Probabilidad: Conceptos Básicos

Notas de Clase

Asignatura: Matemática T

Carrera: TU Programación

Profesor: Lic. Pablo Girollet

Basado en: Webster (2000), Lind, Marchal & Wathen (2019),
Wackerly, Mendenhall & Scheaffer (2010), DeGroot & Schervish (2012)

Índice

Introducción

La incertidumbre es inherente a casi toda situación de toma de decisiones en los negocios, la ingeniería, la medicina y las ciencias sociales. Cuando un médico decide un tratamiento sin conocer con exacteza la respuesta del organismo, cuando un programador estima la probabilidad de que un módulo de software falle, o cuando una empresa proyecta ventas en un mercado incierto, todos ellos están —consciente o inconscientemente— razonando en términos de probabilidades.

La **Teoría de la Probabilidad** es la rama de las matemáticas que cuantifica y formaliza la incertidumbre. Proporciona el lenguaje preciso para describir experimentos cuyos resultados no pueden predecirse de antemano, y entrega herramientas para calcular cuán verosímiles son esos resultados. Desde el punto de vista estadístico, la probabilidad constituye el *puente* entre los datos muestrales y las conclusiones que extraemos sobre la población: sin ella, la inferencia estadística carecería de fundamento matemático.

En esta unidad estudiaremos los conceptos fundamentales de la probabilidad: el experimento aleatorio, el espacio muestral y los eventos; los tres enfoques para asignar probabilidades (clásico, frecuentista y subjetivo); las relaciones entre eventos; la probabilidad condicional; las reglas de la suma y la multiplicación; y dos resultados fundamentales: el Teorema de la Probabilidad Total y el Teorema de Bayes. Cada concepto se ilustra con ejemplos resueltos en detalle.

1. Importancia de la Probabilidad

1.1. ¿Por qué estudiar probabilidad?

Cuando analizamos datos reales, obtenemos una *muestra* de una población mucho más grande. Para poder generalizar —es decir, para hacer **inferencia estadística**— necesitamos evaluar cuán confiables son nuestras conclusiones. Esa evaluación depende directamente de la teoría de la probabilidad.

Existe también una razón de naturaleza más profunda: el mundo físico, social y económico está repleto de fenómenos que no pueden predecirse con certeza absoluta. El tiempo meteorológico, la demanda de un producto, el tiempo de vida de un componente electrónico, la respuesta de un usuario a una interfaz, el rendimiento de una red de comunicaciones: todos son fenómenos **aleatorios** o estocásticos. Modelarlos matemáticamente requiere la Teoría de la Probabilidad.

Definición informal de probabilidad

La **probabilidad** de un evento es un número comprendido entre 0 y 1 que expresa el grado de certeza de que dicho evento ocurra. Un valor de 0 indica imposibilidad; un valor de 1, certeza absoluta.

1.2. Modelos determinísticos vs. modelos estocásticos

Todo modelo matemático de un fenómeno observable es, en esencia, de una de estas dos clases:

Determinístico. El resultado queda completamente determinado por las condiciones iniciales. En física clásica, por ejemplo, si conocemos la fuerza aplicada sobre un cuerpo y su masa, podemos predecir su aceleración con precisión absoluta usando la Segunda Ley de Newton.

Estocástico (o aleatorio). Las condiciones iniciales no determinan unívocamente el resultado: hay variabilidad intrínseca. Aun repitiendo el experimento bajo las mismas condiciones, los resultados pueden diferir. Para estos fenómenos se construyen **modelos probabilísticos**.

2. Conceptos Previos Fundamentales

2.1. Experimento aleatorio

Definición — Experimento aleatorio

Un **experimento aleatorio** E es todo proceso de observación o medición que satisface las tres condiciones siguientes:

i) El resultado no puede conocerse con certeza antes de realizar el experimento.

ii) El conjunto de todos los resultados posibles puede describirse de antemano.

iii) El experimento puede, en principio, repetirse bajo las mismas condiciones.

Ejemplos de experimentos aleatorios:

1. Lanzar un dado equilibrado y anotar el número que aparece en la cara superior.
2. Lanzar una moneda tres veces consecutivas y registrar la secuencia de caras (C) y cruces (K).
3. Seleccionar al azar un componente de un lote de producción y determinar si es defectuoso o no.
4. Medir el tiempo (en horas) que tarda en fallar una lámpara de determinada marca.
5. Registrar el número de consultas que recibe un servidor web en un intervalo de un minuto.
6. Seleccionar aleatoriamente un estudiante de una carrera y anotar su nota final en la asignatura.

Nota

Nótese que los experimentos 1 a 3 tienen un número *finito* de resultados posibles, mientras que los experimentos 4 y 5 pueden tener un número infinito (o al menos muy grande) de resultados. Esta distinción será importante al momento de definir el espacio muestral.

2.2. Espacio muestral

Definición — Espacio muestral

El **espacio muestral** Ω (también llamado espacio de resultados) de un experimento aleatorio E es el conjunto de *todos* los resultados posibles de E.

El espacio muestral puede ser:

- **Finito**: contiene un número contable y acotado de elementos. Ejemplo: $\Omega = \{1, 2, 3, 4, 5, 6\}$ al lanzar un dado.
- **Infinito numerable**: infinito pero enumerable. Ejemplo: número de consultas a un servidor, $\Omega = \{0, 1, 2, 3, \ldots\}$.

- **Infinito no numerable**: continuo. Ejemplo: tiempo de vida de un componente, $\Omega = [0, +\infty)$.

Ejemplos de espacios muestrales:

Ejemplo 2.1 — Dado y monedas

a) **Lanzar un dado:** $\Omega = \{1, 2, 3, 4, 5, 6\}$, con $n = 6$ elementos.

b) **Lanzar una moneda dos veces:** $\Omega = \{(CC), (CK), (KC), (KK)\}$, con $n = 4$ elementos, donde C representa cara y K cruz.

c) **Lanzar una moneda tres veces:** $\Omega = \{CCC, CCK, CKC, CKK, KCC, KCK, KKC, KKK\}$, con $n = 8$ elementos.

d) **Control de calidad**: al extraer un artículo y clasificarlo, $\Omega = \{\text{defectuoso}, \text{ no defectuoso}\}$, con $n = 2$.

