis estadística (con $\alpha = 0{,}05$):

$$H_0 : \theta_1 = \theta_2 \text{ versus } H_1 : \theta_2 > \theta_1.$$

(b) Asumiendo que $\alpha = 0{,}05$ y utilizando el test desarrollado en (a), obtener el valor de mínimo de n para asegurar una potencia del test de al menos un 90 %.

6. Las variables aleatorias $Y_1, Y_2, \cdots, Y_n$ son independientes, con función de densidad:

$$f(y_i) = \frac{1}{\beta x_i} \exp\left\{-\frac{1}{\beta x_i} y_i\right\}, \quad y_i > 0, \ x_i \text{ constantes positivas }, \beta > 0.$$

Construir un test de razón de verosimilitudes para: $H_0 : \beta = 50$ versus $H_1 : \beta \neq 50$.

7. Sean $Y_1, Y_2, \cdots, Y_n$ una muestra aleatoria, con $Y_i \sim N(\beta, \beta^2)$. Se tiene la hipótesis $H_0 : \beta = 1$ versus $H_1 : \beta \neq 1$. Construir un test de razón de verosimilitudes para la hipótesis mencionada.

8. Sean $Y_1, Y_2, \cdots, Y_n$ una muestra aleatoria con $Y_i \sim N(\mu, \mu\beta)$. Se tienen las hipótesis compuestas:

$$H_0 : \beta = 1 \text{ versus } H_1 : \beta \neq 1.$$

Construir el test de razón de verosimilitud para la hipótesis y comentar cómo usar este test para aceptar o rechazar H_0.

9. Se desea comparar las tasas de empleo de Temuco y Concepción. Para ello, se tomaron 2 muestras aleatorias de 100 personas en Temuco y 100 en Concepción. A partir de la muestra tomada en Temuco se encontró que 35 personas estaban desempleadas, mientras que en Concepción 17 estaban sin empleo.

   (a) Determinar la región crítica para las hipótesis:

$$H_0 : p_1 = p_2 \quad \text{versus} \quad H_1 : p_1 > p_2,$$

   donde p_1 es la proporción de desempleo en Temuco y p_2 es la proporción de desempleo en Concepción. Utilizar $\alpha = 0{,}05$.

   (b) Utilizando la información de la muestra, dar conclusiones de los resultados encontrados.

10. Un banco desea comparar las medias de los saldos de tarjetas de crédito internacionales de dos sucursales. Muestras aleatorias de clientes en ambas sucursales dieron como resultado lo siguiente:

- Sucursal 1: $n_1 = 32$, $\bar{X} =$ US 500, $S_1 =$ US 150.
- Sucursal 2: $n_2 = 36$, $\bar{Y} =$ US 375, $S_2 =$ US 130.

Asumiendo que ambas poblaciones son normales e independientes y que las varianzas son iguales pero desconocidas, se pide:

(a) Construir un test para probar si la media de los saldos de tarjetas de crédito es mayor en la sucursal 1 (H_1). Construir la región crítica (usar $\alpha = 0{,}05$).

(b) Calcular la potencia del test, si la diferencia de ambas medias poblacionales es igual a 50 dólares.

11. Se desean comparar la capacidad para frenado de automóviles de dos marcas. Muestras aleatorias de 31 vehículos de cada marca fueron probadas. La medida usada fue la distancia requerida para detenerse cuando se aplica el freno a una velocidad de 90 km/h. Las medias y varianzas muestrales fueron:

- Marca 1: $\bar{X}_1 = 118$, $S_1^2 = 102$.
- Marca 2: $\bar{X}_2 = 109$, $S_2^2 = 87$.

Suponiendo que la distancia de frenado tiene distribución normal y que ambas poblaciones tienen la misma varianza:

(a) Determinar si existe evidencia para concluir que la distancia para detenerse es diferente para las dos marcas de automóviles.

(b) Encontrar un intervalo de confianza para la varianza común a ambas poblaciones.

12. En cierto día se cambia el aceite lubricante en una máquina de avión; el nuevo aceite contiene 30 ppm de plomo. Después de 25 horas de vuelo, se sacan 11 muestras pequeñas de aceite y se queman en un espectrómetro para determinar el nivel de contaminación de plomo presente. Las lecturas observadas en el espectrómetro fueron de:

34,9	37,4	40,1	39,2	34,4	25,1
40,7	34,5	30,6	33,2	34,0.	

Plantear y probar las siguientes hipótesis:

(a) El contenido medio de plomo es de 30 ppm.

(b) La desviación estándar es a lo más de 4 ppm.

13. Una empresa compra lingotes de acero a una siderúrgica, exigiendo en las especificaciones que el peso medio sea de 100 kg con una desviación estándar de 4 kg. Al recibir una partida grande de lingotes, se toma una muestra al azar de 25 lingotes y se aceptará la partida si el peso medio observado es superior o igual a 98 kg. Determinar:

    (a) El nivel de significación (α) que implica el criterio utilizado.

    (b) La probabilidad de error tipo II si la verdadera media es de 97 kg.

    (c) La región crítica, para un nivel de significación $\alpha = 0{,}04$, una muestra de tamaño 16 y la hipótesis alternativa $\mu < 100$.

14. Se está interesado en comparar la resistencia a la tensión de dos tipos de acero producidos por una empresa siderúrgica. Para este efecto, se consideran muestras de tamaño 40 y 32, para los tipos 1 y 2, cuyas medias fueron 18,12 y 16,87 kg/cm^2, respectivamente.

    (a) Si $\sigma_1^2 = 1{,}6$ y $\sigma_2^2 = 1{,}4$, determinar si hay diferencias en la resistencia media para estos tipos de acero. Usar $\alpha = 0{,}01$.

    (b) Determinar la probabilidad de error tipo II si $\mu_1 - \mu_2 = 1$.

    (c) Para un nivel de significación $\alpha = 0{,}05$ y $\beta = 0{,}1$ cuando $\mu_1 - \mu_2 = 1$. Si $n_1 = 40$, determinar el valor de n_2.

    (d) Recalcular (a) sabiendo que $s_1 = 1{,}6$ y $s_2 = 1{,}4$.

15. La cantidad de nicotina contenida en cigarrillos de cierta marca sigue una distribución normal. Se seleccionan al azar 6 de estos cigarrillos, se mide el contenido de nicotina en mg y se registran los siguientes valores:

    20,2 19,8 18,0 17,2 18,3 18,8.

(a) Determinar si la cantidad promedio de nicotina en estos cigarrillos es superior a 18 mg.

(b) Si en (a) el valor crítico de la media es de 18,5. Determinar el error tipo I y tipo II si la verdadera media es de 18,3 mg.

16. En un estudio sobre enfermedades infantiles, se desea investigar si la incidencia de casos de contaminación por parásitos es afectada por la edad. Dos grupos de niños, con edades de 2 a 4 años (grupo 1) y de 7 a 9 años (grupo 2) fueron escogidos para ser examinados por ocurrencia de parásitos. Los datos se presentan a continuación:

Grupo	Muestra	Proporción con parasitosis
Grupo 1	120	0,083
Grupo 2	260	0,104

(a) Proponer un test de hipótesis para saber si los dos rangos de edades tienen el mismo comportamiento en cuanto a la incidencia de la enfermedad.

(b) Construir el test propuesto, determinando la región crítica.

(c) Decidir acerca de la hipótesis propuesta, con un nivel de confianza del 95 %. Interpretar.

17. Para determinar el nivel de tensión ocasionada por los exámenes escolares, 12 estudiantes fueron escogidos y se les midió la pulsación antes (A) y después (D) del examen. A continuación, se presentan los datos obtenidos por cada estudiante:

	1	2	3	4	5	6	7	8	9	10	11	12
A	87	78	85	93	76	80	82	77	91	74	76	79
D	83	84	79	88	75	81	74	71	78	73	76	71

Construir el test, para un nivel de significación del 1 %, para ver si existe mayor tensión (mayor pulsación) antes de la realización de los exámenes. Indicar los supuestos necesarios.

18. Sean $X_1, X_2, \cdots, X_n$ e $Y_1, Y_2, \cdots, Y_n$ muestras aleatorias independientes, donde la distribución de la variables aleatorias es $X_i \sim$ Poisson(λ) e $Y_i \sim$ Poisson$(\lambda(1-\theta))$. Construir un test de razón de verosimilitud para las hipótesis:

    (a) $H_0 : \binom{\lambda}{\theta} = \binom{\lambda_0}{\theta_0}$ versus $H_1 : \binom{\lambda}{\theta} \neq \binom{\lambda_0}{\theta_0}$.

    (b) $H_0 : \lambda = \theta$ versus $H_1 : \lambda \neq \theta$.

Capítulo 7

Tópicos complementarios

A continuación, se presentan tópicos denominados *complementarios*, pues se extienden más allá del tradicional curso de inferencia estadística, aunque buena parte de la teoría necesaria para su desarrollo ya fue abordada en los capítulos anteriores. En este capítulo, se exponen los fundamentos de las técnicas de análisis de datos utilizadas comúnmente.

Los tópicos tienen amplia aplicación práctica y son: correlación, regresión lineal simple y análisis de varianza (anova).

7.1. Correlación lineal

La correlación lineal es un método que se utiliza cuando se busca determinar la asociación entre dos variables cuantitativas. Tal como se dijo en el capítulo en que se trató el tema de la correlación lineal (véase la definición 4.7, para observaciones poblacionales en ese caso), esta es una medida de relación *lineal* entre dos variables en que, para el caso de datos muestrales, la expresión que determina el coeficiente r es la siguiente:

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} = \frac{S_{xy}}{\sqrt{S_{xx} \cdot S_{yy}}}. \tag{7.1}$$

Este valor r es el coeficiente de correlación muestral, entonces interesa saber si esto también ocurre en la población. Es decir, si estas dos variables bajo investigación son realmente correlacionadas (lo cual ocurrirá si $\rho \neq 0$), lo que implica que la hipótesis sea: $H_0 : \rho = 0$ y $H_1 : \rho \neq 0$. Para esta

hipótesis el estadístico de prueba es:

$$T_0 = r\sqrt{\frac{n-2}{1-r^2}}, \tag{7.2}$$

donde la variable aleatoria $T \sim \text{TS}(n-2)$. Como la prueba es de dos colas, se rechazará H_0 si $|T_0| > T_{1-\frac{\alpha}{2}}$.

En esta subsección es indistinto usar X o Y para las variables que se desea correlacionar. Sin embargo, en la subsección siguiente se buscará establecer una relación de causalidad, es decir, cuantificar cuál es el cambio observado en una de las variables cuando se cambian los valores de la otra; en tal caso, la variable que produce el cambio es X y la variable en la que se observa el cambio es Y.

