Departamento de Matemáticas y Estadística

Guía resumida de Teoría de probabilidad

Dr. rer. nat. Humberto Llinás Solano

Universidad del Norte
Barranquilla - Colombia

Contenido

El autor

Humberto LLinás Solano.

Licenciado en Ciencias de la Educación, con énfasis en Matemáticas, Física y Estadística de la Universidad del Atlántico. Magister en Matemáticas, convenio Universidad del Valle-Universidad del Norte. Doctor en Estadística (Dr. rer. nat.) de la Universidad Johannes Gutenberg de Mainz (Alemania). Desde 1998 se desempeña como profesor de tiempo completo de la Universidad del Norte y pertenece a diferentes grupos de investigación.

Información relacionada con la producción intelectual se encuentra en: `https://rpubs.com/hllinas/toc`

Más detalles sobre su hoja de vida pueden visualizarse en:
`https://rpubs.com/hllinas/Bio_Sketch`

CAPÍTULO **1**

Distribuciones de probabilidad

1.1 Experimentos y espacios muestrales

La validez de la mayoría de las teorías científicas está basada, en gran parte, en que los experimentos, sobre los cuales las teorías se fundamentan, suministran esencialmente el mismo resultado cuando estos experimentos se repiten. Un ejemplo de la física es la ley de la caída libre, $s = \frac{1}{2}gt^2$. Las condiciones, bajo las cuales tal experimento se lleva a cabo, determinan ya su resultado. Por lo tanto, lo llamamos EXPERIMENTO DETERMINÍSTICO. Sin embargo, hay experimentos cuyos resultados no son determinados, si las condiciones de los intentos se mantienen constante. Ellos se llaman EXPERIMENTOS ALEATORIOS O ESTOCÁSTICOS.

Definición 1.1.1 Ω, *el conjunto de todos los posibles resultados de un experimento aleatorio, se llama* ESPACIO MUESTRAL. *Los elementos* $\omega \in \Omega$ *(también, las posibles salidas) se llaman* PUNTOS MUESTRALES.

Definición 1.1.2 *El espacio muestral* Ω *se llama*

(a) DISCRETO *si es finito o numerable. Un experimento aleatorio se llama* FINITO *(*DISCRETO*) si su espacio muestral es finito (discreto).*

(b) CONTINUO *si es un intervalo.*

1.2 σ-álgebras

Es importante resaltar que no todo subconjunto de un espacio muestral es un evento. Para que puedan ser catalogados así, dicho evento debe ser un elemento de un conjunto $\mathfrak{F}$ que tiene la estructura de σ-álgebra, concepto que se explicará a continuación.

Definición 1.2.1 *Un sistema $\mathfrak{F}$ de subconjuntos de un conjunto $\Omega \neq \emptyset$ se llama σ-*ÁLGEBRA *(en Ω) si posee las siguientes propiedades:*

(a) $\Omega \in \mathfrak{F}$

(b) Si $A \in \mathfrak{F}$, entonces $\overline{A} := \Omega \setminus A \in \mathfrak{F}$

(c) Si $A_1, A_2, \ldots \in \mathfrak{F}$, entonces $\bigcup\limits_{n=1}^{\infty} A_n \in \mathfrak{F}$

La dupla $(\Omega, \mathfrak{F})$ se llama ESPACIO MEDIBLE *y los conjuntos de $\mathfrak{F}$ se llaman* CONJUNTOS MEDIBLES. *Los elementos de $\mathfrak{F}$ se llaman* EVENTOS. *Todo evento con un sólo elemento se llama* EVENTO ELEMENTAL.

Observación

Obviamente, se cumple que $\emptyset = \overline{\Omega} \in \mathfrak{F}$. Con $A_1, A_2, \ldots \in \mathfrak{F}$, la intersección $\bigcap\limits_{n=1}^{\infty} A_n = \overline{\bigcup\limits_{n=1}^{\infty} \overline{A_n}}$ también está en $\mathfrak{F}$. Con $A, B \in \mathfrak{F}$, los conjuntos $A \cap B,\ A \cup B,\ A \setminus B = A \cap \overline{B}$ y la conocida diferencia simétrica $A \triangle B := (A \setminus B) \cup (A \setminus B)$ también están en $\mathfrak{F}$.

Ejemplo 1.2.2 *Sea $\Omega \neq \emptyset$ dado.*

*(a) El conjunto potencia $\mathfrak{P}(\Omega) := \{A / A \subseteq \Omega\}$ de Ω es la σ-álgebra más grande en Ω. También se llama la "σ-*ÁLGEBRA TOTAL*".*

(b) Supongamos que Ω es discreto y que una σ-álgebra $\mathfrak{F}$ en Ω contiene a todos los subconjuntos unitarios. Entonces $\mathfrak{F}$ debe ser el conjunto potencia de Ω, ya que para cada $A \subseteq \Omega$ se tiene (ya que con Ω también A es discreto)

$$A = \bigcup_{\omega \in A} \{\omega\} \in \mathfrak{F}$$

(c) *El conjunto* $\{\emptyset, \Omega\}$ *es la* σ*-álgebra más pequeña en* Ω*. También es conocida como "*σ*-ÁLGEBRA TRIVIAL" o "VACÍA".* ◄

1.2.1 σ-álgebra generada

Ahora, sea $I \neq \emptyset$ cualquier conjunto de índices y $\mathfrak{F}_i$ una σ-álgebra en Ω para cada $i \in I$. De la definición se sigue que el conjunto

$$\bigcap_{i \in I} \mathfrak{F}_i = \{A \,/\, A \in \mathfrak{F}_i \text{ para todo } i \in I\}$$

también es una σ-álgebra en Ω. Ahora, sea $\mathcal{E}$ cualquier sistema de subconjuntos de Ω y Σ el sistema de todas las σ-álgebras $\mathfrak{F}$ en Ω con $\mathcal{E} \subseteq \mathfrak{F}$. Entonces

$$\mathfrak{F}(\mathcal{E}) := \bigcap_{\mathfrak{F} \in \Sigma} \mathfrak{F} \tag{1.1}$$

es la σ-álgebra más pequeña que contiene a $\mathcal{E}$, es decir, $\mathcal{E} \subseteq \mathfrak{F}(\mathcal{E})$.

Definición 1.2.3 *Sea* $\mathcal{E}$ *un sistema de subconjuntos de un conjunto* $\Omega \neq \emptyset$*. Entonces la* σ*-álgebra* $\mathfrak{F}(\mathcal{E})$*, definida en* (1.1)*, se llama la* σ-ÁLGEBRA GENERADA *por* $\mathcal{E}$ *(en* Ω*). El sistema* $\mathcal{E}$ *se llama* GENERADOR *de una* σ*-álgebra* $\mathfrak{F}$ *si se tiene que* $\mathfrak{F} = \mathfrak{F}(\mathcal{E})$*.*

1.2.2 σ-álgebra de Borel

Definición 1.2.4 *La menor* σ*-álgebra sobre* $\mathbb{R}$ *que contine todos los intervalos de la forma* $(-\infty, a]$ *con* $a \in \mathbb{R}$ *se llama* σ*-álgebra de Borel y se denota por* $\mathfrak{B}$*. Los elementos de* $\mathfrak{B}$ *se llaman* CONJUNTOS DE BOREL*.*

Ya que $\mathfrak{B}$ es una σ-álgebra, entonces los siguientes conjuntos también pertenecen a $\mathfrak{B}$: (a, ∞), $(a, b]$, $(-\infty, a)$, $[a, \infty)$, (a, b), $[a, b]$, $\{a\}$, $\mathbb{N}$ (conjunto de los números naturales), $\mathbb{Q}$ (conjunto de los números racionales) y $\overline{\mathbb{Q}}$ (conjunto de los números irracionales), donde $a, b \in \mathbb{R}$. La verificación de lo anterior se deja como ejercicio.

Es importante recalcar que no todos los subconjuntos de $\mathbb{R}$ son conjuntos borelianos. A continuación, se presenta la siguiente definición de σ-álgebra de Borel en $\mathbb{R}^n$.

Definición 1.2.5 *Sean $a = (a_1, \dots, a_n)$ y $b = (b_1, \dots, b_n)$ elementos de $\mathbb{R}^n$, con $a \leq b$, es decir, $a_i \leq b_i$ para todo $i = 1, \dots, n$. La σ-álgebra $\mathfrak{B}_n$ generada por todos los intervalos de la forma*

$$(a, b] := \{x = (x_i, \dots, x_n) \ / \ a_i < x_i < b_i, \ i = 1, \dots, n\}$$

se llama σ-ÁLGEBRA DE BOREL (EN $\mathbb{R}^n$). *Los elementos de $\mathfrak{B}_n$ se llaman* CONJUNTOS DE BOREL *(en $\mathbb{R}^n$).*

1.3 Espacios de probabilidad

Definición 1.3.1 $(\Omega, \mathfrak{F}, P)$ *se llama un* ESPACIO DE PROBABILIDAD *si $\Omega \neq \emptyset$, $\mathfrak{F}$ es una σ álgebra en Ω (compárese con la definición 1.2.1) y P es una* MEDIDA DE PROBABILIDAD, *o simplemente una* PROBABILIDAD *sobre el espacio medible $(\Omega, \mathfrak{F})$, es decir, para la función $P : \mathfrak{F} \longrightarrow \mathbb{R}$ se cumplen los 3* AXIOMAS DE KOLMOGOROV:

(K1) $P(A) \geq 0$, *para todo* $A \in \mathfrak{F}$.

(K2) $P(\Omega) = 1$.

(K3) Para cada sucesión de eventos $A_1, A_2, \ldots \in \mathfrak{F}$, disyuntos dos a dos, se cumple la llamada σ-ADITIVIDAD, *es decir,* $P\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \sum_{n=1}^{\infty} P(A_n)$.

Cualquier evento A con probabilidad 0 se llama EVENTO NULO.

La serie anterior converge porque el primer axioma asegura que $P(A_n) \geq 0$ y el segundo, que $P(\bigcup_{n=1}^{\infty} A_n) \leq 1$. Es decir, $\sum_{n=1}^{\infty} P(A_n) \leq 1 < \infty$.

Ejemplo 1.3.2 *Sea $\Omega = \{1, 2, 3\}$ y $\mathfrak{F} = \{\emptyset, \Omega, \{1\}, \{2, 3\}\}$. Entonces la siguiente aplicación P definida sobre $\mathfrak{F}$ es una medida de probabilidad.* ◀

$$P(A) = \begin{cases} 0, & \text{si } 3 \notin A; \\ 1, & \text{si } 3 \in A \end{cases}$$

Al igual que en la geometría, se evitan definir conceptos como "punto" y "recta", se ha demostrado que, para una construcción de la teoría de la probabilidad, no es necesaria una definición de conceptos como "evento" y "probabilidad". Esto es con el fin de evitar dificultades lógicas y, en vista de una posible extensa y fácil aplicación de la teoría no es deseable esta definición. Sólo toca trabajar en la teoría de la probabilidad con las formales propiedades de estos conceptos.

Teorema 1.3.3 *Sea $\Omega \neq \emptyset$ y $\mathfrak{F}$ una σ-álgebra en Ω. Para A, B, $(A_n)_{n\in\mathbb{N}}$, $(B_n)_{n\in\mathbb{N}} \in \mathfrak{F}$ se tiene:*

(a) $P(\emptyset) = 0$.

(b) ADITIVIDAD*: Si los eventos A_i, para $i = 1, \ldots, n$, son disyuntos dos a dos, entonces,* $P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$.

(c) $P(\overline{A}) = 1 - P(A)$.

(d) MONOTONÍA*: Si $A \subseteq B$, entonces, $P(A) \leq P(B)$. En especial, $P(A) \leq 1$.*

(e) TEOREMA DE ADICIÓN PARA 2 EVENTOS *o* FÓRMULA DE SILVESTER*:*

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

(g) Si $A \subseteq B$, entonces, $P(B \setminus A) = P(B) - P(A)$.

(h) CONTINUIDAD DESDE ABAJO*: Si $A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots$, entonces, se cumple que* $P\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} P(A_n)$.

(i) CONTINUIDAD DESDE ARRIBA*: Si $B_1 \supseteq B_2 \supseteq B_3 \supseteq \cdots$, entonces, se cumple que* $P\left(\bigcap_{n=1}^{\infty} B_n\right) = \lim_{n\to\infty} P(B_n)$.

DEMOSTRACIÓN:
Los incisos (b), (c), (e), (g) e (i) se dejan como ejercicios para el lector.

(a) Teniendo en cuenta los axiomas de Kolmogorov (ver definición 1.3.1), tenemos

$$\begin{aligned} 1 &\overset{(K2)}{=} P(\Omega) = P(\Omega \cup \emptyset \cup \emptyset \cdots) \overset{(K3)}{=} P(\Omega) + P(\emptyset) + P(\emptyset) + \cdots \\ &\overset{(K1)}{\geq} P(\Omega) + P(\emptyset) \overset{(K2)}{=} 1 + P(\emptyset). \end{aligned}$$

Por tanto, $0 \geq P(\emptyset)$. Teniendo en cuenta el axioma (K1), obtenemos $P(\emptyset) = 0$.

(d) Debido a que $A \subseteq B$, se tiene que $B = A \cup (B \setminus A)$, con $A \cap (B \setminus A) = \emptyset$. Por tanto,

$$P(B) \ = \ P(A \cup (B \setminus A)) \ \stackrel{(b)}{=} \ P(A) + P(B \setminus A) \ \geq \ P(A).$$

(h) Sea $A_0 := \emptyset$ y defínase $C_i := A_i \setminus A_{i-1}$, para todo $i \in \mathbb{N}$. Si $i < j$, entonces, $\overline{A}_{j-1} \subseteq \overline{A}_{i-1}$. Por tanto,

$$C_i \cap C_j \ = \ (A_i \cap A_j) \cap (\overline{A}_{i-1} \cap \overline{A}_{j-1}) \ \stackrel{i<j}{=} \ A_i \cap \overline{A}_{j-1} \ = \ \emptyset.$$

Ahora, con inducción sobre n, es fácil demostrar que $A_n = \bigcup\limits_{i=1}^{n} C_i$, para todo $n \in \mathbb{N}$. Para $n = 1$, la afirmación es clara. Supongamos que la igualdad se cumple para $n = k$ y la demostraremos para $n = k + 1$. Por tanto,

$$\bigcup_{i=1}^{k+1} C_i \ = \ C_{k+1} \cup A_k \ = \ (A_{k+1} \cup A_k) \cap (\overline{A}_k \cup A_k) \ = \ A_{k+1}.$$

Por consiguiente, $\bigcup\limits_{n=1}^{\infty} A_n = \bigcup\limits_{n=1}^{\infty} \bigcup\limits_{i=1}^{n} C_i = \bigcup\limits_{i=1}^{\infty} C_i$. Por lo tanto,

$$\begin{aligned} P\left(\bigcup_{i=1}^{\infty} A_i\right) \ &= \ P\left(\bigcup_{i=1}^{\infty} C_i\right) \ \stackrel{(K3)}{=} \ \sum_{i=1}^{\infty} P(C_i) \\ &\stackrel{(g)}{=} \ \lim_{n\to\infty} \sum_{i=1}^{n} [P(A_i) - P(A_{i-1})] \\ &= \ \lim_{n\to\infty} [P(A_n) - P(\emptyset)] \ \stackrel{(a)}{=} \ \lim_{n\to\infty} P(A_n). \end{aligned}$$

Y con esto encontramos el resultado deseado. ■

Ejemplo 1.3.4 *Sean A, B y C eventos tales que $P(A) = 0,50$, $P(B) = 0,26$, $P(C) = 0,55$, $P(A \cap B) = 0,15$, $P(A \cap C) = 0,25$, $P(B \cap C) = 0,15$ y $P(A \cap B \cap C) = 0,05$. Calcule las siguientes probabilidades: (a) $P(A \cup B)$, (b) $P(A \cap \overline{C})$, (c) $P(\overline{A} \cup C)$ y (d) $P(A \cup B \cup C)$.* ◀

1.4 Espacios de probabilidad discretos y lapacianos

1.4.1 Espacios de probabilidad discretos

Sea $(\Omega, \mathfrak{F}, P)$ un espacio de probabilidad, donde Ω es discreto, es decir, $\Omega = \{\omega_1, \omega_2, \ldots\}$ y $\mathfrak{F}$ es el correspondiente conjunto potencia del conjunto Ω, es decir, $\mathfrak{F} = \{A \,/\, A \subseteq \Omega\}$. Entonces,

(a) Cada $A \in \mathfrak{F}$ es finito o enumerable, es decir, $A = \bigcup\limits_{\omega \in A} \{\omega\}$ es a lo más una unión enumerable de eventos disyuntos dos a dos.

(b) Debido a la σ-aditividad (ver axioma (K3) en la definición 1.3.1) o a la aditividad (ver el teorema 1.3.3) se tiene

$$P(A) = \sum_{\omega \in \Omega} P(\omega),$$

es decir, P es únicamente determinada a través de sus valores[1] $p_j := P(\omega_j)$, para cada $j \in \mathbb{N}$.

(c) El vector $|\Omega|$-dimensional $p = (p_1, p_2, \ldots)$ cumple las condiciones:

- $p_j \geq 0$, para todo $j \in \mathbb{N}$ y
- $\sum\limits_{j} p_j \;=\; \sum\limits_{\omega \in \Omega} P(\omega) \;=\; P(\Omega) \;=\; 1.$

Un vector p que satisface las condiciones anteriores se llama VECTOR DE PROBABILIDAD. En general, toda sucesión de probabilidades de un sistema de eventos se llama también DISTRIBUCIÓN DE PROBABILIDAD.

Ejemplo 1.4.1 *Sea $(\Omega, \mathfrak{F}, P)$ un espacio de probabilidad con:*

$$\begin{aligned}
&\Omega = \{1, 2, 3, 4\} \\
&\mathfrak{F} = \{\emptyset, \Omega, \{1\}, \{2, 3\}, \{4\}, \{1, 2, 3\}, \{2, 3, 4\}, \{1, 4\}\} \\
&P(\{1\}) = \frac{1}{4}, \quad P(\{2, 3\}) = \frac{1}{2}, \quad y \quad P(\{4\}) = \frac{1}{4}
\end{aligned}$$

Entonces $P(\{1, 2, 3\}) = \frac{3}{4}$, $P(\{2, 3, 4\}) = \frac{3}{4}$ y $P(\{1, 4\}) = \frac{1}{2}$. ◀

[1] Para cada $\omega \in \Omega$, defínase $P(\omega) := P(\{\omega\})$.

1.4.2 Espacios de probabilidad lapacianos

Definición 1.4.2 *Un espacio de probabilidad* $(\Omega, \mathfrak{F}, P)$ *se llama* ESPACIO DE PROBABILIDAD LAPLACIANO, *si* Ω *es finito,* $\mathfrak{F}$ *es el conjunto potencia de* Ω *y cada evento elemental* $\omega \in \Omega$ *tiene la misma probabilidad. A la probabilidad* P *se le llama* DISTRIBUCIÓN LAPLACIANA *o* DISTRIBUCIÓN CLÁSICA *en* $(\Omega, \mathfrak{F})$.

Observaciones:

(a) Ya que cada evento elemental tiene la misma probabilidad, se tiene $P(\omega) = \frac{1}{|\Omega|}$, para todo $\omega \in \Omega$. Por consiguiente, para cada evento $A \in \mathfrak{F}$, tenemos

$$P(A) \;=\; P\left(\bigcup_{\omega\in A}\{\omega\}\right) \;=\; \sum_{\omega\in A} P(\omega) \;=\; \sum_{\omega\in A}\frac{1}{\#\Omega} \;=\; \frac{\#A}{\#\Omega}.$$

Por tanto,

$$P(A) = \frac{\#A}{\#\Omega} = \frac{\text{“Números de casos favorables para A”}}{\text{“Número de casos posibles”}}. \tag{1.2}$$

(b) Como se observa en (1.2), en un espacio de probabilidad laplaciano se ve el cálculo de probabilidades sobre el conteo de conjuntos y conduce, por tanto, a problemas de combinatoria. Para más detalles al respecto, se puede consultar el texto de LLINÁS [3].

Ejemplo 1.4.3 *Un estante tiene 6 libros de matemáticas y 4 de física. Halle la probabilidad de que 3 libros determinados de matemáticas estén juntos, si todos*

(a) los libros de matemáticas son diferentes y los libros de física también;

(b) los libros de matemáticas son diferentes y 3 de los libros de física iguales. ◀

Ejemplo 1.4.4 *Una caja de doce lapiceros tiene dos defectuosos. Se extraen tres lapiceros sin reemplazo. ¿Cuál es la probabilidad de que dos salgan defectuosos?* ◀

1.5 Probabilidades condicionales

Está dada por

$$P(B/A) = \frac{P(A\cap B)}{P(A)}, \quad \text{si } P(A) > 0 \tag{1.3}$$

o, equivalentemente a ello, el llamado TEOREMA DE MULTIPLICACIÓN

$$P(B \cap A) = P(B/A)P(A). \tag{1.4}$$

El teorema de multiplicación vale, ilimitadamente, si hacemos uso de la convención $P(B/A)P(A) = 0$ para $P(A) = 0$, ya que, en este caso, siempre se cumple $P(B \cap A) = 0$.

Teorema 1.5.1 *Sea $(\Omega, \mathfrak{F}, P)$ un espacio de probabilidad.*

(a) Para cada $A \in \mathfrak{F}$ con $P(A) > 0$, $P(\cdot/A)$ es una medida de probabilidad sobre $(\Omega, \mathfrak{F})$ que está concentrada en A, es decir, $P(A/A) = 1$. En especial, $P(\cdot/A)$ es una medida de probabilidad sobre $(A, \mathfrak{F} \cap A)$, en donde $\mathfrak{F} \cap A$ es una σ-álgebra en A, definida por $\mathfrak{F} \cap A := \{B \cap A \,/\, B \in \mathfrak{F}\}$.

(b) Para $A_1, \ldots, A_n \in \mathfrak{F}$ se cumple

$$P(A_1 \cap \cdots \cap A_n) = P(A_n) \prod_{i=1}^{n-1} P\Big(A_i \Big/ \bigcap_{j=i+1}^{n} A_j\Big),$$

con las correspondientes convenciones si $P(A_1 \cap \cdots \cap A_n) = 0$.

DEMOSTRACIÓN:

(a) Para cualquier $B \in \mathfrak{F}$, tenemos

$$P(B/A) = \frac{P(B \cap A)}{P(A)} \geq \frac{P(\emptyset)}{P(A)} = 0 \quad \text{y}$$

$$P(\Omega/A) = \frac{P(\Omega \cap A)}{P(A)} = \frac{P(A)}{P(A)} = 1 = \frac{P(A \cap A)}{P(A)} = P(A/A).$$

Si $A_1, A_2, \ldots \in \mathfrak{F}$ es una sucesión de eventos disyuntos dos a dos, entonces, $A \cap A_1, A \cap A_2, \ldots$ también es una sucesión de eventos disyuntos dos a dos. Por consiguiente,

$$P\left(\bigcup_{i=1}^{\infty} A_i \Big/ A\right) = \frac{P\left(\bigcup_{i=1}^{\infty} (A_i \cap A)\right)}{P(A)} = \sum_{i=1}^{\infty} \frac{P(A_i \cap A)}{P(A)} = \sum_{i=1}^{\infty} P(A_i/A).$$

Por tanto, la aplicación $P(\cdot/A) : \mathfrak{F} \longrightarrow \mathbb{R}$ satisface los 3 axiomas de Kolmogorov. O sea, que $P(\cdot/A)$ es una medida de probabilidad sobre $(\Omega, \mathfrak{F})$, que está concentrada sobre A. Observando $P(\cdot/A)$ restringida sobre $\mathfrak{F} \cap A$, se obtiene que $P(\cdot/A)$ es una medida de probabilidad sobre $(A, \mathfrak{F} \cap A)$.

(b) Si $P(A_i \cap \cdots \cap A_n) = 0$, entonces, hacemos uso de la convección

$$P(A_n) \prod_{i=1}^{n-1} P\left(A_i \Big/ \bigcap_{j=i+1}^{n} A_j\right) = 0.$$

Sea ahora $P(A_1 \cap \cdots \cap A_n) > 0$. Entonces,

$$\prod_{i=1}^{n-1} P\left(A_i \Big/ \bigcap_{j=i+1}^{n} A_j\right) \;=\; \prod_{i=1}^{n-1} \frac{P\left(\bigcap\limits_{j=i}^{n} A_j\right)}{P\left(\bigcap\limits_{j=i+1}^{n} A_j\right)} \;=\; \frac{P(A_1 \cap \cdots \cap A_n)}{P(A_n)}.$$

Multiplicando por $P(A)$, obtenemos la conclusión. ■

Ejemplo 1.5.2 *Supongamos que una caja tiene diez bolas, de las cuales tres están defectuosas. Se sacan dos bolas, una detrás de la otra y sin reemplazo. ¿Cuál es la probabilidad de sacar una bola defectuosa seguida de otra defectuosa.* ◀

Ejemplo 1.5.3 *Una caja contiene 6 fichas rojas, 4 blancas y 5 azules. Halle la probabilidad de que se extraigan en el orden roja, blanca y azul si las fichas (a) se reemplazan, (b) no se reemplazan.* ◀

Teorema 1.5.4 (Teorema de la probabilidad total) *Sea dada una descomposición finita o enumerable $A_1, A_2, \ldots$ de Ω, es decir, $A_1, A_2, \ldots \in \mathfrak{F}$ son disyuntos dos a dos y $\bigcup_n A_n = \Omega$. Entonces, para cada $B \in \mathfrak{F}$, se tiene que*

$$P(B) = \sum_n P(B/A_n)\, P(A_n). \tag{1.5}$$

DEMOSTRACIÓN:
Debido a que los conjuntos $A_1, A_2, \ldots$ son disyuntos dos a dos, entonces, los conjuntos $B \cap A_1, B \cap A_2, \ldots$ también son disyuntos dos a dos. Además,

$$\bigcup_n (B \cap A_n) \;=\; B \cap \bigcup_n A_n \;=\; B \cap \Omega \;=\; B.$$

Por tanto,

$$P(B) \;=\; P\left(\bigcup_n (B \cap A_n)\right) \;=\; \sum_n P(B \cap A_n) \;=\; \sum_n P(B/A_n)\,P(A_n).$$

Y así queda demostrado el teromea. ■

Teorema 1.5.5 (Regla de Bayes) *Sea $A_1, A_2, \ldots$ una descomposición finita o enumerable de Ω. Entonces, para cada evento B con $P(B) > 0$ y para todo $n \in \mathbb{N}$, se tiene*

$$P(A_n/B) = \frac{P(B/A_n)\,P(A_n)}{\sum_j P(B/A_j)\,P(A_j)}. \tag{1.6}$$

DEMOSTRACIÓN:
Teniendo en cuenta el teorema de la probabilidad total (teorema 1.5.4), se tiene que

$$P(A_n/B) \;=\; \frac{P(B \cap A_n)}{P(B)} \;=\; \frac{P(B/A_n)\,P(A_n)}{P(B)} \;=\; \frac{P(B/A_n)\,P(A_n)}{\sum_j P(B/A_j)\,P(A_j)}.$$

Y así queda demostrado el teorema. ■

Ejemplo 1.5.6 *Un editor envía propaganda de un libro de estadística al 70% de aquellos profesores que están a cargo de esa materia. El 40% de aquellos que recibieron la propaganda se decidieron a utilizar el libro, inclusive, el 20% de los que no recibieron la propaganda también utilizarán el libro. ¿Cuál es la probabilidad de utilizar el libro?* ◀

1.6 Independencia

Sean $A, B \in \mathfrak{F}$ con $P(A), P(B) > 0$. Entonces, se prueba fácilmente que

$$P(A/B) > P(A) \iff P(B/A) > P(B), \tag{1.7}$$

$$P(A/B) < P(A) \iff P(B/A) < P(B), \tag{1.8}$$

$$P(A/B) = P(A) \iff P(B/A) = P(B). \tag{1.9}$$

Observación

La relación (1.9) significa que el suceso de uno de los dos eventos no tiene ninguna influencia sobre la probabilidad de que suceda el otro evento. En este caso, por el teorema de multiplicación, se tiene que

$$P(A \cap B) \;=\; P(A/B)\, P(B) \;=\; P(A)P(B). \tag{1.10}$$

Para la igualdad (1.10) no se necesitan más las condiciones $P(A), P(B) > 0$. Por consiguiente, definimos

Definición 1.6.1 *Dos eventos $A, B \in \mathfrak{F}$ se llaman (estocásticamente)* INDEPENDIENTES, *si*

$$P(A \cap B) = P(A)\, P(B). \tag{1.11}$$

Se dice que n eventos $A_1, \ldots, A_k \in \mathfrak{F}$ son (completamente) INDEPENDIENTES, *si para todo $1 \leq j_1 < j_2 < \cdots < j_k \leq n$, con $2 \leq k \leq n$, se tiene que*

$$P(A_{j_1} \cap A_{j_2} \cap \cdots \cap A_{j_k}) \;=\; P(A_{j_1}) \cap P(A_{j_2}) \cap \cdots \cap P(A_{j_k}). \tag{1.12}$$

Si $P(A_{j_1} \cap A_{j_2} \cap \cdots \cap A_{j_k}) > 0$, entonces, por (1.12), $P\Big(A_{j_i} \Big/ \bigcap_{\substack{m=1 \\ m \neq i}}^{k} A_{j_m}\Big) = P(A_{j_i})$, para todo $i \in \{1, \ldots, n\}$. La independencia completa implica que

(a) $P(A_1 \cap \cdots A_n) = P(A_1) \cdots P(A_n)$,

(b) $P(A_i \cap A_j) = P(A_i)\, P(A_j)$, *para todo $i, j \in \{1, \ldots, n\}$ con $i \neq j$. Esta es la llamada* INDEPENDENCIA DOS A DOS.