2.3. Eventos o sucesos

Definición — Evento

Un **evento** (o suceso) A es cualquier subconjunto del espacio muestral Ω. Los eventos se denotan con letras mayúsculas: $A, B, C, \ldots$

Se distinguen los siguientes tipos:

Evento simple. Contiene exactamente un resultado. Ej.: $A = \{3\}$ al lanzar un dado.

Evento compuesto. Contiene dos o más resultados. Ej.: $B = \{2, 4, 6\}$ (número par).

Evento seguro. Ω mismo. Siempre ocurre.

Evento imposible. El conjunto vacío $\emptyset$. Nunca ocurre.

Ejemplo 2.2 — Eventos al lanzar un dado

Experimento: lanzar un dado equilibrado. $\Omega = \{1, 2, 3, 4, 5, 6\}$.

- $A = \{1\}$: salga el número 1. *Evento simple.*
- $B = \{2, 4, 6\}$: salga un número par. *Evento compuesto.*
- $C = \{5, 6\}$: salga un número mayor que 4.
- $D = \{1, 2, 3, 4, 5, 6\} = \Omega$: salga algún número. *Evento seguro.*
- $E = \emptyset$: salga el número 7. *Evento imposible.*

2.4. Representación mediante diagramas de Venn

Los diagramas de Venn representan el espacio muestral Ω como un rectángulo, y los eventos como regiones (generalmente elipses o círculos) en su interior. Cada punto del rectángulo representa un resultado posible. La zona de solapamiento entre dos regiones corresponde a los resultados que pertenecen a ambos eventos a la vez. Las operaciones formales entre eventos (unión, intersección, complemento) se definirán en la Sección 5; el diagrama de Venn es la herramienta gráfica para visualizarlas.

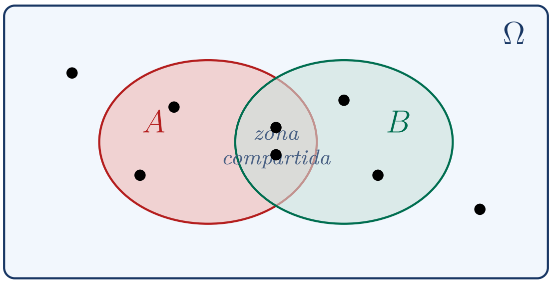

Figura 1: Diagrama de Venn: el rectángulo representa Ω; cada punto, un resultado posible; las regiones coloreadas, los eventos A y B. La zona de solapamiento contiene los resultados que pertenecen a ambos eventos.

3. Axiomas de Probabilidad

El fundamento matemático riguroso de la probabilidad se basa en tres axiomas formulados por el matemático soviético Andréi Kolmogórov en 1933. Estos axiomas definen qué es una función de probabilidad válida sobre un espacio muestral.

Definición — Función de probabilidad (axiomas de Kolmogórov)

Sea E un experimento aleatorio y Ω su espacio muestral. A cada evento $A \subseteq \Omega$ se le asigna un número real $P(A)$, llamado **probabilidad de** A, que satisface los siguientes axiomas:

Axioma 1. $0 \leq P(A) \leq 1$ (*No negatividad y acotamiento*)

Axioma 2. $P(\Omega) = 1$ (*Normalización*)

Axioma 3. Si $A_1, A_2, \ldots$ son eventos **mutuamente excluyentes** (disjuntos dos a dos, es decir, no pueden ocurrir simultáneamente; la definición precisa se da en la Sección 5), entonces:

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

(*Aditividad*)

De estos tres axiomas se derivan todas las demás propiedades de la probabilidad.

Propiedades derivadas de los axiomas

$$P(\emptyset) = 0 \tag{P1}$$

$$P(\bar{A}) = 1 - P(A) \tag{P2}$$

$$\text{Si } A \subseteq B, \text{ entonces } P(A) \leq P(B) \tag{P3}$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \tag{P4}$$

$$0 \leq P(A) \leq 1 \quad \forall\, A \subseteq \Omega \tag{P5}$$

Interpretación práctica de P2: La probabilidad del complemento de un evento es el complemento de su probabilidad. Esta propiedad es muy útil porque a veces es más sencillo calcular la probabilidad del evento contrario.

Ejemplo 3.1 — Verificación de axiomas

Se lanza una moneda equilibrada dos veces. $\Omega = \{CC, CK, KC, KK\}$. Asignamos $P(\{CC\}) = P(\{CK\}) = P(\{KC\}) = P(\{KK\}) = 1/4$.

- **Axioma 1:** Cada probabilidad vale $1/4 \in [0, 1]$. ✓
- **Axioma 2:** $P(\Omega) = 4 \times 1/4 = 1$. ✓
- **Axioma 3:** Los cuatro eventos simples son disjuntos y su suma es 1. ✓

Sea $A = \{\text{al menos una cara}\} = \{CC, CK, KC\}$. Entonces $P(A) = 3/4$ y $P(\bar{A}) = P(\{KK\}) = 1/4 = 1 - 3/4$. ✓

4. Enfoques para Asignar Probabilidades

Existen tres enfoques principales para asignar valores numéricos de probabilidad a los eventos de un experimento. Los tres son compatibles con los axiomas de Kolmogórov.

4.1. Enfoque clásico (o *a priori*)

Definición — Probabilidad clásica (Laplace)

Si un experimento aleatorio tiene n resultados **igualmente posibles** y mutuamente excluyentes, y el evento A está formado por n_A de esos resultados, entonces:

$$\boxed{P(A) = \frac{n_A}{n} = \frac{\text{N}^\circ \text{ de casos favorables a } A}{\text{N}^\circ \text{ total de resultados posibles}}}$$

Este enfoque fue desarrollado originalmente por Pierre-Simon Laplace (1749–1827) para analizar problemas de juegos de azar. El supuesto fundamental es la **equiprobabilidad**: todos los resultados elementales tienen la misma probabilidad de ocurrir.

Cuándo es adecuado:

- Juegos de azar con monedas, dados, cartas o ruletas correctamente construidas.
- Sorteos o extracciones al azar de poblaciones bien definidas.

Limitaciones:

- No aplicable cuando los resultados no son equiprobables.
- Inaplicable cuando el espacio muestral es infinito.
- En contextos de negocios, la equiprobabilidad es raramente válida.

