Índice general

Práctica 1

Espacios de Probabilidad Equiprobables

1. Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y sea $(A_i)_{i\in\mathbb{N}}$ una sucesión de eventos en $\mathcal{F}$. Demostrar las siguientes propiedades:

   (i) **Aditividad finita.** Si $A_1, \ldots, A_n$ son disjuntos entonces $\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}(A_i)$.

   (ii) **σ-subaditividad.** $\mathbb{P}\left(\bigcup_{i\in\mathbb{N}} A_i\right) \leq \sum_{i=1}^{n} \mathbb{P}(A_i)$.

   (iii) $\mathbb{P}(A_i) = 0$ para todo $i \in \mathbb{N}$, entonces $\mathbb{P}\left(\bigcup_{i\in\mathbb{N}} A_i\right) = 0$.

   (iv) $\mathbb{P}(A_i) = 1$ para todo $i \in \mathbb{N}$, entonces $\mathbb{P}\left(\bigcap_{i\in\mathbb{N}} A_i\right) = 1$.

   (v) **Continuidad a izquierda de $\mathbb{P}$.** $(A_i)_{i\in\mathbb{N}}$ creciente, entonces $\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \lim_{i\to\infty} \mathbb{P}(A_i)$.

   (vi) **Continuidad a derecha de $\mathbb{P}$.** $(A_i)_{i\in\mathbb{N}}$ decreciente, entonces $\mathbb{P}\left(\bigcap_{i=1}^{n} A_i\right) = \lim_{i\to\infty} \mathbb{P}(A_i)$.

   Definición: Una sucesión $(A_i)_{i\in\mathbb{N}}$ de eventos se dice creciente (respectivamente decreciente) si $A_i \subset A_{i+1}$ para todo $i \in \mathbb{N}$ (respectivamente $A_{i+1} \subset A_i$ para todo $i \in \mathbb{N}$).

Solución:

(i) Sea $A_1, \ldots, A_k$ disjuntos, entonces $A_1, \ldots, A_k, A_{k+1} = \emptyset, \ldots, A_{n+k} = \emptyset, \ldots$ es una sucesión disjunta infinita de eventos en $\mathcal{F}$, entonces

$$\mathbb{P}\left(\bigcup_{i=1}^{k} A_i\right) = \mathbb{P}\left(\bigcup_{i\in\mathbb{N}} A_i\right) = \sum_{i\in\mathbb{N}} \mathbb{P}(A_i) = \sum_{i=1}^{k} \mathbb{P}(A_i) + \underbrace{\sum_{i=k+1}^{\infty} \mathbb{P}(A_i)}_{\substack{=0,\, pues \\ \mathbb{P}(A_i)=0, \forall i \geq k+1}} = \sum_{i=1}^{k} \mathbb{P}(A_i).$$

(ii) Sean los conjuntos A_1 y A_2.

$$\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2) \Rightarrow \mathbb{P}(A_1 \cup A_2) \leq \mathbb{P}(A_1) + \mathbb{P}(A_2).$$

Sea, ahora, $(A_i)_{i\in\mathbb{N}}$ tal que $\mathbb{P}\left(\bigcup_{i=1}^{j} A_i\right) \leq \sum_{i=1}^{j} \mathbb{P}(A_i)$ para todo $2 \leq j \leq n$ y queremos ver que vale para $j = n+1$.

$$\mathbb{P}\left(\bigcup_{i=1}^{n+1} A_i\right) = \mathbb{P}\left(\bigcup_{i=1}^{n} A_i \cup A_{n+1}\right)$$

por hipótesis inductiva, sabemos que definiendo $A = \bigcup_{i=1}^{n} A_i$ se tiene que $\mathbb{P}(A \cup A_{n+1}) \leq \mathbb{P}(A) + \mathbb{P}(A_{n+1})$. Luego,

$$\mathbb{P}\left(\bigcup_{i=1}^{n+1} A_i\right) = \mathbb{P}\left(\bigcup_{i=1}^{n} A_i \cup A_{n+1}\right) = \mathbb{P}(A) + \mathbb{P}(A_{n+1}) = \mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) + \mathbb{P}(A_{n+1})$$

entonces, por hipótesis inductiva,

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

Por lo tanto, de manera inductiva

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

Otra manera de probar este enunciado es notando que

$$\bigcup_{i\in\mathbb{N}} A_i = A_1 \cup \left(\bigcup_{i=2}^{\infty} A_i\right) \Rightarrow \mathbb{P}\left(\bigcup_{i\in\mathbb{N}} A_i\right) \leq \mathbb{P}(A_1) + \mathbb{P}\left(\bigcup_{i=2}^{\infty} A_i\right) \leq \mathbb{P}(A_1) + \mathbb{P}(A_2) + \mathbb{P}\left(\bigcup_{i=3}^{\infty} A_i\right) \Rightarrow$$

$$\mathbb{P}\left(\bigcup_{i\in\mathbb{N}} A_i\right) \leq \mathbb{P}(A_1) + \mathbb{P}(A_2) + \ldots + \mathbb{P}(A_j) + \mathbb{P}\left(\bigcup_{i=j+1}^{\infty} A_i\right) \leq \underbrace{\ldots\ldots}_{\substack{\text{siguiendo}\\ \text{inductivamente}}} \leq \sum_{i\in\mathbb{N}} \mathbb{P}(A_i).$$

(iii) Sea $(A_i)_{i\in\mathbb{N}} \subset \mathcal{F}$ tal que $\forall i \in \mathbb{N}$ se tiene que $\mathbb{P}(A_i) = 0$. Entonces recordando que $\mathbb{P}(A_i) \in [0,1]$ para todo $i \in \mathbb{N}$,

$$0 \leq \mathbb{P}\left(\bigcup_{i\in\mathbb{N}} A_i\right) \leq \sum_{i\in\mathbb{N}} \underbrace{\mathbb{P}(A_1)}_{=0} = 0 \Rightarrow \mathbb{P}\left(\bigcup_{i\in\mathbb{N}} A_i\right) = 0.$$

(iv) Sea $(A_i)_{i\in\mathbb{N}} \subset \mathcal{F}$ tal que $\forall i \in \mathbb{N}$ se tiene que $\mathbb{P}(A_i) = 1$. Entonces recordando que $\mathbb{P}(A_i) \in [0,1]$ para todo $i \in \mathbb{N}$,

$$1 \geq \mathbb{P}\left(\bigcap_{i\in\mathbb{N}} A_i\right) = 1 - \mathbb{P}\left(\bigcup_{i\in\mathbb{N}} A_i^c\right) \geq 1 - \sum_{i\in\mathbb{N}} \underbrace{\mathbb{P}(A_i^c)}_{=0} = 1 \Rightarrow \mathbb{P}\left(\bigcap_{i\in\mathbb{N}} A_i\right) = 1.$$

(v) Sea $(A_i)_{i\in\mathbb{N}} \subset \mathcal{F}$ creciente, definimos

$$\left.\begin{aligned} F_1 &= A_1 \\ F_2 &= A_2 \cap A_1^c \\ \vdots\ &\quad \vdots \\ \vdots\ &\quad \vdots \\ F_i &= A_i \cap A_{i-1}^c \end{aligned}\right\} \text{ conjuntos disjuntos tal que } \bigcup_{i\in\mathbb{N}} A_i = \bigcup_{i\in\mathbb{N}} F_i \Rightarrow$$

$$\mathbb{P}\left(\bigcup_{i\in\mathbb{N}} A_i\right) = \mathbb{P}\left(\bigcup_{i\in\mathbb{N}} F_i\right) = \sum_{i\in\mathbb{N}} \mathbb{P}(F_i) = \lim_{n\to\infty} \sum_{i=1}^{n} \mathbb{P}(F_i) = \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{i=1}^{n} F_i\right) = \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) \Rightarrow$$

$$\mathbb{P}\left(\bigcup_{i\in\mathbb{N}} A_i\right) = \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \lim_{i\to\infty} \mathbb{P}(A_i).$$

(vi) Sea $(A_i)_{i\in\mathbb{N}}$ decreciente, entonces $(A_i^c)_{i\in\mathbb{N}}$ resulta creciente. Luego,

$$1 - \mathbb{P}\left(\bigcap_{i\in\mathbb{N}} A_i\right) = \mathbb{P}\left(\bigcup_{i\in\mathbb{N}} A_i^c\right) = \lim_{i\to\infty} \mathbb{P}(A_i^c) = 1 - \lim_{i\to\infty} \mathbb{P}(A_i) \Rightarrow \mathbb{P}\left(\bigcap_{i\in\mathbb{N}} A_i\right) = \lim_{i\to\infty} \mathbb{P}(A_i).$$

2. Sean $\Omega = \{\omega_i\}_{i\in\mathbb{N}}$ un espacio numerable y una aplicación $P : \mathcal{P}(\Omega) \to [0,1]$. Probar que P es una medida de probabilidad si y sĺo si existe una sucesiń $(p_i)_{i\in\mathbb{N}}$ de números reales no negativos tal que $\sum_{i\in\mathbb{N}} p_i = 1$ y, para todo $A \subseteq \Omega$, $P(A) = \sum_{\substack{i\in\mathbb{N}| \\ w_i\in A}} p_i$.

Solución:

$(\Rightarrow)$:

$(\Leftarrow):$ Veamos que P es una medida de probabilidad. Sea $(p_i)_{i\in\mathbb{N}}$ una sucesiń tal que $p_i = P(\{w_i\})$, entonces

- $\mathbb{P}(\Omega) = \mathbb{P}\left(\bigcup_{i\in\mathbb{N}}\{w_i\}\right) = \sum\limits_{i\in\mathbb{N}} \mathbb{P}(w_i) = 1.$
- $\mathbb{P}(A) = \mathbb{P}\left(\bigcup\limits_{\substack{i\in\mathbb{N}| \\ \{w_i\}\subset A}} \{w_i\}\right) = \sum\limits_{\substack{i\in\mathbb{N}| \\ w_i\in A}} p_i \leq 1.$
- Sean $(A_i)_{i\in\mathbb{N}}$ una sucesión infinita de eventos disjuntos, entonces

$$A_1 = \bigcup_{\substack{i\in\mathbb{N}| \\ \{w_i\}\subset A_1}} \{w_i\}, A_2 = \bigcup_{\substack{i\in\mathbb{N}| \\ \{w_i\}\subset A_2}} \{w_i\}, \ldots, A_k = \bigcup_{\substack{i\in\mathbb{N}| \\ \{w_i\}\subset A_k}} \{w_i\}, \ldots$$

Por lo tanto,

$$\bigcup_{i\in\mathbb{N}} A_i = \left(\bigcup_{\substack{i\in\mathbb{N}| \\ \{w_i\}\subset A_1}} \{w_i\}\right) \cup \left(\bigcup_{\substack{i\in\mathbb{N}| \\ \{w_i\}\subset A_2}} \{w_i\}\right) \cup \ldots \cup \left(\bigcup_{\substack{i\in\mathbb{N}| \\ \{w_i\}\subset A_j}} \{w_i\}\right) \cup \ldots \cup \ldots \Rightarrow$$

$$\mathbb{P}\left(\bigcup_{i\in\mathbb{N}} A_i\right) = \sum_{\substack{i\in\mathbb{N}| \\ w_i\in A_1}} p_i + \sum_{\substack{i\in\mathbb{N}| \\ w_i\in A_2}} p_i + \ldots + \sum_{\substack{i\in\mathbb{N}| \\ w_i\in A_j}} p_i + \ldots = \mathbb{P}(A_1) + \mathbb{P}(A_2) + \ldots = \sum_{i\in\mathbb{N}} \mathbb{P}(A_i).$$

3. Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y sean $A_1, \ldots, A_n$ eventos en $\mathcal{F}$.
   (a) Probar la siguiente fórmula de inclusión-exclusión.
   $$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{k=1}^{n} \sum_{\substack{\mathcal{J}\subseteq\{1,\ldots,n\} \\ \sharp(\mathcal{J})=k}} (-1)^{k+1} \mathbb{P}\left(\bigcap_{i\in\mathcal{J}} A_i\right).$$
   (b) Concluir que si existen números reales no negativos $b_1, \ldots, b_k$ tales que para $k = 1, \ldots, n$ vale que $\mathbb{P}\left(\bigcap\limits_{i\in\mathcal{J}} b_k\right) = b_k$ para todo $\mathcal{J}$ con $\sharp(\mathcal{J}) = k$ entonces la fórmula anterior se reduce a
   $$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{k=1}^{n} \binom{n}{k} (-1)^{k+1} b_k.$$

4. Se arroja dos veces un dado equilibrado
   (i) Exhibir un espacio muestral que describa dicho experimento.
   (ii) Sean A, B y C los eventos
   $$A = \{\text{La suma de los resultados es par}\}, B = \{\text{La suma de los resultados es 8}\}$$
   $$C = \{\text{Ambos resultados son distintos}\}$$
   Calcular las probabilidades de A, B, C, $A\cap C$, $A\cup C$, $A\setminus B$ y $A\setminus C$.