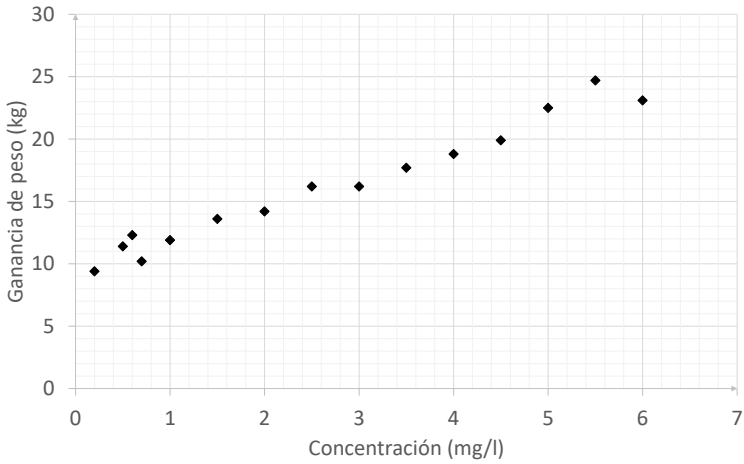
Ejemplo **7.1** En una región ganadera de Chile, el ganado alimentado con un determinado tipo de pasto tiene una ganancia de peso mayor a lo usual. Estudios de laboratorio detectaron una sustancia en el pasto y se quiere verificar si ella puede ser utilizada para mejorar la ganancia de peso de los bovinos. Para el estudio fueron escogidas 15 reces de la misma raza y edad, y cada animal recibió una determinada concentración de la sustancia (medida en mg/l). La ganancia de peso (en kg) después de 30 días fue registrada y es la siguiente:

Concentración	0,2	0,5	0,6	0,7	1,0	1,5	2,0	2,5
Ganancia	9,4	11,4	12,3	10,2	11,9	13,6	14,2	16,2
Concentración	3,0	3,5	4,0	4,5	5,0	5,5	6,0	
Ganancia	16,2	17,7	18,8	19,9	22,5	24,7	23,1	

Solución: En la tabla de datos se puede notar que, a medida que aumenta la concentración de la sustancia, aumenta la ganancia de peso. Además, al graficar estas dos variables en un sistema cartesiano, considerando cada observación de concentración de la sustancia y ganancia de peso como un par ordenado (x, y), como se aprecia en la figura 7.1, la relación que se establece entre las variables se ajusta a una relación lineal o de variación constante.

Según la expresión 7.1, el coeficiente de correlación lineal entre concentración de la sustancia (X) y ganancia de peso (Y) es $r = 0{,}985$. Posteriormente, la prueba de hipótesis para determinar si esta relación es

Figura 7.1: Ganancia de peso y concentración de sustancia



Fuente: elaboración propia.

significativa mediante el cálculo del estadístico de prueba (según la expresión 7.2) es $T_0 = 20{,}36 > T_{0,975}$, el cual lleva a rechazar H_0, lo que implica que el coeficiente de correlación en la población es distinto de cero; por lo tanto, las variables están relacionadas significativamente.

Finalmente, nótese que en el gráfico los puntos tienden a alinearse sobre una recta, cuya inclinación (pendiente) refleja el signo positivo del coeficiente de correlación calculado. □

7.2. Regresión lineal simple (RLS)

El método de la correlación entre variables cuantitativas, se utiliza cuando se busca establecer la asociación entre las variables (que no requiere proponer una dependencia entre estas). Para establecer una relación de dependencia entre las variables ya correlacionadas linealmente y cuantificar el efecto que tendría un cambio en la variable independiente (causa del

efecto) en otra variable (dependiente), el método que se utiliza se llama regresión lineal.

En el caso en estudio, la relación de dependencia entre dos variables se llama *regresión lineal simple* (RLS) y, en caso de agregar más de una variable independiente, se habla de *regresión lineal múltiple* (RLM). Para establecer una relación de dependencia entre dos variables de tipo cuantitativas y utilizar el método de RLS, se requiere que se cumplan ciertos supuestos, que consisten en que gráficamente la asociación entre las variables sea lineal y de variabilidad constante sin patrón de comportamiento.

La definición matemática del modelo de RLS, en términos generales y considerando un conjunto de valores (X_i, Y_i), $i = 1, 2, \cdots, n$, es:

$$Y_i = g(X_i) + \epsilon_i.$$

Es decir, el comportamiento de Y_i es explicado por X_i, a través de la función $g(X_i)$, y por ϵ_i. Se pueden utilizar varias opciones para la función $g(X_i)$, pero la que define la RLS es:

$$g(X_i) = \alpha + \beta X_i.$$

Además, los supuestos antes descritos se traducen en ϵ_i, $i = 1, 2, \cdots, n$, son independientes y $\epsilon_i \sim N(0, \sigma^2)$. De esta forma, se tendrá que:

$$Y_i \sim N(\alpha + \beta x_i, \sigma^2), \quad i = 1, 2, \cdots, n.$$

En regresión lineal, a la variable Y_i se la llama *variable respuesta* o *dependiente*, mientras que a la variable X_i, *variable independiente*, *explicativa* o *covariable*. Los parámetros de interés son α y β, pues tienen interpretaciones que son muy útiles en la práctica; en efecto, el parámetro α es el valor esperado para Y_i cuando X_i es igual a cero.

Para la interpretación del parámetro β, se consideran dos valores para X_i, dados por x y su sucesor $x+1$ y, siendo $E(Y_i|X_i = x)$ el valor esperado de Y_i, cuando $X_i = x$. Entonces:

$$E(Y_i|X_i = x + 1) = \alpha + \beta(x + 1) = (\alpha + \beta x) + \beta = E(Y_i|X_i = x) + \beta.$$

Luego, despejando β:

$$\beta = E(Y_i|X_i = x + 1) - E(Y_i|X_i = x).$$

De la última expresión, se tiene que β representa el cambio esperado en la variable respuesta, cuando la covariable crece en una unidad. Esto nos entrega una idea respecto de la intensidad con la cual la covariable actúa en la variable respuesta.

Estimación de parámetros de RLS

Estos valores α y β corresponden a parámetros (valores establecidos en la población), por lo que se requiere estimar sobre la base de una muestra aleatoria para tener una interpretación de ellos. Como $Y = \alpha + \beta X$, entonces, si se usan los estimadores de α y β, se tendrá la recta ajustada dada por:

$$\widehat{Y} = \widehat{\alpha} + \widehat{\beta} X_i.$$

Propiedad **7.1** Los estimadores de los parámetros de la recta de regresión se calculan usando el método de mínimos cuadrados, con lo cual se obtiene:

$$\begin{aligned}
\widehat{\alpha} =& \bar{Y} - \widehat{\beta}\bar{x}. \\
\widehat{\beta} =& \frac{\sum_{i=1}^n x_i Y_i - n\bar{x}\bar{Y}}{\sum_{i=1}^n x_i^2 - n\bar{x}^2} = \frac{\sum_{i=1}^n (x_i - \bar{x})(Y_i - \bar{Y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{S_{xy}}{S_{xx}}.
\end{aligned}$$

Demostración: Según el método de estimación planteado, la función que hay que minimizar es:

$$Q(\alpha, \beta) = \sum_{i=1}^n (y_i - E(Y_i))^2 = \sum_{i=1}^n (y_i - \alpha - \beta x_i)^2.$$

Realizando las derivadas parciales de la función Q respecto de α y β, igualando a cero ambas expresiones resultantes y resolviendo el sistema de ecuaciones, se obtienen los valores estimados para los parámetros. Por su parte, la equivalencia dada para $\widehat{\beta}$ se obtiene considerando que $\sum (x_i - \bar{x}) = 0$ y que:

$$\begin{aligned}
\sum_{i=1}^n x_i Y_i - n\bar{x}\bar{Y} &= \sum_{i=1}^n (x_i - \bar{x})(Y_i - \bar{Y}) = \sum_{i=1}^n (x_i - \bar{x}) Y_i - \bar{Y} \underbrace{\sum_{i=1}^n (x_i - \bar{x})}_{0} \\
&= \sum_{i=1}^n (x_i - \bar{x}) Y_i.
\end{aligned}$$

Gráficamente, la recta resultante al considerar los estimadores anteriores: $\widehat{y}_i = \widehat{\alpha} + \widehat{\beta} x_i$ representa la mejor recta para mostrar la asociación entre las variables X e Y sobre la base de la muestra aleatoria.

Ejemplo **7.2** A partir del ejemplo 7.1, se propone cuantificar el efecto que tiene la concentración de sustancia en el pasto en la ganancia de peso mediante el método de RLS.

Solución: Para responder a este problema, se requiere calcular los estimadores de los parámetros que definen el modelo de asociación de RLS. Posteriormente, la interpretación de los mismos nos ayudará a responder la pregunta.

En efecto, para obtener el ajuste de RLS, el cálculo de los estimadores de α y β se hará mediante el resultado dado por la propiedad 7.1. Los datos obtenidos del ejemplo son:

$$n = 15;\ \ \sum_{i=1}^{n} x_i y_i = 785{,}55;\ \ \sum_{i=1}^{n} x_i^2 = 163{,}39;\ \ \bar{x} = 2{,}70;\ \ \bar{y} = 16{,}14.$$

Luego, los estimadores de α y β son:

$$\widehat{\alpha} = 9{,}55 \text{ y } \widehat{\beta} = 2{,}44.$$

Por lo tanto, la recta de regresión ajustada es:

$$\widehat{y}_i = 9{,}55 + 2{,}44\, x_i.$$

El gráfico de la recta ajustada se observa en la figura 7.2, que muestra que la recta representa los puntos correspondientes a los datos. La interpretación de los parámetros estimados es:

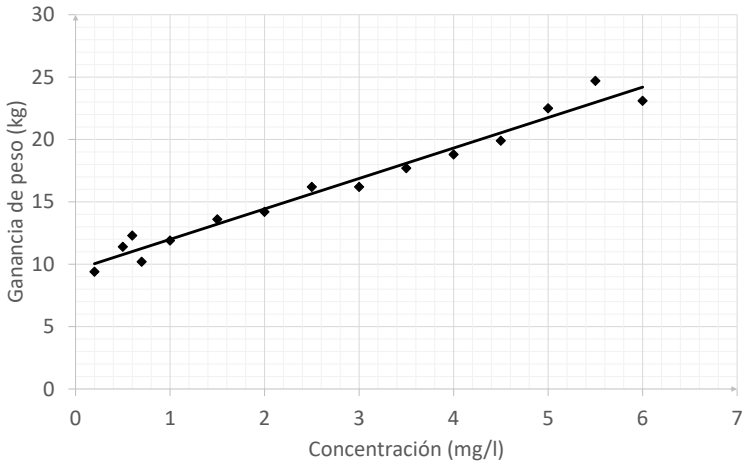
($\widehat{\alpha}$) La ganancia de peso esperada para los bovinos que no reciben sustancia, es decir, $X = 0$, es de 9,55 kg; por otra parte:

($\widehat{\beta}$) El aumento en un 1 mg/l en la concentración de la sustancia implica una ganancia de peso esperada de 2,44 kg. □

Ejercicio propuesto **7.1** Demostrar los siguientes resultados acerca de los estimadores de mínimos cuadrados de α y β (según la propiedad 7.1):

(1) Son estimadores insesgados.

Figura 7.2: Gráfico de concentración y ganancia de peso, con línea de regresión



Fuente: elaboración propia.

(2) La varianza de ambos está dada, respectivamente, por:

$$Var[\hat{\alpha}] = \frac{\sigma^2}{n}\frac{\sum x_i^2}{\sum (x_i - \bar{x})^2} \text{ y } Var[\hat{\beta}] = \frac{\sigma^2}{\sum (x_i - \bar{x})^2}.$$

(3) Estos estimadores están correlacionados, a menos que $\bar{x} = 0$, en efecto:

$$Cov(\hat{\alpha}, \hat{\beta}) = \frac{-\bar{x}\sigma^2}{\sum (x_i - \bar{x})^2}.$$

(4) Si $e_i \sim N(0, \sigma^2)$, entonces $\hat{\beta}$ sigue una distribución normal (bajo el mismo supuesto y bajo la misma manera, se puede mostrar que $\hat{\alpha}$ sigue una distribución normal).

(5) $S^2 = \frac{1}{n-2}\sum (Y_i - \hat{Y})^2$ es un estimador insesgado para σ^2 y $\frac{n-2}{\sigma^2}S^2 \sim \chi^2(n-2)$.

Indicaciones: En (2), utilizar la segunda expresión para $\hat{\beta}$ de la propiedad 7.1. En (3), para calcular la covarianza entre $\bar{Y}$ y $\hat{\beta}$, definir $c_i = \frac{x_i - \bar{x}}{S_{xx}}$ y utilizar el resultado del ejercicio 7 (sección 6.10, capítulo 4). En (4), definir c_i como en (3) y utilizar la función generadora de momentos.