Ni (a) ni (b) solas, con $n > 2$, o (a) y (b) juntas, con $n > 3$, conducen a la independencia completa.

Teorema 1.6.2 *Los eventos $A_1, \ldots, A_n \in \mathfrak{F}$ son (completamente) independientes si y solo si*

$$P(B_1 \cap \cdots \cap B_n) = P(B_1) \cdots P(B_n),$$

para cada posible escogencia de $B_1 \ldots, B_n$, donde $B_j = A_j$ o $B_j = \overline{A_j}$, con $1 \leq j \leq n$.

DEMOSTRACIÓN:
Se deja al lector. ■

Observación:

(a) Si n eventos en un experimento aleatorio son independientes físicamente (es decir, de la disposición de los ensayos), entonces, el modelo teórético-probabilístico descrito es realístico, sólo, si estos eventos también son (estocásticamente) independientes en el modelo.

Por ejemplo, consideremos el doble lanzamiento de un dado. Con respecto al modelo, podemos hacer las siguientes consideraciones: El espacio de probabilidad laplaciano es $\Omega = \{1, 2, \ldots, 6\}^2$, es decir, $\#\Omega = 36$. Los eventos $A :=$"primer lanzamiento es un 2" y $B :=$"segundo lanzamiento es un 5", óbviamente, son físicamente independientes. También, son (estocásticamente) independientes, porque $\#A = \#B = 6$, es decir, $P(A) = P(B) = \frac{1}{6}$ y, por consiguiente, $P(A \cap B) = P(\{(2, 5)\}) = \frac{1}{36} = P(A)P(B)$.

(b) Por otro lado, una independencia física en la realidad no necesariamente necesita corresponder con una independencia estocástica en el modelo. Consideremos, nuévamente, el ejemplo del doble lanzamiento de un dado: Sean A y B como antes y $C :=$"la suma de ambos lanzamientos es 7". Por lo tanto,

$$\#C \;=\; \#\{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\} \;=\; 6,$$

es decir, $P(C) = \frac{6}{36} = \frac{1}{6}$. Por consiguiente,

$$P(B \cap C) \;=\; \frac{\#(B \cap C)}{\#\Omega} \;=\; \frac{\#\{(2, 5)\}}{36} \;=\; \frac{1}{36} \;=\; P(B)\, P(C).$$

O sea, que B y C son (estocásticamente) independientes. En forma semejante, se puede demostrar que A y C son (estocásticamente) independientes. Por lo tanto, A, B y C son independientes dos a dos, pero no completamente independientes, puesto que

$$P(A \cap B \cap C) \;=\; P(\{(2, 5)\}) \;=\; \frac{1}{36} \neq \frac{1}{6^3} \;=\; P(A)\, P(B)\, P(C).$$

Ejemplo 1.6.3 *En su sistema de funcionamiento, una represa tiene cuatro puertas de seguridad idénticas. La probabilidad de que una puerta en particular se abra cuando sea necesario es 0,97. Si las puertas funcionan independientemente, calcule la probabilidad de que (a) al menos una puerta se abra, (b) al menos una puerta no se abra.* ◀

1.7 Variables aleatorias

Definición 1.7.1 *(a) Sean* $(\Omega, \mathfrak{F}, P)$ *un espacio de probabilidad y* $(\Omega', \mathfrak{F}')$ *un espacio medible. Una función* $X : \Omega \to \Omega'$ *se llama* $\mathfrak{F}$-$\mathfrak{F}'$)-VARIABLE ALEATORIA *o, simplemente,* VARIABLE ALEATORIA[2] *si*

$$X^{-1}(A') := \{\omega \in \Omega \,/\, X(\omega) \in A'\} \in \mathfrak{F}, \quad \forall\, A' \in \mathfrak{F}'. \tag{1.13}$$

(b) En el caso $(\Omega', \mathfrak{F}') = (\mathbb{R}, \mathfrak{B})$*, se habla de variables aleatorias* REALES.

(d) Para variables aleatorias reales o numéricas X *y* Y *se define*

$$\{X < Y\} := \{\omega \in \Omega \,/\, X(\omega) < Y(\omega)\}.$$

Análogamente, se definen los conjuntos $\{X \leq Y\}$, $\{X > Y\}$, $\{X \geq Y\}$, $\{X = Y\}$ *y* $\{X \neq Y\}$.

Ejemplo 1.7.2 *Sea* $\Omega = \{1, 2, 3\}$, $\mathfrak{F} = \big\{\emptyset, \{1\}, \{2, 3\}, \Omega\big\}$, P *una medida de probabilidad arbitraria definida sobre* $\mathfrak{F}$*. Supóngase que* $(\Omega', \mathfrak{F}')$ *es un espacio medible con* $\Omega' = \{a, b\}$ *y* $\mathfrak{F}'$ *el conjunto potencia. Entonces, la aplicación* $X : \Omega \to \Omega'$ *definida por*

$$X(\omega) = \begin{cases} a, & \text{si } \omega = 1; \\ b, & \text{si } \omega = 2 \text{ o } \omega = 3; \end{cases}$$

es una $\mathfrak{F}$-$\mathfrak{F}'$*-variable aleatoria.* ◀

A continuación, el siguiente teorema presenta algunas caracterizaciones de variables aleatorias. No se revisará su demostración.

Teorema 1.7.3 *Sean* $X, Y, X_1, X_2, \ldots : \Omega \longrightarrow \mathbb{R}$ *variables aleatorias reales sobre una espacio de probabilidad* $(\Omega, \mathfrak{F}, P)$*. Entonces,*

(a) X *es una variable aleatoria si y sólo si una de las siguientes condiciones (equivalentes) se cumple:*

(i) $\{X \geq a\} \in \mathfrak{F}$*, para todo* $a \in \mathbb{R}$.

[2] Generalmente, las variables aleatorias se simbolizan con letras mayúsculas, en especial, con X, Y, Z.

(ii) $\{X > a\} \in \mathfrak{F}$, *para todo* $a \in \mathbb{R}$.

(iii) $\{X \leq a\} \in \mathfrak{F}$, *para todo* $a \in \mathbb{R}$.

(iv) $\{X < a\} \in \mathfrak{F}$, *para todo* $a \in \mathbb{R}$.

(b) $X \pm Y$ *(si está definida),* XY *(si está definida),*

$$\sup_{n\in\mathbb{N}} X_n, \quad \inf_{n\in\mathbb{N}} X_n, \quad \limsup_{n\to\infty} X_n \quad y \quad \liminf_{n\to\infty} X_n$$

son también variables aleatorias reales.

(c) $\max(X_1, \ldots, X_n)$ *y* $\min(X_1, \ldots, X_n)$ *son también variables aleatorias reales.*

(d) Si $g : \mathbb{R} \longrightarrow \mathbb{R}$ *es una función continua, entonces,* $g \circ X : \Omega \longrightarrow \mathbb{R}$ *es una variable aleatoria real.*

(e) Una función $X = (X_1, \ldots, X_n)$ *es una* $(\mathbb{R}^n, \mathfrak{B}_n)$*-variable aleatoria sobre un espacio de probabilidad* $(\Omega, \mathfrak{F}, P)$ *si y sólo si todas las funciones coordenadas* $X_1, \ldots, X_n$ *son* $(\mathbb{R}^n, \mathfrak{B}_n)$*-variables aleatorias sobre* $(\Omega, \mathfrak{F}, P)$*.*

DEMOSTRACIÓN:
Se deja al lector. ■

Teorema 1.7.4 *Sea* $(\Omega, \mathfrak{F}, P)$ *en espacio de probabilidad y* X *una variable aleatoria real. Entonces, para cada medida de probabilidad sobre* $\mathfrak{F}$*, se define una medida de probabilidad* P_X *sobre* $(\mathbb{R}, \mathfrak{B})$ *a través de*

$$P_X(B) = P\left(X^{-1}(B)\right) := P(X \in B), \quad \forall\, B \in \mathfrak{B} \tag{1.14}$$

DEMOSTRACIÓN:
Se debe verificar que P_X satisface las tres condiciones que definen una medida de probabilidad. Para la tercera condición sólo hay que tener en cuenta que con cada sucesión $(B_n)_{n\in\mathbb{N}}$ de conjuntos disyuntos dos a dos de $\mathfrak{B}$, también, $(X^{-1}(B_n))_{n\in\mathbb{N}}$ es una sucesión de conjuntos disyuntos dos a dos (de $\mathfrak{F}$) y que $X^{-1}\left(\bigcup_{n=1}^{\infty} B_n\right) = \bigcup_{n=1}^{\infty} X^{-1}(B_n)$. ■

Definición 1.7.5 *Consideremos la situación del teorema* 1.7.4*. Entonces, la medida de probabilidad* P_X *se llama la* DISTRIBUCIÓN DE LA VARIABLE ALEATORIA X*. Escribiremos* $X \stackrel{d}{=} D$*, si* D *es la distribución de* X*. La notación* $X \stackrel{d}{=} Y$ *significa que las variables aleatorias* X *y* Y *tienen la misma distribución.*

Ejemplo 1.7.6 *Sea* $\Omega = \{1, 2, 3\}$, $\mathfrak{F} = \big\{\emptyset, \{1\}, \{2, 3\}, \Omega\big\}$, P *dada por:*

$$P(\emptyset) = 0, \quad P(\{1\}) = \frac{1}{5}, \quad P(\{2,3\}) = \frac{4}{5}, \quad P(\Omega) = 1$$

Si $\Omega' = \{a, b\}$, $\mathfrak{F}'$ *es el conjunto potencia de* Ω' *y* $X : \Omega \to \Omega'$ *está definida por*

$$X(\omega) = \begin{cases} a, & \text{si } \omega = 1; \\ b, & \text{si } \omega = 2 \text{ o } \omega = 3; \end{cases}$$

Halle la distribución P_X *de* X. ◀

1.8 Funciones de distribución y densidad

Definición 1.8.1 *Una función* $F : \mathbb{R} \to \mathbb{R}$ *se llama* FUNCIÓN DE DISTRIBUCIÓN *(unidimensional), si tiene las siguientes tres propiedades:*

(a) F *es monotonamente creciente. Es decir,* $F(x) \leq F(y)$, *para todo* $x, y \in \mathbb{R}$ *con* $x \leq y$.

(b) F *es continua por la derecha. Es decir,* $F(x^+) := \lim\limits_{x \searrow y} F(y) = F(x)$, *para todo* $x \in \mathbb{R}$.

(c) $\lim\limits_{x \searrow -\infty} F(x) = 0$ *y* $\lim\limits_{x \nearrow +\infty} F(x) = 1$.

En el caso n-dimensional, tenemos la siguiente definición.

Definición 1.8.2 *Una función* $F : \mathbb{R}^n \to \mathbb{R}$ *se llama* FUNCIÓN DE DISTRIBUCIÓN *(n-dimensional), si tiene las siguientes tres propiedades:*

(a) $\triangle_x^y F \geq 0$, *para todo* $x, y \in \mathbb{R}^n$ *con* $x \leq y$, *donde* $\triangle_x^y F \in \mathbb{R}$ *está definida por*

$$\triangle_x^y F := \sum_{(\varepsilon_1,\ldots,\varepsilon_n)\in\{0,1\}^n} (-1)^{\varepsilon_1+\cdots+\varepsilon_n} F\big(\varepsilon_1 x_1 + (1-\varepsilon_1) y_1, \ldots, \varepsilon_n x_n + (1-\varepsilon_n) y_n\big)$$

(b) F *es continua por la derecha en cada componente.*

(c) $\lim\limits_{(x_1,\ldots,x_n)\nearrow(+\infty,\ldots,+\infty)} F((x_1,\ldots,x_n)) = 1$ *y, para cada subconjunto* $I \subseteq \{1,\ldots,n\}$ *y cada sucesión* $(x_i)_{i\in\{1,\ldots,n\}\setminus I}$ *fija, se tiene*

$$\lim_{(x_i)_{i\in I}\searrow(-\infty)_{i\in I}} F(x_1,\ldots,x_n) = 0$$

Sin demostración, se presenta el siguiente teorema de correspondencia.

Teorema 1.8.3 (Teorema de correspondencia) *Con*

$$F(x) = P(](-\infty,\ldots,-\infty),x]),$$

para $x \in \mathbb{R}^n$*, está definida una biyección entre el conjunto de las medidas de probabilidad* P *sobre* $(\mathbb{R}^n, \mathfrak{B}_n)$ *y el conjunto de las funciones de distribución* $F : \mathbb{R}^n \to \mathbb{R}$ *(n-dimensionales).*

DEMOSTRACIÓN:
Se deja al lector. ■

Definición 1.8.4 *Sea* X *una variable aleatoria real. La función* F_X *definida sobre* $\mathbb{R}$ *por medio de*

$$F_X(x) \;:=\; P_X((-\infty,x]) \;=\; P(X \le x)$$

se llama FUNCIÓN DE DISTRIBUCIÓN ACUMULADA *de* X*. Analogamente se puede presentar una definición para el caso en* $\mathbb{R}^n$*.*

Definición 1.8.5 *Sean* P *una medida de probabilidad sobre* $(\mathbb{R}^n, \mathfrak{B}_n)$ *y* F *la función de distribución correspondiente a* P *(ver teorema* 1.8.3*). Decimos que* P *posee una* DENSIDAD f*, si hay una aplicación* $f : \mathbb{R}^n \to [0,\infty)$*, con*

$$F(x) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f(t)\, d^n t, \quad \forall\, x = (x_1,\ldots,x_n) \in \mathbb{R}^n. \tag{1.15}$$

Si f es una densidad de una medida de probabilidad, entonces, trivialmente se tiene que

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(t)\, d^n t \;=\; \lim_{(x_1,\ldots,x_n)\nearrow(\infty,\ldots,\infty)} F((x_1,\ldots,x_n)) \;=\; 1.$$

Definición 1.8.6 *(a) Una probabilidad P sobre $(\mathbb{R}, \mathfrak{B})$ se llama* DISCRETA*, si existe un conjunto finito o enumerable $\Omega \subseteq \mathbb{R}$ con $P(\Omega) = 1$ y* CONTINUA*, si la correspondiente función de distribución $F_P : \mathbb{R} \longrightarrow \mathbb{R}$ es continua.*

(b) Una variable aleatoria real se llama DISCRETA *(* CONTINUA*) si la correspondiente distribución P_X es discreta (continua).*

Ejemplo 1.8.7 *Sea T la duración de vida de una clase de animales. De la experiencia, se sabe que, en casos ideales, tales duraciones de vida poseen una determinada distribución, la llamada* DISTRIBUCIÓN EXPONENCIAL *con parámetro $\alpha > 0$. En símbolos, $T \stackrel{d}{=} \exp(\alpha)$. Esta distribución está definida por la correspondiente función de distribución*

$$F_T(x) = P(T \leq x) = \begin{cases} 1 - e^{-\alpha x}, & \text{si } x \geq 0, \\ 0, & \text{de otro modo}, \end{cases} \tag{1.16}$$

para todo $x \in \mathbb{R}$. Además, T es una variable aleatoria continua que es, incluso, diferenciable en cada $x \neq 0$ y, por consiguiente, posee una densidad[3] *(compárese con la definición 1.8.5)*

$$f_T(x) := F_T'(x) = \begin{cases} \alpha e^{-\alpha x}, & \text{si } x > 0, \\ 0, & \text{si } x < 0. \end{cases} \tag{1.17}$$

Para $x \in \mathbb{R}$, se tiene que $P(T \leq x) = F_T(x) = \int\limits_{-\infty}^{x} f_T\, dt$. ◀

Ejemplo 1.8.8 *Consideremos el lanzamiento de dos monedas y X la variable aleatoria "número de caras que resultan". Halle la función de probabilidad de X y la función de distribución acumulada F de X.* ◀

Ejemplo 1.8.9 *Sea X una variable aleatoria discreta con valores 0, 1, 2 y 3 y con*

[3] Se puede definir arbitráriamente $f(0) \geq 0$, por ejemplo, f(0)=0.

función de distribución acumulada F definida por

$$F(t) = \begin{cases} 0, & si \;\; t < 0; \\ \\ \frac{1}{7}, & si \;\; 0 \leq t < 1; \\ \\ \frac{1}{3}, & si \;\; 1 \leq t < 2; \\ \\ \frac{3}{4}, & si \;\; 2 \leq t < 3; \\ \\ 1, & si \;\; t \geq 3 \end{cases}$$

Halle la función de probabilidad f de X. ◄

Ejemplo 1.8.10 *Sea dada la siguiente densidad de una variable aleatoria continua X :*

$$f(x) = \begin{cases} e^{-x}, & si\ x \geq 0, \\ 0, & de\ otro\ modo, \end{cases}$$

siendo e la base del logaritmo natural. Construya la función de distribución F correspondiente. ◄

Es importante anotar que si g es una función continua y X es una variable aleatoria, entonces, $g(X)$ también es una variable aleatoria (véase el teorema 1.7.3). Por esta razón, queremos expresar, para algunas funciones especiales f, la función de distribución de $Y := g(X)$ a través de la función de distribución de X.

Teorema 1.8.11 *Sea X una variable aleatoria continua con densidad de probabilidad f_X. Si $g : \mathbb{R} \longrightarrow \mathbb{R}$ es continuamente diferenciable sobre $\mathbb{R}$ y $g(x) \neq 0$ para todo $x \in \mathbb{R}$, y $h := g^{-1}$ es la inversa de g, entonces la función de densidad de la variable $Y = g(X)$ está dada por*

$$f_Y(x) = \begin{cases} f_X\big(h(x)\big)|h'(x)|, & para\ todo\ x\ en\ el\ rango\ de\ g; \\ 0, & de\ otro\ modo \end{cases}$$

DEMOSTRACIÓN:
Se deja al lector. ■

Ejemplo 1.8.12 *Sea X una variable aleatoria real con función de densidad dada por:*

$$f_X(x) = \begin{cases} 1, & \text{si } 0 < x < 1; \\ 0, & \text{en otro caso} \end{cases}$$

Si $Y = e^X$, hallar f_Y y F_Y. ◄

1.9 Esperanza, varianza, momentos y función generadora de momentos

1.9.1 Esperanza y varianza

Sean $(\Omega, \mathfrak{F}, P)$ un espacio de probabilidad y X una variable aleatoria numérica. Las propiedades teóricas-probabilísticas de X se dan de nuevo a través de su distribución P_X. Sin embargo, a menudo, es deseable reunir estas propiedades en un par de valores típicos. Por ejemplo, se quiere definir un valor medio (o valor central) para X y dar una cantidad, que mida la variación de los valores de X alrededor de este valor medio. Estos producen, por ejemplo, el llamado valor esperado y la varianza de X. El valor esperado corresponde al centro de gravedad y la varianza, al momento de inercia de un sistema físico.

Definición 1.9.1 *Sea $(\Omega, \mathfrak{F}, P)$ un espacio de probabilidad y X una variable aleatoria real.*

(a) Si X es una variable aleatoria discreta con valores $x_1, x_2, \ldots$ se dice que la esperanza de X existe si

$$\sum_{k=1}^{\infty} |x_k| \, P(X = k) < \infty$$

En este caso, se define la ESPERANZA *o* VALOR ESPERADO *de X como*

$$E(X) \; = \int_{-\infty}^{\infty} |x| \, f_X(x) \, dx < \infty$$

(b) Si X es una variable aleatoria continua con función de densidad f_X, se dice que la esperanza de X existe si

$$\int_{-\infty}^{\infty} |x|\, f_X(x)\, dx < \infty$$

En este caso, se llama la ESPERANZA *o* VALOR ESPERADO *de X como*

$$E(X) \;=\; \int_{-\infty}^{\infty} x\, f_X(x)\, dx$$

(c) Si $E(|X|) < \infty$, entonces,

$$Var(X) := E([X - E(X)]^2) \tag{1.18}$$

se llama la VARIANZA *de X. El número $\sigma(X) := \sqrt{Var(X)}$ se llama la* DESVIACIÓN ESTÁNDAR *de X.*

Teorema 1.9.2 (Propiedades de la esperanza) *Sean $(\Omega, \mathfrak{F}, P)$ un espacio de probabilidad y X, X_1, X_2 sean variables aleatorias reales. Entonces,*

(a) $E(1_A) = P(A)$, para todo $A \in \mathfrak{F}$, siendo 1_A la variable aleatoria discreta que toma valores 0 y 1 con probabilidades $P(\overline{A})$ y $P(A)$ respectivamente.

(b) $E(a) = a$, para todo $a \in \mathbb{R}$.

(c) Si $X_1 \leq X_2$, entonces, $E(X_1) \leq E(X_2)$.

(d) $E(aX_1 + bX_2) = aE(X_1) + bE(X_2)$, para todo $a, b \in \mathbb{R}$.

(e) $|E(X)| \leq E(|X|)$.

DEMOSTRACIÓN:
Como ejercicio. ■

Para la varianza tenemos el siguiente teorema.

Teorema 1.9.3 (Propiedades de la varianza) *Sea $(\Omega, \mathfrak{F}, P)$ un espacio de probabilidad y X una variable aleatoria real. Entonces,*

$$\begin{aligned} Var(1_A) &= P(A) \cdot (1 - P(A)), \quad \forall\, A \in \mathfrak{F} \\ Var(X) &= E(X^2) - (E(X))^2 \\ Var(aX + b) &= a^2 Var(X), \quad \forall\, a, b \in \mathbb{R}. \end{aligned}$$

DEMOSTRACIÓN:
Tenemos que

$$\begin{aligned} Var(X) &\overset{\text{D. 1.9.1}}{=} E\big(X^2 - 2 \cdot X \cdot E(X) + (E(X))^2\big) \\ &= E(X^2) - 2 \cdot E(X) \cdot E(X) + (E(X))^2 \;=\; E(X^2) - [E(X)]^2. \end{aligned}$$

En especial, para todo $A \in \mathfrak{F}$ y debido a que $1_A^2 = 1_A$, obtenemos

$$Var(1_A) \;=\; E(1_A^2) - [E(1_A)]^2 \;=\; P(A)[1 - P(A)].$$

Para cualesquiera $a, b \in \mathbb{R}$ tenemos

$$\begin{aligned} Var(aX + b) &\overset{\text{def.}}{=} E([aX + b - E(aX + b)]^2) \\ &= E([aX + b - aE(X) + b]^2) \\ &= a^2 E\big([X - E(X)]^2\big) \;=\; a^2 Var(X). \end{aligned}$$

Y de esta forma queda demostrado el teorema. ■

1.9.2 Momentos

Definición 1.9.4 *Sea $(\Omega, \mathfrak{F}, P)$ un espacio de probabilidad y X una variable aleatoria numérica con $E(X) < \infty$ y $k \in \mathbb{N}$. Entonces,*

(a) $E(X^k)$ se llama el k-ÉSIMO MOMENTO de X y a $E(|X^k|)$, el k-ÉSIMO MOMENTO ABSOLUTO de X.

(b) A la esperanza $E\big([X - E(X)]^k\big)$ se le conoce como el k-ÉSIMO MOMENTO CENTRAL de X y a $E\big(|X - E(X)|^k\big)$, el k-ÉSIMO MOMENTO CENTRAL ABSOLUTO[4] de X.

[4]Obsérvese que $E(X)$ es el primer momento de X y $Var(X)$ es el segundo momento centrado de X.

1.9.3 Función generadora de momentos

Definición 1.9.5 *Sea X una variable aleatoria tal que $E(e^{tX})$ es finito para todo $t \in (-\alpha, \alpha)$, con α real positivo. Se define la* FUNCIÓN GENERADORA DE MOMENTOS *de X, denotada por $m_X(\cdot)$ como:*

$$m_X(t) = E(e^{tX}), \qquad t \in (-\alpha, \alpha)$$

Teorema 1.9.6 *S $m_X(t)$ es diferenciable k veces, entonces $m_X^{(k)}(0) = E(X^k)$.*

DEMOSTRACIÓN:
Por inducción sobre k. ■

Ejemplo 1.9.7 *Supongamos que X tiene distribución binomial con parámetros n y p. Al tener en cuenta que*

$$(a+b)^n = \sum_{k=0}^{n} a^k b^{n-k}$$

tenemos que $m_X(t) = (pe^t + q)^n$ donde $q := 1 - p$. Además, $E(X) = m'_X(0) = np$ y $V(X) = m''_X(0) - [m'_X(0)]^2 = np(1-p)$. ◀

Ejemplo 1.9.8 *Si X tiene distribución Poisson con parámetro 3, entonces, teniendo en cuenta que*

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

tenemos que $m_X(t) = \exp\{3(e^t - 1)\}$. ◀

Una propiedad muy importante de la función generatriz de momentos de una variable aleatoria es que, cuando ella existe, caracteriza la distribución de la variable. Por esta razón, presentamos el siguiente teorema, sin demostración porque va más allá del alcance de estas notas.

Teorema 1.9.9 *Sean X y Y variables aleatorias cuyas funciones generadoras de momentos existen. Si $m_X(t) = m_Y(t)$ para todo t, entonces X y Y tienen la misma distribución.*

DEMOSTRACIÓN:
Se deja al lector. ■

1.10 Algunas distribuciones de probabilidad

1.10.1 Distribuciones especiales

1. *Distribuciones especiales discretas*:
   Uniforme discreta, de Bernoulli, binomial, de Poisson, hipergeométrica, binomial negativa, geométrica, de Polya, multinomial, etc.

2. *Distribuciones especiales continuas*:
   Uniforme continua, normal, gamma, exponencial, t de Student, Chi-cuadrada, F de Fisher, Cauchy, Beta, de Laplace, Log-normal, de Rayleigh, Weibull, de Maxwell, del valor extremo, etc.

En las tablas 1.1 y 1.2 (al final de este capítulo) presentamos un resumen de las distribuciones continuas y discretas, respectivamente, más importantes.

1.10.2 Relaciones entre algunas distribuciones

Teorema 1.10.1 (Normal) *Si X tiene distribución normal con parámetros μ y σ^2, es decir, $X \stackrel{d}{=} \mathcal{N}(\mu, \sigma^2)$, entonces, $aX + b \stackrel{d}{=} \mathcal{N}(a\mu + b, a^2\sigma^2)$, para todo $a, b \in \mathbb{R}$.*

DEMOSTRACIÓN:
Como ejercicio. Aplicar el ejercicio 1.36 (que se demuestra del teorema 1.8.11).
■

Teorema 1.10.2 (Normal estándar) *$X \stackrel{d}{=} \mathcal{N}(\mu, \sigma^2)$ si y sólo si $\frac{X-\mu}{\sigma} \stackrel{d}{=} \mathcal{N}(0, 1)$.*

DEMOSTRACIÓN:
Como ejercicio. Aplicar el ejercicio 1.36 (que se demuestra del teorema 1.8.11).
■

Teorema 1.10.3 (Gamma) *Sea X una variable aleatoria real. Sea $\gamma\big(\alpha, \beta\big)$ la distribución gamma con parámetros α y β.*

(a) Si $X \stackrel{d}{=} \mathcal{N}(0, \sigma^2)$, entonces, $X^2 \stackrel{d}{=} \gamma\big(\frac{1}{2}, \frac{1}{2\sigma^2}\big)$.

(b) Si $X \stackrel{d}{=} \gamma(\alpha, \beta)$*, entonces,* $cX \stackrel{d}{=} \gamma\big(\alpha, \frac{\beta}{c}\big)$*, para todo* $c > 0$*.*

DEMOSTRACIÓN:
Como ejercicio. Aplicar los ejercicios 1.37 y 1.36, respectivamente (que se demuestran del teorema 1.8.11). Recordar que $\Gamma(1/2) = \sqrt{\pi}$. ■

Teorema 1.10.4 (Chi-cuadrada) *Supongamos que* $\chi^2(n)$ *representa la distribución chi-cuadrada con* n *grados de libertad.*

(a) $\chi^2(n) = \gamma\left(\frac{n}{2}, \frac{1}{2}\right)$*.*

(b) Si $X \stackrel{d}{=} \mathcal{N}(0,1)$*, entonces,* $X^2 \stackrel{d}{=} \chi^2(1)$*.*

DEMOSTRACIÓN:
Como ejercicio. En la parte (b) aplicar el ejercicio 1.37 (que se demuestra del teorema 1.8.11). ■

✍ Ejercicios

1.1 Dos monedas se lanzan independientemente una detrás de otra y se cuentan cuantas veces ha aparecido “cara”. ¿Se pueden falsificar las monedas de tal forma que los eventos “dos veces sello”, “exactamente una vez cara” y “dos veces cara” aparezcan con la misma probabilidad? **Sugerencia:** Suponga que sí se pueden falsificar en la forma exigida. Defina P_C (P_S) como la probabilidad de que la primera moneda muestre cara (sello). Para la segunda moneda, defina análogamente q_C y q_S. Construya sistemas de ecuaciones, resuélvala.