Solución:

(i) $\Omega = \{(x_1, x_2) : x_1 \in [1,6] \text{ y } x_2 \in [1,6]\}$.

(ii)

$$\mathbb{P}(A) = \mathbb{P}\left(\{x_1 = par \text{ y } x_2 = par\} \sqcup \{x_1 = impar \text{ y } x_2 = impar\}\right) = 2\left(\frac{3}{6}\times\frac{3}{6}\right) = \frac{1}{2}.$$

$$\mathbb{P}(B) = \mathbb{P}\{(2,6);(3,5);(4,4);(5,3);(6,2)\} = \frac{5}{36}$$

$$\mathbb{P}(C) = 1 - \mathbb{P}\left(\bigcup_{i=1}^{6}\{x_i, x_i\}\right) = 1 - \frac{6}{36} = \frac{5}{6}.$$

$$\mathbb{P}(A\cap C) = \frac{2\times 6}{36} = \frac{1}{3} \text{ y } \mathbb{P}(A\cup C) = \mathbb{P}(A) + \mathbb{P}(C) - \mathbb{P}(A\cap C) \Rightarrow \mathbb{P}(A\cup C) = 1$$

$$\mathbb{P}(A\setminus B) \underbrace{=}_{B\subset A} \mathbb{P}(A) - \mathbb{P}(B) = \frac{13}{36} \text{ y } \mathbb{P}(A\setminus C) = \frac{6}{36}.$$

5. Un bolillero contiene N bolillas numeradas desde la 1 hasta la N. Se extraen sucesivamente y sin reposición n bolillas, con $1 \leq n \leq N$. Sean m y k números naturales tales que $1 \leq m \leq N$ y $1 \leq k \leq n$. Hallar la probabilidad de que

   (i) se extraiga la bolilla m en la k-ésima extracción.
   (ii) se extraiga la bolilla m.
   (iii) el máximo número obtenido sea menor o igual a m.
   (iv) el máximo número obtenido sea m.
   (v) dados $a, b \in \mathbb{N}$, $n \leq a < b \leq N$, el máximo número obtenido está entre a y b inclusive.
   (vi) los números de las bolillas extraídas en el orden en que fueron extraídas constituyan una sucesión estrictamente creciente.

   Resolver nuevamente los items anteriores para el caso en que las extracciones se realizan con reposición.

6. Se colocan 4 bolillas indistinguibles en 4 urnas numeradas. Calcular la probabilidad de que la primer urna contenga

   (i) exactamente una bolilla.
   (ii) exactamente dos bolillas.
   (iii) al menos una bolilla.

   Resolver nuevamente el problema para el caso en que las bolillas sean distinguibles.

Solución:

- Bolillas indistinguibles: Sea el espacio muestral

$$\Omega = \{(x_1, x_2, x_3, x_4) : x_i = \sharp(\{\text{bolillas en la irna i}\}) \text{ y } x_1 + x_2 + x_3 + x_4 = 4\} \Rightarrow$$

$$\sharp(\Omega) = \begin{pmatrix} 4+4-1 \\ 4 \end{pmatrix} = 35.$$

donde $\sharp(\Omega)$ representa la cantidad total de formas de ubicar las bolillas.

(i)

$$\mathbb{P}(\{x_1 = 1\}) = \frac{1}{\sharp(\Omega)} \left(1 \times \mathbb{P}\left(\left\{ \begin{array}{c} \text{ubicar las 3 bolillas restantes} \\ \text{en las tres urnas restantes} \end{array} \right\}\right)\right) = \frac{1}{35} \begin{pmatrix} 3+3-1 \\ 3 \end{pmatrix} = \frac{2}{7}.$$

(ii)

$$\mathbb{P}(\{x_1 = 2\}) = \frac{1}{\sharp(\Omega)} \left(1 \times \mathbb{P}\left(\left\{ \begin{array}{c} \text{ubicar las 2 bolillas restantes} \\ \text{en las tres urnas restantes} \end{array} \right\}\right)\right) = \frac{1}{35} \begin{pmatrix} 2+3-1 \\ 2 \end{pmatrix} = \frac{6}{35}.$$

(iii)

$$\mathbb{P}(\{x_1 \geq 1\}) = 1 - \mathbb{P}(\{x_1 = 0\}) = 1 - \frac{1}{\sharp(\Omega)} \left(1 \times \mathbb{P}\left(\left\{ \begin{array}{c} \text{ubicar las 4 bolillas restantes} \\ \text{en las tres urnas restantes} \end{array} \right\}\right)\right) \Rightarrow$$

$$\mathbb{P}(\{x_1 \geq 1\}) = 1 - \frac{1}{35} \begin{pmatrix} 4+3-1 \\ 4 \end{pmatrix} = \frac{4}{7}.$$

- Bolillas distinguibles: Sea el espacio muestral

$$\Omega = \{(x_1, x_2, x_3, x_4) : x_i = \sharp(\{\text{bolillas en la irna i}\})\} \Rightarrow$$

$$\sharp(\Omega) = 4^4.$$

(i)

$$\mathbb{P}(\{x_1 = 1\}) = \frac{1}{\sharp(\Omega)} \begin{pmatrix} 4 \\ 1 \end{pmatrix} (4-1)^{(4-1)} = \frac{27}{64}.$$

(ii)

$$\mathbb{P}(\{x_1 = 2\}) = \frac{1}{\sharp(\Omega)} \begin{pmatrix} 4 \\ 2 \end{pmatrix} (4-1)^{(4-2)} = \frac{27}{128}.$$

(iii)

$$\mathbb{P}(\{x_1 \geq 1\}) = 1 - \mathbb{P}(\{x_1 = 0\}) \Rightarrow$$

$$\mathbb{P}(\{x_1 \geq 1\}) = \frac{1}{\sharp(\Omega)} \left(\begin{pmatrix} 4 \\ 1 \end{pmatrix} (4-1)^{(4-1)} + \begin{pmatrix} 4 \\ 2 \end{pmatrix} (4-1)^{(4-2)} \begin{pmatrix} 4 \\ 3 \end{pmatrix} (4-1)^{(4-3)} + \begin{pmatrix} 4 \\ 4 \end{pmatrix} (4-1)^{(4-1)} \right) = \frac{175}{256}.$$

7. Se toma una muestra al azar con reposición de tamaño r de una población de n elementos.

   (i) Calcular la probabilidad de que en la muestra obtenida no haya elementos repetidos.

   (ii) Se elige un grupo de r personas al azar y se les pregunta su fecha de cumpleaños. Si bien los años no son todos iguales en longitud, y se sabe que los porcentajes de nacimientos no son constantes a través del año, como una primera aproximación podemos suponer que una elección al azar de personas es equivalente a realizar una selección al azar de las fechas de nacimientos y además considerar que el año tiene 365 días. Bajo estas hipótesis, calcular la probabilidad de que entre las r personas elegidas al azar todas tengan distintas fechas de cumpleaños.

   (iii) Calcular la probabilidad de que al distribuir n bolillas distintas en n urnas numeradas se encuentren todas las urnas ocupadas.

Solución:

(i) Sea $x_1, x_2, \ldots, x_r$ una muestra, entonces

$$\Omega = \{(x_1, x_2, \ldots, x_r) : x_i = \{\text{elemento obtenido en la extracción i}\}\} \Rightarrow \sharp(\Omega) = n^r$$

luego,

$$\mathbb{P}(x_i \neq x_j, \forall i \neq j) = \frac{n(n-1)(n-2)\ldots(n-r+1)}{n^r}.$$

(ii) En base al inciso anterior, la probabilidad buscada es

$$\frac{365 \times 364 \times 363 \times \ldots \times (366-r)}{365^r}.$$

(iii) Basta considerar $r = n$; por lo tanto, la probabilidad buscada es

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

8. Se reparten N bolillas distinguibles en n urnas distinguibles de tal modo que cada bolilla tiene la misma probabilidad de llegar a una urna cualquiera sea ésta. Sea el evento $A_i = \{\text{La i-ésima urna no está vacía}\}$.

   (i) Mostrar que para todo $k = 1, \ldots, n$ se tiene

   $$V_k = \mathbb{P}(A_{i_1} \cap \ldots \cap A_{i_k}) = \sum_{j=0}^{k} \binom{k}{j} (-1)^j \left(1 - \frac{j}{n}\right)^N.$$

   Observar que dicha probabilidad no depende de los índices $i_1, \ldots, i_k$ elegidos

   (ii) Probar que si $\lim\limits_{n,N\to\infty} \frac{N}{n} = \lambda$, entonces $\lim\limits_{n,N\to\infty} \frac{N}{n} = V_k = (1-e^{-\lambda})^k$.

   (iii) Probar la identidad

   $$\sum_{j=0}^{n} \binom{n}{j} (-1)^j (n-j)^n = n!$$

Solución:

(i) Sean n urnas distinguibles y como cada bolilla tiene n urnas posibles en donde ubicarse se tiene que $\sharp(\Omega) = n^N$ donde

$$\Omega = \{(x_1, x_2, \ldots, x_N) : x_i = \{\text{urna donde cae la bolilla i}\}\}-$$

Es claro que $A_i^c = \{\text{La i-ésima urna está vacía}\} \Rightarrow \mathbb{P}(A_i^c) = \frac{(n-1)^N}{n^N} = \left(1 - \frac{1}{n}\right)^N$. Luego,

$$1 - V_k = \mathbb{P}(A_{i_1}^c \cup \ldots \cup A_{i_k}^c) = \sum_{j=1}^{k} \binom{k}{j} (-1)^{j+1} \left(\frac{n-j}{n}\right)^N \Rightarrow V_k = 1 - \sum_{j=1}^{k} \binom{k}{j} (-1)^{j+1} \left(1 - \frac{j}{n}\right)^N$$

$$\Rightarrow V_k = 1 + \sum_{j=1}^{k} \binom{k}{j} (-1)^{j+1} \left(1 - \frac{j}{n}\right)^N = \sum_{j=0}^{k} \binom{k}{j} (-1)^j \left(1 - \frac{j}{n}\right)^N.$$

(ii) Si $\lim\limits_{n,N\to\infty} \dfrac{N}{n} = \lambda$, entonces

$$\lim_{n,N\to\infty} V_k = \lim_{n,N\to\infty} \sum_{j=0}^{k} \binom{k}{j} (-1)^j \left(1-\frac{j}{n}\right)^N = \lim_{n,N\to\infty} \sum_{j=0}^{k} \binom{k}{j} (-1)^j \left[\left(1-\frac{j}{n}\right)^{-\frac{n}{j}}\right]^{-j\frac{N}{n}} \Rightarrow$$

$$\lim_{n,N\to\infty} V_k = \sum_{j=0}^{k} \binom{k}{j} (-1)^j (-e^{-\lambda})^j = (1-e^{-\lambda})^k.$$

(iii)

9. Se colocan 6 bolillas en 4.

   (i) ¿Cuál es la probabilidad de que todas las urnas se encuentren ocupadas?

   (ii) ¿Y de que al menotrs tres de ellas lo estén?

Solución:

- Bolillas indistinguibles en urnas numeradas:

  (i) $\mathbb{P}(\{\text{todas las urnas están ocupadas}\}) = \frac{1}{\sharp(\Omega)} \binom{2+4-1}{4-1} = \dfrac{10}{84}.$

  (ii)
  $$\mathbb{P}(\{\text{al menos tres urnas estén ocupadas}\}) = \mathbb{P}\left(\begin{Bmatrix}\text{hay tres urnas ocupadas}\\ \text{y una vacía}\end{Bmatrix}\right) + \mathbb{P}\left(\begin{Bmatrix}\text{todas las urnas}\\ \text{están ocupadas}\end{Bmatrix}\right) \Rightarrow$$
  $$\mathbb{P}(\text{al menos tres urnas estén ocupadas}) = \frac{1}{\sharp(\Omega)} \binom{4}{1} \binom{3+3-1}{3-1} + \frac{10}{84} = \frac{40}{84} + \frac{10}{84} = \frac{50}{84}.$$

- Bolillas distinguibles en urnas numeradas:

10. Se tienen n urnas y n bolillas numeradas desde la 1 hasta la n. Se coloca al azar una bolilla en cada urna. Diremos que se produce un apareamiento cuando alguna de las bolillas es colocada en la urna con su mismo n.

    (i) Sea W_k la probabilidad de que ocurran exactamente k apareamientos. Mostrar que $W_k = \dfrac{1}{k!} \sum\limits_{j=0}^{n-k} \dfrac{(-1)^j}{j!}$.

    Sugerencia: Probar primero la fórmula para el caso $k = 0$.

    (ii) Probar que $\lim\limits_{n\to\infty} W_k = \dfrac{e^{-1}}{k!}$.

Práctica 2

Probabilidad condicional - Independencia

1. Se dispone de dos urnas cuyo contenido es el siguiente:

Urna A : 5 bolillas rojas y 3 blancas.

Urna B : 1 bolilla roja y 2 blancas.

Se arroja un dado y, si el resultado es 3 ó 6, se extrae una bolilla de la urna A y se coloca en la urna B para luego extraer una bolilla de esta última. En caso contrario, el procedimiento se realiza a la inversa.