Pruebas de hipótesis para los parámetros de RLS

Según el modelo de RLS, para que la variable independiente tenga influencia en la dependiente tiene que cumplirse que β sea distinto de cero. De este análisis surge la principal prueba de hipótesis:

$$H_0 : \beta = 0 \text{ versus } H_1 : \beta \neq 0. \tag{7.3}$$

El estadístico de prueba asociado a esta hipótesis se obtiene por:

$$\widehat{\beta} \sim N(\beta, \frac{\sigma^2}{S_{xx}}) \Rightarrow Z = \frac{\widehat{\beta} - \beta}{\sqrt{\frac{\sigma^2}{S_{xx}}}} \sim N(0,1);$$

$$\frac{n-2}{\sigma^2} S^2 \sim \chi^2(n-2), \text{ con } S^2 = \frac{1}{n-2} \sum (Y_i - \bar{Y})^2.$$

Con estos resultados, el estadístico de prueba corresponde a:

$$T|H_0 = T_0 = \frac{\widehat{\beta} - \cancel{\beta}}{S\sqrt{\frac{1}{S_{xx}}}} = \frac{\widehat{\beta}}{S\sqrt{\frac{1}{S_{xx}}}} \sim \text{TS}(n-2).$$

Con lo cual se rechazará H_0 si $|T_0| > T_{1-\frac{\alpha}{2}}$.

También, utilizando las mismas consideraciones, es posible identificar un estadístico de prueba para la hipótesis referida al parámetro α, $H_0 : \alpha = 0$ y $H_1 : \alpha \neq 0$, el cual corresponde a:

$$T|H_0 = T_0 = \frac{\widehat{\alpha} - \cancel{\alpha}}{S\sqrt{\frac{\sum x_i^2}{n\, S_{xx}}}} = \frac{\widehat{\alpha}}{S\sqrt{\frac{\sum x_i^2}{n\, S_{xx}}}} \sim \text{TS}(n-2).$$

Ejemplo **7.3** Considerando el ejemplo 7.2, en el cual se estableció el efecto de la concentración de sustancia (X) en la ganancia de peso (Y) ajustando una recta de regresión. Para verificar la evidencia estadística de la regresión, se realiza la prueba de hipótesis dada en la expresión 7.3.

Para obtener el estadístico de prueba, se mostrará la obtención de los resultados necesarios en la tabla de datos:

i	x_i	y_i	$(Y_i - \widehat{Y})^2$	$(x_i - \bar{x})^2$
1	0,2	9,4	0,41	6,25
2	0,5	11,4	0,40	4,84
3	0,6	12,3	1,65	4,41

4	0,7	10,2	1,12	4,00
5	1	11,9	0,01	2,89
6	1,5	13,6	0,15	1,44
7	2	14,2	0,05	0,49
8	2,5	16,2	0,30	0,04
9	3	16,2	0,45	0,09
10	3,5	17,7	0,15	0,64
11	4	18,8	0,26	1,69
12	4,5	19,9	0,40	3,24
13	5	22,5	0,56	5,29
14	5,5	24,7	2,99	7,84
15	6	23,1	1,19	10,89
Total	40,5	242,1	10,09	54,04

Con estos resultados, se tiene que: $S^2 = 0{,}78$ y $S_{xx} = 54{,}04$; además, se sabe que $\widehat{\beta} = 2{,}44$. Luego, $T_0 = 20{,}33 > T_{0{,}975} = 2{,}160$, con lo cual se rechaza H_0 con un nivel de confianza del 95 %, por lo que se concluye que la concentración de la sustancia afecta significativamente la ganancia de peso de los bovinos. □

Observación **7.1** Intervalo de confianza de $(1 - p_0)100\,\%$ para α y β:

Basándose en el estadístico de prueba T_0 y siguiendo el método dado en el capítulo 5, se puede obtener un intervalo de confianza para los parámetros de la recta de regresión lineal simple, bajo un nivel de confianza de $(1 - p_0)100\,\%$ (con $0 \leqslant p_0 \leqslant 1$). En efecto, el pivote corresponde al estadístico utilizado para la prueba de hipótesis a la que se hace referencia; con esto, los intervalos están dados por:

$$\alpha \in \left(\widehat{\alpha} - t_{1-\frac{p_0}{2}} S \sqrt{\frac{\sum x_i^2}{n\, S_{xx}}}, \widehat{\alpha} + t_{1-\frac{p_0}{2}} S \sqrt{\frac{\sum x_i^2}{n\, S_{xx}}} \right).$$

$$\beta \in \left(\widehat{\beta} - t_{1-\frac{p_0}{2}} S \sqrt{\frac{1}{S_{xx}}}, \widehat{\beta} + t_{1-\frac{p_0}{2}} S \sqrt{\frac{1}{S_{xx}}} \right).$$

Ejemplo **7.4** Los siguientes datos corresponden a dos variables X e Y organizados como pares ordenados (x, y): $(-2, 0), (-1, 0), (0, 1), (1, 1), (2, 3)$.

Considerando esta información:

(a) Ajustar la recta de regresión con el método de mínimos cuadrados.

(b) Determinar la varianza de los estimadores obtenidos en (a) y la covarianza entre estos.

(c) Estimar la varianza σ^2.

(d) Determinar si los datos presentan evidencia suficiente para decir que la pendiente difiere de 0 (usar un nivel de significancia de 0,05).

(e) Obtener un intervalo de confianza del 95 % para α y β.

Solución: (a) Para ajustar la recta, hay que obtener los estimadores de α y β; según la propiedad 7.1, se necesitan obtener algunas funciones de la muestra, las que se presentan a continuación ordenando los datos en una tabla:

x_i	y_i	$x_i y_i$	x_i^2
-2	0	0	4
-1	0	0	1
0	1	0	0
1	1	1	1
2	3	6	4
$\sum x_i = 0$	$\sum y_i = 5$	$\sum x_i y_i = 7$	$\sum x_i^2 = 10$

Con esto, $\widehat{\alpha} = \frac{5}{5} - \frac{7}{10}0 = 1$ y $\widehat{\beta} = \frac{7-0}{10-0} = 0{,}7$. Es decir, $\widehat{Y} = 1 + 0{,}7 \cdot X_i$.

(b) $Var[\widehat{\alpha}] = \frac{1}{5}\sigma^2$, $Var[\widehat{\beta}] = \frac{1}{10}\sigma^2$ $Cov(\widehat{\alpha}, \widehat{\beta}) = 0$; este último resultado es debido a que $\sum x_i = 0$.

(c) Para estimar σ^2 se necesitan conocer algunos resultados: $\sum y_i^2 = 11$, $S_{yy} = \sum y_i^2 - n\bar{y}^2 = 11 - 5 \cdot 1^2 = 6{,}0$, con lo cual se tiene que el estimador de σ^2 es:

$$S^2 = \frac{1}{5-2}(6{,}0 - 0{,}7 \cdot 7) = 0{,}367.$$

(d) La prueba que se pide hace referencia a la hipótesis $H_0 : \beta = 0$ versus $H_1 : \beta \neq 0$. Con los resultados de los puntos anteriores, el estadístico de prueba es:

$$T_0 = \frac{\widehat{\beta} - \beta_0}{s\sqrt{\frac{1}{S_{xx}}}} = \frac{0{,}7 - 0}{0{,}606\sqrt{0{,}10}} = 3{,}65.$$

Por su parte, $t_{0,975}(3) = 3{,}182$, el cual es inferior a T_0, por lo que se rechaza H_0, es decir, hay evidencia suficiente para decir que la pendiente difiere de 0.

(e) Los intervalos de confianza corresponden, respectivamente, a:

α: $1{,}0 \pm 3{,}182 \cdot 0{,}606\sqrt{0{,}20} \Rightarrow \alpha \in (0{,}090; 1{,}310)$.

β: $0{,}7 \pm 3{,}182 \cdot 0{,}606\sqrt{0{,}10} \Rightarrow \beta \in (0{,}138; 1{,}862)$. □

7.3. Análisis de varianza (anova)

Este tipo de problemas, en términos generales, son una extensión del problema estudiado en el capítulo 6 sobre comparación de muestras independientes (sección 6.6), cuando la comparación es de dos o más muestras independientes. Para la aplicación del método, se define una variable Y en la que se estudian los cambios que son atribuibles a una variable X cualitativa con más de dos categorías, llamada factor.

Al proponer que hay diferencias en Y dadas por el factor X, se entiende que cada categoría del factor define una población desde la cual provienen las observaciones Y, de tal manera que puede escribirse Y_{ij}, donde i corresponde a cada uno de los sujetos y j a cada una de las poblaciones o grupos dados por X.

La hipótesis del investigador corresponde a diferencias entre los grupos (al menos uno presenta diferencias respecto a los otros grupos), por lo que la hipótesis nula es que los grupos no presentan diferencias en la mediciones de la variable aleatoria Y, lo cual se puede expresar en función de la media poblacional:

$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$, y

H_1 : por lo menos una de las medias μ_i es diferente de las demás.

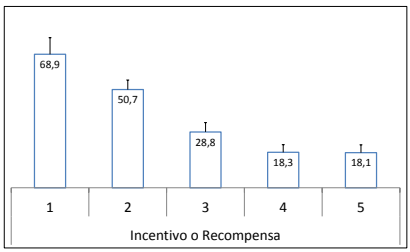
Ejemplo **7.5** Una gran empresa está interesada en elevar la producción, para lo cual desea implementar un sistema de incentivos destinado a sus operarios. La experiencia de otras empresas similares muestra que hay 5 tipos de incentivos que han presentado los mejores resultados en cuanto a elevar la producción. Los incentivos o recompensas son: *dinero*, *cambio de turno* (o de departamento), *descanso remunerado*, *más tiempo para el café* y *una recompensa sorpresa*.

Los profesionales a cargo de la implementación han decidido probar cuál de estos incentivos se adapta mejor a las condiciones de la empresa pensando en el objetivo final: elevar la producción, para lo cual han realizado un diseño experimental que consiste en asignar aleatoriamente a 10 empleados cada incentivo durante una semana, prometiéndoles las recompensas antes señaladas si aumentan la producción; por su parte, para medir el aumento en el volumen de producción, se considera la diferencia entre la producción durante la semana del experimento y la producción promedio por semana durante el mes anterior de cada uno de los empleados de la muestra.

Finalizado el tiempo definido por el experimento, se recoge la información sobre la producción. Los resultados que se observan en la figura 7.3 indican, efectivamente, que hay diferencias en la producción con las distintas recompensas: las recompensas 1 y 2 (dinero y cambio de turno o departamento) son las que generan mayor aumento de la producción.

Figura 7.3: Estadísticos descriptivos para producción por tipo de incentivo

	n	Media	D.E.
1	10	68.9	8.57
2	10	50.7	4.92
3	10	28.8	4.87
4	10	18.3	3.97
5	10	18.1	4.12



Fuente: elaboración propia.

El problema que se presenta a continuación es determinar si las diferencias observadas en los datos descriptivos son significativas al nivel de confianza establecido por la empresa.

Plantear una prueba de hipótesis resuelve este último problema, pues permite probar si la diferencia observada mediante el promedio es significativa. La hipótesis nula es que los 5 incentivos generan la misma producción y la hipótesis alternativa es que al menos 1 de los incentivos genera una producción distinta.