1.2 Sea $(P_n)_{n\in\mathbb{N}}$ una sucesión de medidas de probabilidad que están definidas sobre un σ-álgebra $\mathfrak{F}$ sobre Ω. Sea $(\alpha_n)_n \in \mathbb{N}$ una sucesión de números reales no negativos con $\sum\limits_{n=1}^{\infty} \alpha_n = 1$ y sea $P : \mathfrak{F} \to \mathbb{R}$ con $P(A) := \sum\limits_{n=1}^{\infty} \alpha_n P_n(A)$, para todo $A \in \mathfrak{F}$. Demuestre que P es una medida de probabilidad. **Sugerencia:** Verifique los axiomas correspondientes y tenga en cuenta que, debido a la convergencia absoluta de las series (porque todos los sumandos son no negativos), las series involucradas se pueden intercambiar.

1.3 En una fiesta participan 3 parejas. ¿Cuál es la probabilidad de que ninguno de los señores bailen con su única acompañante, si la pareja del baile se puede determinar a través de la lotería? **Sugerencia:** Sea A_i el evento que representa "el señor i baila unicamente con su única acompañante". Calcule $P(A_1 \cup A_2 \cup A_3)$ aplicando el teorema de adición para 3 eventos (véase el teorema 1.3.3e).

1.4 Demuestre que no hay una distribución laplaciana sobre el conjunto de los números naturales $\mathbb{N}$. **Sugerencia:** Suponga que sí, construya un vector de probabilidad $(p_j)_{j\in\mathbb{N}}$. Luego aplique el tercer axioma de Kolmogorov.

1.5 De un conjunto

$$B \;=\; \{x \in \mathbb{N} \,/\, \forall t \in \mathbb{R} : \frac{1}{4}t^2 - t + 7 - x + \sqrt{68 - x^2} \;\geq\; 0\}$$

se sacan tres números sin reemplazo. Calcule la probabilidad de que éstos formen un número menor que 368.

1.6 Se lanza una moneda al aire hasta que dos veces consecutivas cae del mismo lado. Encuentre la probabilidad de que el experimento se termine en por lo menos n lanzamientos.

1.7 El profesor Probabili, conocido como "Hypergeométrico", acaba de terminar su clase de probabilidad "Cadenas de Markov y sus aplicaciones especiales" y se coloca delante del salón. Una vez que sus estudiantes van dejando el salón en una fila aleatoria, le llama la atención a él que los primeros 5 son varones. "Interesante", dice él a su amigo Estadistis, la probabilidad fue exactamente $\frac{1}{2}$. "Este pregunta": ¿Cuántos varones y hembras han escuchado su clase? Dé una (??) posible cantidad. **Sugerencia:** Sea N el número de estudiantes y m el número de varones. Demuestre que

$$\frac{\binom{m}{5}\binom{N-m}{0}}{\binom{N}{5}} = \frac{1}{2}$$

Escoja, por ejemplo, $N = m + 1$ y resuelva. INQUIETUD: ¿Qué pasa si se escoge $N = m$ o $N = m + k$ para $k \geq 2$?

1.8 Se examina un grupo (grande) de personas con respecto a cierta enferemedad por medio de dos síntomas A y B, respectivamente. Se encuentran, respectivamente: 25% con sólo el síntoma A, 30% con sólo el síntoma B, 15% con ambos y el resto con ninguno de los dos. Para una persona, escogida al azar de este grupo, encuentre las siguientes probabilidades:

(a) Que la persona no presente ningún síntoma.

(b) Que la persona presente al menos uno de los síntomas.

(c) Que la persona presente el síntoma B.

(d) Que la persona presente ambos sítomas, sabiendo que tiene B.

1.9 Un lote de 25 piezas moldeadas por inyección contiene 5 que presentan una contracción excesiva.

(a) Si se seleccionan dos piezas al azar una detrás de otra, y sin reemplazo, ¿cuál es la probabilidad de que la segunda pieza seleccionada sea una con contracción excesiva?

(b) Si se seleccionan tres piezas al azar una detrás de otra, y sin reemplazo, ¿cuál es la probabilidad de que la tercera pieza seleccionada sea una con contracción excesiva?

1.10 Para la distribución de k partículas sobre n celdas diferentes se introducen en la física modelos diferentes que describen la distinta naturaleza de la partícula. Por ejemplo: en el modelo de Maxwell-Boltzmann (MB), las partículas son diferentes. En este modelo se observan los correspondientes posibles estados con igual probabilidad. Calcule la probabilidad de que:

(a) Las celdas $1, \ldots, n$ contienen exactamente $k_1, \ldots, k_n$ partículas, donde $k_1 + \cdots + k_n = k$.

(b) En cualesquiera k celdas fijamente escojidas, se encuentre, en cada una, 1 sola partícula.

(c) En cualesquiera k celdas, se encuentre en cada una, 1 sola partícula.

1.11 Por un canal de comunicaciones afectado por ruido se transmite uno de dos comandos de control en forma de palabras de código 11111 y 00000. Esto se transmite con probabilidad a priori de 0,7 y 0,3 respectivamente. Por causa del ruido, la probabilidad de recepción correcta de cada uno de los símbolos disminuye a 0,6. Se supone que las palabras de código se dañan o distorsionan independientemente. En la salida del receptor se registra la palabra de código 10110. Determine qué comando fue transmitido.

1.12 Supóngase que existen dos métodos excluyentes A y B, respectivamente, para enseñar a los trabajadores cierta habilidad industrial. El porcentaje de un fracaso F es 25% para A y 12,5% para B, respectivamente. Se utiliza el entrenamiento B en un 40% (y A en un 60%).

(a) Reescriba los supuestos como probabilidades ordinarias y condicionales y, luego, resuelva lo que se formula en los incisos siguientes.

(b) Se entrenó a un trabajador según uno de los dos métodos, pero él fracasó. ¿Cuál es la probabilidad de que él haya recibido el entrenamiento A?

(c) Se entrenó a un trabajador con éxito. ¿Cuál es la probabilidad de que él haya recibido el entrenamiento A?

(d) ¿Son independientes los entrenamientos A y B?

(e) ¿Son independientes el entrenamiento A y un fracaso F?

1.13 Se sabe que en un grupo de cuatro componentes hay dos defectuosos. Una inspectora los prueba de uno en uno hasta encontrar las dos piezas defectuosas. Una vez que las localiza interrumpe las pruebas, pero examina la segunda pieza defectuosa por seguridad. Si X es el número de pruebas en la que se detecta la segunda pieza defectuosa, determine la función de probabilidad de X.

1.14 Para verificar la exactitud de sus estados financieros, las empresas a menudo emplean auditores que verifiquen sus ingresos. Los empleados de la empresa se equivocan al registrar los ingresos 5% de las veces. Suponga que un auditor revisa aleatoriamente tres ingresos y que la detección de los errores es independiente. Determine la función de probabilidad del número de errores detectado por el auditor.

1.15 Se lanzan dos dados perfectos y se define la variable aleatoria X de la siguiente manera:

$$X = \begin{cases} 2.000, & \text{si resultan “6” en ambos;} \\ 100, & \text{si resulta “6” sólo en uno;} \\ 0, & \text{si no resulta “6” en ninguno.} \end{cases}$$

(a) Calcule las tres probabilidades que determinan la distribución de X.

(b) Calcule la esperanza de X.

(c) Supóngase que este experimento representa un juego, en el cual los valores de X significan la ganancia (en pesos) del jugador y donde el jugador tiene que pagar 100 (pesos) antes. Interprete el valor de $E(X)$.

1.16 La contaminación del río Magdalena es un problema que se va incrementado cada vez más con el pasar de los años. Sean dadas las siguientes probabilidades:

- La probabilidad de que el río está contaminado es 0,3.
- La probabilidad de que una prueba en una muestra detecta contaminación sabiendo que el río está contaminado es 0,75.

- La probabilidad de que una prueba en una muestra detecta contaminación sabiendo que el río no está contaminado es 0,20.
- La probabilidad de que se permita pesca sabiendo que el río está contaminado y que una prueba en una muestra detecta contaminación es 0,20.
- La probabilidad de que se permita pesca sabiendo que el río no está contaminado y que una prueba en una muestra detecta contaminación es 0,15.
- La probabilidad de que se permita pesca sabiendo que el río está contaminado y que una prueba en una muestra no detecta contaminación es 0,80.
- La probabilidad de que se permita pesca sabiendo que el río no está contaminado y que una prueba en una muestra no detecta contaminación es 0,90.

Encuentre la probabilidad de cada uno de los siguientes eventos:

(a) El río está contaminado, una prueba en una muestra detecta contaminación y se permite pesca.

(b) Una prueba en una muestra no detecta contaminación y se permite pesca.

(c) Se permite pesca.

1.17 Responda las siguientes preguntas. Explique.

(a) Si A, B y C son mutuamente excluyentes, ¿es posible que $P(A) = 0,3$, $P(B) = 0,4$ y $P(C) = 0,5$?

(b) Si $P(A/B) = 1$, ¿se cumple $A = B$?

(c) Si A y B son eventos mutuamente excluyentes, ¿es posible construir un diagrama de Venn que contenga a los tres eventos A, B y C, tales que $P(A/C) = 1$ y $P(B/C) = 0$?

1.18 Sea X una variable aleatoria que tiene distribución exponencial con parámetro $\lambda > 0$.

(a) Demuestre que su función de densidad f es realmente una densidad.

(b) Encuentre una fórmula (y demuéstrela) para los valores $F(t)$ de la función de distribución acumulada F y para $P(X \geq t)$.

(c) Con base en el inciso anterior, halle $F(-3)$ y $P(X \geq -3)$.

(d) Demuestre que $P(X \geq x + z \,/\, X \geq x) = P(X \geq z)$, para todo $x, z > 0$.

1.19 Demuestre las siguientes afirmaciones:

(a) Para cualquier evento A y B con $P(B) > 0$ se cumple que $P(A/B) + P(\overline{A}/B) = 1$.

(b) Si $P(B/A) > P(B)$, entonces $P(\overline{B}/A) < P(\overline{B})$. **Sugerencia**: Sume $P(\overline{B}/A)$ ambos lados de la desigualdad y use el resultado de la parte (a).

(c) Para cualquiera de los tres eventos A, B y C con $P(C) > 0$ se cumple que

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

(d) Si A y B son independientes, entonces $\overline{A}$ y B también lo son.

(e) Si A y B son independientes, entonces también lo son sus complementos.

1.20 Demostrar las fórmulas de la esperanza para algunas distribuciones (a) discretas, (b) continuas.

1.21 Demostrar las fórmulas de la varianza para algunas distribuciones (a) discretas, (b) continuas.

1.22 Demostrar los teoremas (a) 1.10.1, (b) 1.10.2.

1.23 Demostrar el teorema 1.10.3.

1.24 Demostrar el teorema 1.10.4.

1.25 Sea X una variable aleatoria con valores enteros. Suponga que $P(X = 0) = 0$ y $P(X = j) = \frac{c}{j^2}$ si $j \neq 0$, donde $c > 0$ es una constante tal que $\sum\limits_j \frac{c}{j^2} = 1$. Halle $E(X)$.

1.26 Halle la función generadora de momentos de una variable X y aplique el teorema 1.9.6 para hallar $E(X)$ y $V(X)$, si X tiene algunas de las siguientes distribuciones:

(a) Discreta (Poisson, Hipergeométrica, Binomial negativa, Geométrica)

(b) Continua (Normal, Exponencial, Gamma, Chi-cuadrada)

1.27 Sea $\Omega = \{1, 2, 3\}$. Verifique si los siguientes conjuntos forman una σ-álgebra de Ω:

(a) $\mathfrak{F} = \{\emptyset, \{1\}, \{2, 3\}, \Omega\}$

(b) $\mathfrak{R} = \{\emptyset, \{1\}, \{2\}, \{3\}, \Omega\}$

1.28 Si $\mathfrak{B}$ es la σ-álgebra de Borel, demuestre que los siguientes conjuntos también pertenecen a $\mathfrak{B}$: (a, ∞), $(a, b]$, $(-\infty, a)$, $[a, \infty)$, (a, b), $[a, b]$, $\{a\}$, $\mathbb{N}$ (conjunto de los números naturales), $\mathbb{Q}$ (conjunto de los números racionales) y $\overline{\mathbb{Q}}$ (conjunto de los números irracionales).

1.29 Supongamos que se lanzan tres dados. Considere los eventos A=“el resultado del primer lanzamiento es un número primo” y B=“la suma de los resultados obtenidos es menor o igual que 4”. Halle $A \cup B$, $A \cap B$, $A - B$ y $\overline{A}$.

1.30 Sea $\Omega = \{1, 2\}$ y $\mathfrak{F}$ el conjunto potencia de Ω. Demuestre que la siguiente aplicación P definida sobre $\mathfrak{F}$ es una medida de probabilidad:

$$P(A) = \begin{cases} 0, & \text{si } A = \emptyset; \\ 1/3, & A = \{1\}; \\ 2/3, & \text{A=\{2\};} \\ 1, & \text{A=\{1,2\}.} \end{cases}$$

1.31 Sea $\Omega = \{1, 2, 3\}$ y $\mathfrak{F} = \{\emptyset, \Omega, \{2\}, \{1, 3\}\}$. Demuestre que los eventos $\emptyset$ y $\{3\}$ son eventos nulos si P está definida sobre $\mathfrak{F}$ como:

$$P(A) = \begin{cases} 0, & \text{si } 2 \notin A; \\ 1, & \text{si } 2 \in A \end{cases}$$

Verifique que P es una medida de probabilidad.

1.32 Sea $(\Omega, \mathfrak{F}, P)$ un espacio de probabilidad discreto con $\Omega = \{x, y, z\}$, $\mathfrak{F}$ el conjunto potencia y P determinada por el vector de probabilidades $p = (1/7,\ 4/7,\ 2/7)$. Halle $P(\{x, y\})$, $P(\{y, z\})$ y $P(\{x, z\})$.

1.33 Sea $\Omega = \{1, 2, 3\}$, $\mathfrak{F} = \{\emptyset, \{1\}, \{2, 3\}, \Omega\}$, P una medida de probabilidad arbitraria definida sobre $\mathfrak{F}$. Supóngase que $(\Omega', \mathfrak{F}')$ es un espacio medible con $\Omega' = \{a, b\}$ y $\mathfrak{F}'$ el conjunto potencia. Verifique si la aplicación $Y : \Omega \to \Omega'$ definida como sigue es una $\mathfrak{F}$-$\mathfrak{F}'$-variable aleatoria:

$$Y(\omega) = \begin{cases} a, & \text{si } \omega = 2; \\ b, & \text{si } \omega = 1 \text{ o } \omega = 3; \end{cases}$$

1.34 Supóngase que los cumpleaños de las personas pueden ocurrir con igual probabilidad en cualquiera de los 365 días del año (se ignoran años bisiestos y el hecho de que las tasas de natalidad no son uniformes durante el año). Halle la probabilidad de que no haya dos personas, en un grupo de n personas, con el mismo día de cumpleaños.

1.35 En cierta bodega, una caja contiene ocho clavos de 1 pulgada, seis de 1 pulgada y media y cinco de 2 pulgadas. Suponga que se seleccionan cuatro clavos al azar, sin reemplazo y sin orden.

(a) ¿Cuál es la probabilidad de que exactamente tres de los clavos seleccionados sean de 2 pulgadas?

(b) ¿Cuál es la probabilidad de que los cuatro clavos seleccionados sean del mismo tamaño?

(c) ¿Cuál es la probabilidad de que entre los cuatro clavos seleccionados hallan dos de una pulgada?

1.36 Si X es una variable aleatoria continua con función de distribución F_X y densidad de probabilidad f_X, entonces, las siguientes funciones $g : \mathbb{R} \longrightarrow \mathbb{R}$ definen las variables aleatorias $Y = g(X)$ con la función de distribución F_Y y densidad de probabilidad f_Y. Si $Y = aX + b$, para todo $a, b \in \mathbb{R}$, $a \neq 0$, demostrar que $f_Y(x) = \frac{1}{|a|} f_X\big(\frac{x-b}{a}\big)$ y

$$F_Y(t) = \begin{cases} F_X\big(\frac{t-b}{a}\big), & \text{si } a > 0, \\ 1 - F_X\big(\frac{t-b}{a}\big), & \text{si } a < 0. \end{cases}$$

Si $a < 0$, en el caso discreto sería $F_Y(t) = 1 - F_X\big(\frac{t-b}{a}\big) + P\big(X = \frac{t-b}{a}\big)$.

1.37 Si X es una variable aleatoria continua con función de distribución F_X y densidad de probabilidad f_X, entonces, las siguientes funciones $g : \mathbb{R} \longrightarrow \mathbb{R}$ definen las variables aleatorias $Y = g(X)$ con la función de distribución F_Y y densidad de probabilidad f_Y. Si $Y = X^2$, demostrar que

$$f_Y(x) = \begin{cases} 0, & x \leq 0, \\ \frac{1}{2\sqrt{x}}[f_X(\sqrt{x}) + f_X(-\sqrt{x})], & x > 0, \end{cases}$$

$$F_Y(t) = \begin{cases} 0, & t \leq 0, \\ F_X(\sqrt{t}) - F_X(-\sqrt{t}), & t > 0. \end{cases}$$

En el caso discreto, $F_Y(t) = F_X(\sqrt{t}) + F_X(-\sqrt{t}) + P(X = -\sqrt{t})$, si $t > 0$.

1.38 Si X es una variable aleatoria continua con función de distribución F_X y densidad de probabilidad f_X, entonces, las siguientes funciones $g : \mathbb{R} \longrightarrow \mathbb{R}$ definen las variables aleatorias $Y = g(X)$ con la función de distribución F_Y y densidad de probabilidad f_Y. Sea $Y = a\,X^2$, para una constante $a \neq 0$.

(a) Encuentre una fórmula para F_Y en términos de F_X. **Sugerencia**: Haga los cálculos separadamente para los dos casos $a > 0$ y $a < 0$ y, en cada caso, determine para cuáles valores $t \in \mathbb{R}$, se tiene que $F_Y(t) = 0$.

(b) Usando los dos casos de la parte (a), encuentre una fórmula para f_Y en términos de f_X.

1.39 Si X es una variable aleatoria continua con función de distribución F_X. Si $Y = |X|$, demostrar que

$$F_Y(t) = \begin{cases} 0 & \text{si } t \leq 0, \\ F_X(t) - F_X(-t) & \text{si } t > 0. \end{cases}$$

1.40 Sea X una variable aleatoria real con función de densidad dada por:

$$f_X(x) = \begin{cases} e^{-x}, & \text{si } x > 0; \\ 0, & \text{en otro caso} \end{cases}$$

Hallar la densidad de $Y = X^2$.

1.41 Sea X una variable aleatoria real con función de densidad dada por:

$$f_X(x) = \begin{cases} \frac{1}{2}e^{-2x}, & \text{si } x > 0; \\ 0, & \text{en otro caso} \end{cases}$$

Hallar la densidad de $Y = X^3$.

1.42 Sea X una variable aleatoria real con función de densidad dada por

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad x \in \mathbb{R}$$

Hallar la densidad de $Y = X^2$.

1.43 Sea X una variable aleatoria real con función de densidad dada por

$$f(x) = \frac{1}{\pi(x^2 + 1)}, \quad x \in \mathbb{R}$$

Hallar $E(X)$ y la densidad de $Y = \tan^{-1} x$.

NOMBRE	FUNCION	PARA-METROS	$E(X)$	$V(X)$
Uniforme	$f(x) = \frac{1}{b-a}$, $a < x < b$	$a < b$	$\frac{a+b}{2}$	$\frac{(a-b)^2}{12}$
Normal	$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}, e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ $x \in \mathbb{R}$	$\mu \in \mathbb{R}$, $\sigma^2 > 0$	μ	σ^2
Normal estándar	$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, $x \in \mathbb{R}$		0	1
Gamma	$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$, $x > 0$	$\alpha > 0$, $\beta > 0$	α/β	α/β^2
Exponencial	$f(x) = \lambda e^{-\lambda x}$, $x > 0$	$\lambda > 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
t de Student	$f(x) = a_n \left(1 + \frac{x^2}{n}\right)^{-(n+1)/2}$, $a_n := \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{n\pi}}$, $x \in \mathbb{R}$	$n \in \mathbb{N}$	0, $n \geq 2$	$\frac{n}{n-2}$, $n \geq 3$
Chi-cuadrada	$\frac{1}{a_n} x^{\frac{n}{2}-1} e^{-x/2}$, $a_n := 2^{n/2}\Gamma\left(\frac{n}{2}\right)$, $x > 0$	$n > 0$	n	$2n$
F de Fisher	$f(x) = \frac{a_n x^{\frac{m}{2}-1}}{(n+mx)^{(m+n)/2}}$ $a_n := \frac{\Gamma\left(\frac{m+n}{2}\right) m^{m/2} n^{n/2}}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}$, $x > 0$	$m, n \in \mathbb{N}$	$\frac{n}{n-2}$, $n \geq 3$	$\frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$, $n \geq 5$

Tabla 1.1: Resumen de distribuciones continuas

NOMBRE	FUNCION	PARAMETROS	$E(X)$	$V(X)$
Uniforme	$f(x_k) = \frac{1}{n}$, $k = 1, 2, \ldots, n$	$x_1 < x_2 < \cdots < x_n$ $n \in \mathbb{N}$	$\frac{1}{n}\sum\limits_{k=1}^{n} x_k$	$\frac{1}{n}\sum\limits_{k=1}^{n} x_k^2 -$ $-\frac{1}{n}\left(\sum\limits_{k=1}^{n} x_k\right)^2$
De dos puntos	$f(x_1) = p$, $f(x_2) = 1 - p$	$x_1 < x_2$ $0 < p < 1$	$x_1 p + x_2(1-p)$	$(x_1 - x_2)^2 p(1-p)$
Bernoulli	$f(0) = p$, $f(1) = 1 - p$	p	p	$p(1-p)$
Binomial	$f(k) = \binom{n}{k} p^k (1-p)^{n-k}$ $k = 0, 1, 2, \ldots, n$	$0 < p < 1$ $n \in \mathbb{N}$	np	$np(1-p)$
Poisson	$f(k) = \frac{1}{k!} e^{-\lambda} \lambda^k$ $k = 0, 1, 2, 3, \ldots$	$\lambda > 0$	λ	λ
Hiper-geométrica	$\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$ $k \in \mathbb{N}_0,\ k \leq n,\ k \leq M$	$M \in \mathbb{N}_0,\ N \in \mathbb{N},\ n \in N$ $n \leq M \leq N$	$n \cdot \frac{M}{N}$	$\left(\frac{N-n}{N-1}\right) \cdot n \cdot \frac{M}{N} \cdot \left(1 - \frac{M}{N}\right)$
Binomial negativa	$f(k) = \binom{k+r-1}{r-1} p^r (1-p)^k$ $k = 0, 1, 2, \ldots$	$r > 0$, $0 < p < 1$	$\frac{r(1-p)}{p}$	$\frac{r(1-p)}{p^2}$
Geométrica	$f(k) = p(1-p)^k$ $k = 0, 1, 2, \ldots$	$0 < p < 1$	$\frac{1-p}{p}$	$\frac{1-p}{p^2}$

Tabla 1.2: Resumen de distribuciones dicretas

CAPÍTULO **2**

Distribuciones conjuntas y convoluciones

2.1 Vectores aleatorios y distribuciones conjuntas y marginales

Vectores aleatorios son vectores cuyas componentes son variables aleatorias.

Definición 2.1.1 *Sean $X_1, X_2, \ldots, X_n$, n variables aleatorias reales definidas sobre el mismo espacio de probabilidad $(\Omega, \mathfrak{F}, P)$. La $\mathfrak{F}$-$\mathfrak{B}_n$-variable aleatoria $X : \Omega \longrightarrow \mathbb{R}^n$ definida por*

$$X(\omega) \;:=\; (X_1(\omega), \ldots, X_n(\omega))$$

se llama APLICACIÓN PRODUCTO.[1]

A la distribución $P_X = P_{(X_1,\ldots,X_n)}$ de X se le llama la DISTRIBUCIÓN CONJUNTA *de las variables aleatorias $X_1, \ldots, X_n$ y a cada distribución P_{X_i} con $1 \leq i \leq n$, se le llama la* DISTRIBUCIÓN MARGINAL *de X_i.*

[1] A menudo también se simboliza la aplicación producto X con $X_1 \otimes \cdots \otimes X_n$ o con $(X_1, \ldots, X_n)$ y, por esta razón, también es llamado VECTOR ALEATORIO, nombre que también utilizaremos usualmente.

Definición 2.1.2 *Sea* $X := (X_1, \dots, X_n)$ *un vector de* $(\mathbb{R}, \mathfrak{B})$*-variables aleatorias* $X_1, \dots, X_n$ *sobre un espacio de probabilidad* $(\Omega, \mathfrak{F}, P_X)$ *con* $P_X = P_{(X_1,\dots,X_n)}$. *Para todo* $x = (x_1, \dots x_n) \in \mathbb{R}^n$, *la* FUNCIÓN DE DISTRIBUCIÓN CONJUNTA *de* X *se define por*

$$F_X(x) := P_X\big(](-\infty, \dots, -\infty), x]\big) = P_X(X_1 \leq x_1, \dots, X_n \leq x_n)$$

y, para todo $i = 1, \dots, n$, *las* FUNCIONES DISTRIBUCIONES MARGINALES *de* X *están definidas por*

$$\begin{aligned} F_{X_i}(x_i) &:= P_{X_i}(X_i \leq x_i) \\ &= P_X(X_1 < \infty, \dots, X_{i-1} < \infty, X_i \leq x_i, X_{i+1} < \infty, \dots, X_n < \infty). \end{aligned}$$

2.2 Vectores aleatorios discretos y continuos

Como veremos a continuación, un vector aleatorio se puede clasificar en discreto o continuo si sus correspondientes componentes son todas discretas o todas continuas. Inclusive, existe el caso mixto, en que, por ejemplo, una o algunas de las componentes del vector sea(n) discreta(s) y las otras, continuas. Esta última situación la trabajaremos al final de esta sección.

2.2.1 Vectores aleatorios discretos

Definición 2.2.1 *Sea* $X := (X_1, \dots, X_n)$ *un vector de* $(\mathbb{R}, \mathfrak{B})$*-variables aleatorias* $X_1, \dots, X_n$ *sobre un espacio de probabilidad* $(\Omega, \mathfrak{F}, P_X)$ *con* $P_{X_1}, \dots, P_{X_n}$*, respectivamente. El vector aleatorio* X *se llama* DISCRETO *si tiene un número o enumerable de valores, es decir, si todas sus correspondientes componentes* X_i *son discretas.*

A continuación, queremos limitarnos al caso de un vector aleatorio bidimensional (X, Y) discreto, el cual, desde el punto de vista del cálculo de probabilidades, estará determinado cuando todos los valores (x_i, y_k), para $i = 1, \dots, n$ y $k = 1, \dots, m$, del vector aleatorio y las correspondientes PROBABILIDADES CONJUNTAS

$$p_{ik} := P(X = x_i, Y = y_k), \quad i = 1, \dots, n \text{ y } k = 1, \dots, m,$$

con las cuales el vector aleatorio (X,Y) toma estos valores, estén dados. Las PROBABILIDADES MARGINALES son discretas y están dadas por

$$\begin{aligned} p_{i*} &:= P(X=x_i) = P(X=x_i, Y<\infty) = P\Big(X=x_i, \bigcup_{k=1}^{m}\{Y=y_k\}\Big) = \sum_k p_{ik}, \\ p_{*k} &:= P(Y=y_k) = P(X<\infty, Y=y_k) = P\Big(\bigcup_{i=1}^{n}\{X=x_i\}, Y=x_k\Big) = \sum_i p_{ik}, \end{aligned}$$

siendo disyuntas las uniones anteriores. Obsérvese que $p_{ik} \geq 0$ y que, además, se cumple que $p_{**} := \sum_i \sum_k p_{ik} = 1$. Por consiguiente, teniendo en cuenta que las variables aleatorias X y Y pueden ser tomar finitos valores (digamos, n resp. m valores), o infinitos valores (es decir, mayor a n resp. m) o una finitos valores y la otra infinitos valores, entonces, el vector aleatorio (X,Y) se puede caracterizar, también, a través de la llamada TABLA DE DISTRIBUCIÓN:

	$Y=y_1$	y_2	$\cdots$	y_m	$\cdots$	$\sum_k p_{ik}$
$X=x_1$	p_{11}	p_{12}	$\cdots$	p_{1m}	$\cdots$	p_{1*}
x_2	p_{21}	p_{22}	$\cdots$	p_{2m}	$\cdots$	p_{2*}
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
x_n	p_{n1}	p_{n2}	$\cdots$	p_{nm}	$\cdots$	p_{n*}
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
$\sum_i p_{ik}$	p_{*1}	p_{*2}	$\cdots$	p_{*m}	$\cdots$	$p_{**}=1$

Los valores de la función de distribución $F_{(X,Y)}$ resultan de las probabilidades conjuntas de la siguiente manera:

$$F_{(X,Y)}(x,y) = P(X\leq x, Y\leq y) = \sum_{\substack{i \text{ tal que} \\ x_i\leq x}} \sum_{\substack{k \text{ tal que} \\ y_k\leq y}} P(X=x_i, Y=y_k) = \sum_{\substack{i \text{ tal que} \\ x_i\leq x}} \sum_{\substack{k \text{ tal que} \\ y_k\leq y}} p_{ik}$$

Ejemplo 2.2.2 *Suponga que 3 fichas son seleccionadas aleatoriamente (sin reemplazo y al mismo tiempo) de una caja que contiene 3 rojas, 4 blancas y 5 azules. Si X y Y denotan, respectivamente, el número de fichas rojas y blancas escogidas, encuentre:*

(a) La función de probabilidad conjunta de X y Y.