(i) Calcular la probabilidad de que ambas bolillas extraídas sean rojas.

(ii) ¿Cuál es la probabilidad de que ambas bolillas sean blancas si ambas bolillas son del mismo color? Si ambas bolillas son del mismo color, ¿es más probable que sean rojas?

Solución:

(i)

$$\mathbb{P}(b_1 = R \text{ y } b_2 = R) = \mathbb{P}(b_1 = R \text{ y } b_2 = R | dado = 1)\mathbb{P}(dado = 1) + \ldots + \mathbb{P}(b_1 = R \text{ y } b_2 = R | dado = 6)\mathbb{P}(dado = 6) \Rightarrow$$

$$\mathbb{P}(b_1 = R \text{ y } b_2 = R) = 4\left(\frac{1}{3} \times \frac{6}{9} \times \frac{1}{6}\right) + 2\left(\frac{5}{8} \times \frac{2}{4} \times \frac{1}{6}\right).$$

(ii)

$$\mathbb{P}(b_1 = B \text{ y } b_2 = B | b_1 = b_2) = \frac{\mathbb{P}(b_1 = B \text{ y } b_2 = B)}{\mathbb{P}(b_1 = b_2)} = \frac{\mathbb{P}(b_1 = B \text{ y } b_2 = B)}{\mathbb{P}(b_1 = R \text{ y } b_2 = R) + \mathbb{P}(b_1 = B \text{ y } b_2 = B)}.$$

2. Se tienen tres urnas numeradas con veinte bolillas de color en cada una. La primer urna contiene veinte bolillas blancas, la segunda quince y la última diez, siendo todas las bolillas restantes negras. Se elige una de las urnas al azar y se extraen de ella con reposición dos bolillas. La probabilidad de elegir la primer urna es la misma que la de elegir la segunda, mientras que la de elegir la última es igual a su suma. Calcular la probabilidad de haber seleccionado la primer urna sabiendo que las dos bolillas extraídas son blancas

Solución: Sea $p \in [0, 1]$ tal que $p = \mathbb{P}(\{\text{elegir la urna 1}\}) = \mathbb{P}(\{\text{elegir la urna 2}\})$ y $2p = \mathbb{P}(\{\text{elegir la urna 3}\})$.

$$\mathbb{P}\left(\text{elegir la urna 1}\} \mid \{b_1 = B \text{ y } b_2 = B\}\right) = \frac{\mathbb{P}(b_1 = B \text{ y } b_2 = B | \text{elegir la urna 1}\})\mathbb{P}(\text{elegir la urna 1}\})}{\displaystyle\sum_{i=1}^{3} \mathbb{P}(b_1 = B \text{ y } b_2 = B | \text{elegir la urna i}\})\mathbb{P}(\text{elegir la urna i}\})}$$

Por otra parte,

$$\mathbb{P}(b_1 = B \text{ y } b_2 = B | \text{elegir la urna 1}) = \left(\frac{20}{20}\right)^2 = 1$$

$$\mathbb{P}(b_1 = B \text{ y } b_2 = B | \text{elegir la urna 2}) = \left(\frac{15}{20}\right)^2$$

$$\mathbb{P}(b_1 = B \text{ y } b_2 = B | \text{elegir la urna 3}) = \left(\frac{10}{20}\right)^2$$

luego,

$$\mathbb{P}\left(\text{elegir la urna 1}\} \mid \{b_1 = B \text{ y } b_2 = B\}\right) = \frac{p}{1p + \left(\frac{3}{4}\right)^2 p + \left(\frac{1}{2}\right)^2 2p} = \frac{16}{33}$$

3. Cuando se realiza un análisis de laboratorio para diagnosticar una cierta enfermedad en un paciente se pueden cometer dos tipos de errores de diagnóstico: si el análisis da positivo (es decir, éste dice que el paciente está enfermo) pero el paciente está sano se dice que tenemos un falso positivo, mientras que si el análisis da negativo pero el paciente está enfermo se dice que tenemos un falso negativo. Consideremos los eventos $E = \{\text{La persona examinada está enferma}\}$ y $A = \{\text{El resultado del análisis es positivo}\}$. La cantidad $\mathbb{P}(A|E)$ se conoce como sensibilidad del test mientras que $\mathbb{P}(A^c|E^c)$ se denomina la especificidad del test.

   (i) Supongamos que una cierta prueba de laboratorio es tal que $\mathbb{P}(A|E) = \mathbb{P}(A^c|E^c) = 0{,}95$ y además que la probabilidad de que un paciente que se examina padezca la enfermedad es $0{,}005$. ¿Cuál es la probabilidad de que una persona cuyo análisis diagnóstico es positivo esté realmente enferma?

   (ii) Supongamos que $\mathbb{P}(A|E) = \mathbb{P}(A^c|E^c) = p$ y $\mathbb{P}(E) = 0{,}005$. ¿Para qué valor de p es $\mathbb{P}(E|A) = 0{,}95$?

Solución:

(i) $\mathbb{P}(E|A) = \dfrac{\mathbb{P}(A|E)\mathbb{P}(E)}{\mathbb{P}(A)}$ y, además, $\mathbb{P}(A) = \mathbb{P}(A|E)\mathbb{P}(E) + \mathbb{P}(A|E^c)\mathbb{P}(E^c) = \mathbb{P}(A|E)\mathbb{P}(E) + (1-\mathbb{P}(A^c|E^c))(1-\mathbb{P}(E))$. Luego,

$$\mathbb{P}(E|A) = \frac{\mathbb{P}(A|E)\mathbb{P}(E)}{\mathbb{P}(A|E)\mathbb{P}(E) + (1-\mathbb{P}(A^c|E^c))(1-\mathbb{P}(E))} = \frac{0{,}95 \times 0{,}005}{0{,}95 \times 0{,}005 + 0{,}05 \times 0{,}995}.$$

(ii) Sea p tal que $\mathbb{P}(A|E) = \mathbb{P}(A^c|E^c) = p$, entonces

$$0{,}95 = \frac{p \times 0{,}005}{p \times 0{,}005 + (1-p) \times 0{,}005} \Leftrightarrow p(0{,}005 \times 0{,}95 - 0{,}995 \times 0{,}95 - 0{,}005) = -0{,}995 \times 0{,}95$$

4. El color de ojos de una persona está determinado por un sólo par de genes. Si ambos son genes que codifican ojos azules, entonces la persona tendrá ojos azules; si ambos son genes que codifican ojos marrones, la persona tendrá ojos marrones; y si uno de ellos es un gen que codifica ojos azules y el otro codifica ojos marrones, entonces la persona tendrá ojos marrones (es por esto último que se dice que el gen de ojos marrones es dominante por sobre el de ojos azules). A los fines de este ejercicio supondremos que los ojos de una persona sólo pueden ser azules o marrones. Un bebé recién nacido recibe exactamente un gen que codifica el color de sus ojos de cada uno de sus padres mientras que los dos genes de cada padre tienen igual probabilidad de ser transmitidos a su bebé. Supongamos los dos padres de Beto tienen ojos marrones, pero que la hermana de Beto tiene ojos azules.

   (i) ¿Cuál es la probabilidad de que Beto posea un gen que codifica ojos azules?

   (ii) Ahora nos enteramos que Beto tiene ojos marrones. ¿Cambia esta información la probabilidad de que Beto tenga un gen que codifica ojos azules?

   (iii) Beto (de quien ya sabemos que tiene ojos marrones) se casó con Ana, una hermosa mujer de ojos azules. Juntos tuvieron dos hijos, Juan y Pedro. Mostrar que tanto Juan como Pedro tienen la misma probabilidad de tener ojos azules y dar el valor de dicha probabilidad.

   (iv) Si Juan tiene ojos marrones, calcular la probabilidad de que Pedro tenga ojos marrones también.

   (v) ¿Son independientes los eventos $\{\text{Juan tiene ojos marrones}\}$ y $\{\text{Pedro tiene ojos marrones}\}$?

5. **Esquema de Polya.** De un bolillero que contiene B bolillas blancas y R rojas se extraen sucesivamente y al azar n bolillas, devolviendo en cada instancia la bolilla extraída al bolillero junto con otras c bolillas del mismo color. Para cada $n \in \mathbb{N}$ definimos los eventos,

$$R_n = \{\text{La n-ésima bolilla extraída es roja}\} \text{ y } B_n = \{\text{La n-ésima bolilla extraída es blanca}\}.$$

   (i) Probar $\mathbb{P}(R_n) = \dfrac{R}{R+B}$ para todo $n \in \mathbb{N}$.

   (ii) Probar que para todo $m < n$ se tiene

$$\mathbb{P}(R_m \cap R_n) = \frac{R(R+c)}{(R+B)(R+B+c)} \text{ y } \mathbb{P}(R_m \cap B_n) = \frac{RB}{(R+B)(R+B+c)}.$$

   Generalizar a más de dos extracciones.

   (iii) ¿Son los eventos R_m y R_n independientes para $n \neq m$?

   (iv) Calcular $\mathbb{P}(R_3|R_2)$ y $\mathbb{P}(R_1|R_n)$ para $n \geq 2$.

Solución:

(i) Usando inducción en n, si $n = 1$ es claro que $\mathbb{P}(R_1) = \dfrac{R}{R+B}$. Luego,

$$\mathbb{P}(R_1) = \mathbb{P}(R_{n+1}|R_1)\mathbb{P}(R_1) + \mathbb{P}(R_{n+1}|B_1)\mathbb{P}(B_1) \Rightarrow$$

$$\mathbb{P}(R_{n+1}) = \frac{R+c}{R+c+B}\frac{R}{B+R} + \frac{R}{R+B+c}\frac{B}{B+R} = \frac{R(R+c+B)}{(R+B+c)(R+B)} = \frac{R}{R+B}.$$

(ii) Usando inducción en m, si $m = 1$ entonces para $1 < n$

$$\mathbb{P}(R_1 \cap R_n) = \mathbb{P}(R_1)\mathbb{P}(R_n|R_1) = \frac{R}{R+B}\frac{R+c}{R+c+B}.$$

$$\mathbb{P}(R_1 \cap B_n) = \mathbb{P}(R_1)\mathbb{P}(B_n|R_1) = \frac{R}{R+B}\frac{B}{R+c+B}.$$

Sean, ahora $m, n \in \mathbb{N}$ tal que $1 < m < n$,

$$\mathbb{P}(R_{m+1} \cap R_n) = \mathbb{P}(R_{m+1} \cap R_n|R_1)\mathbb{P}(R_1) + \mathbb{P}(R_{m+1} \cap R_n|B_1)\mathbb{P}(B_1) \Rightarrow$$

$$\mathbb{P}(R_{m+1} \cap R_n) = \frac{(R+c)(R+c+c)}{(R+c+B)(R+c+B+c)}\frac{R}{R+B} + \frac{R(R+c)}{(R+B+c)(R+B+c+c)}\frac{B}{R+B} \Rightarrow$$

$$\mathbb{P}(R_{m+1} \cap R_n) = \frac{R(R+c)[(R+c+c)+B]}{(R+c+B)(R+c+B+c)} = \frac{R(R+c)}{(R+c+B)(R+B)}.$$

$$\mathbb{P}(R_{m+1} \cap B_n) = \mathbb{P}(R_{m+1} \cap B_n|R_1)\mathbb{P}(R_1) + \mathbb{P}(R_{m+1} \cap B_n|B_1)\mathbb{P}(B_1) \Rightarrow$$

$$\mathbb{P}(R_{m+1} \cap B_n) = \frac{(R+c)B}{(R+c+B)(R+c+B+c)}\frac{R}{R+B} + \frac{(B+c)R}{(R+B+c)(R+c+B+c)}\frac{B}{R+B} \Rightarrow$$

$$\mathbb{P}(R_{m+1} \cap B_n) = \frac{RB[(R+c)+(B+c)]}{(R+c+B)(R+c+B+c)} = \frac{RB}{(R+c+B)(R+B)}.$$

(iii) Sabemos que $\mathbb{P}(R_n) = \mathbb{P}(R_m) = \dfrac{R}{R+B}$ pero $\mathbb{P}(R_n)\mathbb{P}(R_m) \neq \mathbb{P}(R_{m+1} \cap R_n)$.

Por lo tanto, los eventos R_m y R_n no son independientes.

(iv) $\mathbb{P}(R_3|R_2) = \dfrac{\mathbb{P}(R_3 \cap R_2)}{\mathbb{P}(R_2)} = \dfrac{R(R+c)}{(R+c+B)(R+B)}\dfrac{R+B}{R} = \dfrac{R+c}{R+c+B}$ y

$\mathbb{P}(R_1|R_n) = \dfrac{\mathbb{P}(R_n \cap R_1)}{\mathbb{P}(R_n)} = \dfrac{R(R+c)}{(R+c+B)(R+B)}\dfrac{R+B}{R} = \dfrac{R+c}{R+c+B}$.

6. Se tienen $n+1$ urnas numeradas desde la 0 hasta la n. La urna i contiene i bolillas blancas y $n-i$ negras.

a) Se elige al azar una urna y se extrae de ella una bolilla al azar. Luego se la devuelve a la urna correspondiente. Se repite este procedimiento k veces.