En efecto, sean las muestras aleatorias $Y_{1j}, Y_{2j}, \cdots, Y_{nj}$, $j = 1, 2, 3, 4, 5$,

donde las 5 muestras son independientes, con $Y_{ij} \sim N(\mu_j, \sigma_j^2)$, con $i = 1, 2, \cdots, n$ para $j = 1, 2, 3, 4, 5$, entonces la hipótesis es:

$$\begin{aligned} &H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 \\ &H_1 : \mu_k \neq \mu_l \text{ para algún } k \neq l, 1 \leqslant k, l \leqslant 5. \end{aligned} \tag{7.4}$$

Finalmente, para la solución del problema se requiere un método que permita construir el test para las hipótesis propuestas. □

La construcción del test de hipótesis definido anteriormente requiere aplicar un método conocido como *análisis de varianza de un factor*, donde el factor corresponde a la variable X (en el caso del problema citado son los grupos definidos por cada uno de los incentivos).

Definición **7.1** El análisis de varianza (anova) consiste en la descomposición de las observaciones Y_{ij} (i es el indicador de sujetos y j es el indicador de los grupos) en contribuciones de diferentes fuentes. Se pueden expresar las observaciones en función de:

$$Y_{ij} = \bar{Y} + (\bar{Y}_j - \bar{Y}) + (Y_{ij} - \bar{Y}_j), \quad i = 1, 2, \cdots, n, \quad j = 1, 2, \cdots, k. \tag{7.5}$$

Nótese que cada uno de los k grupos tiene n_j sujetos, es decir, $n = n_1 + \cdots + n_k$.

La expresión 7.5 señala que cada observación se puede descomponer en tres fuentes identificables:

$\bar{Y}$	:	el promedio general.
$\bar{Y}_j - \bar{Y}$	:	la diferencia entre el promedio del grupo y el general (variabilidad intergrupos).
$Y_{ij} - \bar{Y}_j$	:	la diferencia entre cada observación y el promedio del grupo correspondiente (variabilidad intragrupo).

Además, de la expresión 7.5 se puede obtener otra expresión sobre la base de la cual se realiza la prueba de hipótesis, esta es:

$$Y_{ij} - \bar{Y} = (\bar{Y}_j - \bar{Y}) + (Y_{ij} - \bar{Y}_j).$$

Elevando al cuadrado y aplicando doble sumatoria, se obtiene:

$$\sum_{j=1}^{k} \sum_{i=1}^{n_j} \left(Y_{ij} - \bar{Y}\right)^2 = \sum_{j=1}^{k} \sum_{i=1}^{n_j} \left(\bar{Y}_j - \bar{Y}\right)^2 + \sum_{j=1}^{k} \sum_{i=1}^{n_j} \left(Y_{ij} - \bar{Y}_j\right)^2.$$

La expresión anterior se escribe utilizando siglas, estas son:

$$SCT = SCTr + SCE.$$

Donde SCT: suma de cuadrados total, $SCTr$: suma de cuadrados del tratamiento (atribuible al factor) y SCE: suma de cuadrados del error (atribuible a diferencias naturales entre sujetos dentro de un grupo). Estos resultados se organizan en una tabla llamada *tabla de análisis de varianza*, que se compone de la siguiente manera:

Fuente de variación	SC	Gl	CM
Tratamiento	$SCTr$	$k-1$	$CMTr = SCTr/(k-1)$
Error	SCE	$n-k$	$CME = SCE/(n-k)$
Total	SCT	$n-1$	

Luego, el estadístico de prueba para la hipótesis de comparación de medias se obtiene desde la distribución F de Fisher, el cual es:

$$F_0 = \frac{CMTr}{CME} \sim F(k-1, n-k).$$

Con lo cual se rechaza H_0 si $F_0 > F_{1-\alpha}(k-1, n-k)$.

Ejemplo **7.6** Para la situación en estudio (ejemplo 7.5), construir la tabla de análisis de varianza considerando los datos obtenidos que se presentan a continuación para cada incentivo o recompensa:

Recompensa	**Observaciones** (producción)									
1	76	70	59	77	59	69	80	78	61	60
2	52	52	43	48	43	56	52	53	58	50
3	37	26	28	38	25	30	28	25	26	25
4	19	21	16	23	23	23	14	15	13	16
5	11	15	23	15	25	18	20	18	20	16

Solución: Los cálculos necesarios para completar la tabla de análisis de varianza son:

$$\sum_{j=1}^{k}\sum_{i=1}^{n_j}\left(\bar{Y}_j - \bar{Y}\right)^2 = 20180{,}5 \quad \text{y} \quad \sum_{j=1}^{k}\sum_{i=1}^{n_j}\left(Y_{ij} - \bar{Y}_j\right)^2 = 1313{,}2.$$

Con estos resultados se obtiene la tabla de análisis de varianza:

Fuente de variación	SC	Gl	CM	F_0
Tratamiento	20180.5	4	5045.1	172.90
Error	1313.2	45	29.2	
Total	21493.6	49		

Si $\alpha = 0{,}05$ (nivel de confianza del 95 %), entonces $F_{0,95}(4, 45) = 2,58 < F_0 = 172{,}90$, por lo que se rechaza H_0, es decir, con los datos entregados se puede concluir que las recompensas tienen diferentes efectos en la producción. □

Extensión: comparaciones *a posteriori*

Luego de aplicar la técnica de análisis de varianza de un factor, se obtiene la definición acerca de si los grupos tienen efectos diferentes en la variable de interés. En caso de rechazar la hipótesis nula, se llega a la conclusión de que efectivamente se observa un efecto diferente en la variable de interés, entonces surge la pregunta acerca de cuál o cuáles grupos que están produciendo tal diferencia.

Para ayudar a responder esa pregunta, están los test llamados *comparaciones a posteriori*, pues se aplican una vez que se conoce que hay diferencias entre grupos por la prueba anova de un factor. A continuación, se presentan dos técnicas para realizar pruebas *a posteriori*: método de Tukey y método de Bonferroni.

Método de Tukey

Este método se basa en comparar las diferencias (de a pares) entre las medias de cada grupo y un estadístico llamado DHS, que se calcula fijando un nivel de significación y el cuadrado medio del tratamiento (CME en la tabla de análisis de varianza). El criterio es: cada diferencia que supere a DHS indica que son significativamente distintos los grupos; en caso contrario, se dice que no hay evidencia suficiente para señalar que los grupos son significativamente distintos.

En el cálculo de DHS interviene una variable aleatoria llamada *rango estudentizado*, que se denota con $q(r, v)$ y cuya distribución se encuentra

tabulada para distintos valores de r y v. La expresión para DHS es:

$$DHS = q_\alpha(k-1, n-k) \cdot \frac{1}{\sqrt{2}} \cdot \sqrt{\frac{2 \cdot CME}{n_j}}.$$

Nótese que se aplica cuando los grupos tienen igual tamaño n_j (o parecido).

Ejemplo **7.7** A partir del ejemplo 7.6, se determinará cuál de las recompensas ofrece un efecto distinto del resto, es decir, la recompensa que ofrece significativamente mayor producción.

Solución: Para resolver este problema se utilizará el método de Tukey, comenzando por obtener DHS con $\alpha = 0{,}01$:

$$DHS = q_{0,01}(4, 45) \cdot \frac{1}{\sqrt{2}} \cdot \sqrt{\frac{2 \cdot 29{,}18}{10}} = 6{,}87.$$

Las diferencias de medias entre pares de grupos son:

	$\bar{Y_2}$	$\bar{Y_3}$	$\bar{Y_4}$	$\bar{Y_5}$
$\bar{Y_1}$	18,2	40,7	51,2	51,4
$\bar{Y_2}$		21,9	32,4	32,6
$\bar{Y_3}$			10,5	10,7
$\bar{Y_4}$				0,2

Al realizar la comparación con DHS, se observa que la única diferencia de grupos que no lo supera es entre los grupos 4 y 5, es decir, que las recompensas 4 y 5 ofrecen efectos similares en la producción, mientras que las demás recompensas ofrecen efectos distintos entre sí, con un nivel de significación de 0,01. □

Intervalos múltiples t de Bonferroni

Este método consiste en construir intervalos de confianza para las diferencias de pares de grupos $\mu_i - \mu_j$, con $1-\alpha$ de nivel de confianza, utilizando la distribución t-Student. Cada intervalo de confianza es:

$$\bar{Y_i} - \bar{Y_j} \pm t_{1-\frac{\alpha}{2m}}(n-k) \cdot \sqrt{CME\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}.$$

En la expresión anterior, m representa la cantidad de intervalos de confianza simultáneos y los índices i y j corresponden a los grupos, es decir, $1 \leqslant i, j \leqslant k$, con $i \neq j$. Además, la expresión que genera los límites inferior y superior de los intervalos se denomina *error de estimación*.

Para decidir se evalúa si el intervalo obtenido contiene el valor 0. Si es así, entonces esos grupos no presentan diferencias significativas; en caso contrario, se dirá que los grupos son significativamente distintos.

Ejemplo **7.8** Para el ejemplo 7.6 se desea determinar cuál de las recompensas ofrece un efecto distinto del resto, esta vez aplicando el método de los intervalos múltiples de Bonferroni.

Solución: Para utilizar el método de intervalos múltiples t de Bonferroni, se tiene que $m = 10$, luego se requiere construir esta cantidad de intervalos y si $\alpha = 0{,}05$, se tiene que $t_{0{,}9975}(45) = 3{,}1513$, con lo cual se tiene que el error de estimación, EE, es:

$$EE = 3{,}1513 \cdot \sqrt{29{,}18\left(\frac{1}{10} + \frac{1}{10}\right)} = 7{,}613.$$

	Diferencia de medias		EE	Intervalo de confianza	
$\mu_1 - \mu_2$	18,2	$\pm$	7,613	10,587	25,813
$\mu_1 - \mu_3$	40,7	$\pm$	7,613	33,087	48,313
$\mu_1 - \mu_4$	51,2	$\pm$	7,613	43,587	58,813
$\mu_1 - \mu_5$	51,4	$\pm$	7,613	43,787	59,013
$\mu_2 - \mu_3$	21,9	$\pm$	7,613	14,287	29,513
$\mu_2 - \mu_4$	32,4	$\pm$	7,613	24,787	40,013
$\mu_2 - \mu_5$	32,6	$\pm$	7,613	24,987	40,213
$\mu_3 - \mu_4$	10,5	$\pm$	7,613	2,887	18,113
$\mu_3 - \mu_5$	10,7	$\pm$	7,613	3,087	18,313
$\mu_4 - \mu_5$	0,2	$\pm$	7,613	$-7{,}413$	7,813

Con esto, se obtiene que la única comparación que no ofrece diferencias significativas es la de los grupos 4 y 5, es decir, las recompensas 4 y 5 son las únicas que no generan diferencias en la producción. □

7.4. Asociación entre variables cualitativas

Otra de las posibilidades en el estudio de la asociación entre variables está dada por dos variables de tipo cualitativo con r y s categorías de respuesta. Teniendo como hipótesis de investigación que las variables se encuentran relacionadas o no son independientes.

Para contextualizar, consideremos el siguiente ejemplo.

Ejemplo **7.9** Se cree que hay relación entre los resultados obtenidos por los estudiantes de enseñanza media en las asignaturas de Matemáticas y de Física. Para comprobarlo, se revisó el rendimiento de 528 estudiantes y se categorizó el rendimiento en *bajo*, *medio* o *alto*; luego se tiene el resultado para los estudiantes en las dos variables ya definidas. Para este problema, describir las variables y definir la hipótesis.

Solución: Las variables involucradas, ambas cualitativas, son *rendimiento en Matemáticas (bajo, medio o alto)* y *rendimiento en Física (bajo, medio o alto)*. Según los logros observados en los 528 estudiantes, cada uno se clasificó en alguna de las nueve combinaciones posibles de acuerdo con las categorías conjuntas de ambas variables. Por su parte, como el estudio es de relación, la hipótesis es:

H_0: El rendimiento en ambas asignaturas es independiente.

H_1: Existe relación en el rendimiento de ambas asignaturas. □

En general, el experimento consiste en ensayos que cumplen con:

(1) Se realizan n ensayos idénticos (n corresponde al tamaño de muestra).