(b) La función de probabilidad marginal de X.

(c) La función de probabilidad marginal de Y.

(d) La función de distribución marginal de X.

La respuesta de (a) aparece en la tabla 2.1.

Tabla 2.1: Tabla de distribución conjunta para las variables del ejemplo 2.2.2

◀

$P(X = i, Y = j)$	Y= 0	1	2	3
X=0	10/220	40/220	30/220	4/220
1	30/220	60/220	18/220	0
2	15/220	12/220	0	0
3	1/220	0	0	0

2.2.2 Vectores aleatorios continuos

Ahora, analizaremos los llamados vectores aleatorios continuos.

Definición 2.2.3 *Un vector aleatorio finito se llama* CONTINUO *si todas sus correspondientes componentes son continuas.*

Como antes, consideremos ahora, el caso bidimensional: Sea (X, Y) un vector de $(\mathbb{R}, \mathfrak{B})$-variables aleatorias X, Y sobre un espacio de probabilidad $(\Omega, \mathfrak{F}, P_{(X,Y)})$ con distribuciones P_X, P_Y, respectivamente. Entonces, por la definición 2.2.3, (X, Y) será un VECTOR ALEATORIO CONTINUO si existe una función continua $f_{(X,Y)} \geq 0$ tal que

$$F_{(X,Y)}(x, y) \;=\; P_{(X,Y)}(X \leq x, Y \leq y) \;=\; \int\limits_{-\infty}^{x} \int\limits_{-\infty}^{y} f_{(X,Y)}(u, v)\, dv\, du,$$

para todo $x, y \in \mathbb{R}$. La función $F_{(X,Y)}$ se llama FUNCIÓN DE DISTRIBUCIÓN CONJUNTA de (X, Y). Debido a que debe cumplirse $\int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} f_{(X,Y)}(u, v)\, du\, dv = 1$, entonces, $f_{(X,Y)}$ es llamada la DENSIDAD DE PROBABILIDAD CONJUNTA de

(X,Y). La relación entre las funciones de distribución y de densidad conjuntas se puede expresar también en la forma

$$\frac{\partial^2 F_{(X,Y)}(x,y)}{\partial x \partial y} = f_{(X,Y)}(x,y), \quad x, y \in \mathbb{R}.$$

Las FUNCIONES DE DISTRIBUCIONES MARGINALES de X y Y están dadas por

$$\begin{aligned} F_X(x) &= \lim_{y\to\infty} F_{(X,Y)}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{\infty} f_{(X,Y)}(u,v)\,dv\,du, = \int_{-\infty}^{\infty}\int_{-\infty}^{x} f_{(X,Y)}(u,v)\,du\,dv, \\ F_Y(y) &= \lim_{x\to\infty} F_{(X,Y)}(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{y} f_{(X,Y)}(u,v)\,dv\,du = \int_{-\infty}^{y}\int_{-\infty}^{\infty} f_{(X,Y)}(u,v)\,du\,dv, \end{aligned}$$

respectivamente. Derivando con respecto a x resp. y se obtienen las llamadas DENSIDADES DE PROBABILIDADES MARGINALES de X resp. Y:

$$f_X(x) = \int_{-\infty}^{\infty} f_{(X,Y)}(x,y)\,dy, \quad \text{resp.} \quad f_Y(y) = \int_{-\infty}^{\infty} f_{(X,Y)}(x,y)\,dx,$$

que son las densidades de X resp. Y.

Ejemplo 2.2.4 *Para prestar un mejor servicio a sus clientes, durante un intervalo de tiempo de una hora un banco ha facilitado dos ventanillas. En un día seleccionado al azar, sean X la variable aleatoria que representa "la fracción del tiempo que la primera ventanilla permanece ocupada por un cliente" y Y, la variable que representa "la fracción del tiempo que la segunda ventanilla permanece ocupada por otro cliente". Entonces, tanto X como Y toman valores en el intervalo cerrado $[0,1]$. Si la función de densidad conjunta f del vector (X,Y) está dada por*

$$f(x,y) = \begin{cases} \frac{6}{5}(x+y^2), & \textit{si } 0 \le x \le 1, \quad 0 \le y \le 1; \\ 0, & \textit{de otra manera,} \end{cases}$$

(a) Verifique que f es, en realidad, una función de densidad conjunta.

(b) Calcule la probabilidad de que ninguna de las ventanillas esté ocupada más de una cuarta parte del tiempo. (R: 7/640)

(c) La función de densidad marginal de X.

(d) La función de densidad marginal de Y.

(e) La función de distribución marginal de X. ◀

Finalizaremos esta sección con la llamada *distribución normal bidimensional*, la cual es una distribución de un vector aleatorio bidimensional continuo y que muy útil para las aplicaciones.

Definición 2.2.5 *El vector aleatorio (X,Y) continuo tiene una* DISTRIBUCIÓN NORMAL *bidimensional con los parámetros $\mu_1, \mu_2 \in \mathbb{R}$, $\sigma_1^2, \sigma_2^2 > 0$ y $\varrho \in \mathbb{R}$ con $|\varrho| \leq 1$, en símbolos $(X,Y) \stackrel{d}{=} \mathcal{N}(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \varrho)$, si para todo $x, y \in \mathbb{R}$, su densidad conjunta está dada por*

$$f_{(X,Y)}(x,y) := \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\varrho^2}} \exp\Big\{ -\frac{1}{2(1-\varrho^2)}\Big[\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\frac{\varrho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\Big]\Big\}.$$

La distribución normal bidimensional se debe a LAPLACE quien, en 1811, la encuentra al estudiar problemas de estimación lineal con varias variables. El parámetro ϱ tiene con el grado con el "grado" de dependencia entre las variables X y Y. El es conocido como el coeficiente de correlación entre X y Y y se analizará con más detalles en la sección 2.3.

El siguiente teorema caracteriza la distribución marginal de una distribución normal bidimensional.

Teorema 2.2.6 *Si $(X,Y) \stackrel{d}{=} \mathcal{N}(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \varrho)$, entonces, $X \stackrel{d}{=} \mathcal{N}(\mu_1, \sigma_1^2)$ y $Y \stackrel{d}{=} \mathcal{N}(\mu_2, \sigma_2^2)$, es decir, la distribución normal bidimensional tiene distribuciones marginales normales.*

DEMOSTRACIÓN:
Como ejercicio. Usar la identidad

$$\frac{(x-\mu_1)^2}{\sigma_1^2} + \frac{(y-\mu_2)^2}{\sigma_2^2} - \frac{2\varrho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} = \frac{(x-\mu_1)^2}{\sigma_1^2}(1-\varrho^2) + \left(\frac{y-\mu_2}{\sigma_2} - \varrho\frac{(x-\mu_1)}{\sigma_1}\right)^2$$

Posteriormente, al resolver la integral que nos queda utilizamos la sustitución

$$u = \frac{1}{\sqrt{1-\varrho^2}}\left(\frac{y-\mu_2}{\sigma_2} - \varrho\frac{(x-\mu_1)}{\sigma_1}\right)$$

De esta forma queda demostrado el teorema. ■

2.2.3 Variables aleatorias independientes

La noción de independencia de más de dos variables aleatorias es similar a la noción de independencia de más de dos eventos.

Definición 2.2.7 *Las variables $X_1, X_2, \ldots, X_n$ son* INDEPENDIENTES *si y sólo si*

$$P(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n) = P(X_1 \leq x_1)\, P(X_2 \leq x_2) \cdots P(X_n \leq x_n)$$

Lo anterior es equivalente a: Si f es la función de distribución conjunta del vector aleatorio $(X_1, X_2, \ldots, X_n)$ y si f_i es la función de distribución marginal de X_i, $i = 1, 2, \ldots, n$, entonces $X_1, X_2, \ldots, X_n$ son independientes si y sólo si

$$f(x_1, x_2, \ldots, x_n) = f_1(x_1)\, f_2(x_2) \cdots f_n(x_n)$$

Ejemplo 2.2.8 *Suponga que el tiempo de vida, en años, de un cierto producto alimenticio perecedero empacado en cajas de cartón es una variable aleatoria distribuida exponencialmente con parámetro $\lambda = 1$. Si X_1, X_2 y X_3 representan las vidas de tres de estas cajas seleccionadas independientemente, entonces encuentre $P(X_1 < 2,\ 1 < X_2 < 3,\ X_3 > 2)$.*

SOLUCIÓN:
Dado que las cajas fueron seleccionadas independientemente, puede asumirse que las variables X_1, X_2 y X_3 son independientes y que tienen función de densidad conjunta dada por

$$f(x_1, x_2, x_3) = f_{X_1}(x_1)\, f_{X_2}(x_2)\, f_{X_3}(x_3) = e^{-x_1}\, e^{-x_2}\, e^{-x_3} = e^{-x_1 - x_2 - x_3}$$

para todo $x_1, x_2, x_3 > 0$ y $f(x_1, x_2, x_3) = 0$ en otro caso. Por lo tanto,

$$P(X_1 < 2,\ 1 < X_2 < 3,\ X_3 > 2) = 0,0376$$ ◀

2.3 Varianza de sumas, covarianza y correlación

Teorema 2.3.1 (Teorema de multiplicación) *Sean $X_1, \ldots, X_n$ variables aleatorias reales e independientes sobre un espacio de probabilidad $(\Omega, \mathfrak{F}, P)$. Si todas* *las X_i son no negativas o que $E(X_i)$ existe, entonces, se tiene*

$$E(X_1 \cdots X_n) = E(X_1) \cdots E(X_n).$$

Entonces en el segundo caso, la esperanza de $X_1 \cdots X_n$ también existe.

DEMOSTRACIÓN:
Se deja al lector. ■

Ejemplo 2.3.2 *Sean X y Y variables aleatorias discretas con la distribución de probabilidades conjuntas dada en la tabla 2.2. Muestre que $E(XY) = E(X)\,E(Y)$ (o, que es equivalente, que $Cov(X, Y) = 0$), pero que X y Y no son independientes.*

Tabla 2.2: Tabla de probabilidades conjuntas para las variables del ejemplo 2.3.2 ◀

	x= -1	0	1
y=-1	1/16	3/16	1/16
0	3/16	0	3/16
1	1/16	3/16	1/16

Definición 2.3.3 *Dos variables aleatorias reales X, Y sobre un espacio de probabilidad $(\Omega, \mathfrak{F}, P)$, tales que sus respectivas esperanzas existan, se llaman* INCORRELADAS, *cuando el producto $E(XY)$ existe con $E(XY) = E(X)E(Y)$.*

Teorema 2.3.4 (Igualdad de Bienaymé) *Sean $X_1, \ldots, X_n$ variables aleatorias, reales, incorreladas dos a dos y sobre un espacio de probabilidad $(\Omega, \mathfrak{F}, P)$. Si $E(X_i)$ existe, entonces, $Var(X_1 + \cdots + X_n) = Var(X_1) + \cdots + Var(X_n)$.*

DEMOSTRACIÓN:
Para cada $i \in \{1, \ldots, n\}$, sea $Y_i := X_i - E(X_i)$. Por tanto, $E(Y_i) = 0$. Ya que las variables aleatorias $X_1, \ldots, X_n$ son incorreladas dos a dos, tenemos que

$$\begin{aligned} E(Y_iY_j) &= E\big([X_i - E(X_i)][X_j - E(X_j)]\big) \\ &= E\big(X_iX_j - X_iE(X_j) - X_jE(X_i) + E(X_i)E(X_j)\big) \\ &= E(X_iX_j) - E(X_i)E(X_j) - E(X_i)E(X_j) + E(X_i)E(X_j) \overset{\text{hip.}}{=} 0, \end{aligned}$$

para todo $i, j \in \{1, \dots, n\}$ con $i \neq j$. Entonces, las variables aleatorias $Y_1, \dots, Y_n$ también son incorreladas dos a dos. Además,

$$\begin{aligned} Var(X_1 + \cdots + X_n) &\overset{\text{def.}}{=} E\big([X_1 + \cdots + X_n \; - \; E(X_1 + \cdots + X_n)]^2\big) \\ &= E\big([Y_1 + \cdots + Y_n]^2\big) \\ &= E\Big(\sum_{i,j=1}^{n} Y_i Y_j\Big) \; = \; \sum_{i,j=1}^{n} E(Y_i Y_j). \end{aligned}$$

Ya que $E(Y_i Y_j) = 0$, para todo $i, j \in \{1, \dots, n\}$ con $i \neq j$, con lo anterior se tiene que

$$Var(X_1 + \cdots + X_n) \; = \; E(Y_1^2) + \cdots + E(Y_n^2) \; = \; Var(X_1) + \cdots + Var(X_n).$$

De eta forma hemos demostrado el teorema. ■

Definición 2.3.5 *Sean X, Y variables aleatorias reales sobre un espacio de probabilidad $(\Omega, \mathfrak{F}, P)$ con esperanza y varianza finita. Entonces, la cantidad*

$$Cov(X, Y) \; := \; E\big([X - E(X)][Y - E(Y)]\big) \; \in \; \mathbb{R}, \tag{2.1}$$

es la llamada COVARIANZA *de X y Y. El* COEFICIENTE DE CORRELACIÓN *de X y Y está definido por*

$$\varrho := \varrho(X, Y) := \frac{Cov(X, Y)}{\sqrt{Var(X) Var(Y)}}.$$

Algunas propiedades para la covarianza son las siguientes:

Teorema 2.3.6 *Sean $(\Omega, \mathfrak{F}, P)$ un espacio de probabilidad y X, Y variables aleatorias reales sobre $(\Omega, \mathfrak{F}, P)$ con varianza finita. Entonces,*

(a) $Cov(X, X) = Var(X)$.

(b) $Cov(X, Y) = Cov(Y, X)$.

(c) $Cov(aX + b, \, Y) = a \, Cov(X, Y)$ *para todo* $a, b \in \mathbb{R}$.

(d) $\big[Cov(X, Y)\big]^2 \leq Var(X) \, Var(Y)$.

(e) $Cov(X,Y) = E(XY) - E(X)E(Y)$, *es decir,* X *y* Y *son incorreladas si y sólo si* $Cov(X,Y) = 0$.

(f) *Si* X *y* Y *son independientes, entonces,* $Cov(X,Y) = 0$.

DEMOSTRACIÓN:
Se deja al lector. ■

Teorema 2.3.7 *Sean* $(\Omega, \mathfrak{F}, P)$ *un espacio de probabilidad y* $X_1, \ldots, X_n$ *variables aleatorias reales sobre* $(\Omega, \mathfrak{F}, P)$ *con varianza finita. Entonces,*

$$Var\Big(\sum_{i=1}^{n} X_i\Big) \;=\; \sum_{i,j=1}^{n} Cov(X_i, X_j) \;=\; \sum_{i=1}^{n} Var(X_i) \;+\; 2\sum_{\substack{i,j=1\\ i<j}}^{n} Cov(X_i, X_j).$$

DEMOSTRACIÓN:
Teniendo en cuenta que $\Big(\sum_{i=1}^{n} a_i\Big)^2 = \sum_{j=1}^{n}\sum_{i=1}^{n} a_i\, a_j$, válida para todo $n \in \mathbb{N}$ y todo $a_i \in \mathbb{R}$ con $i = 1, \ldots, n$, se tiene que

$$Var\Big(\sum_{i=1}^{n} X_i\Big) \;=\; E\Big(\Big[\sum_{i=1}^{n} X_i - E\Big(\sum_{i=1}^{n} X_i\Big)\Big]^2\Big) \;=\; E\Big(\Big(\sum_{i=1}^{n} [X_i - E(X_i)]\Big)^2\Big) \tag{1}$$

$$= \; E\Big(\sum_{i,j=1}^{n} [X_i - E(X_i)]\,[X_j - E(X_j)]\Big) \;=\; \sum_{i,j=1}^{n} Cov(X_i, X_j).$$

De esta forma queda demostrada la primera igualdad. La segunda igualdad se obtiene trivialmente de (1) y teniendo en cuenta que

$$\Big(\sum_{i=1}^{n} a_i\Big)^2 = \sum_{i=1}^{n} a_i^2 + 2\sum_{\substack{i,j=1\\ i<j}}^{n} a_i\, a_j$$

Con lo anterior queda completamente demostrado el teorema. ■

2.4 Distribuciones, esperanza y varianzas condicionales

2.4.1 Distribuciones condicionales

Definición 2.4.1 *Sea f la función de probabilidad conjunta de un vector aleatorio $(X_1, X_2, \ldots, X_n)$. Entonces, la* FUNCIÓN DE PROBABILIDAD (DENSIDAD) CONDICIONAL CONJUNTA *de $X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_{j-1}, X_{j+1}, \ldots, X_n$, dado que $X_i = x_i$ y $X_j = x_j$, con $i < j$, está dada por*

$$h(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n \,/\, x_i, x_j) \;=\; \frac{f(x_1, \ldots, x_n)}{f_{ij}(x_i, x_j)}$$

siendo f_{ij} la función de probabilidad (densidad) marginal conjunta de las variables aleatorias X_i y X_j.

Ejemplo 2.4.2 *Supongamos que X_1 es una variable aleatoria cuya función de densidad f_1 es la siguiente:*

$$f_1(x) \;=\; \begin{cases} e^{-x}, & \text{si } x > 0, \\ 0, & \text{en otro caso.} \end{cases}$$

Supongamos, además, que para cualquier valor concreto $X_1 = x_1$ ($x_1 > 0$), otras dos variables aleatorias, X_2 y X_3, son independientes, que ambas variables tienen la misma distribución y que la función de densidad condicional de cada una de estas variables, dado $X_1 = x_1$, es la siguiente:

$$g(t/x_1) \;=\; \begin{cases} x_1\, e^{-x_1 t}, & \text{si } t > 0, \\ 0, & \text{en otro caso.} \end{cases}$$

Determine:

(a) La función de densidad condicional conjunta de X_2 y X_3 dado $X_1 = x_1$.

(b) La función de densidad conjunta de X_1, X_2 y X_3.

(c) La función de distribución marginal conjunta de X_2 y X_3.

(d) La función de densidad condicional de X_1 dado que $X_2 = x_2$ y $X_3 = x_3$ ($x_2 > 0$, $x_3 > 0$).

(e) $P(X_1 \leq 1 \,/\, X_2 = 1, X_3 = 4)$. ◄

2.4.2 Teorema de la probabilidad total y regla de Bayes

Teorema 2.4.3 *Si X y Y son dos variables aleatorias discretas, entonces:*

(a) El teorema de la probabilidad total:

$$P(Y=y) \ = \ \sum_x P(Y=y/X=x)\,P(X=x)$$

(b) La regla de Bayes:

$$P(X=x/Y=y) \ = \ \frac{P(X=x,Y=y)}{P(Y=y)} \ = \ \frac{P(Y=y/X=x)\,P(X=x)}{\sum_x P(Y=y/X=x)\,P(X=x)}$$

DEMOSTRACIÓN:
Se deja al lector. ■

Ejemplo 2.4.4 *Sea Y el número de caras en X lanzamientos de una moneda no falsa, donde X es el número generado por el lanzamiento de un dado. Encuentre:*

(a) La distribución condicional de Y dado $X=x$, para $x=1,\dots,6$.

(b) La distribución de Y.

(c) La distribución conjunta de las variables X y Y.

(d) La distribución condicional de X dado $Y=y$, para $y=0,1,\dots,6$.

SOLUCIÓN:

(a) Es la binomial con parámetros $n=x$ y $p=1/2$.

(b) Se muestra en la tabla 2.3.

Tabla 2.3: Distribución de probabilidad para la variable Y del ejemplo 2.4.4

y	0	1	2	3	4	5	6
$P(Y=y)$	$\frac{63}{384}$	$\frac{120}{384}$	$\frac{99}{384}$	$\frac{64}{384}$	$\frac{29}{384}$	$\frac{8}{384}$	$\frac{1}{384}$

Tabla 2.4: Distribución conjunta de las variables X y Y del ejemplo 2.4.4

	X =1	2	3	4	5	6	$f_Y(y)$
Y =0	$\frac{1}{6}\cdot\frac{1}{2}$	$\frac{1}{6}\cdot\frac{1}{4}$	$\frac{1}{6}\cdot\frac{1}{8}$	$\frac{1}{6}\cdot\frac{1}{16}$	$\frac{1}{6}\cdot\frac{1}{32}$	$\frac{1}{6}\cdot\frac{1}{64}$	$\frac{63}{384}$
1	$\frac{1}{6}\cdot\frac{1}{2}$	$\frac{1}{6}\cdot\frac{2}{4}$	$\frac{1}{6}\cdot\frac{3}{8}$	$\frac{1}{6}\cdot\frac{4}{16}$	$\frac{1}{6}\cdot\frac{5}{32}$	$\frac{1}{6}\cdot\frac{6}{64}$	$\frac{120}{384}$
2	0	$\frac{1}{6}\cdot\frac{1}{4}$	$\frac{1}{6}\cdot\frac{3}{8}$	$\frac{1}{6}\cdot\frac{6}{16}$	$\frac{1}{6}\cdot\frac{10}{32}$	$\frac{1}{6}\cdot\frac{15}{64}$	$\frac{99}{384}$
3	0	0	$\frac{1}{6}\cdot\frac{1}{8}$	$\frac{1}{6}\cdot\frac{4}{16}$	$\frac{1}{6}\cdot\frac{10}{32}$	$\frac{1}{6}\cdot\frac{20}{64}$	$\frac{64}{384}$
4	0	0	0	$\frac{1}{6}\cdot\frac{1}{16}$	$\frac{1}{6}\cdot\frac{5}{32}$	$\frac{1}{6}\cdot\frac{15}{64}$	$\frac{29}{384}$
5	0	0	0	0	$\frac{1}{6}\cdot\frac{1}{32}$	$\frac{1}{6}\cdot\frac{6}{64}$	$\frac{8}{384}$
6	0	0	0	0	0	$\frac{1}{6}\cdot\frac{1}{64}$	$\frac{1}{384}$
$f_X(x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	1

(c) Se muestra en la tabla 2.4.

(d) La distribución condicional de X dado $Y = 2$ se muestra en la tabla 2.5. Así, dadas dos caras, el número de monedas lanzadas es igualmente probables en 3 o en 4 y estos valores son los más probables.

Tabla 2.5: Distribución condicional de X dado $Y = 2$ en el ejemplo 2.4.4 ◀

x	1	2	3	4	5	6
$P(X = x/Y = 2)$	0	$\frac{16}{384}$	$\frac{24}{99}$	$\frac{24}{99}$	$\frac{20}{99}$	$\frac{15}{99}$

2.4.3 Esperanza condicional

Las esperanzas condicionales son simplemente esperanzas relacionadas con distribuciones condicionales.

Definición 2.4.5 *Sean X y Y variables aleatorias discretas o continuas y $h(y/x)$ la función de probabilidad (de densidad) condicional de Y dado $X = x$. Entonces, la* ESPERANZA CONDICIONAL DE Y DADO X*, denotada por $E(Y/X)$, es la función de X cuyo valor es $E(Y/X = x)$ si $X = x$. Al valor $E(Y/X = x)$ se le llama la* ESPERANZA CONDICIONAL DE Y DADO QUE $X = x$ *y se define como*

$$E(Y/X = x) = \begin{cases} \int_{-\infty}^{\infty} y\, h(y/x)\, dy, & \text{si } X \text{ y } Y \text{ son continuas;} \\ \\ \sum_y y\, h(y/x), & \text{si } X \text{ y } Y \text{ son discretas.} \end{cases}$$

Es importante recalcar que la $E(Y/X)$ es una variable aleatoria. Por consiguiente, tiene sentido considerar su esperanza.

Teorema 2.4.6 *Sean X y Y variables aleatorias discretas o continuas y f_X la función de probabilidad (de densidad) marginal de X. Entonces,*

$$E(Y) = E\big(E(Y/X)\big) = \begin{cases} \int_{-\infty}^{\infty} E(Y/X = x)\, f_X(x)\, dx, & X \text{ y } Y \text{ continuas;} \\ \\ \sum_x E(Y/X = x)\, f_X(x), & X \text{ y } Y \text{ discretas.} \end{cases}$$

DEMOSTRACIÓN:
Como ejercicio. ■

Ejemplo 2.4.7 (Continuación del ejemplo 2.4.4) *Sea Y el número de caras en X lanzamientos de una moneda no falsa, en donde X es generado por el lanzamiento de un dado. Encuentre (a) $E(Y/X = x)$, (b) $E(Y)$, (c) $E(X/Y = 2)$.* ◀

Varianza condicional

Definición 2.4.8 *Sean X y Y variables aleatorias discretas o continuas y $h(y/x)$ la función de probabilidad (de densidad) condicional de Y dado $X = x$.*

Entonces, la varianza condicional de Y, dado X, notado por $V(Y/X)$, es una función de X cuyo valor es $V(Y/X = x)$. Al valor $V(Y/X = x)$ se le llama la VARIANZA CONDICIONAL DE Y DADO QUE $X = x$ *y se define como*

$$V(Y/X=x) \;=\; \begin{cases} \int\limits_{-\infty}^{\infty} \big(y - E(Y/X=x)\big)^2 \, h(y/x)\, dy, & \textit{si } X \textit{ y } Y \textit{ son continuas;} \\ \\ \sum\limits_{y} \big(y - E(Y/X=x)\big)^2 \, h(y/x), & \textit{si } X \textit{ y } Y \textit{ son discretas,} \end{cases}$$

Es importante recalcar que $V(Y/X)$ también es una variable aleatoria. Para ella se cumplen las siguientes propiedades:

Teorema 2.4.9 *Sean X y Y variables aleatorias discretas o continuas. Entonces,*

(a) $V(Y/X) \;=\; E\Big(\big[Y - E(Y/X)\big]^2\Big) \;=\; E(Y^2/X) \;-\; \big[E(Y/X)\big]^2$

(b) $E\big(V(Y/X)\big) \;=\; E(Y^2) \;-\; E\big([E(Y/X)]^2\big)$

(c) $V\big(E(Y/X)\big) \;=\; E\big([E(Y/X)]^2\big) \;-\; [E(Y)]^2$

(d) $V(Y) \;=\; E\big(V(Y/X)\big) \;-\; V\big(E(Y/X)\big)$

DEMOSTRACIÓN:
Como ejercicio. ■

Ejemplo 2.4.10 *Suponga que en cualquier tiempo t el número de personas que han llegado a una estación de trenes es una variable aleatoria de Poisson con media λt. Si el tren inicial llega a la estación en un tiempo (independiente del tiempo en que llegan los pasajeros) que está uniformemente distribuido sobre $(0, T)$, ¿cuál es la media y la varianza del número de pasajeros que entran al tren?* ◀

2.5 Convoluciones de medidas de probabilidad

Definición 2.5.1 *Sean $X_1, \ldots, X_n$ $(\mathbb{R}^n, \mathfrak{B}_n)$-variables aleatorias independientes sobre un espacio de probabilidad $(\Omega, \mathfrak{F}, P)$. La distribución de la variable aleatoria $X_1 + \cdots + X_n$ se llama la* CONVOLUCIÓN PRODUCTO *o, brévemente, la* CONVOLUCIÓN *de las medidas de probabilidad $P_{X_1}, \ldots, P_{X_n}$ y se simboliza con $P_{X_1} * \cdots * P_{X_n} := P_{X_1+\cdots+X_n}$.*

Ya que la adición " + " es conmutativa y asociativa, también lo es la convolución " $*$ ". Por consiguiente, es suficiente limitarse al estudio del producto convolución de dos factores.

A continuación, presentaremos fórmulas para calcular la distribución de una suma, resta, producto y división de variables aleatorias independientes. Nos limitaremos al caso de dos variables aleatorias reales.