(i) Hallar la probabilidad de que la primera bolilla extraída sea blanca.

(ii) ¿Cuál es la probabilidad de que la primera bolilla extraída provenga de la urna i si ésta es blanca?

(iii) ¿Cuál es la probabilidad de que las k bolillas extraídas sean blancas?

(iv) ¿Son los eventos B_m y B_h independientes para $m \neq h$?

b) Se elige al azar una urna y luego se realizan k extracciones con reposición de la urna elegida.

(i) Hallar la probabilidad de que las k bolillas extraídas sean blancas.

(ii) Si las k bolillas extraídas son blancas, ¿cuál es la probabilidad de que al realizar una nueva extracción de la misma urna esta última bolilla sea blanca?

(iii) ¿Son los eventos B_1 y B_2 independientes?

Solución:

Sean los eventos

$$N_i = \{\text{La i-ésima bolilla extraída es negra}\},\ B_i = \{\text{La i-ésima bolilla extraída es blanca}\} \text{ y}$$

$$U_i = \{\text{La bolilla se extraída de la urna i}\}.$$

a) (i) $\mathbb{P}(B_1) = \displaystyle\sum_{i=0}^{n} \mathbb{P}(B_1|U_i)\mathbb{P}(U_i) = \sum_{i=0}^{n} \frac{1}{n+1}\frac{i}{n-i+i} = \frac{1}{n(n+1)}\sum_{i=0}^{n} i = \frac{1}{2}.$

(ii) $\mathbb{P}\left(\left\{\begin{matrix}\text{la primera bolilla extraída}\\ \text{proviene de la urna i}\end{matrix}\right\}\middle| B_1\right) = \dfrac{\mathbb{P}\left(\left\{\begin{matrix}\text{la primera bolilla extraída}\\ \text{proviene de la urna i}\end{matrix}\right\}\cap B_1\right)}{\mathbb{P}(B_1)}$, entonces

$$\mathbb{P}\left(\left\{\begin{matrix}\text{la primera bolilla extraída}\\ \text{proviene de la urna i}\end{matrix}\right\}\middle| B_1\right) = p\left(B_1 | \left\{\begin{matrix}\text{la primera bolilla extraída}\\ \text{proviene de la urna i}\end{matrix}\right\}\right)\frac{\mathbb{P}(U_i)}{\mathbb{P}(B_1)} = 2\frac{i}{n}\frac{1}{n+1}$$

(iii) $\mathbb{P}\left(\left\{\begin{matrix}\text{las k bolillas extraídas}\\ \text{son blancas}\end{matrix}\right\}\right) = \mathbb{P}\left(\bigcap_{i=1}^{k} B_i\right) = \mathbb{P}\left(B_k | \bigcap_{i=1}^{k-1} B_i\right)\ldots\mathbb{P}\left(B_3 | \bigcap_{i=1}^{2} B_i\right)\mathbb{P}(B_2|B_1)\mathbb{P}(B_1).$

b) (i) $\mathbb{P}\left(\left\{\begin{matrix}\text{las k bolillas extraídas}\\ \text{son blancas}\end{matrix}\right\}\right) = \left(\frac{i}{n}\right)^k.$

(ii) $\mathbb{P}\left(B_{k+1}|B_k \cap B_{k-1}\cap\ldots\cap B_2\cap B_1\right) = \dfrac{\mathbb{P}(B_{k+1}\cap B_k\cap\ldots\cap B_2\cap B_1)}{\mathbb{P}(B_k\cap\ldots\cap B_2\cap B_1)} = \dfrac{i}{n}.$

(iii) Como las extracciones se realizan con reposición, es claro que B_1 y B_2 son independientes.

7. Se extrae al azar una bolilla de una urna que contiene 9 bolillas, de las cuales 3 son blancas, 3 son negras y 3 son rojas. Además, las bolillas de cada color tienen la numeración 1, 2 y 3. Por último, la primer bolilla blanca, la segunda negra y la tercera roja son rayadas. Consideremos los eventos

$$A = \{\text{La bolilla es número 1}\},\ B = \{\text{La bolilla es blanca}\} \text{ y } C = \{\text{La bolilla es rayada}\}.$$

¿Son independientes los eventos A, B y C? ¿Y de a pares?

Solución: Es claro que $\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = \frac{1}{3}$ y $\mathbb{P}(A\cap B\cap C) = \frac{1}{9}$, por lo que los eventos no resultan independientes.

Por otra parte, $\mathbb{P}(A\cap B) = \frac{1}{9} = \mathbb{P}(A)\mathbb{P}(B)$, $\mathbb{P}(A\cap C) = \frac{1}{9} = \mathbb{P}(A)\mathbb{P}(C)$ y $\mathbb{P}(B\cap C) = \frac{1}{9} = \mathbb{P}(B)\mathbb{P}(C)$. Es decir, los eventos son independientes de a pares.

8. Tres jugadores a, b y c se disponen a jugar un torneo de metegol bajo las siguientes reglas: la modalidad de juego es por turnos bajo la consigna de que el ganador de un partido jugará el siguiente con el jugador que se encontraba afuera y aquel que logre dos victorias consecutivas será el ganador del torneo. En cada partido ambos jugadores involucrados tienen la misma probabilidad de ganar. Comienzan jugando a y b.

(i) ¿Cuál es la probabilidad de que el torneo continue indefinidamente? Es decir, hallar la probabilidad de que no haya dos victorias consecutivas de un mismo jugador.

(ii) Calcular la probabilidad de ganar el campeonato para cada uno de los tres jugadores a, b y c. ¿Conviene jugar el primer partido?

Solución:

Sean los eventos

$$A_n = \{\text{a ganó el n-ésimo partido}\},\ B_n = \{\text{b ganó el n-ésimo partido}\} \text{ y } C_n = \{\text{c ganó el n-ésimo partido}\}.$$

por lo tanto, cada torneo se puede expresar como una sucesión $(g_n)_{n\in\mathbb{N}}$ de manera tal que $g_i = A_i, B_i$ o C_i dependiendo de quién haya ganado el partido i-ésimo.

(i)

$$\mathbb{P}(\{\text{2 victorias consecutivas}\}) = \mathbb{P}(g_2 = g_1) + \mathbb{P}(g_3 = g_2|g_2\neq g_1) + \mathbb{P}(g_4 = g_3|g_3\neq g_2 \text{ y } g_2\neq g_1) + \ldots\ldots \Rightarrow$$

$$\mathbb{P}(\{\text{2 victorias consecutivas}\}) = \mathbb{P}(g_2 = g_1) + \mathbb{P}(g_3 = g_2|g_2\neq g_1)\sum_{i=3}^{\infty}\mathbb{P}(g_{i+1} = g_i|g_i\neq g_{n-1} \text{ y } g_{n-1}\neq g_{n-2}) \Rightarrow$$

$$\mathbb{P}(\{\text{2 victorias consecutivas}\}) = \frac{1}{2} + \frac{1}{4} + \sum_{i=3}^{\infty}\left(\frac{1}{2}\right)^n = \sum_{i=0}^{\infty}\left(\frac{1}{2}\right)^n - 1 = \frac{1}{1-\frac{1}{2}} - 1 = 1.$$

De esta manera se deduce que $\mathbb{P}(\{\text{no hay 2 victorias consecutivas}\}) = 0$ y, entonces, alguien va a ganar.

9. (i) Probar que el evento A es independiente de cualquier evento B si y sólo si $\mathbb{P}(A) = 0$ ó 1.

(ii) Mostrar que si A y B son eventos independientes y $A\subset B$ entonces $\mathbb{P}(A) = 0$ ó $\mathbb{P}(B) = 1$.

(iii) Probar que si $A_1,\ldots,A_n$ son eventos independientes y $B_1,\ldots,B_n$ son tales que para cada $i = 1,\ldots,n$ se tiene $B_i = A_i$ ó $B_i = A_i^c$ entonces los eventos $B_1,\ldots,B_n$ también resultan independientes.
Sugerencia: Hacer inducción en la cantidad de complementos involucrados.

(iv) Deducir del ítem anterior que si $A_1, \ldots, A_n$ son eventos independientes entonces se tiene la fórmula

$$\mathbb{P}\left(\bigcup_{k=1}^{n} A_k\right) = 1 - \prod_{k=1}^{n} \left(1 - \mathbb{P}(A_k)\right).$$

Solución:

(i) $\underline{(\Rightarrow)}$: Dado que A es independiente de cualquier evento B, entonces $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$ para todo B.
En particular, si $B = A$ se tiene que $\mathbb{P}(A \cap A) = \mathbb{P}(A)\mathbb{P}(A) = \mathbb{P}^2(A)$ con $\mathbb{P}(A) \geq 0$, entonces $\mathbb{P}(A) = 0$ ó $\mathbb{P}(A) = 1$.

$\underline{(\Leftarrow)}$: Si $\mathbb{P}(A) = 1$, entonces $A = \Omega \Rightarrow \mathbb{P}(A \cap B) = \mathbb{P}(\Omega \cap B) = \mathbb{P}(B) = \mathbb{P}(B)\mathbb{P}(A)$ para todo B.
Por otra parte, si $\mathbb{P}(A) = 0$, entonces $A = \emptyset \Rightarrow \mathbb{P}(A \cap B) = \mathbb{P}(\emptyset \cap B) = \mathbb{P}(\emptyset) = 0 = 0 \times \mathbb{P}(B) = \mathbb{P}(A)\mathbb{P}(B)$ para todo B.
Por lo tanto, A es independiente de cualquier evento B.

(ii) $\mathbb{P}(A) \underbrace{=}_{A \subset B} \mathbb{P}(A \cap B)$ y como A y B son eventos independientes se tiene que $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$. En resumen,

$$\mathbb{P}(A) = \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B) \Rightarrow \mathbb{P}(A) = \mathbb{P}(A)\mathbb{P}(B) \Rightarrow \text{ si } \mathbb{P}(A) \neq 0, 1 = \mathbb{P}(B).$$

Observar que, de la misma forma,

$$\mathbb{P}(A) = \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B) \Rightarrow \mathbb{P}(A) = \mathbb{P}(A)\mathbb{P}(B) \Rightarrow \text{ si } \mathbb{P}(B) \neq 1, \mathbb{P}(A)(1 - \mathbb{P}(B)) = 0 \Rightarrow \mathbb{P}(A).$$

(iii) Sea $B_1, \ldots, B_n$ tales que se pueden ordenar de la siguiente manera

$$B_1 = A_1, B_2 = A_2, \ldots, B_j = A_j, B_{j+1} = A_{j+1}^c, \ldots, B_n^c = A_n^c.$$

Luego,

$$\mathbb{P}\left(\bigcap_{i=1}^{n} B_i\right) = \mathbb{P}\left(\bigcap_{i=1}^{j} B_i \cap \bigcap_{i=j+1}^{n} B_i\right) = \mathbb{P}\left(\bigcap_{i=1}^{j} A_i \cap \bigcap_{i=j+1}^{n} A_i\right)$$

Veamos, primero, que $A_1, A_2, \ldots, A_j, A_{j+1}^c$ son independientes.

$$\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_j \cap A_{j+1}^c) = \mathbb{P}(A_{j+1}^c | A_1 \cap A_2 \cap \ldots \cap A_j)\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_j) \Rightarrow$$

$$\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_j \cap A_{j+1}^c) = [1 - \mathbb{P}(A_{j+1} | A_1 \cap A_2 \cap \ldots \cap A_j)]\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_j) = [1 - \mathbb{P}(A_{j+1})]\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_j)$$

$$\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_j \cap A_{j+1}^c) = \mathbb{P}(A_{j+1}^c)\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_j) = \mathbb{P}(A_{j+1}^c)\mathbb{P}(A_1)\mathbb{P}(A_2) \ldots \mathbb{P}(A_j).$$

Supongamos ahora que $A_1, A_2, \ldots, A_j, A_{j+1}^c, \ldots, A_{j+k}^c$ son independientes. Entonces

$$\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_j \cap A_{j+1}^c \cap \ldots \cap A_{j+k}^c \cap A_{j+k+1}^c) =$$

$$= \mathbb{P}(A_{j+k+1}^c | A_1 \cap A_2 \cap \ldots \cap A_j \cap A_{j+1}^c \cap \ldots \cap A_{j+k}^c)\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_j \cap A_{j+1}^c \cap \ldots \cap A_{j+k}^c) =$$

$$= [1 - \mathbb{P}(A_{j+k+1} | A_1 \cap A_2 \cap \ldots \cap A_j \cap A_{j+1}^c \cap \ldots \cap A_{j+k}^c)]\mathbb{P}(A_1)\mathbb{P}(A_2) \ldots \mathbb{P}(A_j)\mathbb{P}(A_{j+1}^c) \ldots \mathbb{P}(A_{j+k}^c) \Rightarrow$$

$$\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_j \cap A_{j+1}^c \cap \ldots \cap A_{j+k}^c \cap A_{j+k+1}^c) = [1 - \mathbb{P}(A_{j+k+1})]\mathbb{P}(A_1)\mathbb{P}(A_2) \ldots \mathbb{P}(A_j)\mathbb{P}(A_{j+1}^c) \ldots \mathbb{P}(A_{j+k}^c) \Rightarrow$$

$$\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_j \cap A_{j+1}^c \cap \ldots \cap A_{j+k}^c \cap A_{j+k+1}^c) = \mathbb{P}(A_{j+k+1}^c)\mathbb{P}(A_1)\mathbb{P}(A_2) \ldots \mathbb{P}(A_j)\mathbb{P}(A_{j+1}^c) \ldots \mathbb{P}(A_{j+k}^c).$$

(iv) $\mathbb{P}\left(\bigcup_{k=1}^{n} A_k\right) = 1 - \mathbb{P}\left(\bigcap_{k=1}^{n} A_k^c\right) = 1 - \prod_{k=1}^{n} \mathbb{P}(A_k^c) = 1 - \prod_{k=1}^{n} (1 - \mathbb{P}(A_k)).$

10. De un bolillero que contiene n bolillas numeradas de la 1 hasta la n se extrae una al azar. Para cada $1 \leq k \leq n$ definimos el evento $A_k = \{$El número de la bolilla elegida es divisible por k$\}$.