(2) El resultado de cada ensayo corresponde a una de k celdas dadas por todas las combinaciones posibles en las categorías de ambas variables.

(3) La probabilidad de que un solo ensayo se corresponda con una celda particular es p_{ij}, donde (i, j) representa la celda, $i = 1, 2, \cdots, r$ y $j = 1, 2, \cdots, s$. Luego se observa que:

$$p_{11} + p_{12} + \cdots + p_{rs} = \sum\sum p_{ij} = 1.$$

(4) Los ensayos son independientes.

(5) Las cantidades n_{ij} representan el número de ensayos que se corresponden con la celda (i, j). Nótese también que:

$$n_{11} + n_{12} + \cdots + n_{rs} = \sum\sum n_{ij} = n.$$

Este tipo de experientos genera resultados que pueden organizarse en una tabla de contingencia o tabla de doble entrada, tal como se presenta a continuación:

Variable 1	Variable 2				
	Cat. 1	Cat. 2	$\cdots$	Cat. s	Total fila
Cat. 1	n_{11}	n_{12}		n_{1s}	$\sum n_{1i}$
Cat. 2	n_{21}	n_{22}		n_{2s}	$\sum n_{2i}$
$\vdots$				$\ddots$	
Cat. r	n_{r1}	n_{r2}		n_{rs}	$\sum n_{2i}$
Total columna	$\sum n_{i1}$	$\sum n_{i2}$		$\sum n_{is}$	n

Observar que $p_{ij} = \frac{n_{ij}}{n}$ corresponde a la probabilidad de ocurrencia del evento de la celda ij.

Ejemplo **7.10** Del ejemplo 7.9, se tiene que al realizar la clasificación de acuerdo con el experimento descrito se han obtenido los siguientes resultados:

Física	Matemática			
	Alto	Medio	Bajo	Total
Alto	56	71	12	139
Medio	47	163	38	248
Bajo	14	42	85	141
Total	117	276	135	528

Identificar en la tabla los elementos descritos anteriormente y graficar aportando evidencia a la hipótesis planteada.

Solución: En este caso, $r = s = 3$ y $k = r \cdot s = 9$, $n = 528$. Además, se tiene que, por ejemplo, $p_{23} = \frac{42}{528} = 0{,}0795$. □

Nótese que, si las variables son independientes, entonces se espera que la probabilidad de ocurrencia de la celda ij corresponda al producto entre el total de la fila y el total de la columna respectiva. Luego, bajo el mismo

supuesto de que se cumpla la hipótesis nula, el número esperado de ensayos correspondientes a la celda ij es:

$$E[n_{ij}] = (p_{ri} \cdot p_{is})n = \frac{\sum n_{ri}}{n} \cdot \frac{\sum n_{is}}{n} n = \frac{\sum n_{ri} \cdot \sum n_{is}}{n}$$
$$= \frac{\text{Total fila} \cdot \text{Total columna}}{\text{Total general}}.$$

Con esto, se tendrá que el estadístico de prueba está dado por:

$$Q_0^2 = \sum_{i=1}^{r} \sum_{j=1}^{s} \frac{(n_{ij} - E[n_{ij}])^2}{E[n_{ij}]}, \tag{7.6}$$

donde $Q^2 \sim \chi^2((r-1)(s-1))$.

Para la demostración de la expresión 7.6 y la distribución respectiva, se utiliza la distribución binomial y su aproximación a la distribución normal estándar a través del teorema central del límite. Recuérdese que la distribución normal estándar se puede transformar en una distribución chi-cuadrado. Para detalles sobre esta demostración, véase Dennis D. Wackerly, William Mendenhall, Richard L. Scheaffer y Jorge H. Muñoz. *Estadística matemática con aplicaciones*. Thomson, 2002.

El criterio de rechazo establece que se rechaza H_0 para valores grandes de Q_0^2. Es decir, basándose en un nivel de significancia α, se rechazará H_0 si $Q_0^2 \geqslant \chi_{1-\alpha}^2$.

Ejemplo **7.11** A partir del ejemplo 7.10, probar si la relación entre el rendimiento en Matemáticas y Física es estadísticamente significativa.

Solución: Para esta hipótesis, el estadístico de prueba corresponde a la expresión 7.6, para lo cual se requiere obtener las frecuencias esperadas, las cuales se incorporan en la tabla de contingencia siguiente (en paréntesis).

Física	Matemáticas			Total
	Alto	Medio	Bajo	
Alto	56 (30,80)	71 (72,66)	12 (35,54)	139
Medio	47 (54,95)	163 (129,64)	38 (63,41)	248
Bajo	14 (31,24)	42 (73,70)	85 (36,05)	141
Total	117	276	135	528

A partir de estos resultados, $Q_0^2 = 145,78$ y, por otra parte, $\chi^2(4) = 13{,}28$, por lo que se rechaza H_0 y se concluye que las variables están

relacionadas, es decir, el rendimiento en Matemáticas y en Física no es independiente. □

7.5. Ejercicios resueltos

1. En un intento por mejorar la calidad de las cintas de grabación, se les aplican diferentes tinturas a iguales tipos de dispositivos (cintas). Se desea saber si los niveles de distorsión son independientes o no según el color de la cinta (4 tipos de tintura: A, B, C y D); para esto, se dispone de las siguientes mediciones.

Color	Niveles de distorsión						
A	10	15	8	12	15		
B	14	18	21	15			
C	17	16	14	15	17	17	18
D	12	15	17	15	16	15	

Solución: Nótese que los tipos de tintura o color definen grupos para los cuales se cuenta con las mediciones del nivel de distorsión; según lo que se pide, se pueden plantear las hipótesis:

H_0: Los niveles de distorsión son iguales entre los distintos colores.

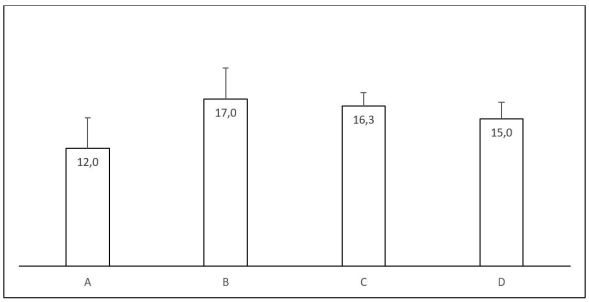
H_1: Los niveles de distorsión son diferentes entre los colores.

Para comparar los grupos se puede utilizar el promedio ($\bar{x}$) y la desviación estándar (S):

Color	N	Promedio ($\bar{x}$)	D.E. (S)
A	5	12,0	3,1
B	4	17,0	3,2
C	7	16,3	1,4
D	6	15,0	1,7
Total	22	15,1	2,8

A partir de la tabla y el gráfico anteriores, se observa que efectivamente hay evidencia de diferencia en los niveles de distorsión para las distintas tinturas (colores). Se verá ahora si estas diferencias

Figura 7.4: Gráfico para el nivel de distorsión según color de la tintura



Fuente: elaboración propia.

son significativas, para lo cual se expresan las hipótesis utilizando parámetros estadísticos:

$$H_0 : \mu_A = \mu_B = \mu_C = \mu_D \text{ y } H_1 : \mu_i = \mu_j, \text{ para algún } i \neq j.$$

En este caso, se utiliza el anova, de lo que resulta:

	Tratamiento	Error	Total
SC (gl)	72,4 (3)	93,4 (18)	165,8 (21)

Con lo cual el estadístico de prueba es $F_0 = 4,65 > F_{0,95}(3, 18) =$ 3,16, por lo tanto, se rechaza H_0, es decir, hay evidencia significativa de que hay diferencia entre los distintos tipos de tintura (colores).

i	j	$\bar{x}_i - \bar{x}_j$	EE_{ij}	IC: Lím. inf.	IC: Lím. sup.
A	B	-5,00	1,53	-9,53	-0,47
A	C	-4,29	1,33	-8,24	-0,33
C	D	-3,00	1,38	-7,09	1,09
B	C	0,71	1,43	-3,52	4,94
B	D	2,00	1,47	-2,36	6,36
C	D	1,29	1,27	-2,47	5,04

Desde la tabla se observa que hay diferencias significativas entre las tinturas A y B y entre A y C, siendo A la que presenta niveles

menores de distorsión. Entonces, si hay que elegir entre estas tres opciones, la elegida sería la A. Por su parte, las otras opciones no presentan diferencias significativas, por lo que, al elegir entre estas otras combinaciones, las alternativas producirían resultados similares.

2. Un fabricante de partículas de madera para la construcción está interesado en la resistencia de las partículas como una función de la temperatura de cocción. A partir de los datos obtenidos al realizar el experimento (en la tabla siguiente), resolver los puntos que se señalan.

Temperatura (C^{O})	Resistencias		
40	66,30	64,84	64,36
45	69,70	66,26	72,06
50	73,23	71,40	68,85
55	75,72	72,57	76,64
60	78,87	77,37	75,94
65	78,82	77,13	77,09

   (a) Realizar el diagrama de dispersión para encontrar evidencia descriptiva del problema.

   (b) Plantear la relación mediante un esquema de regresión lineal simple y encontrar los estimadores.

   (c) Obtener el estadístico de prueba T_0, a partir del cual obtener el valor-p para decidir sobre la hipótesis de interés de la relación planteada en el problema.

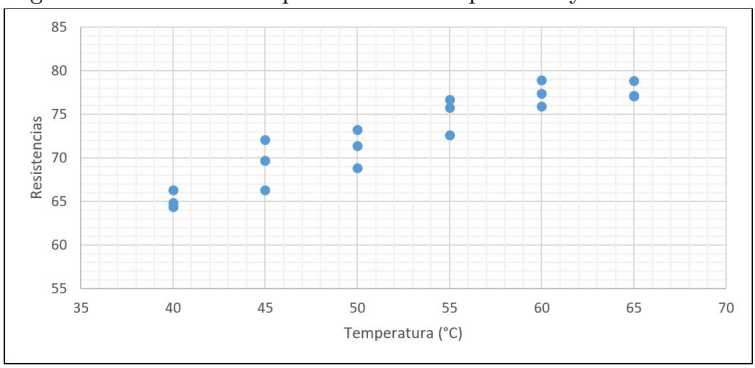
   (d) Interpretar el parámetro β en función del problema.

   Solución: El problema asocia dos variables cuantitativas, que son la temperatura y la resistencia; además, pide ajustar un modelo de regresión lineal para X: temperatura e Y: resistencia de las partículas.

   (a) El diagrama de dispersión de la figura 7.5 muestra una asociación lineal entre las variables. En efecto, a medida que aumenta la temperatura lo hace también la resistencia de las partículas. Además,

el coeficiente de correlación es $r_{XY} = 0{,}924$, el que indica que existe una asociación entre las variables, relación que es significativa ($T_0 = 9{,}69$, $p < 0{,}001$).

Figura 7.5: Gráfico de dispersión de la temperatura y las resistencias



Fuente: elaboración propia.

(b) El modelo teórico es $Y = \alpha + \beta X$; realizando el ajuste de mínimos cuadrados se obtiene: $\widehat{Y} = 45{,}4 + 0{,}52 \cdot X$ ($\widehat{\alpha} = 45{,}4$y $\widehat{\beta} = 5{,}52$).

(c) La hipótesis de interés es $H_0 : \beta = 0$ y $H_1 : \beta \neq 0$, para la cual $T_0 = 6{,}38 > T_{0{,}95}(16) = 2{,}12$; por lo tanto, se rechaza H_0.

(d) Por cada grado centígrado que aumenta la temperatura, la resistencia aumenta en 0,52 unidades.