Teorema 2.5.2 *Si X y Y variables aleatorias continuas, reales e independientes con las densidades de probabilidad f_X resp. f_Y, entonces, las correspondientes densidades de probabilidades $f_{X\pm Y}$, f_{XY} y $f_{X/Y}$ de las variables aleatorias $X \pm Y$, XY y X/Y están dadas por*

$$f_{X+Y}(t) = \int_{-\infty}^{\infty} f_X(x)\, f_Y(t-x)\, dx = \int_{-\infty}^{\infty} f_X(t-y)\, f_Y(y)\, dy \qquad (2.2)$$

$$f_{X-Y}(t) = \int_{-\infty}^{\infty} f_X(x)\, f_Y(x-t)\, dx = \int_{-\infty}^{\infty} f_X(t+y)\, f_Y(y)\, dy, \qquad (2.3)$$

$$f_{XY}(t) = \int_{-\infty}^{\infty} \tfrac{1}{|x|}\, f_X(x)\, f_Y\!\left(\tfrac{t}{x}\right) dx = \int_{-\infty}^{\infty} \tfrac{1}{|y|}\, f_X\!\left(\tfrac{t}{y}\right) f_Y(y)\, dy, \qquad (2.4)$$

$$f_{X/Y}(t) = \tfrac{1}{t^2} \int_{-\infty}^{\infty} |x|\, f_X(x)\, f_Y\!\left(\tfrac{x}{t}\right) dx = \int_{-\infty}^{\infty} |y|\, f_X(yt)\, f_Y(y)\, dy, \qquad (2.5)$$

respectivamente, para todo $t \in \mathbb{R}$. En el caso en que X y Y sean discretas independientes,

$$\begin{aligned} P(X+Y=n) &= \sum_{x=0}^{n} P(X=x, Y=n-x) \\ &= \sum_{x=0}^{n} P(X=x)\, P(Y=n-x) \qquad (2.6) \end{aligned}$$

DEMOSTRACIÓN:
Se deja al lector. ■

Ejemplo 2.5.3 (Convolución de distribuciones binomiales) *Sean $0 \le p \le 1$ y $n, m \in \mathbb{N}$ dados.*

$$\mathcal{B}(n,p) * \mathcal{B}(m,p) = \mathcal{B}(n+m,p). \qquad (2.7)$$

Es decir, la convolución de dos binomiales es también una binomial. ◄

Ejemplo 2.5.4 (Convolución de distribuciones de Poisson) *Sean $\alpha, \beta > 0$ dados. Por consiguiente,*

$$\mathcal{P}(\alpha) * \mathcal{P}(\beta) = \mathcal{P}(\alpha + \beta). \tag{2.8}$$

Es decir, la convolución de dos distribuciones de Poisson es también una Poisson. ◀

Ejemplo 2.5.5 (Convolución de distribuciones normales) *Sean $\mu_1, \mu_2 \in \mathbb{R}$ y $\sigma_1, \sigma_2 > 0$ dados. Para $i = 1, 2$, sea f_i la densidad de la distribución normal $\mathcal{N}(\mu_i, \sigma_i^2)$, es decir,*

$$f_i(x) := \frac{1}{\sigma_i\sqrt{2\pi}} \exp\left(-\frac{(x-\mu_i)^2}{2\sigma_i^2}\right), \quad \forall\, x \in \mathbb{R}.$$

La convolución $\mathcal{N}(\mu_1, \sigma_1^2) * \mathcal{N}(\mu_2, \sigma_2^2)$ tiene la densidad h, que es dada por

$$\begin{aligned} h(x) &= \int f_1(x-y) f_2(y)\, dy \\ &= \frac{1}{2\pi\sigma_1\sigma_2} \int \exp\left(-\frac{(x-y-\mu_1)^2}{2\sigma_1^2} - \frac{(y-\mu_2)^2}{2\sigma_2^2}\right) dy. \end{aligned}$$

Sustituyendo $u = \frac{y-\mu_2}{\sigma_2}$ y haciendo $\mu := \mu_1 + \mu_2$ y $\sigma := +\sqrt{\sigma_1^2 + \sigma_2^2}$, tenemos

$$\begin{aligned} h(x) &= \frac{1}{2\pi\sigma_1} \int \exp\left(-\frac{(x-\sigma_2 u-\mu_1-\mu_2)^2}{2\sigma_1^2} - \frac{u^2}{2}\right) du \\ &= \frac{1}{2\pi\sigma_1} \int \exp\left(-\frac{1}{2\sigma_1^2}[(x-\mu-\sigma_2 u)^2 + \sigma_1^2 u^2]\right) du. \end{aligned}$$

Facilmente, se prueba que

$$\sigma^2 \left(u - \frac{\sigma_2(x-\mu)}{\sigma^2}\right)^2 + \frac{\sigma_1^2(x-\mu)^2}{\sigma^2} = (x-\mu-\sigma_2 u)^2 + \sigma_1^2 u^2.$$

Sustituyendo esto, obtenemos

$$h(x) = \frac{1}{2\pi\sigma_1} \int \exp\left(-\frac{\sigma^2}{2\sigma_1^2}\left[u - \frac{\sigma_2(x-\mu)}{\sigma^2}\right]^2 - \frac{(x-\mu)^2}{2\sigma^2}\right) du.$$

Reemplazando ahora $v = \left(u - \frac{\sigma_2(x-\mu)}{\sigma^2}\right)\frac{\sigma}{\sigma_1}$, se tiene

$$\begin{aligned} h(x) &= \frac{1}{2\pi\sigma_1} \int \exp\left(-\frac{v^2}{2} - \frac{(x-\mu)^2}{2\sigma^2}\right) \frac{\sigma_1}{\sigma} dv \\ &= \frac{1}{2\pi\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \int \exp\left(-\frac{v^2}{2}\right) dv \\ &= \frac{1}{2\pi\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \end{aligned}$$

en donde se ha utilizado que $\int \exp\left(-\frac{v^2}{2}\right) dv = 1$. *Esto demuestra que la convolución* $\mathcal{N}(\mu_1, \sigma_1^2) * \mathcal{N}(\mu_2, \sigma_2^2)$ *está distribuida normálmente con los parámetros* $\mu = \mu_1 + \mu_2$ *y* $\sigma^2 = \sigma_1^2 + \sigma_2^2$, *es decir,*

$$\mathcal{N}(\mu_1, \sigma_1^2) * \mathcal{N}(\mu_2, \sigma_2^2) \ = \ \mathcal{N}(\mu_1 + \mu_2,\, \sigma_1^2 + \sigma_2^2). \tag{2.9}$$

Es decir, la convolución de dos normales es también una normal. ◄

Teorema 2.5.6 *Sean* $X_1, \ldots, X_n$ *variables aleatorias independientes.*

(a) *Si* $X_i \stackrel{d}{=} \mathcal{N}(\mu_i, \sigma_i^2)$ *para cada* $i = 1, \ldots, n$, *entonces,* $X_1 + \cdots + X_n \stackrel{d}{=} \mathcal{N}(\mu_1 + \cdots + \mu_n, \sigma_1^2 + \cdots + \sigma_n^2)$.

(b) *Si* $X_i \stackrel{d}{=} \gamma(\alpha_i, \beta)$ *para cada* $i = 1, \ldots, n$, *entonces,* $X_1 + \cdots + X_n \stackrel{d}{=} \gamma(\alpha_1 + \cdots + \alpha_n, \beta)$.

(c) *Si* $X_i \stackrel{d}{=} \mathcal{N}(0,1)$ *para cada* $i = 1, \ldots, n$, *entonces,* $X_1^2 + \cdots + X_n^2 \stackrel{d}{=} \chi^2(n)$.

DEMOSTRACIÓN:
Como ejercicio. ■

Ahora, analizaremos dos conceptos importantes en la matemática estadística: las media y varianza empíricas de variables aleatorias.

Definición 2.5.7 *Sean* $X_1, \ldots, X_n$ *variables aleatorias. Entonces, la variable aleatoria* $\overline{X}_{(n)} := \frac{1}{n}(X_1 + \cdots + X_n)$ *se llama la* MEDIA ARITMÉTICA *o* MEDIA EMPÍRICA *de* $X_1, \ldots, X_n$ *y a* $S_{(n)}^2 := \frac{1}{n-1} \sum_{k=1}^{n} (X_k - \overline{X}_{(n)})^2$ *se le llama* VARIANZA EMPÍRICA.

Para la media aritmética, obtenemos el siguiente teorema.

Teorema 2.5.8 *Sean* $X_1, \ldots, X_n$ *variables aleatorias independientes tales que* $X_i \stackrel{d}{=} \mathcal{N}(\mu, \sigma^2)$ *para cada* $i = 1, \ldots, n$. *Entonces,* $\overline{X}_{(n)} \stackrel{d}{=} \mathcal{N}\big(\mu, \frac{\sigma^2}{n}\big)$.

DEMOSTRACIÓN:
Como ejercicio. ■

Ahora, demostraremos algunos resultados básicos relacionados con la varianza empírica.

Teorema 2.5.9 (Distribución Chi-cuadrada) *Sean $X_1, \ldots, X_n$ variables aleatorias independientes con $E(X_k) = \mu$ y $Var(X_k) = \sigma^2$, para cada $k = 1, \ldots, n$. Además, sean $S^2_{(n)}$ y $\overline{X}_{(n)}$ la varianza y media empírica de $X_1, \ldots, X_n$, respectivamente.*

(a) Se cumple que $E(S^2_{(n)}) = \sigma^2$.

(b) Sea $Y_k := \left(X_k - \overline{X}_{(k-1)}\right)\sqrt{\frac{k-1}{k}}$ para $k = 2, \ldots, n$. Entonces, Y_k y Y_k^2 son independientes y se cumple que $\sum\limits_{k=2}^{n} Y_k^2 = \sum\limits_{k=1}^{n} \left(X_k - \overline{X}_{(k-1)}\right)^n = (n-1)S^2_{(n)}$.

(c) Si $X_k \overset{d}{=} \mathcal{N}(\mu, \sigma^2)$ para todo $k = 1, \ldots, n$, entonces, $\frac{n-1}{\sigma^2} S^2_{(n)} \overset{d}{=} \chi^2(n-1)$.

DEMOSTRACIÓN:

(a) Demuestre, primero, que

$$E(S^2_{(n)}) \;=\; \frac{1}{n-1}\left[\sum_{k=1}^{n} E(X_k^2) - n\,E\left(\overline{X}_{(n)}\right)^2\right]$$

(b) Por inducción sobre n. El caso $n = 2$ es claro. Ahora, supongamos que el teorema se cumple para $s = n-1$ y demostraremos para $s = n$.

$$\begin{aligned}
\sum_{k=1}^{n-1} Y_k^2 &= \sum_{k=1}^{n-1} X_k^2 - (n-1)\overline{X}^2_{(n-1)} + (X_n - \overline{X}_{(n-1)})\frac{n-1}{n} \\
&= \sum_{k=1}^{n} X_k^2 - \frac{1}{n}X_n^2 + \frac{1-n}{n(n-1)}\left(\sum_{k=1}^{n-1} X_k\right)^2 - \frac{2}{n}X_n\sum_{k=1}^{n-1} X_k \\
&= \sum_{k=1}^{n} X_k^2 - \frac{1}{n}\left(X_n + \sum_{k=1}^{n-1} X_k\right)^2 \\
&= \sum_{k=1}^{n} X_k^2 - n\overline{X}^2_{(n)} \\
&= (n-1)S^2_{(n)}
\end{aligned}$$

(c) Primero demuestre que $X_k - \overline{X}_{(k-1)} \overset{d}{=} \mathcal{N}\left(0, \sigma^2\left(\frac{k}{k-1}\right)\right)$. Si Y_k es como en (b), entonces, $Y_k \overset{d}{=} \mathcal{N}(0, \sigma^2)$. Por esta razón, $Y_k^2 \overset{d}{=} \gamma\left(\frac{1}{2}, \frac{1}{2\sigma^2}\right)$ y esto implica que $\frac{n-1}{\sigma^2} S^2_{(n)} \overset{d}{=} \gamma\left(\frac{n-1}{2}, \frac{1}{2}\right)$. ■

Ahora, definiremos otras variables que obedecen la distribución t de Student $\mathcal{T}(n)$ con n grados de libertad.

Teorema 2.5.10 (Distribución t de Student) *Sean X, Y, $X_1, \ldots, X_n$ y $Y_1, \ldots, Y_m$ variables aleatorias. Además, sean $S^2_{(n)}$ y $\overline{X}_{(n)}$ resp. $S^2_{(m)}$ y $\overline{Y}_{(m)}$ la varianza y media empírica de $X_1, \ldots, X_n$ y de $Y_1, \ldots, Y_m$, respectivamente. Supongamos que se tiene la independencia, por un lado, entre todas las X_i; por otro lado, entre todas las Y_j; y también entre X y Y. Si $\mathcal{T}(n)$ representa la distribución t de Student con n grados de libertad, entonces:*

(a) Si $X \stackrel{d}{=} \mathcal{N}(0,1)$ y $Y \stackrel{d}{=} \chi^2(n)$, entonces, $t := \frac{X}{\sqrt{Y/n}} \stackrel{d}{=} \mathcal{T}(n)$.

(b) Si $X_i \stackrel{d}{=} \mathcal{N}(\mu, \sigma^2)$ para cada $i = 1, \ldots, n$, entonces, se cumple que $t := \frac{\overline{X}_{(n)} - \mu}{S_{(n)}/\sqrt{n}} \stackrel{d}{=} \mathcal{T}(n-1)$.

(c) Si $X_i \stackrel{d}{=} \mathcal{N}(\mu_1, \sigma^2)$ para cada $i = 1, \ldots, n$ y $Y_j \stackrel{d}{=} \mathcal{N}(\mu_2, \sigma^2)$ para cada $j = 1, \ldots, m$, entonces, $t := \frac{\left(\overline{X}_{(n)} - \overline{Y}_{(m)}\right) - (\mu_1 - \mu_2)}{\sqrt{\frac{S^2_{(n,m)}}{n} + \frac{S^2_{(n,m)}}{m}}} \stackrel{d}{=} \mathcal{T}(m+n-2)$, siendo $S^2_{(n,m)} := \frac{(n-1)S^2_{(n)} + (m-1)S^2_{(m)}}{m+n-2}$ la llamada VARIANZA MUESTRAL COMBINADA.

DEMOSTRACIÓN:
Sólo demostraremos (a). Recuerde que la densidad t de Student con n grados de libertad es

$$\mathcal{T}(n) \sim \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{t^2}{n}\right)^{-(n+1)/2}$$

Se tiene que $\frac{Y}{n} \stackrel{\mathrm{d}}{=} \gamma\left(\frac{n}{2}, \frac{n}{2}\right)$. Debido a que $f_{\sqrt{X}}(x) = 2\,|x|\, f_X(x^2)$, tenemos que

$$\sqrt{\frac{Y}{n}} \sim \frac{\left(\frac{n}{2}\right)^{n/2}}{\Gamma\left(\frac{n}{2}\right)} \cdot (x^2)^{(n/2)-1} \cdot 2x \cdot e^{-nx^2/2}, \quad x > 0$$

Teniendo en cuenta lo anterior, aplicando la segunda igualdad de (2.5), haciendo la sustitución $u = \frac{1}{2}x^2\,(t^2 + n)$ y sabiendo que $\Gamma(x) = \int\limits_0^\infty u^{x-1} e^{-u}\, du$

para $x > 0$, tenemos

$$\begin{aligned}
\frac{X}{\sqrt{Y/n}} &\sim \frac{2 \cdot \left(\frac{n}{2}\right)^{n/2}}{\sqrt{2\pi}\,\Gamma\left(\frac{n}{2}\right)} \int_0^\infty x\,(x^2)^{(n-1)/2} \exp\left\{-\frac{1}{2}\,x^2\,(t^2+n)\right\} dx \\
&= \frac{2 \cdot \left(\frac{n}{2}\right)^{n/2}}{\sqrt{2\pi}\,\Gamma\left(\frac{n}{2}\right)} \int_0^\infty \frac{2^{(n-1)/2}\,u^{(n-1)/2}}{(t^2+n)^{(n-1)/2}}\,e^{-u}\,\frac{du}{t^2+n} \\
&= \frac{2^{-1/2} \cdot 2 \cdot \left(\frac{n}{2}\right)^{n/2} \cdot 2^{(n-1)/2}}{\sqrt{\pi}\,\Gamma\left(\frac{n}{2}\right)\,(t^2+n)^{(n+1)/2}} \int_0^\infty u^{-1+(n+1)/2}\,e^{-u}\,du \\
&\sim \mathcal{T}(n)
\end{aligned}$$

La demostración de (b) y (c) se dejan como ejercicio. ■

Ahora, presentaremos algunos resultados con respecto a la distribución F de Fisher $\mathcal{F}(m,n)$ con m y n grados de libertad.

Teorema 2.5.11 (Distribución F de Fisher) *Sean $X, Y, X_1, \ldots, X_n$ y $Y_1, \ldots, Y_m$ variables aleatorias. Además, sean $S^2_{(n)}$ y $\overline{X}_{(n)}$ resp. $S^2_{(m)}$ y $\overline{Y}_{(m)}$ la varianza y media empírica de $X_1, \ldots, X_n$ resp. $Y_1, \ldots, Y_m$. Supongamos que se tiene la independencia, por un lado, entre todas las X_i; por otro lado, entre todas las Y_j; y también entre X y Y. Si $\mathcal{F}(m,n)$ representa la la distribución F de Fisher con m y n grados de libertad, entonces:*

(a) Si $X \overset{d}{=} \chi^2(m)$ y $Y \overset{d}{=} \chi^2(n)$, entonces, $F := \frac{X/m}{Y/n} = \frac{nX}{mY} \overset{d}{=} \mathcal{F}(m,n)$.

(b) Si $X_i \overset{d}{=} \mathcal{N}(\mu_1, \sigma_1^2)$ para cada $i = 1, \ldots, n$ y $Y_j \overset{d}{=} \mathcal{N}(\mu_2, \sigma_2^2)$ para cada $j = 1, \ldots, m$, entonces, $F := \frac{S^2_{(n)}/\sigma_1^2}{S^2_{(m)}/\sigma_2^2} \overset{d}{=} \mathcal{F}(n-1, m-1)$.

DEMOSTRACIÓN:
Sólo demostraremos (a). Recuerde que

$$\mathcal{F}(m,n) \sim \frac{\Gamma\left(\frac{m+n}{2}\right)\,m^{m/2}\,n^{n/2}\,t^{\frac{m}{2}-1}}{\Gamma\left(\frac{m}{2}\right)\,\Gamma\left(\frac{n}{2}\right)\,(n+mt)^{(m+n)/2}}, \quad t > 0$$

$$\gamma(\alpha,\beta) \sim \frac{\beta^\alpha}{\Gamma(\alpha)}\,t^{\alpha-1}\,e^{-\beta t}, \quad t > 0$$

Se tiene que $nX \stackrel{\mathrm{d}}{=} \gamma\left(\frac{m}{2}, \frac{1}{2n}\right)$ y $mY \stackrel{\mathrm{d}}{=} \gamma\left(\frac{n}{2}, \frac{1}{2m}\right)$. Por tanto, teniendo en cuenta lo anterior, aplicando la segunda igualdad de (2.5), haciendo la sustitución $u = \frac{x}{2}\left(\frac{t}{n} + \frac{1}{m}\right)$ y sabiendo que $\Gamma(x) = \int\limits_0^\infty u^{x-1} e^{-u}\, du$, tenemos

$$\begin{aligned} F &= \frac{1}{(2n)^{m/2}(2m)^{n/2}\,\Gamma\left(\frac{m}{2}\right)\,\Gamma\left(\frac{n}{2}\right)} \int\limits_0^\infty |x|\, x^{-2+(m+n)/2}\, t^{-1+m/2} \exp\left\{-\frac{tx}{2n} - \frac{x}{2m}\right\} dx \\ &\sim \mathcal{F}(m,n) \end{aligned}$$

La parte (b) se deja como ejercicio. ■

2.6 Teoremas de transformación

Sea $X = (X_1, \ldots, X_n)$ un vector aleatorio y sean $g_1, \ldots, g_k$ funciones definidas sobre $A \subseteq \mathbb{R}^n$ y de valor real. Supóngase que $Y_1 := g_1(X_1, \ldots, X_n)$, ..., $Y_k := g_k(X_1, \ldots, X_n)$ son variables aleatorias reales. Se desea determinar la distribución conjunta de las variables aleatorias $Y_1, \ldots, Y_k$, en términos de la distribución conjunta de las variables aleatorias $X_1, \ldots, X_n$. Primero, supondremos que $X_1, \ldots, X_n$ son discretas y que se conoce su distribución conjunta.

Ejemplo 2.6.1 *Sean X_1 y X_2 variables aleatorias discretas con distribución conjunta dada en la tabla* 2.6.

Tabla 2.6: Tabla de distribución conjunta para las variables del ejemplo 2.6.1

	$X_1 = 0$	1	2
$X_2 = 0$	3/28	9/28	3/28
1	6/28	6/28	0
2	1/28	0	0

Sean $g_1(x_1, x_2) = x_1 - x_2$ y $g_2(x_1, x_2) = x_1 + x_2$. Es obvio que las variables aleatorias

$$Y_1 = g_1(X_1, X_2) = X_1 - X_2 \quad y \quad Y_2 = g_2(X_1, X_2) = X_1 + X_2$$

toman los valores $-2, -1, 0, 1$ y $0, 1, 2, 3, 4$, respectivamente. La distribución conjunta de Y_1 y Y_2 se muestra en la tabla 2.7.

Tabla 2.7: Tabla de distribución conjunta para las variables Y_1 y Y_2 del ejemplo 2.6.1 ◀

	Y_1= -2	-1	0	1	2
Y_2= 0	0	0	3/28	0	0
1	0	6/28	0	9/28	0
2	1/28	0	6/28	0	3/28
3	0	0	0	0	0
4	0	0	0	0	0

Ahora presentaremos dos teoremas que nos permitirán determinar la distribución de una función de un vector aleatorio para el caso en que las variables $X_1, \ldots, X_n$ sean continuas. Antes, la siguiente definición.

Definición 2.6.2 (Jacobiano) *Sea $h : R^n \to R^m$ una función definida por*

$$h(x_1, \ldots, x_n) = \big(h_1(x_1, \ldots, x_n), \ldots, h_m(x_1, \ldots, x_n)\big)$$

Las derivadas parciales de estas funciones h_i (si existen) pueden ser organizadas en una matriz de m por n, llamada la MATRIZ JACOBIANA *de h:*

$$\frac{\partial h}{\partial x} := \frac{\partial(h_1, \ldots, h_m)}{\partial(x_1, \ldots, x_n)} := \begin{pmatrix} \frac{\partial h_1}{\partial x_1} & \cdots & \frac{\partial h_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial h_m}{\partial x_1} & \cdots & \frac{\partial h_m}{\partial x_n} \end{pmatrix}$$

siendo $x := (x_1, \ldots, x_n)$. Si $m = n$, entonces la matriz Jacobiana será cuadrada. Su determinante lo llamaremos DETERMINANTE JACOBIANO *o, simplemente,* JACOBIANO *y lo simbolizaremos por $J_h(x)$.*

Ejemplo 2.6.3 *El Jacobiano de la función $h : R^3 \to R^3$ definida por*

$$h(x_1, x_2, x_3) \;=\; (2x_1, x_2 - x_3, x_3^2)$$

es

$$J_h(x) \;=\; \det \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 2x_3 \end{pmatrix} \;=\; 4x_3$$

Esto indica que no necesariamente el jacobiano es un número. ◀

Teorema 2.6.4 *Sea $X = (X_1, \ldots, X_n)$ un vector aleatorio con densidad conjunta f_X. Sea $g : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ una aplicación inyectiva. Supóngase que tanto g como su inversa $h : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^n$ son continuas. Si las derivadas parciales de h existen y son continuas y si su jacobiano $J_h(y)$ es diferente de cero, entonces, el vector aleatorio $Y = g(X)$ tiene función de densidad conjunta f_Y dada por*

$$f_Y(y) = \begin{cases} |J_h(y)| \cdot f_X(h(y)) & \textit{si } y \textit{ está en el rango de } g, \\ 0, & \textit{de otro modo.} \end{cases}$$

DEMOSTRACIÓN:
Se deja al lector. ■

Ejemplo 2.6.5 *Sea $X = (X_1, X_2)$ un vector aleatorio con función de densidad de probabilidad conjunta dada por:*

$$f_X(x_1, x_2) = \begin{cases} 1 & \textit{si } 0 < x_1, x_2 < 1, \\ 0, & \textit{de otro modo.} \end{cases}$$

Hallar la función de densidad de probabilidad conjunta de $Y = (Y_1, Y_2)$, donde $Y_1 = X_1 + X_2$ y $Y_2 = X_1 - X_2$. ◀

Teorema 2.6.6 *Sea $X = (X_1, \ldots, X_n)$ un vector aleatorio con densidad conjunta f_X. Sea $g : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ una aplicación. Supóngase que $\mathbb{R}^n$ se puede particionar en k conjuntos disyuntos $A_1, \ldots, A_k$ de tal manera que la aplicación g restringida a A_i, para $i = 1, \ldots, k$, es una aplicación inyectiva con inversa h_i. Si las primeras derivadas parciales de h_i existen y son continuas y si los jacobianos $J_i(y) := J_{h_i}(y)$ son diferentes de cero en el rango de la transformación, para $i = 1, \ldots, k$, entonces, el vector aleatorio $Y = g(X)$ tiene función de densidad conjunta f_Y dada por*

$$f_Y(y) = \sum_{i=1}^{k} |J_i(y)| \, f_X(h_i(y))$$

DEMOSTRACIÓN:
Se deja al lector. ■

Ejemplo 2.6.7 *Sea X una variable aleatoria con función de densidad de probabilidad f_X. Utilize el teorema de transformación 2.6.6 para encontrar la función de densidad de la variable aleatoria $Y = X^4$.* ◀

✍ Ejercicios

2.1 La función de densidad conjunta de dos variables aleatorias continuas X y Y está dada por

$$f(x,y) \ = \begin{cases} 2\,e^{-x}\,e^{-2y}, & \text{si } x>0, \quad y>0; \\ 0, & \text{de otra manera.} \end{cases}$$

(a) Calcule $P(X>1, Y<1)$

(b) Halle $P(X<Y)$

(c) Halle la función de distribución acumulada marginal F_X de X.

2.2 Sean X, Y variables aleatorias con densidad conjunta f, dada por

$$f(x,y) = \begin{cases} 4, & 0 \le x \le 1,\ 0 \le y \le 1,\ 2y \le x \\ 0, & \text{en otros casos.} \end{cases}$$

(a) Calcule la densidad marginal f_X de X y la f_Y de Y.

(b) Determine si X y Y son independientes.

(c) Calcule $P\left(X \ge \frac{1}{2}\right)$.

(d) Halle $E(3X)$.

(e) Encuentre la densidad condicional de Y dado $X = x$.

2.3 Considere un círculo de radio R y suponga que un punto dentro del círculo es escogido aleatoriamente de tal manera que todas las regiones dentro del círculo de igual área contienen con la misma probabilidad al punto (en otras palabras, el punto está uniformemente distribuido dentro del círculo). Si el centro es el origen, si X y Y denotan las coordenadas del punto escogido y ya que (X,Y) es igualmente probable estar cerca a cada punto en el círculo, entonces la función de densidad conjunta de X y Y está dada por

$$f(x,y) \ = \begin{cases} k & \text{si } x^2+y^2 \le R^2, \\ 0 & \text{de otro modo,} \end{cases}$$

para algún valor de k.

(a) Determine k y haga un bosquejo de la gráfica de f.

(b) Encuentre la función de densidad marginal de X.

(c) Construya las función de densidad marginal de Y.

(d) Encuentre la probabilidad de que D, la distancia desde el origen al punto seleccionado, es menor o igual que t.

2.4 La función de densidad conjunta de dos variables aleatorias continuas X y Y está dada por

$$f(x,y) = \begin{cases} e^{-(x+y)} & \text{si } x > 0, \quad y > 0, \\ 0 & \text{de otro modo.} \end{cases}$$

Encuentre la función de densidad de la variable aleatoria $\frac{X}{Y}$.

2.5 Demuestre el teorema 2.2.6.

2.6 Sean X, Y variables aleatorias con densidad conjunta f, dada por

$$f(x,y) = \begin{cases} x+y, & 0 \le x \le 1,\ 0 \le y \le 1 \\ 0, & \text{en otros casos.} \end{cases}$$

(a) Calcule $P(X + Y \le 1)$.

(b) Halle $P\left(X < \frac{1}{4} / Y > \frac{1}{4}\right)$.

(c) Calcule la densidad marginal f_X de X y la f_Y de Y.

(d) Determine si X y Y son independientes.

(e) Halle la covarianza $Cov(X,Y)$.

2.7 Sea X una variable aleatoria discreta tal que

$$P(X=0) = P(X=1) = P(X=-1) = \frac{1}{3}$$

y defina

$$Y = \begin{cases} 0, & \text{si } X \neq 0, \\ 1, & \text{si } X = 0. \end{cases}$$

(a) Muestre que $Cov(X,Y) = 0$, pero que X y Y no son independientes.

(b) Halle la función de probabilidad conjunta de las variables aleatorias $U_1 = X + Y$ y $U_2 = X - Y$.

2.8 Sea $\Omega = \{a, b, c\}$, $\mathfrak{F}$ el conjunto potencia y $P(\omega) = 1/3$ para todo $\omega \in \Omega$. Considere las variables aleatorias X y Y definidas como sigue:

$$X(\omega) = \begin{cases} 1, & \text{si } \omega = a, b; \\ 0, & \text{si } \omega = c. \end{cases} \qquad Y(\omega) = \begin{cases} \pi, & \text{si } \omega = a; \\ 1/2, & \text{si } \omega = b. \\ -1, & \text{si } \omega = c. \end{cases}$$

Halle: (a) $E(X/Y = \pi)$, (b) $E(X/Y = 1/2)$, (c) $E(X/Y = -1)$, (d) $E(X/Y)$, (e) $E(E(X/Y))$.