(i) Calcular la probabilidad del evento A_k para todo número natural k divisor de n.

(ii) Probar que los eventos $A_{k_1}, \ldots, A_{k_m}$ son independientes si $k_1, \ldots, k_m$ son divisores de n; coprimos dos a dos.

(iii) Sea φ la función de Euler de la teoría de números definida para cada número natural según la fórmula

$$\varphi(n) = \sharp\{m \in \{1, \ldots, n\} : m \text{ es coprimo con } n\}.$$

Probar la identidad

$$\varphi(n) = n \prod_{\substack{p, \, primo \\ p | n}} \left(1 - \frac{1}{p}\right).$$

Solución:

(i) Si k es divisor de n, entonces existe $t \in \mathbb{N}$ tal que $n = kt$; por lo tanto, hay t bolillas divisibles por k. Luego,

$$\mathbb{P}(A_k) = \frac{t}{n} = \frac{1}{k}.$$

(ii) $A_{k_1}, \ldots, A_{k_m}$ son independientes si y sólo si $\mathbb{P}(A_{k_1} \cap A_{k_2} \cap \ldots \cap A_{k_m}) = \mathbb{P}(A_{k_1})\mathbb{P}(A_{k_2}) \ldots \mathbb{P}(A_{k_m})$, donde $\mathbb{P}(A_{k_i}) = \dfrac{1}{k_i}$ puesto que $k_1, \ldots, k_m$ son divisores de n.

Por otra parte, si $k_1, \ldots, k_m$ son coprimos entonces no tienen divisores en común. Por lo tanto, el número de la bolilla elegida debe ser divisible por cada $k_1, \ldots, k_m$. Luego,

$$\mathbb{P}(A_{k_1} \cap A_{k_2} \cap \ldots \cap A_{k_m}) = \frac{1}{k_1 \ldots k_m} = \frac{1}{k_1}\frac{1}{k_2} \ldots \frac{1}{k_m} = \mathbb{P}(A_{k_1})\mathbb{P}(A_{k_2}) \ldots \mathbb{P}(A_{k_m}).$$

es decir, los eventos $A_{k_1}, \ldots, A_{k_m}$ son independientes.

Práctica 3

Variables Aleatorias Discretas

1. Sea X una variable aleatoria. Calcular $\mathbb{P}(X \in A)$, $\mathbb{P}(X \in B)$, $\mathbb{P}(X \in A \cap B)$ y $\mathbb{P}(X \in B \setminus A)$ en cada uno de los siguientes casos

   (i) X tiene función de distribución acumulada

$$F_X(x) = \begin{cases} 0 & x < -3 \\ \frac{1}{4} & -3 \leq x < 1 \\ \frac{3}{4} & 1 \leq x < 2 \\ 1 & x \geq 2 \end{cases}$$

   y los eventos son $A = [-3, 1]$ y $B = (-2, 2)$

   (ii) X tiene función de distribución acumulada

$$F_X(x) = \begin{cases} 0 & x \leq 0 \\ \frac{x}{4} & 0 < x < 1 \\ \frac{1}{3} & 1 \leq x \leq 2 \\ \frac{x}{6} & 2 < x < 4 \\ \frac{x}{8} + \frac{1}{4} & 4 \leq x \leq 6 \\ 1 & x > 2 \end{cases}$$

   y los eventos son $A = [1, 5]$ y $B = \left(\dfrac{1}{2}, 3\right)$

   (iii) Decidir en cada caso si X es una variable aleatoria discreta.

<u>Solución:</u>

(i) $\mathbb{P}(X \in [-3, 1]) = \mathbb{P}(X \in (-\infty, 1]) - \mathbb{P}(X \in (-\infty, -3]) = F_X(1) - F_X(-3^-) = \dfrac{3}{4} - 0 = \dfrac{3}{4}.$

$\mathbb{P}(X \in (-2, 2)) = \mathbb{P}(X \in (-\infty, 2)) - \mathbb{P}(X \in (-\infty, -2]) = F_X(2^-) - F_X(-2) = \dfrac{3}{4} - \dfrac{1}{4} = \dfrac{1}{2}.$

$\mathbb{P}(X \in A \cap B) = \mathbb{P}(X \in (-2, 1]) = \mathbb{P}(X \in (-\infty, 1]) - \mathbb{P}(X \in (-\infty, -2]) = F_X(1) - F_X(-2) = \dfrac{3}{4} - \dfrac{1}{4} = \dfrac{1}{2}.$

$\mathbb{P}(X \in A \setminus B) = \mathbb{P}(X \in (1, 2)) = \mathbb{P}(X \in (-\infty, 2)) - \mathbb{P}(X \in (-\infty, 1]) = F_X(2^-) - F_X(1) = \dfrac{3}{4} - \dfrac{3}{4} = 0.$

(ii) $\mathbb{P}(X \in [1, 5]) = \mathbb{P}(X \in (-\infty, 5]) - \mathbb{P}(X \in (-\infty, 1]) = F_X(5) - F_X(1^-) = \left(\dfrac{5}{8} + \dfrac{1}{4}\right) - \lim\limits_{x \to 1^-} \dfrac{x}{4} = \dfrac{5}{8} + \dfrac{1}{4} - \dfrac{1}{4} = \dfrac{5}{8}.$

$\mathbb{P}\left(X \in \left(\dfrac{1}{2}, 3\right)\right) = \mathbb{P}(X \in (-\infty, 3)) - \mathbb{P}(X \in (-\infty, \frac{1}{2}]) = F_X(3^-) - F_X\left(\dfrac{1}{2}\right) = \lim\limits_{x \to 3^-} \dfrac{x}{6} - \dfrac{1}{8} = \dfrac{1}{2} - \dfrac{1}{8} = \dfrac{3}{8}.$

$\mathbb{P}(X \in A \cap B) = \mathbb{P}(X \in [1, 3)) = \mathbb{P}(X \in (-\infty, 3)) - \mathbb{P}(X \in (-\infty, 1)) = F_X(3^-) - F_X(1^-) = \lim\limits_{x \to 3^-} \dfrac{x}{6} - \lim\limits_{x \to 1^-} \dfrac{x}{4} \Rightarrow$
$\mathbb{P}(X \in A \cap B) = \dfrac{1}{2} - \dfrac{1}{4} = \dfrac{1}{4}.$

$\mathbb{P}(X \in A \setminus B) = \mathbb{P}\left(X \in \left(\dfrac{1}{2}, 1\right)\right) = \mathbb{P}(X \in (-\infty, 1)) - \mathbb{P}(X \in (-\infty, \frac{1}{2}]) = F_X(1^-) - F_X\left(\dfrac{1}{2}\right) = \lim\limits_{x \to 1^-} \dfrac{x}{4} - \dfrac{1}{8} = \dfrac{1}{8}.$

(iii) En el caso (i), X es discreta. Mientas que en el caso (ii), X es mixta.

2. Se dispone de un tiro al blanco, cuyo blanco está compuesto por un círculo de radio 3. Podemos pensar al resultado del tiro como un experimento aleatorio, por simplicidad supondremos que el tiro siempre impactará en el círculo. Si ponemos el centro en el origen de $\mathbb{R}^2$, el espacio muestral del experimento es $\Omega = \{(x,y) : x^2 + y^2 < 9\}$. Sea $\mathcal{F}$ la σ-álgebra de los borelianos en $\mathbb{R}^2$. Asumamos que la probabilidad de que un dardo caiga en una cierta región A es proporcional a su área $|A|$. Es decir,
$$\mathbb{P}(A) = \frac{|A \cap \Omega|}{9\pi}.$$
Luego disponemos de un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$.

   (i) Se establece un sistema de puntajes para este juego del siguiente modo. El blanco se divide en tres círculos concéntricos C_1, C_2 y C_3 centrados en el origen y de radios 1, 2 y 3. Estos círculos dividen al blanco en tres anillos A_1, A_2 y A_3 donde
   $$A_k = \{(x,y) : k-1 \leq \sqrt{x^2+y^2} < k\}.$$
   Un jugador tira su dardo. Si el dardo cae en A_i entonces se lleva $4-i$ puntos (cuánto más cerca del centro está el tiro, más puntaje se lleva). Hallar la función de distribución acumulada de la variable aleatoria $X = \{\text{puntaje asignado a un tiro}\}$. ¿Es X una variable aleatoria discreta?

   (ii) El modo de asignar los puntajes se cambia. Ahora el puntaje obtenido será tres menos la distancia entre el punto (x,y) donde el dardo impacta y el centro del blanco. Si llamamos Y a la variable aleatoria que da el puntaje nuevo, tenemos
   $$Y = 3 - \sqrt{x^2+y^2}.$$
   Hallar la función de distribución acumulada de la variable aleatoria Y. ¿Es Y una variable aleatoria discreta? ¿Es Y una variable aleatoria continua?

   (iii) Finalmente, se propone un tercer modo de asignar puntaje para el juego de los dardos, que denotaremos por la variable aleatoria Z, y que consiste en lo siguiente: si el dardo cae en $A_1 = \{(x,y) : 0 \leq \sqrt{x^2+y^2} < 1\}$, entonces el jugador recibe 3 puntos, sino, el puntaje se define por 3 menos la distancia entre el punto (x,y) donde el dardo impacta y el centro del blanco. O sea,
   $$Z = \begin{cases} 3 & \sqrt{x^2+y^2} < 1 \\ 3 - \sqrt{x^2+y^2} & \text{en caso contrario.} \end{cases}$$
   Hallar la función de distribución acumulada de la variable aleatoria Z. ¿Es Z una variable aleatoria discreta? ¿Es Z una variable aleatoria continua?

Solución:

(i) Es claro que $R_X = \{1,2,3\}$ puesto que si el dado cae en A_1 el puntaje es 3, si el dado cae en A_2 el puntaje es 2 y, por último, si cae en A_3 el puntaje es 1.

$$\mathbb{P}(X \leq 1) = \underbrace{\mathbb{P}(X < 1)}_{=0} + \mathbb{P}(X = 1) = \mathbb{P}((x,y) \in A_3)) = \frac{|A_3 \cap \Omega|}{9\pi} = \frac{\pi(9-4)}{9\pi} = \frac{5}{9}$$

$$\mathbb{P}(X < 2) = \mathbb{P}(X \leq 1) = \frac{5}{9} \text{ y } \mathbb{P}(X = 2) = \mathbb{P}((x,y) \in A_2) = \frac{|A_2 \cap \Omega|}{9\pi} = \frac{\pi(4-1)}{9\pi} = \frac{3}{9}$$

$$\mathbb{P}(X < 2) = \mathbb{P}(X \leq 2) = \frac{5}{9} + \frac{3}{9} = \frac{8}{9} \text{ y } \mathbb{P}(X = 3) = \mathbb{P}((x,y) \in A_1) = \frac{|A_1 \cap \Omega|}{9\pi} = \frac{\pi}{9\pi} = \frac{1}{9}$$

Luego, $F_X(x) = \begin{cases} 0 & x < 1 \\ \frac{5}{9} & 1 \leq x < 2 \\ \frac{8}{9} & 2 \leq x < 3 \\ 1 & x \geq 3 \end{cases}$ y X resulta ser una variable aleatoria discreta.

(ii) Es claro que la variable Y es continua tal que $R_Y = (0,3]$.

$$\mathbb{P}(Y \leq k) = \mathbb{P}(3 - \sqrt{x^2+y^2} \leq k) = \mathbb{P}(\sqrt{x^2+y^2} \geq 3-k) = 1 - \mathbb{P}(\sqrt{x^2+y^2} \leq 3-k).$$

es decir que se busca los puntos (x,y) tales que pertenecen al círculo de radio $3-k$,

$$\mathbb{P}(Y \leq k) = 1 - \frac{\pi(3-k)^2}{9\pi} = 1 - \left(1 - \frac{k}{3}\right)^2 \Rightarrow F_Y(y) = 1 - \left(1 - \frac{y}{3}\right)^2, \text{ si } y \in (0,3)$$

Entonces, $F_Y(y) = \begin{cases} 0 & y \leq 0 \\ 1 - \left(1 - \dfrac{y}{3}\right)^2 & 0 < y < 3 \\ 1 & y \geq 3 \end{cases}$.