3. Un estudio sobre la relación entre el contenido de violencia de los programas de televisión y la edad de los espectadores arrojó los resultados que aparecen en la tabla siguiente; el estudio incluyó a 81 personas (clasificadas como espectadores de programas de alto contenido de violencia o de bajo contenido de violencia).

Contenido de violencia	Edad (en años cumplidos)		
	16-34	35-54	55 o más
Bajo	8	12	21
Alto	18	15	7

(a) Plantear la hipótesis.

(b) Entregar evidencia descriptiva sobre la hipótesis.

(c) Determinar si la evidencia es significativa.

Solución: (a) La hipótesis en este caso es: H_0: *la edad y el tipo de programa de preferencia son independientes* y H_1: *la edad y el tipo de programa de preferencia no son independientes (están asociados).* Por tanto, se trata de una asociación entre dos variables cualitativas.

(b) El 69 % de las personas de entre 16 y 34 años prefiere películas con alto contenido de violencia, mientras que solo el 25 % de las personas de 55 o más años prefiere este tipo de películas, por lo que esta evidencia apunta a que las variables no son independientes.

(c) Para determinar si la evidencia es significativa, se planteará el test de hipótesis de (a). Para calcular el estadístico de prueba Q_0, se requiere obtener las frecuencias esperadas:

	Edad: 16-34	Edad: 35-54	Edad: 55 o más	Total
Bajo	8 (13,2)	12 (13,7)	21 (14,2)	41
Alto	18 (12,8)	15 (13,3)	7 (13,8)	40
Total	26	27	28	81

Luego, $Q_0^2 = \dfrac{(8-13{,}2)^2}{13{,}2} + \cdots + \dfrac{(7-13{,}8)^2}{13{,}8} = 11{,}17 > \chi^2_{0{,}95}(2) = 5{,}99$, por lo que se rechaza H_0, es decir, existe evidencia significativa de asociación (no son independientes) entre las variables.

7.6. Ejercicios propuestos

1. Un estudio desea evaluar el efecto de un determinado entrenamiento en el tiempo de reacción de atletas sometidos a un cierto estímulo. El entrenamiento consiste en la repetición de un movimiento y es utilizada una muestra de 37 atletas. Para cada atleta se asigna un cierto número de repeticiones (X) y, entonces, se mide el tiempo de reacción (Y) en milisegundos. Una recta de regresión mediante el

método de mínimos cuadrados se ajusta a los datos y se obtiene la ecuación:

$$\widehat{y}_i = 80{,}5 - 0{,}90\, x_i, \quad i = 1, 2, \cdots, 37.$$

Interpretar las estimaciones de (a) α y (b) β.

2. Para verificar el efecto de la una variable X sobre la variable Y, se realiza un experimento que resulta en los pares (x_i, y_i) siguientes: $(3; 13{,}3)$, $(7; 24{,}3)$, $(5; 15{,}9)$, $(2; 12{,}8)$, $(7; 29{,}5)$, $(3; 14{,}5)$, $(5; 23{,}3)$, $(8; 32{,}6)$, $(2; 12{,}0)$ y $(1; 4{,}6)$:

   (a) Obtener la recta ajustada.

   (b) Construir el diagrama de dispersión basándose en los pares de valores entregados.

   (c) Basándose solo en el gráfico, indicar si el ajuste es el adecuado.

3. Un estudio busca determinar el efecto de la obesidad en la presión sanguínea. Para esto, se registran los pesos de 6 individuos y se construye la variable X, que representa la razón entre los pesos reales y el ideal. Estudios realizados con anterioridad muestran que el método de regresión lineal simple es adecuado para esta situación. Los datos obtenidos son:

Razón (X)	1,23	1,42	1,35	1,67	1,65	1,56
Presión sistólica (Y)	129	130	133	139	136	134

   (a) Ajustar la recta: $y = \alpha + \beta\, d$, donde $d = x - \bar{x}$.

   (b) Verificar si el parámetro β es estadísticamente significativo.

   (c) ¿Cuál es la interpretación para α en la recta obtenida en (a)?

   (d) ¿Cuál es la presión sistólica esperada para individuos con razón de pesos real/ideal igual a 1,25?

4. La cantidad de lluvia es un factor importante en la productividad agrícola. Para medir ese efecto, se registra el índice pluviométrico (lluvia en mm) y la producción (en ton) del último año en ocho regiones productoras de soya.

Lluvia	120	140	122	150	115	190	130	118
Producción	40	46	45	37	25	54	33	30

(a) Graficar identificando la relación entre las variables. Calcular el coeficiente de correlación y realizar la prueba para la verdadera correlación entre las variables.

(b) Ajustar la recta de regresión. Interpretar el parámetro β.

(c) Utilizando el resultado de (b), encontrar la producción esperada para una región con índice pluviométrico igual a 160 mm.

5. Un estudio desea determinar la influencia que tiene el peso de nacimiento (en oz) en el incremento del peso entre los días 70 y 100 de vida (expresado como la tasa respecto del peso de nacimiento), para lo cual se cuenta con información de 12 niños (tabla siguiente).

Peso de nacimiento	112	111	107	119	92	80
Tasa de incremento	63	66	72	52	75	118
Peso de nacimiento	81	84	118	106	103	
Tasa de incremento	120	114	42	72	90	

(a) Identificar las variables (independiente y dependiente).

(b) Graficar, identificando los pares (x, y) según las variables del problema, para definir el tipo de relación.

(c) Calcular el coeficiente de correlación y probar si en la población las variables están correlacionadas.

(d) Ajustar la recta de regresión lineal simple para el problema.

(e) Por ejemplo, si para un niño el peso de nacimiento es de 95 oz, predecir cuál podría ser su tasa de crecimiento entre los días 70 y 100 de vida.

6. Se desea determinar si la edad puede afectar la presión sistólica de la sangre (en mmHg) en mujeres. Para esto, se recoge información de 15 personas.

Edad	42	46	42	71	80	74	70	80	85
Presión	130	115	148	100	156	162	151	156	162
Edad	72	64	81	41	61	75			
Presión	158	155	160	125	150	165			

(a) Identificar las variables (independiente y dependiente).

(b) Graficar, identificando los pares (x, y) según las variables del problema, para definir el tipo de relación.

(c) Calcular el coeficiente de correlación y probar si en la población las variables están correlacionadas.

(d) Ajustar la recta de regresión lineal simple para el problema.

(e) Predecir cuál podría ser la presión sistólica para una mujer de 50 años de edad.

7. Se realizó un estudio para conocer la eficiencia de una nueva vacuna contra la gripe. Para esto, se contó con 1000 personas, de las cuales un grupo recibió dos vacunas; otro, una vacuna, y otro grupo no recibió ninguna vacuna. Para estas personas (todas) se registró el estado de salud: con gripe y sin gripe. Al respecto, los investigadores están interesados en saber si hay relación entre las categorías de vacunas (*no se les aplicó vacuna o sin vacuna*, *una vacuna* y *dos vacunas*) y el estado de salud (*contrajeron gripe* y *no contrajeron gripe*), para lo cual cuentan con la siguiente información que se registra en la tabla de contingencia:

Estado	Sin vacuna	Una vacuna	Dos vacunas	Total
Gripe	24	9	13	46
Sin gripe	289	100	565	954
Total	313	109	578	1000

(a) Plantear la hipótesis.

(b) Entregar evidencia descriptiva sobre la hipótesis.

(c) Determinar si la evidencia es significativa.

Apéndice

A. Fórmulas

A.1. Expresiones matemáticas

1. Teorema del binomio: $(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i}$, con $\binom{n}{i} = \frac{n!}{(n-i)!i!}$.

2. Serie geométrica: $\frac{1}{1-p} = \sum_{i=0}^{\infty} p^i, \ |p| < 1$.

3. Expansión en serie de Taylor: $e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!}$.

4. Función gamma: $\Gamma(x) = \int_0^{\infty} t^{x-1} e^t \, dt, \ x > 0$.

Propiedades: $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$; $\Gamma(x+1) = x\Gamma(x)$; $\Gamma(n) = (n-1)!$, si $n \in \mathbb{N}$.

5. Función beta: $B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1} \, dx = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$.

A.2. Resumen de distribuciones continuas

Distribución	Función de densidad	FGM: $M_X(t)$	Valor esperado y varianza
Normal	$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\}$	$\exp\left\{t\mu + \frac{1}{2}t^2\sigma^2\right\}$	$E[X] = \mu$
$X \sim N(\mu, \sigma^2)$	$-\infty < x < \infty,\ -\infty < \mu < \infty,\ \sigma > 0$		$V[X] = \sigma^2$.
Gamma	$f(x) = (\Gamma(\alpha)\beta^\alpha)^{-1} x^{\alpha-1} \exp\left\{-\frac{x}{\beta}\right\}$	$(1-\beta t)^{-\alpha}$	$E[X] = \alpha\beta$
$X \sim \text{Gamma}(\alpha, \beta)$	$x > 0,\ \alpha > 0,\ \beta > 0$		$V[X] = \alpha\beta^2$.
Exponencial	$f(x) = \frac{1}{\beta} \exp\left\{-\frac{x}{\beta}\right\}$	$(1-\beta t)^{-1}$	$E[X] = \beta$
$X \sim \text{Exp}(\beta)$	$x > 0,\ \beta > 0$		$V[X] = \beta^2$.
Chi-cuadrado	$f(x) = \left(\Gamma(\frac{n}{2})2^{\frac{n}{2}}\right)^{-1} x^{\frac{n}{2}-1} \exp\left\{-\frac{x}{2}\right\}$	$(1-2t)^{-\frac{n}{2}}$	$E[X] = n$
$X \sim \chi^2(n)$	$x > 0,\ n > 0$		$V[X] = 2n$.
Uniforme	$f(x) = \frac{1}{b-a}$	$\frac{\exp\{tb\}-\exp\{at\}}{t(b-a)}$	$E[X] = \frac{a+b}{2}$
$X \sim \text{Uniforme}(a, b)$	$a < x < b\ \ (a, b \in \mathbb{R})$		$V[X] = \frac{(b-a)^2}{12}$.
Weibull	$f(x) = \alpha \cdot \theta^{-\alpha} x^{\alpha-1} \exp\left\{-(\frac{x}{\theta})^\alpha\right\}$		$E[X] = \theta\,\Gamma(1+\alpha^{-1})$
$X \sim \text{Weibull}(\alpha, \theta)$	$x > 0,\ \alpha > 0,\ \theta > 0$		$V[X] = \theta^2[\Gamma(1+\frac{2}{\alpha}) - \Gamma^2(1+\frac{1}{\alpha})]$.
Beta	$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$		$E[X] = \alpha(\alpha+\beta)^{-1}$
$X \sim \text{Beta}(\alpha, \beta)$	$0 < x < 1,\ \alpha > 0,\ \beta > 0$		$V[X] = \alpha\beta[(\alpha+\beta)^2(\alpha+\beta+1)]^{-1}$.