2.9 Demuestre las expresiones (2.4) y (2.5).

2.10 Demuestre la expresión (2.7).

2.11 Supongamos que $X_1 \sim \mathcal{P}(\lambda_1)$ y $X_2 \sim \mathcal{P}(\lambda_2)$ son variables de Poisson independientes.

(a) Demuestre que $X_1 + X_2 \sim \mathcal{P}(\lambda_1 + \lambda_2)$.

(b) Calcule las probabilidades condicionales $P(X_1 = x/X_2 = k)$ e interprete el resultado en términos de una distribución conocida.

(c) Calcule las probabilidades condicionales $P(X_2 = x/X_1 = k)$.

2.12 Sean X y Y variables aleatorias independientes con distribución de Poisson de parámetros λ_1 y λ_2, respectivamente. Sea n un entero no negativo fijo.

(a) Demuestre que X tiene, bajo la condición de que $X + Y = n$, tiene una distribución binomial con parámetros n y $p := \frac{\lambda_1}{\lambda_1+\lambda_2}$.

(b) Calcule la esperanza de X bajo la condición de que $X + Y = n$.

2.13 Demuestre los teoremas (a) 2.5.6, (b) 2.5.8, (c) 2.5.9, (d) 2.5.10, (e) 2.5.11.

2.14 Demuestre los siguientes resultados:

(a) Si X_1 y X_2 variables aleatorias independientes. Si $X_i \overset{\mathrm{d}}{=} \mathcal{N}(\mu_i, \sigma_i^2)$ para $i = 1, 2$, entonces $X_1 - X_2 \overset{\mathrm{d}}{=} \mathcal{N}(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$.

(b) Sean X_i, Y_i, $S^2_{(n)}$ y $S^2_{(m)}$ como en el teorema 2.5.11. Si $X_i \overset{\mathrm{d}}{=} \mathcal{N}(\mu_1, \sigma^2)$ para cada $i = 1, \ldots, n$ y $Y_j \overset{\mathrm{d}}{=} \mathcal{N}(\mu_2, \sigma^2)$ para cada $j = 1, \ldots, m$, entonces, $F := \frac{S^2_{(n)}}{S^2_{(m)}} \overset{\mathrm{d}}{=} \mathcal{F}(n-1, m-1)$.

2.15 Bajo un Experimento de Bernoulli se entiende un experimento estocástico con sólo dos salidas “éxito” (E) o “fracaso” (F). Un éxito sucede con probabilidad p. Este experimento de Bernoulli se ejecuta varias veces. Para $n \in \mathbb{N}$, sea K_n el número de éxitos entre los primeros n intentos. Sean T_1 el número de fracasos hasta el primer éxito y T_2 el total de fracasos hasta el segundo éxito.

(a) Calcule las distribuciones de K_n, T_1, T_2, $X := T_2 - T_1$, así como, $E(T_1)$.

(b) ¿Son X y T_1 independientes? (Argumentos)

(c) Demuestre que $P(T_1 = k/T_2 = n) = \frac{1}{n+1}$ para todo $0 \leq k \leq n$.

2.16 Sea $\mathcal{G}(p)$ la distribución geométrica con el parámetro p, con $0 \leq p \leq 1$. Supongamos que $X_1, X_2, \ldots$ es una sucesión de variables aleatorias, independientes y distribuidas según $\mathcal{G}(p)$. Es decir, $P(X_i = x) = p\,q^x$ para $x = 0, 1, 2, \ldots$, siendo $q = 1 - p$. Para cada n, defínase $S_n = X_1 + \cdots + X_n$.

(a) Calcule $P(S_2 = k)$ y con ello demuestre que S_2 tiene distribución binomial negativa con los parámetros 2 y p.

(b) Calcule $P(S_3 = k)$ y con ello demuestre que S_3 tiene distribución binomial negativa con los parámetros 3 y p.

(c) Demuestre por inducción sobre n que $P(S_n = k) = \binom{k+n-1}{n-1} p^n\, q^k$ y con ello demuestre que S_n tiene distribución binomial negativa con los parámetros n y p.

2.17 Demuestre: Si $(X, Y) \stackrel{\mathrm{d}}{=} \mathcal{N}(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \varrho)$, entonces,

$$X + Y \stackrel{\mathrm{d}}{=} \mathcal{N}(\mu_1 + \mu_2,\ \sigma_1^2 + \sigma_2^2 + 2\varrho\sigma_1\sigma_2)$$

2.18 Supongamos que el nivel de agua de cierto río puede modelarse por medio de una variable aleatoria X con función de distribución acumulada $F(t) = 1 - \frac{1}{x^2}$, para $1 \leq t < \infty$, interpretando $t = 1$ como el nivel más bajo posible.

(a) Calcule la densidad, esperanza (si existe) y la varianza (si existe) de X.

(b) Supongamos que se hacen n observaciones independientes de X, lo que conduce a $X_1, \ldots, X_n$ y que es de interés la variable $Y := \min_i X_i$. Calcule la función de distribución acumulada F_Y, la densidad f_Y y la esperanza (si existe) de Y.

2.19 Sean dadas dos variables aleatorias X y Y continuas con función de densidad conjunta f definida por

$$f(x, y) \;=\; \begin{cases} \frac{x(1+3y^2)}{4}, & \text{si } 0 < x < 2 \text{ y } 0 < y < 1; \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Encuentre $f_X(x)$, $f_Y(y)$ y determine si X y Y son independientes.

2.20 Sean X_1 y X_2 variables aleatorias con la distribución de probabilidad conjunta que se da en la tabla 2.6.

(a) Determine si X_1 y X_2 son independientes.

(b) Encuentre la función de distribución conjunta de las variables aleatorias $Y_1 = X_1 + X_2$ y $Y_2 = X_1 X_2$.

2.21 Suponga que X y Y son independientes y uniformemente distribuidas sobre el intervalo $(0,1)$. Calcule: (a) $P(X^2 + Y^2 \leq 1)$, (b) $P(Y \leq X^2)$.

2.22 Un hombre y una mujer deciden encontrarse en cierto lugar. Si cada persona llega independientemente en un tiempo uniformemente distribuido entre las 12:00 p.m. y 1:00 p.m., encontrar la probabilidad de que el primero en llegar tenga que esperar más de 10 minutos.

2.23 Supóngase que la duración X de un dispositivo electrónico es una variable aleatoria continua con función de densidad de probabilidad dada por:

$$f(x) = \begin{cases} 1000/x^2 & \text{si } x > 1000, \\ 0 & \text{de otro modo.} \end{cases}$$

Sean X_1 y X_2 dos determinaciones independientes de la anterior variable. Hallar la función de densidad de probabilidad de la variable $Z = \frac{X_1}{X_2}$.

2.24 Sean X e Y variables aleatorias. Sean $Z := X + Y$ y $W := X - Y$.

(a) Hallar la función de densidad conjunta de Z y W.

(b) Deducir, luego, una función de densidad de Z y una de W.

(c) Si las variables X y Y son independientes, ¿cómo está dada la función de densidad para la variable Z?

2.25 Sean X e Y variables aleatorias independientes con funciones de densidad dadas por:

$$f_X(x) = \begin{cases} 1 & \text{si } 1 < x < 2, \\ 0 & \text{de otro modo} \end{cases}, \qquad f_Y(y) = \begin{cases} 1/2 & \text{si } 3 < y < 5, \\ 0 & \text{de otro modo} \end{cases}$$

Determine la función de densidad de $Z = X + Y$.

2.26 Sean X e Y variables aleatorias con función de densidad conjunta dada por:

$$f(x,y) = \begin{cases} 3x & \text{si } 0 \leq y \leq x \leq 1, \\ 0 & \text{de otro modo} \end{cases}$$

Determine la función de densidad de $W = X - Y$.

2.27 Sean X e Y variables aleatorias. Sean $Z := XY$ y $W := Y$.

(a) Hallar la función de densidad conjunta de Z y W.

(b) Deducir, luego, una función de densidad de Z.

2.28 Sean X e Y variables aleatorias. Sean $Z := X/Y$ (la cual está definida si $P(Y = 0) = 0$) y $W := Y$.

(a) Hallar la función de densidad conjunta de Z y W.

(b) Deducir, luego, una función de densidad de Z.

2.29 Sean X e Y variables aleatorias independientes y distribuidas uniformemente sobre el intervalo $(0, 1)$. Halle la función de densidad de (a) $Z = XY$ (b) $Z = X/Y$.

2.30 Sea Y el número de caras en cuatro lanzamientos de una moneda no falsa. Calcule la esperanza condicional de Y sabiendo que han resultado dos o menos caras. Interprete el resultado obtenido.

2.31 Si X_1, X_2 y X_3 son variables aleatoria continuas con función de densidad conjunta

$$f(x_1, x_2, x_3) = \begin{cases} x_1^2\, e^{-x_1(1+x_2+x_3)}, & \text{si } x_1 > 0,\ x_2 > 0 \text{ y } x_3 > 0, \\ 0, & \text{en otro caso,} \end{cases}$$

determine:

(a) La función de densidad marginal conjunta de X_2 y X_3.

(b) $P(X_2 + X_3 < 4)$.

2.32 Suponga que las variables aleatorias X, Y y Z tienen la función de densidad conjunta f, definida por $f(x, y, z) = kxyz$ si $x, y, z \in (0, 1)$ y $f(x, y, z) = 0$, de otro modo. Halle el valor de k.

2.33 Suponga que en cinco puntos del océano medimos la intensidad del sonido causado por ciertos ruidos. Sean X_1, X_2, X_3, X_4 y X_5 variables aleatorias que representan la intensidad del sonido en cada uno de los cinco puntos. Supongamos que la función de densidad conjunta de estas cinco variables viene dada por

$$f(x_1, x_2, \ldots, x_5) = \begin{cases} x_1 x_2 \cdots x_5\, e^{-(x_1^2+x_2^2+\cdots+x_5^2)/2}, & \text{si } x_1, x_2, \ldots, x_5 \geq 0; \\ 0, & \text{de otro modo.} \end{cases}$$

Definamos Y como la intensidad máxima, es decir, $Y = \max\{X_1, X_2, \ldots, X_5\}$. Determine la probabilidad de que Y sea menor o igual que $t \in \mathbb{R}$.

2.34 Si la función de probabilidad conjunta de tres variables aleatorias, X, Y y Z, cada una con posibles valores 1, 2, 3 y 4, es:

$$f(x,y,z) = \begin{cases} \frac{1}{64}, & \text{si } x,y,z \in \{1,2,3,4\}; \\ 0, & \text{de otro modo} \end{cases}$$

(a) Determine $P(X = 2, Y = 3, Z \leq 2)$.

(b) ¿Son X, Y y Z independientes?

2.35 La función de probabilidad conjunta de dos variables aleatorias discretas X y Y está dada por

$$f(x,y) = \begin{cases} cxy, & \text{si } x,y \in \{1,2,3,4\}; \\ 0, & \text{de otro modo} \end{cases}$$

(a) Halle la constante c.

(b) Encuentre las funciones de probabilidad marginal de X y de Y.

(c) Determine si X y Y son independientes.

(d) Construya las funciones de distribución acumulada marginal de X y de Y.

(e) Halle la función de probabilidad condicional de X dada $Y = k$.

(f) Halle la función de probabilidad condicional de Y dada $X = k$.

(g) Encuentre la función de distribución conjunta de las variables $U = X + Y$ y $V = X - Y$.

2.36 Sean X y Y variables aleatorias continuas que tienen función de densidad conjunta

$$f(x,y) = \begin{cases} c(x^2 + y^2), & \text{si } x,y \in [0,1]; \\ 0, & \text{de otro modo} \end{cases}$$

(a) Halle la constante c.

(b) Halle las funciones de probabilidad marginal de X y de Y.

(c) Calcule $P(Y < 1/2)$.

(d) Determine si X y Y son independientes.

(e) Halle las funciones de distribución acumulada marginal de X y de Y.

(f) Halle la función de probabilidad condicional de X dada $Y = s$ y la de Y dada $X = t$.

2.37 Si X y Y son variables aleatorias independientes con función de densidad conjunta

$$f(x, y) \;=\; \begin{cases} e^{-(x+y)}, & \text{si } x \geq 0,\ y \geq 0; \\ 0, & \text{de otro modo} \end{cases}$$

y $U = \frac{X}{Y}$, $V = X + Y$, halle:

(a) La función de densidad de U.

(b) La función de densidad de V.

(c) La función de densidad conjunta de U y V.

2.38 Una urna contiene tres fichas blancas y dos azules. Sea selecciona una muestra aleatoria de tamaño 2 (con reemplazo). Sean

$$X \;=\; \begin{cases} 1, & \text{si la primera ficha seleccionada es blanca.} \\ 0, & \text{si la primera ficha seleccionada es azul.} \end{cases}$$

$$Y \;=\; \begin{cases} 1, & \text{si la segunda ficha seleccionada es blanca.} \\ 0, & \text{si la segunda ficha seleccionada es azul.} \end{cases}$$

(a) Halle la distribución de probabilidad conjunta de las variables X y Y.

(b) Calcule la esperanza de X bajo la condición de que $Y = y$.

2.39 Si X y Y tienen función de densidad conjunta

$$f(x, y) \;=\; \begin{cases} xy/96, & \text{si } 0 < x < 4,\ 1 < y < 5; \\ 0, & \text{de otro modo} \end{cases}$$

y $U = XY^2$, $V = X^2Y$, halle la función de densidad conjunta de U y V.

2.40 Sean X y Y variables aleatorias que tienen función de densidad conjunta

$$f(x, y) = \frac{1}{2\pi}\, e^{-(x^2+y^2)}, \quad \text{si } x, y \in \mathbb{R}$$

Si R y θ son nuevas variables aleatorias tales que $x = R\cos\theta$ e $y = R\sin\theta$, demostrar que la función de densidad de R es

$$f_R(r) \;=\; \begin{cases} r\,e^{-r^2}, & \text{si } r \geq 0 \\ 0, & \text{de otro modo} \end{cases}$$

Sugerencia: Halle primero la densidad conjunta de R y θ.

CAPÍTULO **3**

Algunos teoremas de convergencia

Contenido

3.1 Convergencia de sucesiones de variables aleatorias

Sean $X, X_1, X_2, \ldots$ variables aleatorias reales sobre un mismo espacio de probabilidad. En esta sección estudiaremos dos tipos de convergencia. Sin embargo, es importante resaltar que hay otros tipos de convergencias diferentes a los tratado acá.

Definición 3.1.1 (Convergencia en probabilidad) *Sean $X, X_1, X_2, \ldots$ variables aleatorias reales sobre un espacio de probabilidad $(\Omega, \mathfrak{F}, P)$. La sucesión*

*$(X_n)_{n\in\mathbb{N}}$ se llama (P)-*ESTOCÁSTICAMENTE CONVERGENTE *o que* CONVERGE EN PROBABILIDAD *hacia una variable aleatoria real X sobre $(\Omega, \mathfrak{F}, P)$, si*

$$\lim_{n\to\infty} P(|X_n - X| \geq \varepsilon) = 0$$

para todo $\varepsilon > 0$, y escribiremos $X_n \xrightarrow[n\,\to\,\infty]{P} X$.

Ejemplo 3.1.2 *Sean $X_1, X_2, \dots$ variables aleatorias reales tales que $P(X_n = 0) = 1 - \frac{1}{n}$ y $P(X_n = n) = \frac{1}{n}$ para $n = 1, 2, \dots$ Sea $\varepsilon > 0$. Entonces,*

$$P(|X_n| > \varepsilon) = \begin{cases} \frac{1}{n}, & \text{si } n > \varepsilon; \\ 0 & \text{si } n \leq \varepsilon. \end{cases}$$

Por lo tanto, $X_n \xrightarrow[n\,\to\,\infty]{P} 0$ ◀

Ejemplo 3.1.3 *Sean $X_1, X_2, \dots$ variables aleatorias reales sobre un espacio de probabilidad $(\Omega, \mathfrak{F}, P)$. Supongamos que X_n tiene distribución de Bernoulli con parámetro $\left(\frac{1}{2}\right)^n$. Ya que $P(X_n = 1) = \left(\frac{1}{2}\right)^n$, se cumple que $X_n \xrightarrow[n\,\to\,\infty]{P} 0$. Observe que como*

$$V(X_n) = \left(\frac{1}{2}\right)^n \left[1 - \left(\frac{1}{2}\right)^n\right]$$

puede notarse el decrecimiento de la varianza a medida que n se incrementa. Es decir, que X_n va perdeindo el carácter de variable aleatoria porque su varianza va tendiendo a cero, la variable va asumiendo rasgos de una constante. ◀

Finalmente, presentamos el otro tipo de convergencia, conocida como *convergencia en distribución*.

Definición 3.1.4 (Covergencia en distribución) *Sea $X, X_1, X_2, \dots$ variables aleatorias reales sobre un espacio de probabilidad $(\Omega, \mathfrak{F}, P)$ con funciones de distribución $F, F_1, F_2, \dots$ respectivamente. Se dice que $(X_n)_n$* CONVERGE EN DISTRIBUCIÓN *a X, y se escribe $X_n \xrightarrow[n\,\to\,\infty]{d} X$ si*

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$$

para todo x punto de continuidad de F.

Observe que la convergencia en distribución, al igual que en los otros casos de convergencia considerados anteriormente hace referencia a la convergencia de una sucesión de números reales y no a la convergencia de una sucesión de eventos.

Ejemplo 3.1.5 *Sean $X, X_1, X_2, \ldots$ variables aleatorias reales sobre un espacio de probabilidad $(\Omega, \mathfrak{F}, P)$ definidas por $X_n(\omega) = \frac{1}{n}$ y $X = 0$. Como*

$$F_n(t) = \begin{cases} 1, & \text{si } \frac{1}{n} \leq t; \\ 0 & \text{en otro caso,} \end{cases} \qquad y \qquad F(t) = \begin{cases} 1, & \text{si } t \geq 0; \\ 0 & \text{en otro caso,} \end{cases}$$

Entonces, $X_n \xrightarrow[n \to \infty]{d} 0$. ◀

Ejemplo 3.1.6 *Sean $X_1, X_2, \ldots, X_n$ variables aleatorias continuas, independientes e identicamente distribuidas con distribución uniforme en $(0, \theta)$. Sea $Y_{(n)} := \max_i X_i$. La función de distribución acumulada de $Y_{(n)}$ viene dada por*

$$F_{Y_{(n)}}(t) = \begin{cases} 0, & t \leq 0 \\ \left(\frac{t}{\theta}\right)^n, & 0 < t \leq \theta \\ 1, & t > \theta \end{cases}$$

Ahora consideremos la variable aleatoria $U_n = n(\theta - Y_{(n)})$. Para todo $t \in (0, n\theta)$, la función de distribución acumulada de U_n es

$$F_{U_n}(t) \; = \; 1 - \left(1 - \frac{t/\theta}{n}\right)^n \quad \xrightarrow[n \to \infty]{} \quad 1 - e^{-t/\theta}$$

Por consiguiente, $U_n \xrightarrow[n \to \infty]{d} U$, donde U es una variable aleatoria que tiene distribución exponencial con parámetro $1/\theta$. ◀

Existen relaciones y propiedades que tienen que ver con los diferentes tipos de convergencia, como se muestra en el siguiente teorema.

Teorema 3.1.7 *Sean $X, X_1, X_2, \ldots$ y $Y_1, Y_2, \ldots$ variables aleatorias reales sobre un espacio de probabilidad $(\Omega, \mathfrak{F}, P)$.*

(a) Si $X_n \xrightarrow[n \to \infty]{P} X$, entonces $X_n \xrightarrow[n \to \infty]{d} X$. El recíproco no es cierto (véase ejercicio 3.7).

(b) Si $X_n \xrightarrow[n\to\infty]{P} c$, *entonces* $X_n \xrightarrow[n\to\infty]{d} c$, *siendo c una constante. En este caso, el recíproco también es cierto.*

(c) Si $Y_n \xrightarrow[n\to\infty]{d} X$ *y* X_n *es* ASINTOTICAMENTE EQUIVALENTE[1] *a* Y_n *, entonces* $X_n \xrightarrow[n\to\infty]{d} X$.

(d) Si $X_n \xrightarrow[n\to\infty]{P} X$ *y g es una función continua, entonces* $g(X_n) \xrightarrow[n\to\infty]{P} g(X)$. *El resultado es análogo para la convergencia en distribución.*

DEMOSTRACIÓN:
Se deja al lector. ■

En el siguiente teorema presentamos relaciones entre las convergencias y algunas operaciones aritméticas.

Teorema 3.1.8 *Sean* $X, X_1, X_2, \ldots$ *y* $Y_1, Y_2, \ldots$ *variables aleatorias reales sobre un espacio de probabilidad* $(\Omega, \mathfrak{F}, P)$.

(a) Si $X_n \xrightarrow[n\to\infty]{d} X$, *entonces* $aX_n + b \xrightarrow[n\to\infty]{d} aX + b$.

(b) Si $X_n \xrightarrow[n\to\infty]{d} X$ *y* $Y_n \xrightarrow[n\to\infty]{P} 0$, *entonces* $X_nY_n \xrightarrow[n\to\infty]{P} 0$.

(c) Si $X_n \xrightarrow[n\to\infty]{d} X$ *y* $Y_n \xrightarrow[n\to\infty]{d} c$, *entonces* $X_n + Y_n \xrightarrow[n\to\infty]{d} X + c$.

(d) Si $X_n \xrightarrow[n\to\infty]{d} X$ *y* $Y_n \xrightarrow[n\to\infty]{P} c$ *con* $c \neq 0$, *entonces* $X_nY_n \xrightarrow[n\to\infty]{d} cX$.

(e) Si $X_n \xrightarrow[n\to\infty]{d} X$ *y* $Y_n \xrightarrow[n\to\infty]{P} c$ *con* $c \neq 0$, *entonces* $\frac{X_n}{Y_n} \xrightarrow[n\to\infty]{d} \frac{X}{c}$.

DEMOSTRACIÓN:
Se deja al lector. ■

---

[1]Es decir, si $X_n - Y_n \xrightarrow[n\to\infty]{P} 0$ y escribiremos también $X_n \underset{n\to\infty}{\overset{a}{=}} Y_n$.

3.2 Ley de los grandes números

Cuando se conoce la distribución de una variable aleatoria X, se puede encontrar su esperanza y varianza. Sin embargo, cuando se conocen estos dos valores, no se puede calcular probabilidades del tipo $P(|X-c|>\varepsilon)$, pero se puede encontrar una cota con ayuda de la llamada *desigualdad de Chevischev*, inclusive cuando la distribución de X sea desconocida.

Teorema 3.2.1 (Desigualdad de Chevischev) *Sea X una variable aleatoria real sobre un espacio de probabilidad $(\Omega, \mathfrak{F}, P)$ con $E(X)=\mu$ y $V(X)=\sigma^2$. Entonces para cada $\varepsilon>0$, se cumple*

$$P(|X-\mu| \geq \varepsilon) \ \leq \ \frac{\sigma^2}{\varepsilon^2} \tag{3.1}$$

DEMOSTRACIÓN:
Sea A un evento de Ω. Considere la variable aleatoria I_A definida por

$$I_A(\omega) := \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{de otro modo} \end{cases}$$

Teniendo en cuenta que $1_A \geq 0$ y que $E(1_A)=P(A)$, encontramos que

$$\sigma^2 = E\left((X-\mu)^2\right) \ \geq E\left((X-\mu)^2\, 1_{\{|X-\mu|>\varepsilon\}}\right) \geq \ \varepsilon^2\, P(|X-\mu|>\varepsilon)$$

Con ello queda demostrado el teorema. ■

Ejemplo 3.2.2 *Una empresa considera dos inversiones posibles. Como aproximación inicial asigna probabilidades (subjetivas) a cada uno de los siguientes eventos: perder un 20% por cada dólar invertido, perder un 10%, ni ganar ni perder, ganar un 10% y ganar un 20%. Sea X el rendimiento por cada dólar invertido en el primer proyecto y Y el rendimiento por cada dólar invertido en el segundo. Las probabilidades asignadas son:*

x	-0,20	-0,10	0	+0,10	+0,20
$P(X=x)$	0,1	0,2	0,4	0,2	0,1

y	-0,20	-0,10	0	+0,10	+0,20
$P(Y=y)$	0,01	0,04	0,10	0,50	0,35

(a) *Calcule los rendimientos esperados por cada dólar invertido en cada proyecto. ¿Cuál proyecto le parece a usted que representa la inversión más atractiva?*

(b) *Halle* $P(|X-\mu| \leq \sigma)$ *y* $P(|X-\mu| \leq 2\sigma)$.

(c) *Estime las probabilidades anteriores utilizando la desigualdad de Chevischev.*

SOLUCIÓN:

(a) $E(X) = 0$ *y* $E(Y) = 0,114$.

(b) *Ya que* $\sigma_X = 0,110$, *tenemos que* $P(|X-\mu| \leq \sigma) = 0,80$ *y* $P(|X-\mu| \leq 2\sigma) = 1$.

(c) *La regla de Tchebychev indica que estas probabilidades deben ser al menos* 0 *y* $0,75$ *respectivamente.* ◀

Ejemplo 3.2.3 *Una variable aleatoria* X *continua tiene una media* $\mu = 8$, *varianza* $\sigma^2 = 9$ *y distribución de probabilidad desconocida. Entonces,*

$$P(-4 < X < 20) \;=\; P(|X-8| \leq (4)(3)) \;\geq\; \frac{15}{16} = 0,9375 \qquad ◀$$

Como una aplicación de la desigualdad de Chevischev se obtiene la *ley débil de los grandes números.*

Definición 3.2.4 *Una sucesión* $(X_n)_{n\in\mathbb{N}}$ *de variables aleatorias numéricas sobre un espacio de probabilidad* $(\Omega, \mathfrak{F}, P)$, *se llama* IDÉNTICAMENTE DISTRIBUIDA, *si* $P_{X_1} = P_{X_2} = \cdots$, *es decir, si todas las variables aleatorias* X_n *tienen la misma distribución.*

Teorema 3.2.5 (ley débil de los grandes números) *Sean* $X_1, X_2, \ldots$ *variables aleatorias independientes y reales sobre un espacio de probabilidad* $(\Omega, \mathfrak{F}, P)$ *con* $\mu_i := E(X_i)$ *y* $\sigma_i^2 := Var(X_i) < \infty$, *para todo* $i \in \mathbb{N}$. *Si la condición* $\lim\limits_{n\to\infty} \frac{1}{n^2} \sum\limits_{i=1}^{n} \sigma_i^2 = 0$ *(*CONDICIÓN DE MARKOV*) se cumple, entonces*

$$\lim_{n\to\infty} P(|\overline{X}_{(n)} - \overline{\mu}_{(n)}| \geq \varepsilon) = 0$$

para todo $\varepsilon > 0$, *siendo* $\overline{\mu}_{(n)} = \frac{1}{n}\sum\limits_{i=1}^{n} \mu_i$. *Es decir,* $\overline{X}_{(n)} - \overline{\mu}_{(n)} \xrightarrow[n\to\infty]{P} 0$. *O sea, la sucesión* $(\overline{X}_{(n)} - \overline{\mu}_{(n)})_{n\in\mathbb{N}}$ *converge estocásticamente (en probabilidad) hacia cero.*

Es un error escribir $\overline{X}_{(n)} \xrightarrow[n\to\infty]{P} \overline{\mu}_{(n)}$ *porque* $\overline{\mu}_{(n)}$ *también depende de* n.

DEMOSTRACIÓN:
Verifique que $E(\overline{X}_{(n)}) = \overline{\mu}_{(n)}$ y $V(\overline{X}_{(n)}) = \frac{1}{n^2}\sum_{i=1}^{n}\sigma_i^2$. Luego, aplique la desigualdad de Chevischev y téngase en cuenta la condición de Markov. ■

Corolario 3.2.6 (ley débil de los grandes números, particular) *Sean $X_1, X_2, \ldots$ variables aleatorias reales, independientes e identicamente distribuidas sobre un espacio de probabilidad $(\Omega, \mathfrak{F}, P)$ con $\mu := E(X_n) < \infty$ y $\sigma^2 := Var(X_n) < \infty$. Entonces, $\overline{X}_{(n)} \underset{n \to \infty}{\overset{P}{\longrightarrow}} \mu$.*

La varianza σ^2 no necesariamente debe existir, pero como caso particular del teorema 3.2.5, sí debe existir. Es más, existe una demostración de la ley débil de los grandes números (debido al matemático ruso Alexander Khinchin) que no exige el supuesto de varianza finita.