Ejemplo 4.1 — Dado equilibrado

Se lanza un dado equilibrado. $\Omega = \{1, 2, 3, 4, 5, 6\}$, $n = 6$.

a) A: salga un número par. $A = \{2, 4, 6\}$, $n_A = 3$.

$$P(A) = \frac{3}{6} = \frac{1}{2} = 0{,}5$$

b) B: salga un número mayor que 4. $B = \{5, 6\}$, $n_B = 2$.

$$P(B) = \frac{2}{6} = \frac{1}{3} \approx 0{,}333$$

c) C: salga el número 7. $C = \emptyset$, $n_C = 0$.

$$P(C) = \frac{0}{6} = 0$$

d) $D = \Omega$: salga cualquier número. $n_D = 6$.

$$P(D) = \frac{6}{6} = 1$$

Ejemplo 4.2 — Baraja de 40 cartas

Se extrae una carta al azar de una baraja española de 40 cartas (4 palos: espadas, bastos, copas, oros; 10 cartas por palo: as, 2, 3, 4, 5, 6, 7, sota, caballo, rey). $n = 40$.

a) A: salga un as. $n_A = 4$ (uno por cada palo).

$$P(A) = \frac{4}{40} = \frac{1}{10} = 0{,}1$$

b) B: salga una espada. $n_B = 10$ (10 cartas de espada).

$$P(B) = \frac{10}{40} = \frac{1}{4} = 0{,}25$$

c) C: salga un rey. $n_C = 4$.

$$P(C) = \frac{4}{40} = 0{,}1$$

d) D: salga una figura (sota, caballo o rey). $n_D = 12$.

$$P(D) = \frac{12}{40} = 0{,}3$$

Ejemplo 4.3 — Extracción de una carta con condición

Un estudiante de programación prepara una tómbola con 20 tarjetas numeradas del 1 al 20. Calcula las siguientes probabilidades usando el enfoque clásico:

a) A: la tarjeta extraída tiene un número par. $A = \{2, 4, 6, 8, 10, 12, 14, 16, 18, 20\}$, $n_A = 10$.

$$P(A) = \frac{10}{20} = \frac{1}{2} = 0{,}5$$

b) B: la tarjeta tiene un número mayor que 15. $B = \{16, 17, 18, 19, 20\}$, $n_B = 5$.

$$P(B) = \frac{5}{20} = \frac{1}{4} = 0{,}25$$

c) C: la tarjeta tiene un número múltiplo de 3. $C = \{3, 6, 9, 12, 15, 18\}$, $n_C = 6$.

$$P(C) = \frac{6}{20} = \frac{3}{10} = 0{,}3$$

d) $\bar{A}$: la tarjeta tiene un número impar. $n_{\bar{A}} = 10$.

$$P(\bar{A}) = \frac{10}{20} = 0{,}5 = 1 - P(A) = 1 - 0{,}5$$

4.2. Enfoque de frecuencia relativa (*a posteriori*)

Definición — Probabilidad frecuentista

Si un experimento se repite n veces bajo las mismas condiciones y el evento A ocurre n_A veces, entonces la **probabilidad frecuentista** de A se estima como:

$$\boxed{P(A) \approx \frac{n_A}{n} = \frac{\text{N° de veces que ocurrió } A}{\text{N° total de ensayos}}}$$

donde la aproximación mejora conforme $n \to \infty$ (Ley de los Grandes Números).

Este enfoque usa **datos históricos empíricos** para estimar probabilidades futuras. Es el más utilizado en los negocios, la ingeniería y la medicina.

Condiciones necesarias:

- El número de ensayos debe ser **suficientemente grande** para que la estimación sea confiable.
- Las condiciones del experimento deben poder **reproducirse** al menos aproximadamente.

Ejemplo 4.4 — Control de calidad

Una fábrica de componentes electrónicos inspeccionó 2 000 unidades durante el último mes y encontró 80 defectuosas. Se estima la probabilidad de que una unidad elegida al azar sea defectuosa:

$$P(\text{defectuoso}) = \frac{80}{2000} = 0{,}04 = 4\,\%$$

Esta estimación se utilizará para proyecciones de garantía y planificación de la producción.

Ejemplo 4.5 — Nacimientos en un hospital

Durante el año 2023, en un hospital regional nacieron 500 bebés, de los cuales 262 fueron nenas. Se estima la probabilidad de que el próximo nacimiento sea una nena:

$$P(\text{nena}) = \frac{262}{500} = 0{,}524$$

Nótese que la probabilidad *no es* exactamente 0,5 como podría suponerse desde el enfoque clásico, ya que en la realidad hay ligeras variaciones biológicas.

Ejemplo 4.6 — Fiabilidad de un sistema

Un sistema de servidores registró 365 días de operación, de los cuales en 12 días hubo al menos una caída del servicio. La probabilidad de que un día elegido al azar tenga una caída es:

$$P(\text{caída}) = \frac{12}{365} \approx 0{,}033 \approx 3{,}3\,\%$$

Esta información es clave para planificar mecanismos de redundancia y acuerdos de nivel de servicio (SLA).

4.3. Enfoque subjetivo

Definición — Probabilidad subjetiva

La **probabilidad subjetiva** es una medida del *grado de creencia* personal de un individuo respecto a la ocurrencia de un evento, basada en su experiencia, juicio experto e información disponible, cuando no existe simetría ni datos históricos aplicables.

Se emplea cuando:

- El evento es único o irrepetible (e.g., el éxito de un nuevo producto).
- No hay datos históricos relevantes.
- Las condiciones del experimento no pueden reproducirse.

Ejemplos de probabilidades subjetivas:

- Un analista afirma que hay un 70 % de probabilidad de que la economía entre en recesión el próximo año.
- Un médico estima en 90 % la probabilidad de que una cirugía tenga éxito en un paciente específico.
- Un desarrollador de software estima en 60 % la probabilidad de lanzar su aplicación antes del plazo previsto.

Nota importante

Aunque la probabilidad subjetiva refleja una creencia personal, debe igualmente satisfacer los axiomas de Kolmogórov: cada probabilidad debe estar entre 0 y 1, y si se consideran todos los resultados posibles de manera que cubran completamente el espacio muestral sin solaparse, sus probabilidades deben sumar 1. Esta exigencia garantiza la coherencia interna del juicio experto.

5. Relaciones entre Eventos

Para operar con probabilidades de eventos complejos, necesitamos definir rigurosamente las operaciones sobre eventos, que son análogas a las operaciones sobre conjuntos.