A.3. Resumen de distribuciones discretas

Distribución	Función de densidad	FGM: $M_X(t)$	Valor esperado y varianza
Binomial	$f(x) = \binom{n}{x} p^x (1-p)^{n-x}$	$[pe^t + (1-p)]^n$	$E[X] = np$
$X \sim \text{Binomial}(n, p)$	$x = 0, 1, .., n,\ 0 \leqslant p \leqslant 1,\ n \geqslant 0$		$V[X] = np(1-p)$
Poisson	$f(x) = e^{-\lambda}\lambda^x / x!$	$\exp\{\lambda(e^t - 1)\}$	$E[X] = \lambda$
$X \sim \text{Poisson}(\lambda)$	$x = 0, 1, 2, \cdots,\ \lambda > 0$		$V[X] = \lambda$
Geométrica	$f(x) = (1-p)^{x-1} p$	$\frac{pe^t}{1-qe^t}$	$E[X] = \frac{1}{p}$
$X \sim \text{Geométrica}(p)$	$x = 1, 2, \cdots$		$V[X] = \frac{1-p}{p^2}$
Hipergeométrica	$f(x) = \binom{m}{x}\binom{N-m}{n-x} / \binom{N}{n}$		$E[X] = n\frac{m}{N}$
$X \sim \text{Hiperg}(N, m, n)$	$x = 0, \cdots, \text{mín}(m, n)$		$V[X] = n\frac{m}{N}(1 - \frac{m}{N})\frac{N-n}{N-1}$
Binomial negativa	$f(x) = \binom{x-1}{k-1} p^k q^{x-k}$	$\left(\frac{pe^t}{1-qe^t}\right)^k$	$E[X] = \frac{k}{p},\ k = 2, 3, \cdots$
$X \sim \text{Bneg}(k, p)$	$x = k, k+1, \cdots;\ k = 1, 2, \cdots$		$V[X] = \frac{kq}{p^2},\ k = 2, 3, \cdots$

Apéndice

B. Respuestas de los ejercicios

Capítulo 1

1. 360 360.
2. 360.
3. 365^{10}.
4. (a) 120. (b) 20.
5. (a) 350. (b) 150. (c) 300.
6. 2880.
7. 10.
11. (a) $\binom{r+n-1}{n-1}$. (b) $\frac{n}{\binom{r+n-1}{n-1}}$. (c) $\frac{\binom{r+n-2k-1}{n-k-1}}{\binom{r+n-1}{n-1}}$.
12. 0,0024.
13. 0,427.
14. 0,1157.
15. (a) $\frac{1}{4}$. (b) $\frac{5}{8}$. (c) $\frac{7}{8}$. (d) $\frac{3}{8}$.
16. $\frac{1}{24}$.
17. 0,06.

19. (a) n^d. (b) $1 - \left(1 - \frac{1}{n}\right)^d$. (c) $1 - e^{-1}$.

20. (a) 0,95. (b) 0,15. (c) 0,30.

21. (a) 0,095. (b) 0,855. (c) 0,045. (d) 0,955.

22. 0,2.

23. (a) $\frac{1}{5}$. (b) $\frac{3}{5}$.

24. $\frac{28}{45}$.

25. 0,073.

26. (a) 0,467. (b) 0,643.

27. (a) $\frac{n}{m}\frac{m-1}{n-1}$.

28. (a) 0,260. (b) 0,462.

29. (a) 0,833. (b) 0,941. (c) 0,06.

30. 0,10.

31. 0,022.

32. 0,720.

33. 0,323.

34. $\frac{1}{3}$.

35. (a) 0,22. (b) 0,732. (c) 0,878.

Capítulo 2

1. (b) $[k\,e^t + (1-k)]^n$. (c) np y $n\,k(1-k)$.

2. (a) $Rec_X = \{0, 1, 2\}$. (b) $p(0) = \frac{24}{70}$, $p(1) = \frac{36}{70}$, $p(2) = \frac{10}{70}$.

3. (c) $(1-2t)^{-\frac{n}{2}}$. (d) $2n$.

4. (a) $\frac{13}{27}$. (b) $\frac{1}{3}$. (c) $\frac{22}{27}$.

5. (b) $F(x) = 0$, si $x < 0$; $F(x) = -\frac{5}{2} + \frac{2x}{5} - \frac{x^2}{50}$, si $0 \leqslant x \leqslant 5$; $F(x) = 1$, si $x > 10$. (c) 3,35.

6. (a) $0 \leqslant x \leqslant \frac{1}{2}$ y $b = 1$. (b) $f(x)$ existe si $F(x)$ es continua, luego $a = \frac{1}{2}$.

8. (a) $F(x) = 0$, si $x < 0$; $F(x) = \frac{x^2}{2}$, si $0 \leqslant x < 1$; $F(x) = \frac{3}{4}$, si $1 \leqslant x < 2$; $F(x) = \frac{x+1}{4}$, si $2 \leqslant x < 3$; $F(x) = 1$, si $x \geqslant 3$. (b) No se puede porque $F(x)$ no es continua. (c.1) 0,5, (c.2) 0,5, (c.3) 0,25, (c.4) 0,75 y (c.5) 0,5.

9. (a) $f(x) = \frac{ax}{k}$, si $0 \leqslant x \leqslant k$; $f(x) = \frac{(x-2)a}{k-2}$, si $k < x \leqslant 2$. (b) $E[X] = \frac{k+2}{3}$ y $Var(X) = \frac{c^2-2c+4}{18}$.

10. (a) $a = \frac{18}{5}$; $b = -\frac{12}{5}$. (b) $\frac{7}{20}$. (c) $Var(X) = 0{,}16$.

11. (a) $M_X(t) = 0{,}2e^{-5}e^{5e^t} + 0{,}8\alpha\,(1-(1-\alpha)e^t)^{-1}$. (b) $M_X(t) = 2(2-t)^{-1}$, con $t < 2$.

14. (a) Si $x \leqslant -1$, $F(x) = 0$; si $-1 < x < 1$, $F(x) = 0{,}5(x+1)$ y si $x \geqslant 1$, $F(x) = 1$. (b) $f_Y(y) = \left(2\sqrt{y}\right)^{-1}$, $0 < y < 1$.

15. (a) $a = 2$. (b) $f_T(t) = -2t^{-3}$, $t < -1$.

Capítulo 3

3. 0,034.

4. (a) $X \sim \text{Binomial}(n = 100; p = 0{,}001)$. (b) $0{,}999^{100}$. (c) $0{,}1 \cdot 0{,}999^{99}$. (d) $2e^{-2}$.

5. (a) 0,148. (b) 0,400.

6. (a) $X \sim \text{Binomial}(n, p = \frac{1}{100})$. (b) 229.

7. $ECM(T_1) = \frac{p(1-p)}{n}$; $ECM(T_2) = \frac{p(n-4)-p^2(n+4)-1}{(n+2)^2}$.

10. El vehículo de 1000 kg (con los otros la probabilidad de sobrepasar el máximo es muy alta).

11. (a) 0,683. (b) 0,157.

12. (a) $\alpha = 3$ y $\beta = 2$. (b) 0,062.

13. (a) 0,212. (b) 403 y 162 754, respectivamente.

15. Normal: μ; exponencial: $\alpha \log 2$; log-normal$(\mu = 0; \sigma^2 = 1)$: 1.

16. (b) $F(x) = \frac{1}{b\pi} \arctan\left(\frac{x-a}{b}\right) + \frac{1}{2}$. (c) La media y mediana es a.

17. Exponencial: $R(t) = \exp\left\{-\frac{1}{\alpha}t\right\}$, $Z(t) = \frac{1}{\alpha}$;
Weibull: $R(t) = \exp\left\{-\left(\frac{t}{\theta}\right)^{\alpha}\right\}$, $Z(t) = \frac{\alpha}{\theta^{\alpha}} t^{\alpha-1}$.

Capítulo 4

2. (a) $k = 3$. (b) $f_X(x) = 3(\frac{1}{3} + x^2 - \frac{4}{3}x^3)$, $0 < x < 1$; $f_Y(y) = 4y^3$, $0 < y < 1$. (c) $f_{Y|X}(y) = (x^2 + y^2)(\frac{1}{3} + x^2 - \frac{4}{3}x^3)^{-1}$. (d) La varianza depende de la correlación, la cual requiere de $E[XY] = \frac{3}{8}$. Los demás ejercicios se pueden obtener de los resultados anteriores.

3. (b) 0,155.

7. (d) Véase Dennis D. Wackerly, William Mendenhall, Richard L. Scheaffer y Jorge H. Muñoz. *Estadística matemática con aplicaciones*. Thomson, 2002.

8. Ayuda: Nótese que $X_1 \mid X_2 \sim N(x_2, x_2^2)$ y $X_2 \sim$ Uniforme$(0, 1)$.

9. (a) $\sigma_1^2 + \mu_1^2 + \mu_2$. (b) $\exp\left\{\frac{1}{2}\right\}$.
(c) $M_{X_1+X_2}(t) = \exp\left\{t\,(\mu_1 + \mu_2) + \frac{1}{2}t^2(\sigma_1^2 + \sigma_2^2)\right\}$ y
$X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

15. ρ.

16. Indicación: Definiendo $v = y_1 \cdot (y_1 + y_2)^{-1}$ y $w = y_2$, luego la función solicitada se obtiene resolviendo $f_V(v) = \int_{Rec_W} f_{V,W}(v, w)\, dw$.

17. (a) $\frac{n_2}{n_2 - 2}$. (b) $\frac{2n_2^2(n_1+n_2-2)}{n_1(n_2-2^2)(n_2-4)}$.

18. $c(n_1, n_2) y_1^{\frac{n_1}{2}-1} y_2^{\frac{n_1+n_2}{2}-2} \exp\left\{-\frac{y_2}{2}(y_1 + 1)\right\}$, donde $c(n_1, n_2)$ es una constante que depende de n_1 y n_2.

20. $\frac{1}{4\pi\sigma^2} \exp\left\{-\frac{1}{2\sigma^2}\left(\frac{y_1^2+y_2^2}{2} + 2\mu^2 + 2\mu y_2\right)\right\}$.

Capítulo 5

1. $\widehat{\alpha} = \bar{Y} + \widehat{\beta}\bar{X}$, $\widehat{\beta} = \left(\sum Y_i X_i - \widehat{\alpha}\sum X_i\right)\left(\sum X_i^2\right)^{-1}$ y $\widehat{\sigma}^2 = \bar{Y} - \widehat{\alpha} - \widehat{\beta}\bar{X}$.
2. $\widehat{\theta}_1 = \frac{1}{2n_1}\sum x_i^2$ y $\widehat{\theta}_2 = \frac{1}{2n_2}\sum y_i^2$.
3. $\widehat{\mu} = \sum x_i$ y $\widehat{\lambda} = n\widehat{\mu}^2\left(\sum \frac{(x_i-\widehat{\mu})^2}{x_i}\right)^{-1}$.
4. $\widehat{\theta} = \sqrt{-\frac{1}{y} + \left(\frac{2+y}{2y}\right)^2} - \frac{2+y}{y}$, donde $y = \frac{1}{n}\sum \log x_i$.
5. (a) $\widehat{\beta}_{MV} = \frac{1}{n}\sum \frac{y_i}{a_i}$, $\widehat{\beta}_{MM} = \frac{\sum y_i}{\sum a_i}$, $\widehat{\beta}_{MC} = \frac{\sum a_i y_i}{\sum a_i^2}$. (b) Todos los estimadores son insesgados, $Var[\widehat{\beta}_{MV}] = \frac{\beta^2}{n^2}$, $Var[\widehat{\beta}_{MM}] = \frac{\sum y_i}{\sum a_i}$, $Var[\widehat{\beta}_{MC}] = \beta^2\frac{\sum a_i^2}{(\sum a_i)^2}$. (c) $\widehat{\beta} \pm z_{1-\frac{\alpha}{2}}\frac{\widehat{\beta}}{\sqrt{n}}$.
7. (a) Todos son insesgados. (b) $Var[\widehat{\theta}_1] = \theta^2$, $Var[\widehat{\theta}_2] = 0,5\theta$, $Var[\widehat{\theta}_3] = \frac{5}{9}\theta^2$, $Var[\widehat{\theta}_4] = \theta^2$ y $Var[\widehat{\theta}_5] = \frac{1}{3}\theta^2$.
8. (a) $\frac{n}{-\sum \log x_i}$. (b) $\frac{n}{n-1}\theta$ y $\frac{n^2}{(n-1)^2(n-2)}\theta^2$, valor esperado y varianza, respectivamente. (c) El estimador es asintóticamente insesgado y consistente.
9. (a) $E[T_1] = \lambda$, $Var[T_1] = \frac{\lambda}{n}$ y $E[T_2] = \lambda$, $Var[T_2] = \frac{2n+1}{3n(n+1)}\lambda$. (b) Ambos estimadores son insesgados y consistentes.
10. (a) $k_1 = 2$, $k_2 = \frac{n+1}{n}$. (b) Ambos estimadores son consistentes.
13. (a) $\bar{x}$. (b) El estimador es consistente. (c) $\theta \in [\frac{2\sum X_i}{\chi^2_{0,975}}; \frac{2\sum X_i}{\chi^2_{0,025}}]$.
14. $\bar{X} \pm t_{1-\frac{\alpha}{2}}\sqrt{\frac{S^2}{n}}$.
15. (a) $\left[\frac{n-1}{\chi^2_{1-\frac{\alpha}{2}}}S^2; \frac{n-1}{\chi^2_{\frac{\alpha}{2}}}S^2\right]$. (b) $S^4 \pm z_{1-\frac{\alpha}{2}}2\,S^2\sqrt{\frac{2}{n}}$.
16. (a) $\bar{X} - \bar{Y} \pm z_{1-\frac{\alpha}{2}}S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$, donde $S_p^2 = \frac{(n_1-1)S_x^2 + (n_2-1)S_y^2}{n_1+n_2-2}$. (b) $\left[\frac{S_x^2}{S_y^2 F_{1-\frac{\alpha}{2}}}; \frac{S_x^2}{S_y^2 F_{\frac{\alpha}{2}}}\right]$.
17. (a) $\bar{X} \pm z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{X}(1-\bar{X})}{n}}$. (b) $\bar{X}^2 \pm z_{1-\frac{\alpha}{2}}2\bar{X}\sqrt{\frac{\bar{X}(1-\bar{X})}{n}}$.
18. $\frac{1}{\alpha}\bar{X} \pm z_{1-\frac{p_0}{2}}\frac{1}{\alpha}\sqrt{\frac{\bar{X}}{\alpha n}}$.