DEMOSTRACIÓN:
Aplique la desigualdad de Chevischev. ■

Corolario 3.2.7 (ley débil de los grandes números) *Sean $X_1, X_2, \ldots$ variables aleatorias reales, independientes e identicamente distribuidas sobre un espacio de probabilidad $(\Omega, \mathfrak{F}, P)$ con $\mu := E(X_n) < \infty$ y $\sigma^2 := Var(X_n) < \infty$. Entonces, $\overline{X}_{(n)} \underset{n \to \infty}{\overset{P}{\longrightarrow}} \mu$.*

La varianza σ^2 no necesariamente debe existir, pero como caso particular del teorema 3.2.5, sí debe existir. Es más, existe una demostración de la ley débil de los grandes números (debido al matemático ruso Alexander Khinchin) que no exige el supuesto de varianza finita.

DESMOSTRACIÓN:
Se deja al lector. ■

A continuación presentamos el conocido *teorema* o *ley de los grandes números* de Jakob Bernoulli. Véase applet en [5].

Teorema 3.2.8 (Ley de los grandes números de J. Bernoulli) *Sea X_1, X_2, … variables aleatorias reales, independientes e identicamente distribuidas sobre un espacio de probabilidad $(\Omega, \mathfrak{F}, P)$ con $X_n \sim \mathcal{B}(1, p)$. Entonces, se cumple que $\overline{X}_{(n)} \underset{n \to \infty}{\overset{P}{\longrightarrow}} p$.*

DEMOSTRACIÓN:
Aplique la desigualdad de Chevischev. ■

Observación:

Esta ley de los grandes números se puede interpretar así: Si el número de ensayos n es suficientemente grande, entonces, la probabilidad para que la frecuencia relativa de los eventos se diferencie muy poco con p, es casi 1.

Para la introducción del concepto de la probabilidad hemos atribuido a tales eventos una probabilidad, cuyas frecuencias relativas muestran una determinada estabilidad durante una serie de ensayos grande. Esta convergencia estocástica es un análogo teorético-probabilístico a la estabilidad observada de las frecuencias relativas. Con esto, no está demostrada, de ninguna manera, esta estabilidad, sino que se demuestra símplemente que la teoría de la probabilidad produce resultados, que armonizan con sus bases. Es importante recalcar que la teoría posibilita una descripión precisa de la estabilidad de la frecuencia relativa. Al mismo tiempo, podemos entender el porqué, para la definición de la probabilidad, nosotros no pudimos definir con precisión la estabilidad de las frecuencias relativas. Asimismo, en esta descripción, aparece el concepto de la probabilidad; esta ley de los grandes números sólo dice que grandes desviaciones entre la frecuencia relativa y la probabilidad son *muy improbables* en un serie de intentos muy grande.

Por tanto, aquí se trata de un aparente "circulus vitiosus": Nosotros hemos definido la probabilidad con ayuda de la frecuencia relativa, por otro lado, en la caracterización de la estabilidad de la frecuencia relativa aparece también el concepto de la probabilidad. Si embargo, en realidad, se trata de dos cosas completamente diferentes. La "definición" de la probabilidad como el valor numérico, al rededor del cual oscila la frecuencia relativa, no es una definición matemática, sino una descripción del fondo real del concepto de probabilidad. Sin embargo, el teorema de Bernoulli de los grandes números es derivado en virtud de la *definición matemática* de la probabilidad y, con esto, en realidad no hay circulus vitiosus.

El teorema que acabamos de conocer, es sólo un caso especial de leyes más generales, que se han presentado anteriormente. Del mismo modo, la aproximación de la distribución binomial a través de una dsitribución normal es, también, sólo un caso especial de teoremas de distribuciones límites del cálculo de probabilidades. Esto se tratará más adelante.

3.3 Teorema central del límite

En general, en estadística, Los teoremas de distribuciones límites más importantes son los llamados *teoremas centrales de distribuciones límites*, los cuales expresan el hecho de que la suma de un número grande de variables aleatorias independientes está aproximadamente normalmente distribuida bajo condiciones generales. Estos teoremas desvelan los motivos para que, en las diferentes ramas aplicadas, se encuentren distribuciones normales o casi normales. Un típico ejemplo para ello son las inexactitudes en las medidas; el error de medida total se compone de diferentes errores pequeños. Por tanto, a través de los teoremas centrales del límite se justifica el supuesto de que los errores de medidas están distribuidos normalmente. Por consiguiente, la distribución normal se conoce también con el nombre de *ley de los errores*. El caso más simple del teorema central del límite es el llamado *teorema de Moivre-Laplace*, el cual presentaremos como un caso especial de un teorema más general, *el teorema central del límite de Lindeberg-Lévy*, en el cual se considera el caso de variables aleatorias reales independientes e identicamente distribuidas. Omitimos su demostración.

Teorema 3.3.1 (Teorema central del límite de Lindeberg-Lévy) *Si X_1, X_2, ... son variables aleatorias reales, independientes e identicamente distribuidas con $E(X_n^2) < \infty$. Sean $\mu := E(X_n)$ y $\sigma^2 := Var(X_n)$. Si $\sigma > 0$, entonces*

$$Z_n := \frac{\overline{X}_{(n)} - \mu}{\sigma/\sqrt{n}} \xrightarrow[n\to\infty]{d} \mathcal{N}(0,1).$$

DEMOSTRACIÓN:
Se deja al lector. ■

Ahora, presentaremos el famoso teorema de Moivre-Laplace, el cual es un caso particular del teorema central del límite.

Corolario 3.3.2 (Teorema central del límite de Moivre-Laplace) *Si X_1, X_2, ... son variables aleatorias reales, independientes, con distribución $\mathcal{B}(n,p)$,* *entonces,*

$$Z_n := \frac{\overline{X}_{(n)} - p}{\sqrt{p(1-p)/n}} \xrightarrow[n\to\infty]{d} \mathcal{N}(0,1).$$

DEMOSTRACIÓN:
Es un caso especial del teorema 3.3.1. ■

Ejemplo 3.3.3 *Una moneda se tira 100 veces y supongamos que la probabilidad de que aparezca cara resp. sello es* $\frac{1}{2}$. *Calcularemos la probabilidad para que aparezca cara entre 50 y 60 veces. Sea* X_n *el evento cara, el cual puede tomar los eventos* $0, 1, \ldots, 100$. *Debido a que* $E(X_n) = 50$ *y* $Var(X_n) = 25$, *entonces, por el teorema de Moivre-Laplace (corolario* 3.3.2*), se tiene que*

$$\begin{aligned} P(50 < X_n < 60) &= P\big(\tfrac{50-50}{5} < \tfrac{X_n-50}{5} < \tfrac{60-50}{5}\big) \\ &= P\big(0 < \tfrac{X_n-50}{5} < 2\big) \ = \ P(0 < Z_n < 2) \end{aligned}$$

De una tabla para la distribución normal obtenemos que el valor aproximado de esta proabibilidad es 0.4315. ◀

Para la práctica, el teorema de Moivre-Laplace se formula generalmente de la siguiente manera.

Teorema 3.3.4 (Aproximación de la binomial a la normal) *Sea* X *cualquier variable aleatoria que tiene distribución binomial con parámetros* n *y* p.

(a) Si $n \geq 30$, *entonces la distribución binomial se puede aproximar a la distribución normal con* $\mu = np$ *y* $\sigma^2 = np(1-p)$.

(b) Si $np \geq 5$ *y* $n(1-p) \geq 5$, *entonces también la distribución binomial se puede aproximar a la distribución normal con* $\mu = np$ *y* $\sigma^2 = np(1-p)$.

En cualquiera de los dos casos, se cumple que

$$P(X \leq k) \ = \ B(k;n;p) \ \approx \ \Phi\left(\frac{k + 0,5 - np}{\sqrt{np(1-p)}}\right)$$

DEMOSTRACIÓN:
Se deja al lector. ■

Ejemplo 3.3.5 *Un fabricante sabe por experiencia que de 17.000 productos, el 4% es rechazado por defectos. Si un nuevo lote de 800 unidades va a ser inspeccionado, ¿cuál es la probabilidad aproximada de que menos de 35 unidades sean rechazadas?*

SOLUCIÓN:
Sea X la variable aleatoria que representa el "número de productos rechazados". Entonces, X es una variable binomial con parámetros $n = 800$ y $p = 0,04$. Como $n \geq 30$, entonces[2] la distribución binomial se puede aproximar por la distribución normal con $\mu = np = 32$ y $\sigma^2 = np(1-p) = 30,72$. Por consiguiente, encontramos que a probabilidad aproximada de que menos de 35 unidades sean rechazadas es:

$$\begin{aligned} P(X < 35) &= P(X \leq 34) \;=\; B(34; 800; 0,04) \\ &\approx \Phi\left(\frac{34 + 0,5 - 32}{\sqrt{30,72}}\right) \;=\; \Phi(0,45) \;=\; 0,6736 \end{aligned}$$ ◀

✍ Ejercicios

3.1 Se lanza una moneda corriente una vez. Para $n = 1, 2, \dots$ se define la variable aleatoria X_n como sigue:

$$X_n := \begin{cases} 1, & \text{si el resultado del lanzamiento es sello;} \\ 0, & \text{si el resultado del lanzamiento es cara.} \end{cases}$$

y sea X la variable aleatoria dada por:

$$X := \begin{cases} 1, & \text{si el resultado del lanzamiento es cara;} \\ 0, & \text{si el resultado del lanzamiento es sello.} \end{cases}$$

Demuestre que $X_n \;\underset{n \to \infty}{\overset{P}{\not\longrightarrow}}\; X$.

3.2 Sean $X_1, X_2, \dots$ variables aleatorias reales sobre un espacio de probabilidad $(\Omega, \mathfrak{F}, P)$ definidas por $X_n(\omega) = 1 + \frac{1}{n}$. Demuestre que $X_n \;\underset{n \to \infty}{\overset{P}{\longrightarrow}}\; 1$

3.3 Sean $X_1, X_2, \dots$ variables aleatorias reales sobre un espacio de probabilidad $(\Omega, \mathfrak{F}, P)$, independientes e identicamente distribuidas con distribución normal $\mathcal{N}(\mu, \sigma^2)$. Demuestre que $\overline{X}_{(n)} \;\underset{n \to \infty}{\overset{P}{\longrightarrow}}\; \mu$.

---

[2] Observemos que también podemos aplicar la parte (b) del teorema, puesto que se cumple que $np = 32 \geq 5$ y $n(1-p) = 768 \geq 5$.

3.4 Sean $X_1, X_2, \ldots$ variables aleatorias reales sobre un espacio de probabilidad $(\Omega, \mathfrak{F}, P)$, independientes e identicamente distribuidas con distribución normal estándar. Demuestre que $\frac{Y_n}{n} \xrightarrow[n \to \infty]{P} 1$, siendo $Y_n := \sum_{i=1}^{n} X_i^2$.

3.5 Sean $X_1, X_2, \ldots$ variables aleatorias reales sobre un espacio de probabilidad $(\Omega, \mathfrak{F}, P)$, independientes. Supongamos que cada X_i tiene distribución de Bernoulli con parámetro p_i (no todos iguales). Demuestre que se cumple $\overline{X}_{(n)} - \overline{p}_{(n)} \xrightarrow[n \to \infty]{P} 0$, siendo $\overline{p}_{(n)} := \frac{1}{n} \sum_{i=1}^{n} p_i$.

3.6 Sean $X, X_1, X_2, \ldots$ variables aleatorias reales sobre un espacio de probabilidad $(\Omega, \mathfrak{F}, P)$ independientes, definidas por

$$P(X_n = 1) = 1 - P(X_n = 0) = p_n$$

donde $\lim_{n \to \infty} p_n = 0$. Demuestre que $X_n \xrightarrow[n \to \infty]{P} 0$.

3.7 Considere las variables dadas en el ejercicio 3.1. Verifique que $X_n \xrightarrow[n \to \infty]{d} X$ pero que $X_n \overset{P}{\underset{n \to \infty}{\not\longrightarrow}} X$.

3.8 Sea $X \sim B\left(1, \frac{1}{2}\right)$. Sean $X_1, X_2, \ldots$ variables aleatorias idénticas, definidas por $X_n = X$, para todo n.

(a) ¿Son independientes las variables $X_1, X_2, \ldots$?

(b) Demuestre que $X_n \xrightarrow[n \to \infty]{d} X$.

(c) Si $Y = 1 - X$, demuestre que $X_n \xrightarrow[n \to \infty]{d} Y$

(d) Demuestre que X_n no converge en probabilidad a Y.

3.9 Sean $X_1, X_2, \ldots$ variables aleatorias reales sobre un espacio de probabilidad $(\Omega, \mathfrak{F}, P)$, independientes, definidas por $P(X_n = n^3) = \frac{1}{n^2}$ y $P(X_n = 0) = 1 - \frac{1}{n^2}$. Demuestre que $X_n \xrightarrow[n \to \infty]{P} 0$.

3.10 Sean $X_1, X_2, \ldots$ variables aleatorias reales sobre un espacio de probabilidad $(\Omega, \mathfrak{F}, P)$, independientes, con $P(X_n = 1) = \frac{1}{n}$ y $P(X_n = 0) = 1 - \frac{1}{n}$. Demuestre que $X_n \xrightarrow[n \to \infty]{P} 0$.

3.11 Sean $X_1, X_2, \ldots$ variables aleatorias reales sobre un espacio de probabilidad $(\Omega, \mathfrak{F}, P)$, independientes y con función de densidad f_n como se define abajo. Demuestre que $X_n \xrightarrow[n \to \infty]{P} 0$

$$f_n(x) = \frac{n}{\pi(1 + n^2 x^2)}, \quad n \geq 1$$

Sugerencia: $\int \frac{1}{1+x^2} = \arctan x + c$, donde $\arctan : \mathbb{R} \to \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ es una función continua y estrictamente creciente con $\lim\limits_{x \to +\infty} \arctan x = \frac{\pi}{2}$.

3.12 Sean $Y_1, Y_2, \ldots$ variables aleatorias reales sobre un espacio de probabilidad $(\Omega, \mathfrak{F}, P)$, independientes e identicamente distribuidas. Supongamos que cada una de estas variables puede tomar valor en $\{0, 1, \ldots, 9\}$ con igual probabilidad $1/10$. Sea $X_n := \sum\limits_{i=1}^{n} \frac{Y_i}{10^i}$.

(a) Encuentre la función generadora de momentos de la variable $U_i := \frac{Y_i}{10^i}$.

(b) Encuentre la función generadora de momentos de X_n.

(c) Demuestre que la función generadora de momentos de X_n converge a $M(t) = \frac{e^t - 1}{t}$ cuando $n \to \infty$, es decir, a la función generadora de momentos de la uniforme en $[0, 1]$.

(d) Demuestre que $X_n \xrightarrow[n \to \infty]{d} Y$, donde Y es una variable aleatoria con distribución uniforme en $[0, 1]$.

3.13 Sea X una variable aleatoria real sobre un espacio de probabilidad $(\Omega, \mathfrak{F}, P)$ con $E(X) = \mu$. Demuestre que para cada $\varepsilon > 0$, se cumple la llamada DESIGUALDAD DE MARKOV: $P(|X| \geq \varepsilon) \leq \frac{\mu}{\varepsilon}$.

3.14 Por experiencia, un profesor sabe que el puntaje obtenido por un estudiante en el examen final de su materia es una variable aleatoria con media 40. Obtener una cota superior para la probabilidad de que el estudiante obtenga un puntaje mayor o igual a 80.

3.15 Suponga que $X \sim \mathcal{B}\left(5, \frac{1}{2}\right)$. Use la desigualdad de Markov (véase el ejercicio 3.13) para encontrar una cota superior para $P(X \geq 3)$. Calcule esa probabilidad de manera exacta y compare sus resultados.

3.16 Suponga que X tiene distribución uniforme sobre el intervalo $(0, 10)$. Use la desigualdad de Chevischev con el fin de encontrar una cota superior para $P(|X - 5| \geq 4)$. Calcule esa probabilidad de manera exacta y compare sus resultados.

3.17 Suponga que es conocido que el número de artículos producidos en una empresa durante una semana es una variable aleatoria con media 50.

(a) ¿Qué se puede decir acerca de la probabilidad de que la producción de esta semana exceda 75 artículos?

(b) Si la varianza de una producción de la semana es 25, entonces, ¿qué se puede decir acerca de la probabilidad de que la producción de esta semana esté entre 40 y 60 artículos?

3.18 ¿Existe una variable aleatoria X con $E(X) = \mu$ y $V(X) = \sigma^2$ para la cual se satisfaga

$$P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = 0,6?$$

3.19 Si $Var(X) = 0$, demuestre que $P(EX = X) = 1$. **Sugerencia**: Aplique la desigualdad de Chevischev y, luego tome límite cuando $n \to \infty$.

3.20 Si 10 dados se lanzan, encontrar la probabilidad aproximada de que la suma de los números de las caras esté entre 30 y 40, inclusive.

3.21 Sean $X_1, X_2, \ldots, X_{10}$ variables aleatorias reales sobre un espacio de probabilidad $(\Omega, \mathfrak{F}, P)$, independientes y con distribución uniforme en el intervalo $(0,1)$. Calcular una aproximación para $P\left(\sum\limits_{i=1}^{10} X_i > 6\right)$.

3.22 Sea $X_1, X_2, \ldots$ una sucesión de variables aleatorias indepedientes e identicamente distribuidas que tienen distribución de Poisson con parámetro 1.

(a) Demuestre que $S_n := X_1 + \cdots + X_n$ tiene distribución de Poisson con parámetro n.

(b) Demuestre que $F(0) = \frac{1}{2}$, donde F es la función de distribución acumulada de una variable aleatoria Z que tiene distribución normal estándar.

(c) Use el teorema central del límite para demostrar que

$$\lim_{n\to\infty} \sum_{k=0}^{n} \frac{e^{-n}\, n^k}{k!} \;=\; \frac{1}{2}$$

A

Apéndice de tablas

A.1 La función de distribución binomial

La tabla muestra la probabilidad $P(X \leq k) = B(k; n, p)$ de que ocurran máximo k éxitos en n ensayos independientes, cada uno con probabilidad de éxito p.

Estas probabilidades se calculan para $n = 5$, 10, 15, 20 y 25.

(a) Tabla binomial para $n = 5$

	p												
k	0,05	0,10	0,20	0,25	0,30	0,40	0,50	0,60	0,70	0,75	0,80	0,90	0,95
0	0,774	0,590	0,328	0,237	0,168	0,078	0,031	0,010	0,002	0,001	0,000	0,000	0,000
1	0,977	0,919	0,737	0,633	0,528	0,337	0,188	0,087	0,031	0,016	0,007	0,000	0,000
2	0,999	0,991	0,942	0,896	0,837	0,683	0,500	0,317	0,163	0,104	0,058	0,009	0,001
3	1,000	1,000	0,993	0,984	0,969	0,913	0,812	0,663	0,472	0,367	0,263	0,081	0,023
4	1,000	1,000	0,999	0,999	0,998	0,990	0,969	0,922	0,832	0,763	0,672	0,410	0,226
5	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000

(b) Probabilidades binomiales acumuladas para $n = 10$

	p												
k	0,05	0,10	0,20	0,25	0,30	0,40	0,50	0,60	0,70	0,75	0,80	0,90	0,95
0	0,599	0,349	0,107	0,056	0,028	0,006	0,001	0,000	0,000	0,000	0,000	0,000	0,000
1	0,914	0,736	0,376	0,244	0,149	0,046	0,011	0,002	0,000	0,000	0,000	0,000	0,000
2	0,988	0,930	0,678	0,526	0,383	0,167	0,055	0,012	0,002	0,000	0,000	0,000	0,000
3	0,999	0,987	0,879	0,776	0,650	0,382	0,172	0,055	0,011	0,004	0,001	0,000	0,000
4	1,000	0,998	0,967	0,922	0,850	0,633	0,377	0,166	0,047	0,020	0,006	0,000	0,000
5	1,000	1,000	0,994	0,980	0,953	0,834	0,623	0,367	0,150	0,078	0,033	0,002	0,000
6	1,000	1,000	0,999	0,996	0,989	0,945	0,828	0,618	0,350	0,224	0,121	0,013	0,001
7	1,000	1,000	1,000	1,000	0,998	0,988	0,945	0,833	0,617	0,474	0,322	0,070	0,012
8	1,000	1,000	1,000	1,000	1,000	0,998	0,989	0,954	0,851	0,756	0,624	0,264	0,086
9	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,994	0,972	0,944	0,893	0,651	0,401

(c) Probabilidades binomiales acumuladas para $n = 15$

	p												
k	0,05	0,10	0,20	0,25	0,30	0,40	0,50	0,60	0,70	0,75	0,80	0,90	0,95
0	0,463	0,206	0,305	0,013	0,005	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
1	0,829	0,549	0,167	0,080	0,035	0,005	0,000	0,000	0,000	0,000	0,000	0,000	0,000
2	0,964	0,816	0,398	0,236	0,127	0,027	0,004	0,000	0,000	0,000	0,000	0,000	0,000
3	0,995	0,944	0,648	0,461	0,297	0,091	0,018	0,002	0,000	0,000	0,000	0,000	0,000
4	0,999	0,987	0,836	0,686	0,515	0,217	0,059	0,009	0,001	0,000	0,000	0,000	0,000
5	1,000	0,998	0,939	0,852	0,722	0,403	0,151	0,034	0,004	0,001	0,000	0,000	0,000
6	1,000	1,000	0,982	0,943	0,869	0,610	0,304	0,095	0,015	0,004	0,001	0,000	0,000
7	1,000	1,000	0,996	0,983	0,950	0,787	0,500	0,213	0,050	0,017	0,004	0,000	0,000
8	1,000	1,000	0,999	0,996	0,985	0,905	0,696	0,390	0,131	0,057	0,018	0,000	0,000
9	1,000	1,000	1,000	0,999	0,996	0,966	0,849	0,597	0,278	0,148	0,061	0,002	0,000
10	1,000	1,000	1,000	1,000	0,999	0,991	0,941	0,783	0,485	0,314	0,164	0,013	0,000
11	1,000	1,000	1,000	1,000	1,000	0,998	0,982	0,909	0,703	0,539	0,352	0,056	0,005
12	1,000	1,000	1,000	1,000	1,000	1,000	0,996	0,973	0,873	0,764	0,602	0,184	0,036
13	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,995	0,965	0,920	0,833	0,451	0,171
14	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,995	0,987	0,965	0,794	0,537

(d) Probabilidades binomiales acumuladas para $n = 20$

	p												
k	0,05	0,10	0,20	0,25	0,30	0,40	0,50	0,60	0,70	0,75	0,80	0,90	0,95
0	0,358	0,122	0,012	0,003	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
1	0,736	0,392	0,069	0,024	0,008	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000
2	0,925	0,677	0,206	0,091	0,035	0,004	0,000	0,000	0,000	0,000	0,000	0,000	0,000
3	0,984	0,867	0,411	0,225	0,107	0,016	0,001	0,000	0,000	0,000	0,000	0,000	0,000
4	0,997	0,957	0,630	0,415	0,238	0,051	0,006	0,000	0,000	0,000	0,000	0,000	0,000
5	1,000	0,989	0,804	0,617	0,416	0,126	0,021	0,002	0,000	0,000	0,000	0,000	0,000
6	1,000	0,998	0,913	0,786	0,608	0,250	0,058	0,006	0,000	0,000	0,000	0,000	0,000
7	1,000	1,000	0,968	0,898	0,772	0,416	0,132	0,021	0,001	0,000	0,000	0,000	0,000
8	1,000	1,000	0,990	0,959	0,887	0,596	0,252	0,057	0,005	0,001	0,000	0,000	0,000
9	1,000	1,000	0,997	0,986	0,952	0,755	0,412	0,128	0,017	0,004	0,001	0,000	0,000
10	1,000	1,000	0,999	0,996	0,983	0,872	0,588	0,245	0,048	0,014	0,003	0,000	0,000
11	1,000	1,000	1,000	0,999	0,995	0,943	0,748	0,404	0,113	0,041	0,010	0,000	0,000
12	1,000	1,000	1,000	1,000	0,999	0,979	0,868	0,584	0,228	0,102	0,032	0,000	0,000
13	1,000	1,000	1,000	1,000	1,000	0,994	0,942	0,750	0,392	0,214	0,087	0,002	0,000
14	1,000	1,000	1,000	1,000	1,000	0,998	0,979	0,874	0,584	0,383	0,196	0,011	0,000
15	1,000	1,000	1,000	1,000	1,000	1,000	0,994	0,949	0,762	0,585	0,370	0,043	0,003
16	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,984	0,893	0,775	0,589	0,133	0,016
17	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,996	0,965	0,909	0,794	0,323	0,075
18	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,992	0,976	0,931	0,608	0,264
19	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,997	0,988	0,878	0,642

(e) Probabilidades binomiales acumuladas para $n = 25$

	p												
k	0,05	0,10	0,20	0,25	0,30	0,40	0,50	0,60	0,70	0,75	0,80	0,90	0,95
0	0,277	0,072	0,004	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
1	0,642	0,271	0,027	0,007	0,002	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
2	0,873	0,537	0,098	0,032	0,009	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
3	0,966	0,764	0,234	0,096	0,033	0,002	0,000	0,000	0,000	0,000	0,000	0,000	0,000
4	0,993	0,902	0,421	0,214	0,090	0,009	0,000	0,000	0,000	0,000	0,000	0,000	0,000
5	0,999	0,967	0,617	0,378	0,193	0,029	0,002	0,000	0,000	0,000	0,000	0,000	0,000
6	1,000	0,991	0,780	0,561	0,341	0,074	0,007	0,000	0,000	0,000	0,000	0,000	0,000
7	1,000	0,998	0,891	0,727	0,512	0,154	0,022	0,001	0,000	0,000	0,000	0,000	0,000
8	1,000	1,000	0,953	0,851	0,677	0,274	0,054	0,004	0,000	0,000	0,000	0,000	0,000
9	1,000	1,000	0,983	0,929	0,811	0,425	0,115	0,013	0,000	0,000	0,000	0,000	0,000
10	1,000	1,000	0,994	0,970	0,902	0,586	0,212	0,034	0,002	0,000	0,000	0,000	0,000
11	1,000	1,000	0,998	0,980	0,956	0,732	0,345	0,078	0,006	0,001	0,000	0,000	0,000
12	1,000	1,000	1,000	0,997	0,983	0,846	0,500	0,154	0,017	0,003	0,000	0,000	0,000
13	1,000	1,000	1,000	0,999	0,994	0,922	0,655	0,268	0,044	0,020	0,002	0,000	0,000
14	1,000	1,000	1,000	1,000	0,998	0,966	0,788	0,414	0,098	0,030	0,006	0,000	0,000
15	1,000	1,000	1,000	1,000	1,000	0,987	0,885	0,575	0,189	0,071	0,017	0,000	0,000
16	1,000	1,000	1,000	1,000	1,000	0,996	0,946	0,726	0,323	0,149	0,047	0,000	0,000
17	1,000	1,000	1,000	1,000	1,000	0,999	0,978	0,846	0,488	0,273	0,109	0,002	0,000
18	1,000	1,000	1,000	1,000	1,000	1,000	0,993	0,926	0,659	0,439	0,220	0,009	0,000
19	1,000	1,000	1,000	1,000	1,000	1,000	0,998	0,971	0,807	0,622	0,383	0,033	0,001
20	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,991	0,910	0,786	0,579	0,098	0,007
21	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,998	0,967	0,904	0,766	0,236	0,034
22	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,991	0,968	0,902	0,463	0,127
23	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,998	0,993	0,973	0,729	0,358
24	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,996	0,928	0,723

A.2 La función de distribución de Poisson

La función tabulada es la función de distribución acumulada

$$P(k;\lambda) = \sum_{k=0}^{n} \binom{n}{k} \frac{\lambda^k}{k!}\, e^{-\lambda}$$

para algunos valores de λ.