(iii) Es claro que $R_Z = (0,2] \cup \{3\}$ puesto que si el dado cae en A_1 el puntaje será 3, mientras que si no cae allí el puntaje está dado por $3 - d((x,y);(0,0))$ (tal que $d((x,y);(0,0)) \in [0,2]$).

Por otra parte, $\mathbb{P}(z=3) = \dfrac{1}{9}$ mientras que dado $k \in (0,2]$ se tiene que $\mathbb{P}(Z=k) = \dfrac{8}{9}$ y $\mathbb{P}(Z=0) = 0$. Puesto que la función de densidad no integra 1, entonces Z no es una variable aleatoria.

3. La probabilidad de acertar al blanco en un juego es $\dfrac{1}{5}$. Se tiran 10 tiros de manera independientes.
   (i) Hallar la probabilidad de que se produzcan al menos dos aciertos.
   (ii) Hallar la probabilidad condicional de que se produzcan al menos dos aciertos, asumiendo que se produjo al menos un acierto.

Solución: Sea $X = \{\text{cantidad de aciertos}\}$ donde $X \sim Bi\left(10, \dfrac{1}{5}\right)$,

(i)

$$\mathbb{P}(\{\text{al menos dos aciertos}\}) = \mathbb{P}(X \geq 2) = 1 - \mathbb{P}(A \leq 1) = 1 - \left(\begin{array}{c} 10 \\ 0 \end{array}\right)\left(\frac{1}{5}\right)^0\left(\frac{4}{5}\right)^{10-0} - \left(\begin{array}{c} 10 \\ 1 \end{array}\right)\left(\frac{1}{5}\right)^1\left(\frac{4}{5}\right)^{10-1}$$

(ii)

$$\mathbb{P}\left(\{\text{al menos 2 aciertos}\} | \begin{Bmatrix} \text{se produjo al} \\ \text{menos un acierto} \end{Bmatrix}\right) = \frac{\mathbb{P}(X \geq 2 \text{ y } X \geq 1)}{\mathbb{P}(X \geq 1)} = \frac{\mathbb{P}(X \geq 2)}{\mathbb{P}(X \geq 1)} = \frac{1 - \mathbb{P}(X \leq 2)}{1 - \mathbb{P}(X \leq 0)}$$

4. (i) Sea $X \sim Bi(n,p)$ e $Y \sim \mathcal{G}(p)$. Demostrar que $\mathbb{P}(X=0) = \mathbb{P}(Y>n)$.
   (ii) Hallar el número de niños que debe tener un matrimonio para que la probabilidad de tener al menos un varón sea mayor o igual a $\dfrac{8}{9}$.

Solución:

(i) $\mathbb{P}(X=0) = \left(\begin{array}{c} n \\ 0 \end{array}\right) p^0(1-p)^{n-0} = (1-p)^n$ y

$$\mathbb{P}(Y>n) = 1 - \mathbb{P}(Y \leq n) = 1 - \sum_{k=1}^{n} p(1-p)^{k-1} = 1 - \left[p\frac{1-(1-p)^{n-1+1}}{1-(1-p)}\right] = (1-p)^n$$

por lo tanto, $\mathbb{P}(X=0) = \mathbb{P}(Y>n)$.

(ii) Sea $X = \begin{cases} 0 & \text{si nace varón} \\ 1 & \text{si nace mujer} \end{cases}$, entonces $X \sim Bi(1,p)$ donde $p = \mathbb{P}(X=1)$ (es decir, la probabilidad de que nazca una mujer). Por lo tanto, $1-p$ es la probabilidad de que nazca un varón y, entonces,

$$\mathbb{P}(Y \geq n) \geq \frac{8}{9} \Rightarrow (1-p)^n \geq \frac{8}{9} \Rightarrow n\ln(1-p) \geq \ln\left(\frac{8}{9}\right) \Rightarrow n \geq \ln\left(\frac{8}{9}\right)\frac{1}{\ln(1-p)}.$$

5. En una corte judicial una mesa de 12 jurados debe decidir el destino de un acusado llevado a juicio en dicha corte. Éste será declarado culpable y condenado si al menos 8 de los 12 jurados lo votan culpable. Supongamos que el acusado tiene probabilidad α de haber cometido el delito del que se lo acusa y que cada jurado decide su voto de manera independiente, con probabilidad θ de tomar la decisión correcta. ¿Cuál es la probabilidad de que el acusado sea condenado?

Solución: Sean los eventos

$$X = \{\text{cantidad de jurados que lo condenan correctamente}\} \sim Bi(12,\theta) \text{ y}$$

$$Y = \{\text{cantidad de jurados que lo condenan incorrectamente}\} \sim Bi(12, 1-\theta).$$

y es claro que

$$\mathbb{P}(\{\text{condenado}\}) = \mathbb{P}(\{\text{condenado}|\text{es culpable}\})\underbrace{\mathbb{P}(\{\text{es culpable}\})}_{=\alpha} + \mathbb{P}(\{\text{condenado}|\text{no es culpable}\})\underbrace{\mathbb{P}(\{\text{no es culpable}\})}_{=1-\alpha}$$

luego,

$$\mathbb{P}(\{\text{condenado}|\text{es culpable}\}) = \mathbb{P}(X \geq 8) = \sum_{i=8}^{12} \binom{12}{i} \theta^i (1-\theta)^{12-i} \text{ y}$$

$$\mathbb{P}(\{\text{condenado}|\text{no es culpable}\}) = \mathbb{P}(Y \geq 8) = \sum_{i=8}^{12} \binom{12}{i} (1-\theta)^i \theta^{12-i}$$

entonces,

$$\mathbb{P}(\{\text{condenado}\}) = \sum_{i=8}^{12} \binom{12}{i} \left[\alpha\theta^i(1-\theta)^{12-i} + (1-\alpha)(1-\theta)^i\theta^{12-i}\right]$$

6. Sea X una variable aleatoria con distribución geométrica de parámetro p.
   (i) Calcular la función de distribución acumulada de X.
   (ii) Hallar la probabilidad de que X sea impar. ¿Es mayor que la probabilidad de que X sea par?

Solución:

(i) $\mathbb{P}_X(X \leq x) = \sum_{k=1}^{x} p(1-p)^{k-1} = p\sum_{k=1}^{x}(1-p)^{k-1} = p\left[\frac{1-(1-p)^{x-1+1}}{1-(1-p)}\right] = 1-(1-p)^x.$

(ii) $$\mathbb{P}(\text{X es impar}) = \sum_{\substack{k=1\\ k\ impar}}^{x} p(1-p)^{k-1} = \begin{cases} p\displaystyle\sum_{j=0}^{\frac{x-1}{2}}(1-p)^{2j+1} & x \text{ es impar} \\ p\displaystyle\sum_{j=0}^{\frac{x}{2}-1}(1-p)^{2j+1} & x \text{ es par}\end{cases}, \text{ entonces}$$

$$\mathbb{P}(\text{X es impar}) = \lim_{x\to\infty}\left[p(1-p)\frac{1-\left((1-p)^2\right)^{\frac{x-1}{2}+1}}{1-(1-p)^2}\right] = p(1-p)\frac{1}{1-(1-p)^2} = \frac{1-p}{2-p}.$$

7. Se dice que una variable aleatoria X a valores en $\mathbb{N}_0$ tiene la propiedad de falta de memoria discreta si para todo par de números $s, t \in \mathbb{N}$ se verifica
$$\mathbb{P}(X \geq s+t|X > t) = \mathbb{P}(X \geq s).$$
Probar que una variable aleatoria X a valores en $\mathbb{N}_0$ posee la propiedad de falta de memoria discreta si y sólo si tiene distribución geométrica de parámetro $p = \mathbb{P}(X = 1)$.
Sugerencia: Una posibilidad es probar primero por inducción que $\mathbb{P}(X > n) = [\mathbb{P}(X > 1)]^n$ para todo $n \in \mathbb{N}$.

Solución:

$\underline{(\Rightarrow)}$: Probemos, primero, que $\mathbb{P}(X > n) = (\mathbb{P}(X > 1))^n$ para todo $n \in \mathbb{N}$. Luego,

$$\text{si } n = 0 \Rightarrow \mathbb{P}(X > 0) = 1 = [\mathbb{P}(X > 1)]^0 = 1.$$

$$\text{si } n = 1 \Rightarrow \mathbb{P}(X > 1) = [\mathbb{P}(X > 1)]^1.$$

$$\mathbb{P}(X > n) = \mathbb{P}(X > n-1)\mathbb{P}(X > 1) \underbrace{=}_{H.I.} [(\mathbb{P}(X > 1))]^{n-1}\mathbb{P}(X > 1) = [\mathbb{P}(X > 1)]^n.$$

Entonces

$$\mathbb{P}(X \geq s+t|X > t) = \mathbb{P}(X \geq s) \Rightarrow \mathbb{P}(X \geq s+t|X > t) = \mathbb{P}(X \geq s)\mathbb{P}(X > t) \Rightarrow X \sim \mathcal{G}(1-\mathbb{P}(X > 1)).$$

$\underline{(\Leftarrow)}$: Sea $X \sim \mathcal{G}(\mathbb{P}(X > 1)) \Rightarrow \mathbb{P}(X > k) = (1-p)^p$, luego

$$\mathcal{P}(X \geq s+t|X > t) = \frac{\mathbb{P}(X \geq s+t \text{ y } X > t)}{\mathbb{P}(X > t)} = \frac{\mathbb{P}(X \geq s+t)}{X > t} = \frac{(1-p)^{s+t+1}}{(1-p)^t} = (1-p)^{s+1} = \mathbb{P}(X > s).$$

8. **El problema de los puntos.** Calcular la probabilidad de obtener k éxitos antes de r fracasos en una sucesión de ensayos Bernoulli independientes con probabilidad de éxito p.

Solución: Sea

$$X = \{\text{cantidad de experimientos que se hicieron hasta que se prozcan k-éxitos}\}$$

entonces $X \sim BN(n,p)$ con $n = k, k+1, \dots$ y

$$\mathbb{P}(k \leq X \leq b + r - 1) = \sum_{n=k}^{k+r-1} \binom{n-1}{k-1} p^k (1-p)^{n-k}.$$

9. Un comprador de componentes eléctricas los adquiere en lotes de tamaño 10. Antes de comprar inspecciona 3 componentes elegidas al azar del lote, y lo aceptará si ninguna de las componentes revisadas está fallada. Encontrar la proporción de lotes que rechazará el comprador si el 30 % de los lotes tiene 4 componentes defectuosas y el 70 % sólo una.

Solución: Sean el evento $A = \{\text{el comprador acepta el lote}\}$, $B = \{\text{en el lote hay 4 componentes dañadas}\}$ y $C = \{\text{en el lote hay una componente dañada}\}$. Entonces

$$\mathbb{P}(A) = \mathbb{P}(A|B).\mathbb{P}(B) + \mathbb{P}(A|C)\mathbb{P}(C) = \frac{\binom{4}{0}\binom{6}{3}}{\binom{10}{3}} 0,30 + \frac{\binom{9}{3}\binom{1}{0}}{\binom{10}{3}} 0,70.$$

Finalmente, la probabilidad buscada es $1 - \mathbb{P}(A)$.

10. Un fabricante ofrece relojes en lotes de 50, en el cual hay buenos, recuperables y desechables. Para ver si adquiere el lote, un comprador toma una muestra de 8 y los revisa. En realidad hay 20 buenos, 25 recuperables y 5 desechables en el lote elegido.

  (i) Hallar la probabilidad de que compre el lote, si el criterio para decidir es que entre los revisados no haya ninguno desechable.

  (ii) Hallar la probabilidad de que compre el lote, si el criterio para decidir es que entre los revisados al menos 5 deben ser buenos y ninguno desechable.

Solución:

(i) $\mathbb{P}(\{\text{se aquiere el lote}\}) = \mathbb{P}(\text{desechables} = 0 \text{ y recuperables} + \text{buenos} = 8)) = \dfrac{\binom{5}{0}\binom{45}{8}}{\binom{50}{8}}$

(ii)

$$\begin{aligned}
&\mathbb{P}(\{\text{se aquiere el lote}\}) = \mathbb{P}(\text{desechables} = 0 \text{ y } 5 \leq \text{buenos} \leq 8)) = \\
&= \mathbb{P}(\text{desechables} = 0, \text{ recuperables} = 3 \text{ y buenos} = 5)) + \mathbb{P}(\text{desechables} = 0, \text{ recuperables} = 2 \text{ y buenos} = 6)) \\
&+ \mathbb{P}(\text{desechables} = 0, \text{ recuperables} = 1 \text{ y buenos} = 7)) + \mathbb{P}(\text{desechables} = 0, \text{ recuperables} = 0 \text{ y buenos} = 8))
\end{aligned}$$

$$= \frac{1}{\binom{50}{8}} \left[\binom{5}{0}\binom{25}{3}\binom{20}{5} + \binom{5}{0}\binom{25}{2}\binom{20}{6} + \binom{5}{0}\binom{25}{1}\binom{20}{7} + \binom{5}{0}\binom{25}{0}\binom{20}{8} \right].$$

11. Un minorista ha verificado que la demanda de cajones es una variable aleatoria con distribución de Poisson de parámetro $\lambda = 2$ cajones por semana. El minorista completa su stock los lunes por la mañana a fin de tener 4 cajones al principio de la semana, y no vuelve a completar su stock sino hasta la semana siguiente. Al efectuar un análisis de la actividad de su negocio, se le plantean las siguientes preguntas:

  (i) ¿Cuál es la probabilidad de vender todo su stock durante una semana?