5.1. Unión, intersección y complemento

Definición — Operaciones entre eventos

Sean A y B dos eventos del espacio muestral Ω:

Operación	Notación	Descripción
Unión	$A \cup B$	Ocurre A, ocurre B, o ambos
Intersección	$A \cap B$	Ocurren A y B simultáneamente
Complemento de A	$\bar{A}$ (o A^c)	No ocurre A

Definición — Eventos mutuamente excluyentes

Dos eventos A y B son **mutuamente excluyentes** (o incompatibles) si no pueden ocurrir simultáneamente, es decir:

$$A \cap B = \emptyset \quad \Longrightarrow \quad P(A \cap B) = 0$$

Definición — Eventos exhaustivos

Un conjunto de eventos $A_1, A_2, \ldots, A_n$ es **exhaustivo** (o **colectivamente exhaustivo**) si entre todos ellos cubren la totalidad del espacio muestral, es decir, si al menos uno de ellos *siempre* debe ocurrir:

$$A_1 \cup A_2 \cup \cdots \cup A_n = \Omega$$

En particular, A y $\bar{A}$ son siempre exhaustivos, pues $A \cup \bar{A} = \Omega$.

Definición — Eventos complementarios

$\bar{A}$ es el complemento de A: el conjunto de todos los resultados de Ω que no pertenecen a A. Se verifica:

$$A \cup \bar{A} = \Omega, \quad A \cap \bar{A} = \emptyset, \quad P(A) + P(\bar{A}) = 1$$

Ejemplo 5.1 — Operaciones con eventos (baraja de 40 cartas)

Experimento: se extrae una carta al azar de una baraja española de 40 cartas (4 palos × 10 cartas cada palo: as, 2, ..., 7, sota, caballo, rey). Se definen:

$$A = \{\text{as}\}, \quad B = \{\text{espada}\}, \quad C = \{\text{rey}\}$$

a) Unión $A \cup B$: la carta es un as, una espada, o ambas cosas a la vez. Enumerando: hay 4 ases (as de espadas, bastos, copas y oros) y 10 espadas; pero el as de espadas ya fue contado entre los ases. En total, $A \cup B$ tiene $4 + 10 - 1 = 13$ elementos distintos.

$$P(A \cup B) = \frac{13}{40}$$

(La fórmula que justifica formalmente este conteo se estudiará en la Sección 7, Regla de la Suma.)

b) Unión $A \cup C$: la carta es un as o un rey. Como ninguna carta puede ser as y rey a la vez, no hay solapamiento: $A \cup C$ tiene $4 + 4 = 8$ elementos.

$$P(A \cup C) = \frac{8}{40} = 0{,}2$$

c) Intersección $A \cap B$: la carta es as y espada al mismo tiempo. Solo existe una carta así: el as de espadas. $n_{A \cap B} = 1$.

$$P(A \cap B) = \frac{1}{40} = 0{,}025$$

d) Intersección $A \cap C$: la carta sería as y rey a la vez, lo cual es imposible. $A \cap C = \emptyset$; por lo tanto A y C son **mutuamente excluyentes**.
e) Complemento $\bar{A}$: la carta no es un as. Los resultados favorables son todas las cartas que no son ases: $40 - 4 = 36$. $n_{\bar{A}} = 36$.

$$P(\bar{A}) = \frac{36}{40} = 0{,}9$$

Verificación con la propiedad del complemento (Sección 3): $1 - P(A) = 1 - \frac{4}{40} = \frac{36}{40} = 0{,}9$. ✓

5.2. Leyes de De Morgan

Las leyes de De Morgan son útiles para trabajar con complementos de uniones e intersecciones:

Leyes de De Morgan

$$\overline{A \cup B} = \bar{A} \cap \bar{B} \qquad \overline{A \cap B} = \bar{A} \cup \bar{B}$$

En palabras: el complemento de la unión es la intersección de los complementos; el complemento de la intersección es la unión de los complementos.

6. Probabilidad Condicional

6.1. Concepto e interpretación

Muchas veces el conocimiento de que cierto evento ha ocurrido modifica la probabilidad de que otro evento ocurra. Formalizamos esto con el concepto de **probabilidad condicional**.

Definición — Probabilidad condicional

Sean A y B dos eventos con $P(B) \neq 0$. La **probabilidad condicional** de A dado que B ha ocurrido se define como:

$$\boxed{P(A \mid B) = \frac{P(A \cap B)}{P(B)}}$$

Análogamente, si $P(A) \neq 0$:

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}$$

Interpretación: al condicionar en B, el nuevo espacio muestral se reduce a B, y dentro de él calculamos la proporción que corresponde a $A \cap B$.

Ejemplo 6.1 — Dado: número par sabiendo que es mayor que 4

Se lanza un dado equilibrado. $\Omega = \{1, 2, 3, 4, 5, 6\}$.

$$A = \{2, 4, 6\} \text{ (par)}, \quad B = \{5, 6\} \text{ (mayor que 4)}$$

$$A \cap B = \{6\}, \quad P(A \cap B) = \frac{1}{6}, \quad P(B) = \frac{2}{6}$$

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{1/6}{2/6} = \frac{1}{2}$$

Intuición: si sabemos que el resultado es 5 o 6 (evento B), solo el 6 es par, luego la probabilidad de par dado que salió mayor que 4 es $1/2$.

Ejemplo 6.2 — Control de calidad con dos líneas de producción

Una planta tiene dos líneas de producción. El 60 % de las piezas proviene de la Línea 1 y el 40 % de la Línea 2. De las piezas de la Línea 1, el 3 % son defectuosas; de las de la Línea 2, el 5 % son defectuosas.

Definamos:

$$\begin{aligned} L_1 &= \text{pieza proviene de Línea 1}, & P(L_1) &= 0{,}60 \\ L_2 &= \text{pieza proviene de Línea 2}, & P(L_2) &= 0{,}40 \\ D &= \text{pieza es defectuosa} \end{aligned}$$

Las probabilidades condicionales que describen la tasa de defecto de cada línea son:

$$P(D \mid L_1) = 0{,}03, \qquad P(D \mid L_2) = 0{,}05$$

Lectura de estas probabilidades:

- $P(D \mid L_1) = 0{,}03$ significa: *dado que* la pieza proviene de la Línea 1, la probabilidad de que sea defectuosa es 3 %.
- $P(D \mid L_2) = 0{,}05$ significa: *dado que* la pieza proviene de la Línea 2, la probabilidad de que sea defectuosa es 5 %.