(a) Tabla de Poisson para $\lambda \leq 1$

	λ									
k	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
0	0,905	0,819	0,741	0,670	0,607	0,549	0,497	0,449	0,407	0,368
1	0,995	0,982	0,963	0,938	0,910	0,878	0,844	0,809	0,772	0,736
2	1,000	0,999	0,996	0,992	0,986	0,977	0,966	0,953	0,937	0,920
3	1,000	1,000	1,000	0,999	0,998	0,997	0,994	0,991	0,987	0,981
4	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,999	0,998	0,996
5	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,999
6	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000

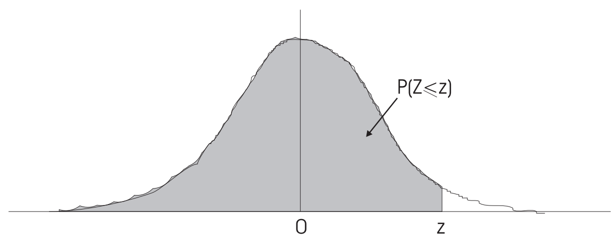
(b) Tabla de Poisson para $2 \leq \lambda \leq 20$

	λ										
k	2	3	4	5	6	7	8	9	10	15	20
0	0,135	0,050	0,018	0,007	0,002	0,001	0,000	0,000	0,000	0,000	0,000
1	0,406	0,199	0,092	0,040	0,017	0,007	0,003	0,001	0,000	0,000	0,000
2	0,677	0,423	0,238	0,125	0,062	0,030	0,014	0,006	0,003	0,000	0,000
3	0,857	0,647	0,433	0,265	0,151	0,082	0,042	0,021	0,010	0,000	0,000
4	0,947	0,815	0,629	0,440	0,285	0,173	0,100	0,055	0,029	0,001	0,000
5	0,983	0,916	0,785	0,616	0,446	0,301	0,191	0,116	0,067	0,003	0,000
6	0,995	0,966	0,889	0,762	0,606	0,450	0,313	0,207	0,130	0,008	0,000
7	0,999	0,988	0,949	0,867	0,744	0,599	0,453	0,324	0,220	0,018	0,001
8	1,000	0,996	0,979	0,932	0,847	0,729	0,593	0,456	0,333	0,037	0,002
9	1,000	0,999	0,992	0,968	0,916	0,830	0,717	0,587	0,458	0,070	0,005
10	1,000	1,000	0,997	0,986	0,957	0,901	0,816	0,706	0,583	0,118	0,011
11	1,000	1,000	0,999	0,995	0,980	0,947	0,888	0,803	0,697	0,185	0,021
12	1,000	1,000	1,000	0,998	0,991	0,973	0,936	0,876	0,792	0,268	0,039
13	1,000	1,000	1,000	0,999	0,996	0,987	0,966	0,926	0,864	0,363	0,066
14	1,000	1,000	1,000	1,000	0,999	0,994	0,983	0,959	0,917	0,466	0,105
15	1,000	1,000	1,000	1,000	0,999	0,998	0,992	0,978	0,951	0,568	0,157
16	1,000	1,000	1,000	1,000	1,000	0,999	0,996	0,989	0,973	0,664	0,221
17	1,000	1,000	1,000	1,000	1,000	1,000	0,998	0,995	0,986	0,749	0,297
18	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,998	0,993	0,819	0,381
19	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,997	0,875	0,470
20	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,998	0,917	0,559
21	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,947	0,644
22	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,967	0,721
23	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,981	0,787
24	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,989	0,843
25	1,000	1,000	1,000	0,994	0,970	0,902	0,586	0,212	0,034	0,994	0,888
26	1,000	1,000	1,000	0,998	0,980	0,956	0,732	0,345	0,078	0,997	0,922
27	1,000	1,000	1,000	1,000	0,997	0,983	0,846	0,500	0,154	0,998	0,948
28	1,000	1,000	1,000	1,000	0,999	0,994	0,922	0,655	0,268	0,999	0,966
29	1,000	1,000	1,000	1,000	1,000	0,998	0,966	0,788	0,414	1,000	0,978
30	1,000	1,000	1,000	1,000	1,000	1,000	0,987	0,885	0,575	1,000	0,987
31	1,000	1,000	1,000	1,000	1,000	1,000	0,996	0,946	0,726	1,000	0,992
32	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,978	0,846	1,000	0,995
33	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,993	0,926	1,000	0,997
34	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,998	0,971	1,000	0,999
35	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,991	1,000	0,999
36	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,998	1,000	1,000

A.3 La función de distribución normal estándar

La tabla muestra la probabilidad $P(Z \leq z)$ de que una variable estándar Z sea menor que el número z.

Por ejemplo, la probabilidad de que una variable aleatoria estándar sea menor que $1,96$ es $P(Z \leq 1,96) = 0,975$.



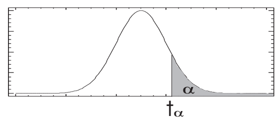
(a) Áreas para valores negativos de Z

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
-3,4	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003
-3,3	0,0005	0,0005	0,0005	0,0004	0,0004	0,0004	0,0004	0,0004	0,0004	0,0004
-3,2	0,0007	0,0007	0,0006	0,0006	0,0006	0,0006	0,0006	0,0005	0,0005	0,0005
-3.1	0,0010	0,0009	0,0009	0,0009	0,0008	0,0008	0,0008	0,0008	0,0007	0,0007
-3,0	0,0013	0,0013	0,0013	0,0012	0,0012	0,0011	0,0011	0,0011	0,0010	0,0010
-2,9	0,0019	0,0018	0,0017	0,0017	0,0016	0,0016	0,0015	0,0015	0,0014	0,0014
-2,8	0,0026	0,0025	0,0024	0,0023	0,0023	0,0022	0,0021	0,0021	0,0020	0,0019
-2,7	0,0035	0,0034	0,0033	0,0032	0,0031	0,0030	0,0029	0,0028	0,0027	0,0026
-2,6	0,0047	0,0045	0,0044	0,0043	0,0041	0,0040	0,0039	0,0038	0,0037	0,0036
-2,5	0,0062	0,0060	0,0059	0,0057	0,0055	0,0054	0,0052	0,0051	0,0049	0,0048
-2,4	0,0082	0,0080	0,0078	0,0075	0,0073	0,0071	0,0069	0,0068	0,0066	0,0064
-2,3	0,0107	0,0104	0,0102	0,0099	0,0096	0,0094	0,0091	0,0089	0,0087	0,0084
-2,2	0,0139	0,0136	0,0132	0,0129	0,0125	0,0122	0,0119	0,0116	0,0113	0,0110
-2,1	0,0179	0,0174	0,0170	0,0166	0,0162	0,0158	0,0154	0,0150	0,0146	0,0143
-2,0	0,0228	0,0222	0,0217	0,0212	0,0207	0,0202	0,0197	0,0192	0,0188	0,0183
-1,9	0,0287	0,0281	0,0274	0,0268	0,0262	0,0256	0,0250	0,0244	0,0239	0,0233
-1,8	0,0359	0,0352	0,0344	0,0336	0,0329	0,0322	0,0314	0,0307	0,0301	0,0294
-1,7	0,0446	0,0436	0,0427	0,0418	0,0409	0,0401	0,0392	0,0384	0,0375	0,0367
-1,6	0,0548	0,0537	0,0526	0,0516	0,0505	0,0495	0,0485	0,0475	0,0465	0,0455
-1,5	0,0668	0,0655	0,0643	0,0630	0,0618	0,0606	0,0594	0,0582	0,0571	0,0559
-1,4	0,0808	0,0793	0,0778	0,0764	0,0749	0,0735	0,0722	0,0708	0,0694	0,0681
-1,3	0,0968	0,0951	0,0934	0,0918	0,0901	0,0885	0,0869	0,0853	0,0838	0,0823
-1,2	0,1151	0,1131	0,1112	0,1093	0,1075	0,1056	0,1038	0,1020	0,1003	0,0985
-1,1	0,1357	0,1335	0,1314	0,1292	0,1271	0,1251	0,1230	0,1210	0,1190	0,1170
-1,0	0,1587	0,1562	0,1539	0,1515	0,1492	0,1469	0,1446	0,1423	0,1401	0,1379
-0,9	0,1841	0,1814	0,1788	0,1762	0,1736	0,1711	0,1685	0,1660	0,1635	0,1611
-0,8	0,2119	0,2090	0,2061	0,2033	0,2005	0,1977	0,1949	0,1922	0,1894	0,1867
-0,7	0,2420	0,2389	0,2358	0,2327	0,2296	0,2266	0,2236	0,2206	0,2177	0,2148
-0,6	0,2743	0,2709	0,2676	0,2643	0,2611	0,2578	0,2546	0,2514	0,2483	0,2451
-0,5	0,3085	0,3050	0,3015	0,2981	0,2946	0,2912	0,2877	0,2843	0,2810	0,2776
-0,4	0,3446	0,3409	0,3372	0,3336	0,3300	0,3264	0,3228	0,3192	0,3156	0,3121
-0,3	0,3821	0,3783	0,3745	0,3707	0,3669	0,3632	0,3594	0,3557	0,3520	0,3483
-0,2	0,4207	0,4168	0,4129	0,4009	0,4052	0,4013	0,3974	0,3936	0,3897	0,3859
-0,1	0,4602	0,4562	0,4522	0,4483	0,4443	0,4404	0,4364	0,4325	0,4286	0,4247
-0,0	0,5000	0,4960	0,4920	0,4880	0,4840	0,4801	0,4761	0,4721	0,4681	0,4641

(b) Áreas para valores positivos de Z

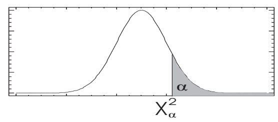
z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9278	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9948	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9961	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9971	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998

A.4 Valores críticos para la distribución t



	α						
ν	0,10	0,05	0,025	0,01	0,005	0,001	0,0005
1	3,078	6,314	12,706	31,821	63,657	318,31	636,620
2	1,886	2,920	4,303	6,965	9,925	22,326	31,598
3	1,638	2,353	3,182	4,541	5,841	10,213	12,924
4	1,533	2,132	2,776	3,747	4,604	7,173	8,610
5	1,476	2,015	2,571	3,365	4,032	5,893	6,869
6	1,440	1,943	2,447	3,143	3,707	5,208	5,959
7	1,415	1,895	2,365	2,998	3,499	4,785	5,408
8	1,397	1,860	2,306	2,896	3,355	4,501	5,041
9	1,383	1.833	2,262	2,821	3,250	4,297	4,781
10	1,372	1,812	2,228	2,764	3,169	4,144	4,587
11	1,363	1,796	2,201	2,718	3,106	4,025	4,437
12	1,356	1,782	2,179	2,681	3,055	3,930	4,318
13	1,350	1,771	2,160	2,650	3,012	3,852	4,221
14	1,345	1,761	2,145	2,624	2,977	3,787	4,140
15	1,341	1,753	2,131	2,602	2,947	3,733	4,073
16	1,337	1,746	2,120	2,583	2,921	3,686	4,015
17	1,333	1,740	2,110	2,567	2,898	3,646	3,965
18	1,330	1,734	2,101	2,552	2,878	3,610	3,922
19	1,328	1,729	2,093	2,539	2,861	3,579	3,883
20	1,325	1,725	2,086	2,528	2,845	3,552	3,850
21	1,323	1,721	2,080	2,518	2,831	3,527	3,819
22	1,321	1,717	2,074	2,508	2,819	3,505	3,795
23	1,319	1,714	2,069	2,500	2,807	3,485	3,767
24	1,318	1,711	2,064	2,492	2,797	3,467	3,745
25	1,316	1,708	2,060	2,485	2,787	3,450	3,725
26	1,315	1,706	2,056	2,479	2,779	3,435	3,707
27	1,314	1,703	2,052	2,473	2,771	3,421	3,690
28	1,313	1,701	2,048	2,467	2,763	3,408	3,674
29	1,311	1,699	2,045	2,462	2,756	3,396	3,659
30	1,310	1,697	2,042	2,457	2,750	3,385	3,646
32	1,309	1,694	2,037	2,449	2,738	3,365	3,622
34	1,307	1,691	2,032	2,441	2,728	3,348	3,601
36	1,306	1,688	2,028	2,434	2,719	3,333	3,582
38	1,304	1,686	2,024	2,429	2,712	3,319	3,566
40	1,303	1,684	2,021	2,423	2,704	3,307	3,551
50	1,299	1,676	2,009	2,403	2,678	3,262	3,496
60	1,296	1.671	2,000	2,390	2,660	3,232	3,460
120	1,282	1,658	1,980	2,358	2,617	3,160	3,373
$\infty(=z)$	1,282	1,645	1,960	2,326	2,576	3,090	3,291

A.5 Distribución chi-cuadrada

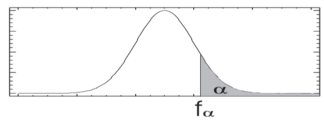


	α									
ν	0,995	0,99	0,98	0,975	0,95	0,90	0,80	0,75	0,70	0,50
1	0,000	0,000	0,000	0,001	0,00393	0,0158	0,0642	0,102	0,148	0,4550
2	0,010	0,0201	0,0404	0,0506	0,103	0,211	0,446	0,575	0,713	1,386
3	0,0717	0,115	0,185	0,216	0,352	0,584	1,005	1,213	1,424	2,366
4	0,207	0,297	0,429	0,484	0,711	1,064	1,649	1,923	2,195	3,357
5	0,412	0,554	0,752	0,831	1,145	1,610	2,343	2,675	3,000	4,351
6	0,676	0,872	1,134	1,237	1,635	2,204	3,070	3,455	3,828	5,348
7	0,989	1,239	1,564	1,690	2,167	2,833	3,822	4,255	4,671	6,346
8	1,344	1,646	2,032	2,180	2,733	3,490	4,594	5,071	5,527	7,344
9	1,735	2,088	2,532	2,700	3,325	4,168	5,380	5,899	6,393	8,343
10	2,156	2,558	3,059	3,247	3,940	4,865	6,179	6,737	7,267	9,342
11	2,603	3,053	3,609	3,816	4,575	5,578	6,989	7,584	8,148	10,341
12	3,074	3,571	4,178	4,404	5,226	6,304	7,807	8,438	9,034	11,340
13	3,565	4,107	4,765	5,009	5,892	7,042	8,634	9,299	9,926	12,340
14	4,075	4,660	5,368	5,629	6,571	7,790	9,467	10,165	10,821	13,339
15	4,601	5,229	5,985	6,262	7,261	8,547	10,307	11,036	11,721	14,339
16	5,142	5,812	6,614	6,908	7,962	9,312	11,152	11,912	12,624	15,338
17	5,697	6,408	7,255	7,564	8,672	10,085	12,002	12,792	13,531	16,338
18	6,844	7,633	8,567	8,907	10,117	11,651	13,716	14,562	15,352	18,338
19	6,844	7,633	8,567	8,907	10,117	11,651	13,716	14,562	15,352	18,338
20	7,434	8,260	9,237	9,591	10,851	12,443	14,578	15,452	16,266	19,337
21	8,034	8,897	9,915	10,283	11,591	13,240	15,445	16,344	17,182	20,337
22	8,643	9,542	10,600	10,982	12,338	14,041	16,314	17,240	18,101	21,337
23	9,260	10,196	11,293	11,688	13,091	14,848	17,187	18,137	19,021	22,337
24	9,886	10,856	11,992	12,401	13,848	15,659	18,062	19,037	19,943	23,337
25	10,520	11,524	12,692	13,120	14,611	16,473	18,940	19,939	20,867	24,337
26	11,160	12,198	13,409	13,844	15,379	17,292	19,820	20,843	21,792	25,336
27	11,808	12,879	14,125	14,573	16,151	18,114	20,703	21,749	22,719	26,336
28	12,461	13,565	14,847	15,308	16,928	18,939	21,588	22,657	23,647	27,336
29	13,121	14,256	15,574	16,047	17,708	19,768	22,475	23,567	24,577	28,336
30	13,787	14,953	16,306	16,791	18,493	20,599	23,364	24,478	25,508	29,336
31	14,457	15,655		17,538	19,280	21,433				
32	15,134	16,362		18,291	20,072	22,271				
33	15,815	17,073		19,046	20,866	23,110				
34	16,501	17,789		19,806	21,664	23,952				
35	17,191	18,508		20,569	22,465	24,796				
36	17,887	19,233		21,336	23,269	25,643				
37	18,584	19,960		22,105	24,075	26,492				
38	19,289	20,691		22,878	24,884	27,343				
39	19,994	21,425		23,654	25695	28,196				
40	20,706	22,164		24,433	26,509	29,050				

(b) Valores críticos $\chi^2_\alpha(\nu)$ (continuación)

	α									
ν	0,30	0,25	0,20	0,10	0,05	0,025	0,02	0,01	0,005	0,001
1	1,074	1,323	1,642	2,706	3,841	5,024	5,412	6,635	7,879	10,827
2	2,408	2,773	3,219	4,605	5,991	7,378	7,824	9,210	10,597	13,815
3	3,665	4,108	4,642	6,251	7,815	9,348	9,837	11,345	12,838	16,268
4	4,878	5,385	5,989	5,779	9,488	11,143	11,668	13,277	14,860	18,465
5	6,064	6,626	7,289	9,236	11,070	12,832	13,388	15,086	16,750	20,517
6	7,231	7,841	8,558	10,645	12,592	14,449	15,033	16,812	18,548	22,457
7	8,383	9,037	9,803	12,017	14,067	16,013	16,622	18,475	20,278	24,322
8	9,524	10,219	11,030	13,362	15,507	17,535	18,168	20,090	21,955	26,125
9	10,656	11,389	12,242	14,684	16,919	19,023	19,679	21,666	23,589	27,877
10	11,781	12,549	13,442	15,987	18,307	20,483	21,161	23,209	25,188	29,588
11	12,899	13,701	14,631	17,275	19,675	21,920	22,618	24,725	26,757	31,264
12	14,011	14,845	15,812	18,549	21,026	23,337	24,054	26,217	28,300	32,909
13	15,119	15,984	16,985	19,812	22,362	24,736	25,472	27,688	29,819	34,528
14	16,222	17,117	18,151	21,064	23,685	26,119	26,873	29,141	31,319	36,123
15	17,322	18,245	19,311	22,307	24,996	27,488	28,259	30,578	32,801	37,697
16	18,418	19,369	20,465	23,542	26,296	28,845	29,633	32,000	34,267	39,252
17	19,511	20,489	21,615	24,769	27,587	30,191	30,995	33,409	35,718	40,790
18	20,601	21,605	22,760	25,989	28,869	31,526	32,346	34,805	37,156	42,312
19	21,689	22,718	23,900	27,204	30,144	32,852	33,687	36,191	38,582	43,820
20	22,775	23,828	25,038	28,412	31,410	34,170	35,020	37,566	39,997	45,315
21	23,858	24,935	26,171	29,615	32,671	35,479	36343	38,932	41,401	46,797
22	24,939	26,039	27,301	30,813	33,924	36,781	37,659	40,289	42,796	48,268
23	26,018	27,141	28,429	32,007	35,172	38,076	38,968	41,638	44,181	49,728
24	27,096	28,241	29,553	33,196	36,415	39,364	40,270	42,980	45,558	51,179
25	28,172	29,339	30,675	34,382	37,652	40,646	41,566	44,314	46,928	52,620
26	29,246	30,434	31,795	35,563	38,885	41,923	42,856	45,642	48,290	54,052
27	30,319	31,528	32,912	36,741	40,113	43,194	44,140	46,963	49,645	55,476
28	31,391	32,620	34,027	37,916	41,337	44,461	45,419	48,278	50,993	56,893
29	32,461	33,711	35,139	39,087	42,557	45,722	46,693	49,588	52,336	58,302
30	33,530	34,800	36,250	40,256	43,773	46,979	47,962	50,892	53,672	59,703
31				41,422	44,985	48,231		52,190	55,000	
32				42,585	46,194	49,480		53,486	56,328	
33				43,745	47,400	50,724		54,774	57,646	
34				44,903	48,602	51,966		56,061	58,964	
35				46,059	49,802	53,203		57,340	60,272	
36				47,212	50,998	54,437		58,619	61,581	
37				48,363	52,192	55,667		59,891	62,880	
38				49,513	53,384	56,896		61,162	64,181	
39				50,660	54,572	58,119		62,426	65,473	
40				51,805	55,758	59,342		63,691	66,766	

A.6 Valores críticos para la distribución F



(a) Valores críticos $F_\alpha(\nu_1, \nu_2)$ para $\alpha = 0,05$

	ν_1								
ν_2	1	2	3	4	5	6	7	8	9
1	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04
120	3,92	3,07	2,68	2,45	2,29	2,17	2,09	2,02	1,96
∞	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88

(b) Valores críticos $F_\alpha(\nu_1, \nu_2)$ para $\alpha = 0,05$

	ν_1									
ν_2	10	12	15	20	24	30	40	60	120	∞
1	241,9	243,9	245,9	248,0	249,1	250,1	251,1	252,2	253,3	254,3
2	19,40	19,41	19,43	19,45	19,45	19,46	19,47	19,48	19,49	19,50
3	8,79	8,74	8,70	8,66	8,64	8,62	8,59	8,57	8,55	8,53
4	5,96	5,91	5,86	5,80	5,77	5,75	5,72	5,69	5,66	5,63
5	4,74	4,68	4,62	4,56	4,53	4,50	4,46	4,43	4,40	4,36
6	4,06	4,00	3,94	3,87	384	3,81	3,77	3,74	3,70	3,67
7	3,64	3,57	3,51	3,44	3,41	3,38	3,34	3,30	3,27	3,23
8	3,35	3,28	3,22	3,15	3,12	3,08	3,04	3,01	2,97	2,93
9	3,14	3,07	3,01	2,94	2,90	2,86	2,83	2,79	2,75	2,71
10	2,98	2,91	2,85	2,77	2,74	2,70	2,66	2,62	2,58	2,54
11	2,85	2,79	2,72	2,65	2,61	2,57	2,53	2,49	2,45	2,40
12	2,75	2,69	2,62	2,54	2,51	2,47	2,43	2,38	2,34	2,30
13	2,67	2,60	2,53	2,46	2,42	2,38	2,34	2,30	2,25	2,21
14	2,60	2,53	2,46	2,39	2,35	2,31	2,27	2,22	2,18	2,13
15	2,54	2,48	2,40	2,33	2,29	2,25	2,20	2,16	2,11	2,07
16	2,49	2,42	2,35	2,28	2,24	2,19	2,15	2,11	2,06	2,01
17	2,45	2,38	2,31	2,23	2,19	2,15	2,10	2,06	2,01	1,96
18	2,41	2,34	2,27	2,19	2,15	2,11	2,06	2,02	1,97	1,92
19	2,38	2,31	2,23	2,16	2,11	2,07	2,03	1,98	1,93	1,88
20	2,35	2,28	2,20	2,12	2,08	2,04	1,99	1,95	1,90	1,84
21	2,32	2,25	2,18	2,10	2,05	2,01	1,96	1,92	1,87	1,81
22	2,30	2,23	2,15	2,07	2,03	1,98	1,94	1,89	1,84	1,78
23	2,27	2,20	2,13	2,05	2,01	1,96	1,91	1,86	1,81	1,76
24	2,25	2,18	2,11	2,03	1,98	1,94	1,89	1,84	1,79	1,73
25	2,24	2,16	2,09	2,01	1,96	1,92	1,87	1,82	1,77	1,71
26	2,22	2,15	2,07	1,99	1,95	1,90	1,85	1,80	1,75	1,69
27	2,20	2,13	2,06	1,97	1,93	1,88	1,84	1,79	1,73	1,67
28	2,19	2,12	2,04	1,96	1,91	1,87	1,82	1,77	1,71	1,65
29	2,18	2,10	2,03	1,94	1,90	1,85	1,81	1,75	1,70	1,64
30	2,16	2,09	2,01	1,93	1,89	1,84	1,79	1,74	1,68	1,62
40	2,08	2,00	1,92	1,84	1,79	1,74	1,69	1,64	1,58	1,51
60	1,99	1,92	1,84	1,75	1,70	1,65	1,59	1,53	1,47	1,39
120	1,91	1,83	1,75	1,66	1,61	1,55	1,50	1,43	1,35	1,25
∞	1,83	1,75	1,67	1,57	1,52	1,46	1,39	1,32	1,22	1,00

(c) Valores críticos $F_\alpha(\nu_1, \nu_2)$ para $\alpha = 0,01$

	ν_1								
ν_2	1	2	3	4	5	6	7	8	9
1	4052	4999,5	5403	5625	5764	5859	5928	5981	6022
2	98,50	99,00	99,17	99,25	99,30	99,33	99,36	99,37	99,39
3	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,35
4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66
5	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16
6	13,75	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98
7	12,25	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72
8	11,26	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91
9	10,56	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35
10	10,04	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94
11	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63
12	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39
13	9,07	6,70	5,74	5,21	4,86	4,62	4,44	4,30	4,19
14	8,86	6,51	5,56	5,04	4,69	4,46	4,28	4,14	4,03
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89
16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78
17	8,40	6,11	5,18	4,67	4,34	4,10	3,93	3,79	3,68
18	8,29	6,01	5,09	4,58	4,25	4,01	3,84	3,71	3,60
19	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52
20	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,46
21	8,02	5,78	4,87	4,37	4,04	3,81	3,64	3,51	3,40
22	7,95	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35
23	7,88	5,66	4,76	4,26	3,94	3,71	3,54	3,41	3,30
24	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,26
25	7,77	5,57	4,68	4,18	3,85	3,63	3,46	3,32	3,22
26	7,72	5,53	4,64	4,14	3,82	3,59	3,42	3,29	3,18
27	7,68	5,49	4,60	4,11	3,78	3,56	3,39	3,26	3,15
28	7,64	5,45	4,57	4,07	3,75	3,53	3,36	3,23	3,12
29	7,60	5,42	4,54	4,04	3,73	3,50	3,33	3,20	3,09
30	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,07
40	7,31	5,18	4,31	3,83	3,51	3,29	3,12	2,99	2,89
60	7,08	4,98	4,13	3,65	3,34	3,12	2,95	2,82	2,72
120	6,85	4,79	3,95	3,48	3,17	2,96	2,79	2,66	2,56
∞	6,63	4,61	3,78	3,32	3,02	2,80	2,64	2,51	2,41

(d) Valores críticos $F_\alpha(\nu_1, \nu_2)$ para $\alpha = 0,01$

	ν_1									
ν_2	10	12	15	20	24	30	40	60	120	∞
1	6056	6106	6157	6209	6235	6261	6287	6313	6339	6366
2	99,40	99,42	99,43	99,45	99,46	99,47	99,47	99,48	99,49	99,50
3	27,23	27,05	26,87	26,69	26,60	26,50	26,41	26,32	26,22	26,13
4	14,55	14,37	14,20	14,02	13,93	13,84	13,75	13,65	13,56	13,46
5	10,05	9,89	9,72	9,55	9,47	9,38	9,29	9,20	9,11	9,02
6	7,87	7,72	7,56	7,40	7,31	7,23	7,14	7,06	6,97	6,88
7	6,62	6,47	6,31	6,16	6,07	5,99	5,91	5,82	5,74	5,65
8	5,81	5,67	5,52	5,36	5,28	5,20	5,12	5,03	4,95	4,86
9	5,26	5,11	4,96	4,81	4,73	4,65	4,57	4,48	4,40	4,31
10	4,85	4,71	4,56	4,41	4,33	4,25	4,17	4,08	4,00	3,91
11	4,54	4,40	4,25	4,10	4,02	3,94	3,86	3,78	3,69	3,60
12	4,30	4,16	4,01	3,86	3,78	3,70	3,62	3,54	3,45	3,36
13	4,10	3,96	3,82	3,66	3,59	3,51	3,43	3,34	3,25	3,17
14	3,94	3,80	3,66	3,51	3,43	3,35	3,27	3,18	3,09	3,00
15	3,80	3,67	3,52	3,37	3,29	3,21	3,13	3,05	2,96	2,87
16	3,69	3,55	3,41	3,26	3,18	3,10	3,02	2,93	2,84	2,75
17	3,59	3,46	3,31	3,16	3,08	3,00	2,92	2,83	2,75	2,65
18	3,51	3,37	3,23	3,08	3,00	2,92	2,84	2,75	2,66	2,57
19	3,43	3,30	3,15	3,00	2,92	2,84	2,76	2,67	2,58	2,49
20	3,37	3,23	3,09	2,94	2,86	2,78	2,69	2,61	2,52	2,42
21	3,31	3,17	3,03	2,88	2,80	2,72	2,64	2,55	2,46	2,36
22	3,26	3,12	2,98	2,83	2,75	2,67	2,58	2,50	2,40	2,31
23	3,21	3,07	2,93	2,78	2,70	2,62	2,54	2,45	2,35	2,26
24	3,17	3,03	2,89	2,74	2,66	2,58	2,49	2,40	2,31	2,21
25	3,13	2,99	2,85	2,70	2,62	2,54	2,45	2,36	2,27	2,17
26	3,09	2,96	2,81	2,66	2,58	2,50	2,42	2,33	2,23	2,13
27	3,06	2,93	2,78	2,63	2,55	2,47	2,38	2,29	2,20	2,10
28	3,03	2,90	2,75	2,60	2,52	2,44	2,35	2,26	2,17	2,06
29	3,00	2,87	2,73	2,57	2,49	2,41	2,33	2,23	2,14	2,03
30	2,98	2,84	2,70	2,55	2,47	2,39	2,30	2,21	2,11	2,01
40	2,80	2,66	2,52	2,37	2,29	2,20	2,11	2,02	1,92	1,80
60	2,63	2,50	2,35	2,20	2,12	2,03	1,94	1,84	1,73	1,60
120	2,47	2,34	2,19	2,03	1,95	1,86	1,76	1,66	1,53	1,38
∞	2,32	2,18	2,04	1,88	1,79	1,70	1,59	1,47	1,32	1,00

Bibliografía & Referencias

[1] Blanco, L., *Probabilidad*. Bogotá: Universidad Nacional de Colombia, 2004.

[2] Grimmet, G.; Stirzacker, D., *Probability and random processes*. New York: Oxford, 1992.

[3] Llinás, H.; Rojas, C., *Estadística descriptiva y distribuciones de probabilidad*. Barranquilla: Ediciones Uninorte, 2005.

[4] Llinás, H., *Medida e integración*. Barranquilla: Ediciones Uninorte, 2007.

[5] Llinás, H., *Applet: La ley de los grandes números*. http://ylang-ylang.uninorte.edu.co/Objetos/Estadistica/LeyDeGrandesNumeros/index.html

[6] Pitman, J., *Probability*. Berlin-Heidelberg-New York: Springer-Verlag, 1993.

[7] Ross, S., *A first course in probability*. México: Prentice Hall, 2002.

Índice