  (ii) ¿Cuál es la probabilidad de vender todo su stock semanal en al menos dos de las semanas del mes?

  (iii) ¿Cuál es la probabilidad de que en una semana sea incapaz de cumplir por lo menos con un pedido?

  (iv) ¿Cuál es el mínimo número de cajones con los que deberá iniciar cada semana para que la probabilidad de cumplir con todos los pedidos sea mayor o igual a 0,99?

(v) ¿Cuál es la distribución del número de cajones vendidos por semana?

Solución:

(i) Sea $X = \{$cantidad de cajones que se demandan por semana$\}$ tal que $X \sim \mathcal{P}(2)$, luego

$$\mathbb{P}(\{\text{vendor todo el stock durante la semana}\}) = \mathbb{P}(X \geq 4) = 1 - \mathbb{P}(X < 4) = 1 - \sum_{k=0}^{3} \frac{2^k}{k!} e^{-2} = 1 - e^{-2}\left[1 + 2 + 2 + \frac{4}{3}\right].$$

(ii) Sean los eventos $X_i = \{$cantidad de cajones que se demandan en la semana i$\}$ con $i = 1, 2, 3, 4$ donde $X_i \sim \mathcal{P}(2)$ y como cada lunes se arranca con un completo, se tiene que, X_1, X_2, X_3 y X_4 son iid.

$$\mathbb{P}(X_i \geq 4 \text{ y } X_j \geq 4) \underbrace{=}_{i \neq j} \mathbb{P}(X_i \geq 4)\mathbb{P}(X_j \geq 4) = [1 - \mathbb{P}(X_i < 4)][\mathbb{P}(X_j < 4)] = [1 - \mathbb{P}(X < 4)]^2.$$

$$\mathbb{P}(X_i \geq 4 \text{ y } X_j \geq 4 \text{ y } X_k \geq 4) \underbrace{=}_{\substack{i \neq j \\ j \neq k \\ k \neq i}} \mathbb{P}(X_i \geq 4)\mathbb{P}(X_j \geq 4)\mathbb{P}(X_k \geq 4) = [1 - \mathbb{P}(X_i < 4)][1 - \mathbb{P}(X_j < 4)][1 - \mathbb{P}(X_k < 4) \Rightarrow$$

$$\mathbb{P}(X_i \geq 4 \text{ y } X_j \geq 4 \text{ y } X_k \geq 4) \underbrace{=}_{\substack{i \neq j \\ j \neq k \\ k \neq i}} [1 - \mathbb{P}(X < 4)]^3.$$

$$\mathbb{P}(X_i \geq 4 \text{ y } X_j \geq 4 \text{ y } X_k \geq 4 \text{ y } X_n \geq 4) \underbrace{=}_{\substack{i \neq j,k,n \\ j \neq k,n \\ k \neq n}} [1 - \mathbb{P}(X < 4)]^4.$$

luego,

$$\mathbb{P}\left(\left\{\begin{matrix} \text{vender todo el stock semanal en} \\ \text{al menos dos de las semana del mes} \end{matrix}\right\}\right) = 6[1 - \mathbb{P}(X < 4)]^2 + 4[1 - \mathbb{P}(X < 4)]^3 + 1[1 - \mathbb{P}(X < 4)]^4.$$

(iii) Sean los eventos $X_i = \{$cantidad de cajones que se demandan en la semana i$\}$ con $i = 1, 2, 3, 4$ donde $X_i \sim \mathcal{P}(2)$ y como cada lunes se arranca con un completo, se tiene que, X_1, X_2, X_3 y X_4 son iid.

$$\mathbb{P}\left(\left\{\begin{matrix} \text{en una semana sea incapaz} \\ \text{de cumplir por lo menos con un pedido} \end{matrix}\right\}\right) = \mathbb{P}(X_i > 4), \quad \text{para algún } i.$$

entonces

$$\mathbb{P}\left(\left\{\begin{matrix} \text{en una semana sea incapaz} \\ \text{de cumplir por lo menos con un pedido} \end{matrix}\right\}\right) = \sum_{k=5}^{\infty} \mathbb{P}(X \leq k) = \sum_{k=5}^{\infty} \frac{2^k}{k!} e^{-2} = 1 - e^{-2}\left[1 + 2 + 2 + \frac{4}{3} + \frac{2}{3}\right].$$

(iv) Sea k la cantidad mínima de cajones que hay en stock, entonces para que pueda cumplir con todos los pedidos semanales es necesario que $X_i \leq k$. Luego,

$$\mathbb{P}(X_i \leq k) \geq 0{,}99 \Rightarrow \sum_{i=0}^{k} \frac{2^i}{i!} e^{-2} \geq 0{,}99.$$

(v) Sea $Y = \{$cantidad de cajones vendidos por semana$\}$, entonces

$$\left.\begin{matrix} si\ X < 4 \Rightarrow Y = X \\ si\ X \geq 4 \Rightarrow Y = 4 \end{matrix}\right\} \Rightarrow \mathbb{P}_Y(y = k) = \begin{cases} \frac{2^k}{k!} e^{-2} & k \leq 3 \\ 1 - e^{-2} \displaystyle\sum_{i=0}^{3} \frac{2^i}{i!} & k = 4 \end{cases}$$

12. Se tienen 3 fuentes radiactivas F_1, F_2 y F_3. El número de partículas que emite cada fuente por hora es una variable con distribución $\mathbb{P}(\lambda_i)$, siendo $\lambda_1 = 2$, $\lambda_2 = 3$ y $\lambda_3 = 4$. Un investigador elige una fuente al azar y observa que ésta emite 4 partículas en una hora. Encontrar la probabilidad de que haya elegido la fuente F_2.

Solución: Sea $X_i = \{$cantidad de partículas emitidas por la fuente $F_i\}$ tal que $X_i \sim \mathcal{P}(\lambda_i)$ donde $\lambda_1 = 2$, $\lambda_2 = 3$ y $\lambda_3 = 4$.

$$\mathbb{P}(\{\text{elegir } F_2 | \text{emite 4 partículas en una hora}\}) = \frac{\mathbb{P}(\{\text{elegir } F_2 \text{ y emite 4 partículas en una hora}\})}{\mathbb{P}(\{\text{emite 4 partículas en una hora}\})} \Rightarrow$$

$$\mathbb{P}(\{\text{elegir } F_2 | \text{emite 4 partículas en una hora}\}) = \frac{\mathbb{P}(X_2 = 4)}{\frac{1}{3}[\mathbb{P}(X_1 = 4 + \mathbb{P}(X_2 = 4) + \mathbb{P}(X_3 = 4))]}.$$

13. Dada X una variable aleatoria discreta y $x_0 \in R_X$ decimos que x_0 es un valor más probable para X si

$$\mathbb{P}_X(x_0) = \sup_{x \in R_X} \mathbb{P}_X(x).$$

(i) Probar que toda variable aleatoria discreta admite al menos un valor más probable.

(ii) Verificar que $[(n+1)p]$ es un valor más probable para la distribución binomial de parámetros n y p.

(iii) Verificar que $[\lambda]$ es un valor más probable para la distribución de Poisson de parámetro λ.

(iv) Verificar que $\left[\dfrac{(m+1)(n+1)}{N+2}\right]$ es un valor más probable para la distribución hipergeométrica $\mathcal{H}(N,m,n)$.

Sugerencia: Para resolver los últimos tres items puede serle útil estudiar los cocientes $\dfrac{\mathbb{P}_X(k)}{\mathbb{P}_X(k-1)}$ con $k \in \mathbb{N}$.

Solución:

(i) Como X es discreta, entonces R_X es un conjunto a lo sumo numerable tal que $0 \leq \mathbb{P}_X(x_i) \leq 1$ para todo $x_i \in R_X$.

Supongamos, entonces, que el conjunto $A = \{\mathbb{P}_X(x_i) | x_i \in R_X\}$ no tiene supremo, entonces no es acotado superiormente puesto que $A \neq \emptyset$. Pero eso es una contradicción puesto que $\mathbb{P}_X(x_i) \leq 1$ para todo $x_i \in R_X$.

Entonces, el conjunto $A = \{\mathbb{P}_X(x_i) | x_i \in R_X\}$ tiene supremo.

(ii) $$\frac{\mathbb{P}_X(k)}{\mathbb{P}_X(k-1)} = \frac{\binom{n}{k} p^k (1-p)^{n-k}}{\binom{n}{k-1} p^{k-1}(1-p)^{n-k+1}} = \frac{p}{1-p}\frac{n-k+1}{k}.$$

Por otra parte, $\mathbb{P}_X(k)$ es creciente para $\dfrac{p}{1-p}\dfrac{n-k+1}{k} \geq 1$ si y sólo si $p(n-k+1) \geq k(1-p)$ si y sólo si $p(n+1) \geq k$.
Y, es decreciente para $\dfrac{p}{1-p}\dfrac{n-k+1}{k} \leq 1$ si y sólo si $p(n-k+1) \leq k(1-p)$ si y sólo si $p(n+1) \leq k$.

Por lo tanto, un valor más probable resulta cuando $k = [p(n+1)]$.

(iii) $$\frac{\mathbb{P}_X(k)}{\mathbb{P}_X(k-1)} = \frac{\lambda^k}{k!}e^{-\lambda}\frac{(k-1)!}{\lambda^{k-1}e^{-\lambda}} = \frac{\lambda}{k}.$$

Por otra parte, $\mathbb{P}_X(k)$ es creciente para $\dfrac{\lambda}{k} \geq 1$ si y sólo si $\lambda \geq k$. Y, es decreciente para $\dfrac{\lambda}{k} \leq 1$ si y sólo si $\lambda \leq k$.

Por lo tanto, un valor más probable resulta cuando $k = [\lambda]$.

(iv) $$\frac{\mathbb{P}_X(k)}{\mathbb{P}_X(k-1)} = \frac{\binom{r}{k}\binom{N-r}{m-k}}{\binom{N}{m}} \frac{\binom{N}{m}}{\binom{r}{k-1}\binom{N-r}{m-k+1}} = \frac{r-k+1}{k}\frac{m-k+1}{N-r-m+k}.$$

Por otra parte, $\mathbb{P}_X(k)$ es creciente para $\dfrac{r-k+1}{k}\dfrac{m-k+1}{N-r-m+k} \geq 1$ si y sólo si $-2k + (r+1)(+1) \geq kN$ si y sólo si $\dfrac{(r+1)(m+1)}{N+2} \geq k$. Y, es decreciente para $\dfrac{r-k+1}{k}\dfrac{m-k+1}{N-r-m+k} \leq 1$ si y sólo si $-2k + (r+1)(+1) \leq kN$ si y sólo si $\dfrac{(r+1)(m+1)}{N+2} \leq k$. .

Por lo tanto, un valor más probable resulta cuando $k = \left[\dfrac{(r+1)(m+1)}{N+2}\right]$.

14. Dada $X : \Omega \to \mathbb{R}$ una variable aleatoria no negativa, probar que existe una sucesión $(X_n)_{n\in\mathbb{N}}$ de variables aleatorias con imagen de cardinal finito definidas sobre el mismo Ω tal que para cada $\omega \in \Omega$ se verifica $0 \leq X_n(\omega) \leq X_{n+1}(\omega)$ para todo $n \in \mathbb{N}$ y $\lim\limits_{n\to\infty} X_n(\omega) = X(\omega)$.

Solución:

Práctica 4

Variables Aleatorias Continuas

1. El diámetro D expresado en decímetros del tronco de una cierta especie de árboles es una variable aleatoria continua (o absolutamente continua) con función de densidad

$$f_D(X) = kx\mathbb{1}_{(0,10)}(x)$$

   (i) Hallar el valor de la constante k.
   (ii) ¿Cuál es la probabilidad de que el diámetro de un árbol de esa especie elegido al azar mida entre 4 y 6 decímetros?
   (iii) ¿Cómo se modifica la probabilidad anterior si se sabe que el diámetro mide más de 5 decímetros?
   (iv) En un área del bosque hay 3 árboles de esa especie. Calcular la probabilidad de que exactamente 2 de ellos tengan diámetro entre 4 y 6 decímetros.
   (v) ¿Cuántos árboles habría que muestrear en el bosque para que la probabilidad de encontrar al menos uno cuyo diámetro mida entre 4 y 6 decímetros sea no menor a $0,99$?