Nota: para calcular la probabilidad de seleccionar al azar una pieza que sea defectuosa (independientemente de su línea de origen) se necesita combinar estas probabilidades condicionales; ese cálculo se realizará en los Ejemplos 7.1 y **??** de las secciones siguientes.

Ejemplo 6.3 — Probabilidad condicional con tabla de contingencia

En una empresa de desarrollo de software, 200 empleados fueron encuestados sobre su lenguaje preferido y su nivel de experiencia:

	Python	**Java**	**Otro**	**Total**
Junior	40	30	20	90
Senior	50	35	25	110
Total	90	65	45	200

a) Probabilidad de que un empleado elegido al azar sea Senior:

$$P(\text{Senior}) = \frac{110}{200} = 0{,}55$$

b) Probabilidad de que prefiera Python dado que es Senior:

$$P(\text{Python} \mid \text{Senior}) = \frac{50/200}{110/200} = \frac{50}{110} \approx 0{,}4545$$

c) Probabilidad de que sea Junior dado que prefiere Java:

$$P(\text{Junior} \mid \text{Java}) = \frac{30/200}{65/200} = \frac{30}{65} \approx 0{,}4615$$

6.2. Tablas de contingencia

6.2.1. ¿Qué es una tabla de contingencia?

En muchos estudios se observan simultáneamente **dos variables categóricas** sobre los mismos sujetos o unidades de análisis. Por ejemplo: el género y la preferencia de producto de un cliente; el sistema operativo y el tipo de error reportado; el nivel de experiencia y el lenguaje de programación preferido de un desarrollador. Una **tabla de contingencia** (también llamada tabla de doble entrada o tabla de clasificación cruzada) organiza y resume esa información de manera que las relaciones entre las variables quedan a la vista de forma inmediata.

Definición — Tabla de contingencia

Una **tabla de contingencia** es una tabla de doble entrada que muestra la distribución de frecuencias *conjuntas* de dos variables categóricas. Las categorías de una variable se disponen en las filas y las de la otra en las columnas. Cada celda interior contiene la **frecuencia conjunta**: el número de observaciones que pertenecen simultáneamente a esa categoría de fila y a esa categoría de columna.

6.2.2. Estructura general de una tabla de contingencia

Sea una muestra de n observaciones clasificadas según dos variables: la variable de fila R con categorías $R_1, R_2, \ldots, R_r$, y la variable de columna C con categorías $C_1, C_2, \ldots, C_c$.

	C_1	C_2	$\cdots$	C_c	**Total fila**
R_1	n_{11}	n_{12}	$\cdots$	n_{1c}	$n_{1\bullet}$
R_2	n_{21}	n_{22}	$\cdots$	n_{2c}	$n_{2\bullet}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
R_r	n_{r1}	n_{r2}	$\cdots$	n_{rc}	$n_{r\bullet}$
Total columna	$n_{\bullet 1}$	$n_{\bullet 2}$	$\cdots$	$n_{\bullet c}$	n

La notación utilizada es la siguiente:

n_{ij}: **Frecuencia conjunta** de la celda (i, j): número de observaciones en la categoría R_i de la variable de fila y en la categoría C_j de la variable de columna.

$n_{i\bullet} = \sum_{j=1}^{c} n_{ij}$: **Total de fila** (o frecuencia marginal de fila): número total de observaciones en la categoría R_i, independientemente de la columna.

$n_{\bullet j} = \sum_{i=1}^{r} n_{ij}$: **Total de columna** (o frecuencia marginal de columna): número total de observaciones en la categoría C_j, independientemente de la fila.

$n = \sum_{i=1}^{r} n_{i\bullet} = \sum_{j=1}^{c} n_{\bullet j}$: **Total general**: número total de observaciones en la muestra.

Por qué se llama ”de contingencia”

El nombre hace referencia a la idea de *contingencia* entre las variables: si conocer el valor de una variable cambia las probabilidades de los valores de la otra, decimos que las variables son contingentes entre sí (es decir, no son independientes). Las tablas de contingencia son precisamente la herramienta para detectar y cuantificar esa dependencia.

6.2.3. Tipos de probabilidad que se extraen de una tabla

A partir de los datos de una tabla de contingencia con n observaciones totales se pueden calcular tres tipos de probabilidades:

Probabilidades obtenibles de una tabla de contingencia

Tipo	Fórmula general	Descripción
Marginal de fila	$P(R_i) = \frac{n_{i\bullet}}{n}$	Prob. de pertenecer a la categoría R_i
Marginal de columna	$P(C_j) = \frac{n_{\bullet j}}{n}$	Prob. de pertenecer a la categoría C_j
Conjunta	$P(R_i \cap C_j) = \frac{n_{ij}}{n}$	Prob. de pertenecer a R_i y C_j simultáneamente
Condicional	$P(C_j \mid R_i) = \frac{n_{ij}}{n_{i\bullet}}$	Prob. de C_j sabiendo que se está en R_i
Condicional	$P(R_i \mid C_j) = \frac{n_{ij}}{n_{\bullet j}}$	Prob. de R_i sabiendo que se está en C_j

Verificación de coherencia con la definición de probabilidad condicional: la fórmula de la condicional desde la tabla es consistente con la definición de la Sección 6.1:

$$P(C_j \mid R_i) = \frac{P(R_i \cap C_j)}{P(R_i)} = \frac{n_{ij}/n}{n_{i\bullet}/n} = \frac{n_{ij}}{n_{i\bullet}}$$

El denominador n se cancela, y el resultado es simplemente el cociente entre la frecuencia conjunta de la celda y el total de fila correspondiente.

6.2.4. Verificación de independencia desde la tabla

Recordando que dos eventos A y B son independientes cuando $P(A \cap B) = P(A) \cdot P(B)$ (Sección 7), en el contexto de tablas de contingencia las variables de fila y columna son **independientes** si y solo si, para toda celda (i, j):

$$P(R_i \cap C_j) = P(R_i) \cdot P(C_j) \quad \Longleftrightarrow \quad \frac{n_{ij}}{n} = \frac{n_{i\bullet}}{n} \cdot \frac{n_{\bullet j}}{n}$$

Equivalentemente, la independencia implica que la probabilidad condicional $P(C_j \mid R_i)$ es igual a la marginal $P(C_j)$ para toda combinación (i, j): conocer la fila no modifica la distribución de la columna.