19. $e^{\bar{X}} \pm z_{1-\frac{\alpha}{2}} e^{\bar{X}} \sqrt{\frac{\sigma_0^2}{n\bar{X}}}$.

21. $\mu \in (-2{,}1; 2{,}85)$.

22. $\sigma^2 \in (0{,}031; 0{,}22)$.

23. (a) Sí, véase el teorema central del límite. (b) El intervalo de confianza es: $(0{,}038; 0{,}166)$.

Capítulo 6

1. (a) $\Phi(-2{,}667)$; (b) $\beta = \mathbb{P}\left(\frac{2{,}92-\mu_0}{0{,}03}\right)$, $\mu_0 > 1$. (c) $1 - \beta = \Phi(2{,}333)$.

2. (a) $\Re = \left\{\bar{X} : \bar{X} > \frac{1}{2n}\beta_0 \chi^2_{1-\alpha}(2n)\right\}$.
(b) Potencia $= \mathbb{P}\left(\chi^2 > n\beta_h^{-1}\beta_0\chi^2_{1-\alpha}(2n)\right)$, con $\beta_h > \beta_0$. (c) Se rechaza H_0 si $X_0 > 46{,}40$ (resultado para $\chi^2_{0{,}950}(72) = 92{,}81$).

3. $\Re = \left\{\frac{\bar{X}}{\bar{Y}} | \frac{\bar{X}}{\bar{Y}} < F_{p_0}(2n, 2k)\right\}$.

4. (a) $\Re = \left\{\bar{X} : \bar{X} > t_{1-\alpha}\sqrt{\frac{S^2}{n}}\right\}$. (b) $n = \sigma^2(z_{0{,}95} - z_{0{,}20})^2$.

5. (a) $\Re = \left\{(\bar{X} - \bar{Y}) : \bar{Y} - \bar{X} > z_{1-\alpha}\sqrt{\frac{1}{n}\left(\bar{X}(1-\bar{X}) + \bar{Y}(1-\bar{Y})\right)}\right\}$.
(b) $n = 8{,}59 \cdot (\theta_0^2)^{-1}\left(\bar{X}(1-\bar{X}) + \bar{Y}(1-\bar{Y})\right)$.

6. $-2\log(\widehat{\lambda}) = -\frac{n}{2}\log\left(\frac{1}{50n}\sum\frac{Y_i}{X_i}\right) + \frac{1}{25}\sum\frac{Y_i}{X_i} - 2n$ y
$\Re = \left\{\widehat{\lambda} : -2\log(\widehat{\lambda}) > \chi^2_{1-\alpha}(1)\right\}$.

7. $\Re = \left\{\lambda | 2\log\lambda > \chi^2_{1-\alpha}(1)\right\}$, donde $2\log\lambda = 2n\log(\bar{y}) - \sum(y_i - 1)^2 + \bar{y}^{-2}\sum(y_i - \bar{y})^2$.

9. (a) $\Re = \{\widehat{p}_1 - \widehat{p}_2 | \widehat{p}_1 - \widehat{p}_2 > 0{,}10\}$. (b) $\widehat{p}_1 - \widehat{p}_2 = 0{,}18 > 0{,}10$, se rechaza H_0 con un 95 % de confianza, es decir, la proporción de desempleados en Temuco es significativamente mayor que la de Concepción.

10. (a) $\Re = \left\{\bar{X} - \bar{Y} > 56{,}5\right\}$. (b) 0,43.

11. (a) $T_0 = 3{,}645$. (b) Con $\alpha = 0{,}05$, el IC es $[68{,}06; 140{,}0]$.

12. (a) $T_0 = 3{,}61 > t_{0{,}975}(10) = 2{,}23$, se rechaza H_0. (b) $\chi^2_0 = 12{,}77 > \chi^2_{0{,}950}(10) = 25{,}2$, se rechaza H_0.

13. (a) 0,01. (b) 0,11. (c) $\Re = \left\{\bar{X} | \bar{X} < 98{,}6\right\}$.

15. (a) Test de comparación de proporciones, $H_0 : p_1 = p_2$ versus $H_1 : p_1 \neq p_2$. (b) $\Re = \{\widehat{p}_1 - \widehat{p}_2 | \widehat{p}_1 - \widehat{p}_2 > 0,062 \vee \widehat{p}_1 - \widehat{p}_2 < -0,062\}$. (c) $\widehat{p}_1 - \widehat{p}_2 = -0,021 \notin \Re$, por lo que no se puede rechazar H_0, es decir, con un nivel del 5 % se concluye que la incidencia de parásitos en los dos grupos puede ser considerada la misma.

16. $H_0 : \mu_D = 0$ versus $H_1 : \mu_D > 0$, $T_0 = 2{,}57 < t_{0{,}99} = 2{,}72$, entonces no se puede rechazar H_0.

Capítulo 7

1. (a) El tiempo de reacción de los individuos sin entrenamiento es de 80,50 ms. (b) Cada repetición contribuye a reducir el tiempo de reacción en 0,90 ms.

2. (a) $\widehat{y}_i = 3{,}58 + 3{,}42\, x_i$. (b) Graficar en plano cartesiano. (c) El gráfico muestra que los datos se ajustan adecuadamente a un modelo lineal.

3. (a) $\widehat{y}_i = 133{,}50 + 18{,}85\, d_i$. (b) Es estadísticamente significativo (95 % de confianza). (c) Presión sistólica esperada para un individuo con razón de pesos igual a 1,48 (que corresponde al promedio de los datos). (d) 129,2.

4. (a) $r = 0{,}725$ y $T_0 = 2{,}58$. (b) $\alpha = 1{,}553$ y $\beta = 0{,}274$. (c) $\widehat{y}(160) = 45{,}4$.

5. (a) X: peso de nacimiento (variable independiente), Y: tasa de incremento en el peso entre los días 70 y 100 de vida (variable dependiente). (b) El gráfico muestra una relación lineal inversa (a mayor peso de nacimiento menor tasa de incremento de peso). (c) $r = -0{,}946$ y $T_0 = -9{,}23$. (d) $\alpha = 256{,}3$, $\beta = -1{,}47$. (e) 90,1 % del peso de nacimiento.

6. (a) X: edad en años (variable independiente), Y: presión sistólica de la sangre en mmHg (variable dependiente). (b) El gráfico muestra una relación lineal directa (a mayor edad también es mayor la presión sistólica observada). (c) $r = 0{,}566$ y $T_0 = 2{,}475$. (d) $\alpha = 99{,}6$, $\beta = 0{,}71$. (e) 135 mmHg.

7. (a) H_0: Las categorías de vacunas y el estado de salud son independientes y H_1: Las categorías de vacunas y el estado de salud no son independientes (están asociadas). (b) El 52 % de quienes no recibieron vacuna contrajeron la enfermedad, mientras que solo el 28 % de quienes recibieron dos vacunas contrajeron la enfermedad, evidencia a favor de H_1. (c) $Q_0^2 = 20,61 > \chi^2(2) = 5{,}99$, se rechaza H_0.

Apéndice

C. Bibliografía

1. Heleno Bolfarine y Mônica Carneiro Sandoval. *Introdução à inferência estatística*, vol. 2. SBM, 2001.

2. George C. Canavos. *Probabilidad y estadística, aplicaciones y métodos*. McGraw-Hill, 1988.

3. George Casella y Roger L. Berger. *Statistical inference*, vol.2. Duxbury Pacific Grove, CA, 2002.

4. Carles M. Cuadras. *Problemas de probabilidades y estadística*, vol. 1. Inferencia estadística. Edicions Universitat Barcelona, 2016.

5. Carles M. Cuadras. *Problemas de probabilidades y estadística*, vol. 2. Inferencia estadística. Edicions Universitat Barcelona, 2016.

6. Carlos A. B. Dantas. *Probabilidade: Um curso introdutório*. Edusp, 2013.

7. Morris H. Degroot. *Probabilidad y estadística*. Addison-Wesley Iberoamericana, 1988.

8. Jay L. Devore. *Probabilidad y estadística para ingenierías y ciencias*. Cengage Learning Editores, 2008.

9. Barry R. James. *Probabilidade: um curso em nível intermediário*, núm. 519.2. 1996.

10. Chap T. Le y Lynn E. Eberly. *Introductory biostatistics*. John Wiley & Sons, 2016.

11. Sharon L. Lohr. *Sampling: Design and Analysis*. Arizona State University, 2009.

12. Marcos N. Magalhães y Antonio Pedroso. *Nocões de Probabilidade e estatística*. Edusp, 1999.

13. Paul L. Meyer. *Probabilidad y aplicaciones estadísticas*. Fondo Educativo Interamericano, 1973.

14. Arturo Mora, Luis Cid y María Valenzuela. *Probabilidades y estadística*. Universidad de Concepción, Facultad de Ciencias Físicas y Matemáticas, Departamento de Estadística, 1996.

15. Pierre P. Romagnoli. *Probabilidades doctas con discos, árboles, bolitas y urnas*. JC Sáez Editor, 2011.

16. Sheldon M. Ross. *A first course in probability*. Pearson Education International, 2009.

17. Eugenio Saavedra. *Cálculo de probabilidades*. Editorial de la Universidad de Santiago de Chile, 2000.

18. Dennis D. Wackerly, William Mendenhall, Richard L. Scheaffer y Jorge R. Muñoz. *Estadística matemática con aplicaciones*. Thomson, 2002.

19. Ronald E. Walpole, Raymond H. Myers y Sharon L. Myers. *Probabilidad y estadística para ingenieros*. Pearson Educación, 1999.

Índice alfabético

Este libro fue compuesto por el equipo de
Ediciones Universidad de la Frontera y Rodrigo Puchi Ancasay
durante el confinamiento del invierno de 2020.
Para los textos del interior se utilizó la fuente CMU,
diseñada por Donald E. Knuth. En la portada se usó Libertad,
del tipógrafo Fernando Díaz.