Tabla de contingencia vs. diagrama de Venn

Ambas son herramientas para visualizar relaciones entre eventos, pero con alcances distintos. El diagrama de Venn es útil para dos o tres eventos con pocas categorías cada uno. La tabla de contingencia es más práctica cuando se trabaja con datos reales provenientes de muestras: permite leer de manera directa todas las probabilidades (marginales, conjuntas y condicionales) y verificar la independencia sin necesidad de construir diagramas.

6.2.5. Tabla de contingencia de probabilidades

Además de trabajar con frecuencias absolutas, es habitual presentar la tabla en términos de **frecuencias relativas** (proporciones), dividiendo cada celda por el total n. Se obtiene así una **tabla de probabilidades conjuntas**, donde cada celda muestra $P(R_i \cap C_j)$, los totales de fila muestran $P(R_i)$ y los totales de columna muestran $P(C_j)$. La suma de todas las celdas interiores es 1.

	C_1	C_2	$\cdots$	**Total (marginal de fila)**
R_1	$P(R_1 \cap C_1)$	$P(R_1 \cap C_2)$	$\cdots$	$P(R_1)$
R_2	$P(R_2 \cap C_1)$	$P(R_2 \cap C_2)$	$\cdots$	$P(R_2)$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
Total (marginal de columna)	$P(C_1)$	$P(C_2)$	$\cdots$	1

6.3. Ejemplos de tablas de contingencia

Ejemplo 6.3 — Lenguaje preferido y nivel de experiencia

En una empresa de desarrollo de software, 200 empleados fueron encuestados sobre su **lenguaje de programación preferido** y su **nivel de experiencia**. Los resultados se resumen en la siguiente tabla de contingencia:

	Python	**Java**	**Otro**	**Total**
Junior	40	30	20	90
Senior	50	35	25	110
Total	90	65	45	200

Definimos los eventos: $J = \{\text{Junior}\}$, $S = \{\text{Senior}\}$, $Py = \{\text{Python}\}$, $Jv = \{\text{Java}\}$, $O = \{\text{Otro}\}$.

a) Probabilidades marginales (se leen en los totales):

$$P(J) = \frac{90}{200} = 0{,}45 \qquad P(S) = \frac{110}{200} = 0{,}55$$
$$P(Py) = \frac{90}{200} = 0{,}45 \qquad P(Jv) = \frac{65}{200} = 0{,}325 \qquad P(O) = \frac{45}{200} = 0{,}225$$

Verificación: $P(J) + P(S) = 0{,}45 + 0{,}55 = 1$. ✓
Verificación: $P(Py) + P(Jv) + P(O) = 0{,}45 + 0{,}325 + 0{,}225 = 1$. ✓

b) Probabilidades conjuntas (se leen en las celdas interiores, dividiendo por $n = 200$):

$$P(J \cap Py) = \frac{40}{200} = 0{,}20 \quad P(J \cap Jv) = \frac{30}{200} = 0{,}15 \quad P(J \cap O) = \frac{20}{200} = 0{,}10$$
$$P(S \cap Py) = \frac{50}{200} = 0{,}25 \quad P(S \cap Jv) = \frac{35}{200} = 0{,}175 \quad P(S \cap O) = \frac{25}{200} = 0{,}125$$

Verificación: $0{,}20 + 0{,}15 + 0{,}10 + 0{,}25 + 0{,}175 + 0{,}125 = 1$. ✓

c) Probabilidades condicionales (se leen dentro de una fila o columna):

Dado que el empleado es Senior, ¿cuál es la probabilidad de que prefiera Python?

$$P(Py \mid S) = \frac{n_{S,Py}}{n_{S\bullet}} = \frac{50}{110} \approx 0{,}4545$$

Dado que el empleado es Junior, ¿cuál es la probabilidad de que prefiera Java?

$$P(Jv \mid J) = \frac{30}{90} = \frac{1}{3} \approx 0{,}333$$

Dado que el empleado prefiere Java, ¿cuál es la probabilidad de que sea Junior?

$$P(J \mid Jv) = \frac{30}{65} \approx 0{,}4615$$

Verificación con la definición formal: $P(Py \mid S) = \dfrac{P(S \cap Py)}{P(S)} = \dfrac{50/200}{110/200} = \dfrac{50}{110} \approx 0{,}4545$. ✓

d) Verificación de independencia entre nivel y lenguaje preferido:
Si las variables fueran independientes, debería cumplirse $P(J \cap Py) = P(J) \cdot P(Py)$ para todas las celdas.

$$P(J) \cdot P(Py) = 0{,}45 \times 0{,}45 = 0{,}2025 \quad \neq \quad P(J \cap Py) = 0{,}20$$

Como $0{,}2025 \neq 0{,}20$, las variables **no son independientes**: el nivel de experiencia y el lenguaje preferido están asociados. Podemos confirmarlo comparando probabilidades condicionales:

$$P(Py \mid J) = \frac{40}{90} \approx 0{,}444 \quad \neq \quad P(Py \mid S) = \frac{50}{110} \approx 0{,}455 \quad \neq \quad P(Py) = 0{,}45$$

Las diferencias son pequeñas, pero estadísticamente la tabla indica asociación.

Ejemplo 6.4 — Sistema operativo y tipo de error reportado

El equipo de soporte técnico de una empresa registró durante un mes $n = 500$ tickets de soporte, clasificados según el **sistema operativo** del usuario y el **tipo de error** reportado:

	Error de red	Error de software	Error de hardware	Total
Windows	80	150	50	280
macOS	40	60	20	120
Linux	60	30	10	100
Total	180	240	80	500

Definimos: W = Windows, M = macOS, L = Linux; R = error de red, Sw = error de software, H = error de hardware.

a) Tabla de probabilidades conjuntas (dividiendo cada celda por 500):

	Red	Software	Hardware	Marginal
Windows	0,160	0,300	0,100	0,560
macOS	0,080	0,120	0,040	0,240
Linux	0,120	0,060	0,020	0,200
Marginal	0,360	0,480	0,160	1,000

b) Probabilidades marginales:

$$P(W) = 0{,}560, \qquad P(M) = 0{,}240, \qquad P(L) = 0{,}200$$
$$P(R) = 0{,}360, \qquad P(Sw) = 0{,}480, \qquad P(H) = 0{,}160$$

c) Probabilidades condicionales (distribución del tipo de error dentro de cada sistema operativo):

Distribución de errores dado Windows:

$$P(R \mid W) = \frac{80}{280} \approx 0{,}286, \quad P(Sw \mid W) = \frac{150}{280} \approx 0{,}536, \quad P(H \mid W) = \frac{50}{280} \approx 0{,}179$$

Distribución de errores dado Linux:

$$P(R \mid L) = \frac{60}{100} = 0{,}600, \quad P(Sw \mid L) = \frac{30}{100} = 0{,}300, \quad P(H \mid L) = \frac{10}{100} = 0{,}100$$

d) Interpretación: los usuarios de Linux reportan errores de red con una frecuencia mucho mayor (60 %) que los de Windows (28,6 %) o macOS. Esto sugiere que el sistema operativo y el tipo de error *no son independientes*, lo que podría orientar al equipo de soporte a investigar problemas de configuración de red específicos del entorno Linux.

e) Verificación de independencia para la celda (Linux, Red):

$$P(L) \cdot P(R) = 0{,}200 \times 0{,}360 = 0{,}072 \quad \neq \quad P(L \cap R) = \frac{60}{500} = 0{,}120$$

La diferencia es clara: las variables **no son independientes**. ×

Ejemplo 6.5 — Verificación de independencia en una tabla

Se desea verificar si, en el ejemplo anterior, las variables son independientes para los usuarios de **macOS** con **error de software**:

$$P(M) \cdot P(Sw) = 0{,}240 \times 0{,}480 = 0{,}1152 \qquad \text{vs} \qquad P(M \cap Sw) = \frac{60}{500} = 0{,}120$$

$0{,}1152 \neq 0{,}120$: tampoco son independientes para esta celda.

Para que las variables fuesen completamente independientes, *todas* las celdas de la tabla deberían satisfacer $P(R_i \cap C_j) = P(R_i) \cdot P(C_j)$. Basta que una sola no lo cumpla para que las variables no sean independientes.

Tabla de frecuencias esperadas bajo independencia (valores que deberían observarse si las variables fuesen independientes):

$$E_{ij} = n \cdot P(R_i) \cdot P(C_j) = \frac{n_{i\bullet} \times n_{\bullet j}}{n}$$

	Red	**Software**	**Hardware**
Windows	$500 \times 0{,}560 \times 0{,}360 = 100{,}8$	134,4	44,8
macOS	$500 \times 0{,}240 \times 0{,}360 = 43{,}2$	57,6	19,2
Linux	$500 \times 0{,}200 \times 0{,}360 = 36{,}0$	48,0	16,0

Comparando con los valores observados (fila Linux: 60, 30, 10), las diferencias son apreciables, lo que confirma la falta de independencia. El análisis formal de si esas diferencias son estadísticamente significativas se realiza mediante el *test chi-cuadrado de independencia*, que se estudia en unidades posteriores.

7. Reglas de la Probabilidad

7.1. Regla de la multiplicación

La regla de la multiplicación permite calcular la probabilidad de la **intersección** de dos eventos.

Regla de la multiplicación — Forma general

Para cualesquiera dos eventos A y B:

$$P(A \cap B) \;=\; P(A) \times P(B \mid A) \;=\; P(B) \times P(A \mid B)$$

Ejemplo 7.1 — Extracción sin reposición

De una urna con 5 bolas rojas y 3 azules se extraen 2 bolas **sin reposición**. ¿Cuál es la probabilidad de que ambas sean rojas?
Sea R_1 = primera bola roja, R_2 = segunda bola roja.

$$P(R_1) = \frac{5}{8}, \qquad P(R_2 \mid R_1) = \frac{4}{7}$$

(después de sacar una roja, quedan 4 rojas y 3 azules = 7 bolas).

$$P(R_1 \cap R_2) = \frac{5}{8} \times \frac{4}{7} = \frac{20}{56} = \frac{5}{14} \approx 0{,}357$$

Ejemplo 7.2 — Líneas de producción (retomando Ejemplo 6.2)

Retomamos el Ejemplo 6.2: $P(L_1) = 0{,}60$, $P(L_2) = 0{,}40$, $P(D \mid L_1) = 0{,}03$, $P(D \mid L_2) = 0{,}05$.

a) ¿Cuál es la probabilidad de que una pieza provenga de la Línea 1 **y además** sea defectuosa?

$$P(L_1 \cap D) = P(L_1) \times P(D \mid L_1) = 0{,}60 \times 0{,}03 = 0{,}018$$

b) ¿Cuál es la probabilidad de que una pieza provenga de la Línea 2 **y además** sea defectuosa?

$$P(L_2 \cap D) = P(L_2) \times P(D \mid L_2) = 0{,}40 \times 0{,}05 = 0{,}020$$

Estos resultados parciales se utilizarán en la Sección **??** para calcular la probabilidad total de defecto y las probabilidades inversas.

Ejemplo 7.3 — Baraja (verificación con dos métodos)

Se extrae una carta de una baraja de 40 cartas. Hallar $P(\text{as} \cap \text{espada})$.

Método 1 (clásico): el único as de espada da $P = 1/40$.

Método 2 (multiplicación):

$$P(A \cap B) = P(A) \times P(B \mid A) = \frac{4}{40} \times \frac{1}{4} = \frac{1}{40}$$

Ambos métodos coinciden. ✓

7.2. Sucesos independientes

Definición — Independencia de eventos

Dos eventos A y B son **independientes** si y solo si la ocurrencia de uno de ellos no modifica la probabilidad del otro:

$$P(A \mid B) = P(A) \quad \text{y} \quad P(B \mid A) = P(B)$$

Teniendo en cuenta esto último y considerando la **Regla de multipicación** en su forma general antes vista, se deduce que:

Regla de la multiplicación — Forma especial

A y B son *independientes* si y solo si:

$$P(A \cap B) \;=\; P(A) \times P(B)$$

Independencia vs. exclusión mutua

Atención: independencia y exclusión mutua son conceptos completamente distintos e incluso, en cierto sentido, opuestos.

- Si A y B son **mutuamente excluyentes** y $P(A) > 0$, $P(B) > 0$, entonces *no pueden* ser independientes, pues $P(A \cap B) = 0 \neq P(A) \cdot P(B) > 0$.
- La independencia implica que los eventos pueden ocurrir simultáneamente; la exclusión mutua implica que no pueden.

Ejemplo 7.4 — Lanzamiento de dos dados independientes

Se lanzan dos dados equilibrados de manera independiente. ¿Cuál es la probabilidad de que ambos muestren un número par?
Sea A = primer dado par ($P(A) = 3/6 = 1/2$) y B = segundo dado par ($P(B) = 1/2$). Por independencia:

$$P(A \cap B) = P(A) \times P(B) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} = 0{,}25$$

Ejemplo 7.5 — Verificación de independencia (baraja)

En el ejemplo de la baraja (40 cartas), definamos A = as y B = espada.

$$P(A) = \frac{4}{40} = 0{,}1, \quad P(B) = \frac{10}{40} = 0{,}25, \quad P(A \cap B) = \frac{1}{40} = 0{,}025$$

¿Son independientes? Verificamos:

$$P(A) \times P(B) = 0{,}1 \times 0{,}25 = 0{,}025 = P(A \cap B)$$

Sí son independientes. El palo de la carta no afecta al valor, ni viceversa. ✓

7.3. Regla de la multiplicación para n eventos independientes

Si $A_1, A_2, \dots, A_n$ son eventos **mutuamente independientes**:

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1) \times P(A_2) \times \cdots \times P(A_n)$$

Ejemplo 7.6 — Fiabilidad de un sistema en serie

Un sistema electrónico tiene tres componentes en serie. El sistema falla si *cualquiera* de los componentes falla. Las probabilidades de que cada componente **funcione** correctamente son: $P(C_1) = 0{,}95$, $P(C_2) = 0{,}98$, $P(C_3) = 0{,}99$. Los componentes fallan de manera independiente.
La probabilidad de que el sistema funcione es:

$$P(\text{sistema funciona}) = P(C_1) \times P(C_2) \times P(C_3) = 0{,}95 \times 0{,}98 \times 0{,}99 \approx 0{,}9212$$

La probabilidad de que el sistema **falle** es:

$$P(\text{sistema falla}) = 1 - 0{,}9212 = 0{,}0788 \approx 7{,}88\,\%$$

7.4. Regla de la suma

La regla de la suma calcula la probabilidad de la **unión** de dos eventos.

Regla de la suma

Para cualesquiera dos eventos A y B:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Si A y B son **mutuamente excluyentes** ($A \cap B = \emptyset$):

$$P(A \cup B) = P(A) + P(B)$$

El término $-P(A \cap B)$ corrige el *doble conteo*: los elementos de $A \cap B$ fueron incluidos tanto en $P(A)$ como en $P(B)$.

Ejemplo 7.7 — Regla de la suma (baraja)

Baraja de 40 cartas. Calcular la probabilidad de que la carta extraída sea:

a) As o rey (eventos mutuamente excluyentes o incompatibles):

$$P(A \cup C) = P(A) + P(C) = \frac{4}{40} + \frac{4}{40} = \frac{8}{40} = 0{,}2$$

b) As o espada (eventos compatibles, pues existe el as de espada):

$$P(A \cup B) = \frac{4}{40} + \frac{10}{40} - \frac{1}{40} = \frac{13}{40} = 0{,}325$$

Ejemplo 7.8 — Uso de navegadores web

En un grupo de usuarios, la probabilidad de que un usuario utilice el navegador Chrome es $P(C) = 0{,}62$ y la de que utilice Firefox es $P(F) = 0{,}25$. Supongamos que $P(C \cap F) = 0{,}07$ (usuarios que usan ambos).

a) Probabilidad de que use Chrome o Firefox:

$$P(C \cup F) = 0{,}62 + 0{,}25 - 0{,}07 = 0{,}80$$

b) Probabilidad de que use **ninguno** de los dos:

$$P(\overline{C \cup F}) = 1 - 0{,}80 = 0{,}20$$

c) Probabilidad de usar Chrome pero no Firefox:

$$P(C \cap \bar{F}) = P(C) - P(C \cap F) = 0{,}62 - 0{,}07 = 0{,}55$$

7.5. La fórmula de *Al menos uno*

Es frecuente necesitar la probabilidad de que ocurra *al menos uno* de varios eventos. Resulta más sencillo calcular el complemento:

Probabilidad de "al menos uno"

$$P(\text{al menos uno}) = 1 - P(\text{ninguno})$$

Ejemplo 7.9 — Al menos un éxito en pruebas unitarias

Un programador somete su código a tres pruebas unitarias independientes. Las probabilidades de que cada prueba detecte un error son $p_1 = 0{,}7$, $p_2 = 0{,}6$, $p_3 = 0{,}5$.
¿Cuál es la probabilidad de que *al menos una* prueba detecte un error?
$P(\text{ninguna detecta}) = (1 - 0{,}7)(1 - 0{,}6)(1 - 0{,}5) = 0{,}3 \times 0{,}4 \times 0{,}5 = 0{,}06$

$$P(\text{al menos una detecta}) = 1 - 0{,}06 = 0{,}94$$

8. Cuadro Resumen de Fórmulas

Resumen de fórmulas fundamentales

Concepto	**Fórmula**
Prob. clásica	$P(A) = n_A/n$
Complemento	$P(\bar{A}) = 1 - P(A)$
Prob. condicional	$P(A \mid B) = P(A \cap B)/P(B)$
Regla multiplicación (forma general)	$P(A \cap B) = P(A) \cdot P(B \mid A)$
Regla multiplicación (forma especial)	$P(A \cap B) = P(A) \cdot P(B)$
Regla de la suma	$P(A \cup B) = P(A) + P(B) - P(A \cap B)$
Excl. mutua	$P(A \cup B) = P(A) + P(B)$
Al menos uno	$P(\geq 1) = 1 - P(\text{ninguno})$

Bibliografía

1. **Webster, A. L.** (2000). *Estadística aplicada a los negocios y la economía* (3.ª ed.). McGraw-Hill Interamericana.

2. **Lind, D. A., Marchal, W. G., & Wathen, S. A.** (2019). *Estadística aplicada a los negocios y la economía* (16.ª ed.). McGraw-Hill Education.

3. **Wackerly, D. D., Mendenhall, W., & Scheaffer, R. L.** (2010). *Estadística matemática con aplicaciones* (7.ª ed.). Cengage Learning.

4. **DeGroot, M. H., & Schervish, M. J.** (2012). *Probability and Statistics* (4.ª ed.). Addison-Wesley.

5. **Ross, S. M.** (2014). *Introducción a la probabilidad y la estadística para ingeniería y ciencias* (9.ª ed.). Cengage Learning.

6. **Devore, J. L.** (2016). *Probabilidad y estadística para ingeniería y ciencias* (9.ª ed.). Cengage Learning.