Solución:

(i) Sabemos que $\int_{\mathbb{R}} f_D(x)dx = 1$, entonces

$$1 = \int_0^{10} kxdx = k\left.\frac{x^2}{2}\right|_0^{10} = k\frac{100}{2} \Rightarrow k = \frac{1}{50}.$$

Por lo tanto,

$$f_D(X) = \frac{1}{50}x\mathbb{1}_{(0,10)}(x).$$

(ii) Buscamos $\mathbb{P}(4 \leq D \leq 6)$,

$$\mathbb{P}(4 \leq D \leq 6) = \int_4^6 \frac{1}{50}x\mathbb{1}_{(0,10)}(x)dx = \frac{1}{50}\int_4^6 xdx = \frac{1}{50}\left.\frac{x^2}{2}\right|_4^6 = \frac{36-16}{100} = \frac{20}{100} = \frac{2}{10}.$$

(iii) Buscamos $\mathbb{P}(4 \leq D \leq 6|D \geq 5)$,

$$\mathbb{P}(4 \leq D \leq 6|D \geq 5) = \frac{\mathbb{P}((4 \leq D \leq 6 \text{ y } D \geq 5)}{\mathbb{P}(D \geq 5)} = \frac{\mathbb{P}(5 \leq D \leq 6)}{\mathbb{P}(D \geq 5)}.$$

$$\mathbb{P}(5 \leq D \leq 6) = \frac{1}{50}\left.\frac{x^2}{2}\right|_5^6 = \frac{36-25}{100} = \frac{11}{100}$$

$$\mathbb{P}(D \geq 5) = \frac{1}{50}\left.\frac{x^2}{2}\right|_5^{10} = \frac{100-25}{100} = \frac{75}{100}.$$

luego, $\mathbb{P}(4 \leq D \leq 6|D \geq 5) = \dfrac{\mathbb{P}(5 \leq D \leq 6)}{\mathbb{P}(D \geq 5)} = \dfrac{11}{75}.$

(iv) $\mathbb{P}(\{\text{dos árboles tengan diámetro entre 4 y 6}\}) = \begin{pmatrix} 3 \\ 2 \end{pmatrix} (\mathbb{P}(4 \leq D \leq 6))^2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} (1 - \mathbb{P}(4 \leq D \leq 6)) \Rightarrow$

$$\mathbb{P}(\{\text{dos árboles tengan diámetro entre 4 y 6}\}) = 3\left(\frac{2}{10}\right)^2\left(1 - \frac{2}{10}\right).$$

(v) $\mathbb{P}(\{\text{al menos un árbol tenga diámetro entre 4 y 6}\}) \geq 0{,}99 \Leftrightarrow \mathbb{P}(\{\text{ningún árbol tiene diámetro entre 4 y 6}\}) < 0{,}01$, luego

$$\mathbb{P}(\{\text{ningún árbol tiene diámetro entre 4 y 6}\}) < 0{,}01 \Rightarrow \left(1-\frac{2}{10}\right)^n < 0{,}01 \Rightarrow$$

$$\left(\frac{8}{10}\right)^n < 0{,}01 \Rightarrow (0{,}8)^n < 0{,}01 \Rightarrow n\ln(0{,}8) < \ln(0{,}01) \Rightarrow n > \frac{\ln(0{,}01)}{\ln(0{,}8)} \Rightarrow n \geq 21.$$

2. El colectivo que toma Felipe para ir al trabajo llega a la parada en algún momento entre las 10 y las 10 : 30 de la mañana con distribución uniforme. Felipe llega a la parada a las 10 de la mañana.

   (i) ¿Cuál es la probabilidad de que tenga que esperar más de 10 minutos?

   (ii) Si el colectivo no ha llegado todavía a las 10 : 15, hallar la probabilidad de que Felipe tenga que esperar por lo menos otros 10 minutos más.

Solución:

Sea $X = \{\text{horario en que llega el colectivo}\}$, entonces $X \sim \mathcal{U}(0,30)$ y, por lo tanto, $f_X(x) = \frac{1}{30}\mathbb{1}_{(0,30)}(x)$.

Notar que X está medido en minutos desde las 10hs.

(i) $\mathbb{P}(\{\text{esperar más de 10 minutos}\}) = \mathbb{P}(X \geq 10) = \int_{10}^{\infty}\frac{1}{30}\mathbb{1}_{(0,30)}(x)dx = \frac{1}{30}\int_{10}^{30}dx = \frac{1}{30}x\,|_{10}^{30} = \frac{2}{5}$.

(ii) $\mathbb{P}(X \geq 20 | X \geq 15) = \frac{\mathbb{P}(X \geq 20 \text{ y } X \geq 15)}{\mathbb{P}(X \geq 15)} = \frac{\mathbb{P}(X \geq 20)}{\mathbb{P}(X \geq 15)}$, luego

$$\left.\begin{array}{l} \mathbb{P}(X \geq 20) = \int_{20}^{30}\frac{1}{30}dx = \frac{30-20}{30} = \frac{10}{30} \\ \mathbb{P}(X \geq 15) = \int_{15}^{30}\frac{1}{30}dx = \frac{30-15}{30} = \frac{15}{30} \end{array}\right\} \Rightarrow \mathbb{P}(X \geq 20|X \geq 15) = \frac{10}{15} = \frac{2}{3}.$$

3. Se dice que una variable aleatoria X tiene distribución simétrica respecto de $\theta \in \mathbb{R}$ si $X - \theta \sim \theta - X$.

   (i) Dar dos ejemplos de variables aleatorias con distribución simétrica, una discreta y otra continua.

   (ii) Sea X una variable aleatoria absolutamente continua. Probar que X tiene distribución simétrica respecto de θ si y sólo si f_X es simétrica respecto de θ, i.e. $f_X(\theta - x) = f_x(\theta + x)$ para todo $x \in \mathbb{R}$.
   Sugerencia: Si encuentra problemas técnicos, puede suponer que la densidad f_X es una función continua.

Solución:

(i)

Discreta: Sea $X \sim Bi\left(4, \frac{1}{2}\right)$ tiene distribución simétrica respecto de $\theta = 2$ puesto que:
Si $Y = X - 2$, entonces

$$\mathbb{P}_Y(Y = y) = \mathbb{P}_X(X - 2 = y) = \mathbb{P}_X(X = y+2) = \binom{4}{y+2}\left(\frac{1}{2}\right)^{y+2}\left(\frac{1}{2}\right)^{4-(y+2)} = \binom{4}{y+2}\left(\frac{1}{2}\right)^2.$$

Si $\widetilde{Y} = 2 - X$, entonces

$$\mathbb{P}_{\widetilde{Y}}(\widetilde{Y} = \widetilde{y}) = \mathbb{P}_X(2 - X = \widetilde{y}) = \mathbb{P}_X(X = 2-\widetilde{y}) = \binom{4}{2-\widetilde{y}}\left(\frac{1}{2}\right)^{2-\widetilde{y}}\left(\frac{1}{2}\right)^{4-(2-\widetilde{y})} = \binom{4}{2-\widetilde{y}}\left(\frac{1}{2}\right)^2.$$

Por lo tanto, $Y \sim \widetilde{Y}$.

Continua: Sea $X \sim \mathcal{U}(0,1)$, es claro qu tiene distribución respecto de $\theta = \frac{1}{2}$.

(ii)($\Rightarrow$) : Por la definición de simetría, resula trivial.

($\Leftarrow$) : Sea f_X continua tal que X tiene distribución simétrica respecto de θ y sea $Y = X - \theta$ y $\widetilde{Y} = \theta - X$ tales que Y y $\widetilde{Y}$ tienen la misma distribución.

$$F_X(\theta - x) = \mathbb{P}(X \leq \theta - x) = \mathbb{P}(X - \theta \leq -x) = \mathbb{P}(Y \leq -x) = F_Y(-x) \Rightarrow$$

$$\frac{\partial F_X(\theta - x)}{\partial x} = \frac{\partial F_Y(-x)}{\partial x} \Rightarrow f_X(\theta - x)(-1) = f_Y(-x)(-1).$$

$$F_X(\theta + x) = \mathbb{P}(X \leq \theta + x) = \mathbb{P}(X - \theta \leq x) = 1 - \mathbb{P}(\widetilde{Y} \leq x) = 1 - F_{\widetilde{Y}}(x) \Rightarrow$$

$$\frac{\partial F_X(\theta + x)}{\partial x} = \frac{\partial(1 - F_{\widetilde{Y}}(x))}{\partial x} \Rightarrow f_X(\theta + x)(1) = f_{\widetilde{Y}}(x).$$

4. Una fábrica produce pilas cuya duración en horas, cuando se las destina para un determinado uso, tiene distribución normal de parámetros $\mu_0 = 53$ y $\sigma_0^2 = 25$. No obstante, un desperfecto en un sector de la fábrica produjo un cambio en la calidad de las pilas: del total que se fabrica una proporción $0,7$ de ellas tiene la duración correcta mientras que las restantes están falladas y tienen una duración en horas con distribución normal de parámetros desconocidos μ_1 y σ_1. Desafortunadamente, no hay forma de distinguir entre una pila común y una fallada a simple vista. Sea D la duración en horas de una pila extraída al azar del lote de producción de la fábrica.

   (i) Calcular μ_1 y σ_1^2 sabiendo que $\mathbb{P}(D \geq 47) = 0,82688$ y $\mathbb{P}(D \geq 60) = 0,5746$.

   (ii) Calcular la función de densidad de la duración en horas de una pila extraída al azar del lote de producción de la fábrica.

   (iii) Si se extrae una pilar al azar del lote de producción y se observa que dura más de 51 horas en funcionamiento, ¿cuál es la probabilidad de que sea una pila fallada?

Solución:

(i) Sea $k \in \mathbb{R}$, entonces

$$\mathbb{P}(D \geq k) = \underbrace{\mathbb{P}\left(D \geq k | \left\{\begin{matrix}\text{la pila tiene la}\\ \text{duración correcta}\end{matrix}\right\}\right)}_{=\mathbb{P}(X_{sana} \geq k)} \underbrace{\mathbb{P}\left(\left\{\begin{matrix}\text{la pila tiene la}\\ \text{duración correcta}\end{matrix}\right\}\right)}_{=0,7}$$
$$+ \underbrace{\mathbb{P}\left(D \geq k | \left\{\begin{matrix}\text{la pila no tiene la}\\ \text{duración correcta}\end{matrix}\right\}\right)}_{=\mathbb{P}(X_{fallada} \geq k)} \underbrace{\mathbb{P}\left(\left\{\begin{matrix}\text{la pila no tiene la}\\ \text{duración correcta}\end{matrix}\right\}\right)}_{=0,3}.$$

donde $X_s = X_{sana} \sim N(53, 25)$ y $X_f = X_{fallada} \sim N(\mu_1, \sigma_1^2)$.
Luego,

$$\left.\begin{matrix}\mathbb{P}(D \geq 47) = 0{,}7\mathbb{P}(X_s \geq 47) + 0{,}3\mathbb{P}(X_f \geq 47) = 0,82688\\ \mathbb{P}(D \geq 60) = 0{,}7\mathbb{P}(X_s \geq 60) + 0{,}3\mathbb{P}(X_f \geq 60) = 0,5746\end{matrix}\right\} \Rightarrow \begin{matrix}\mathbb{P}\left(Z_f \geq \frac{47-\mu_1}{\sigma_1}\right) = \frac{10}{3}\left[0,82688 - 0{,}7\mathbb{P}\left(Z_s \geq \frac{47-53}{5}\right)\right]\\ \mathbb{P}\left(Z_f \geq \frac{60-\mu_1}{\sigma_1}\right) = \frac{10}{3}\left[0,5746 - 0{,}7\mathbb{P}\left(Z_s \geq \frac{60-53}{5}\right)\right]\end{matrix}$$

entonces,

$$\begin{matrix}\mathbb{P}\left(Z_f \geq \frac{47-\mu_1}{\sigma_1}\right) = \frac{10}{3}\left[0,82688 - 0{,}7\mathbb{P}\left(Z_s \geq -1{,}2\right)\right] \approx 0,6915\\ \mathbb{P}\left(Z_f \geq \frac{60-\mu_1}{\sigma_1}\right) = \frac{10}{3}\left[0,5746 - 0{,}7\mathbb{P}\left(Z_s \geq 1,4\right)\right] \approx 0,0030\end{matrix} \Rightarrow \left\{\begin{matrix}\frac{47-\mu_1}{\sigma_1} \approx 0,5\\ \frac{60-\mu_1}{\sigma_1} \approx -2,75\end{matrix}\right.$$

por lo tanto, $\left\{\begin{matrix}\mu_1 + 0,5\sigma_1 = 47\\ \mu_1 - 2,75\sigma_1 = 60\end{matrix}\right.$ y, entonces, $\mu_1 = 49$ y $\sigma_1 = 4$.

(ii) $F_D(d) = \mathbb{P}_D(D \leq d) = 0,7\mathbb{P}_{X_s}(X_s \leq d) + 0,3\mathbb{P}_{X_f}(X_f \leq d) = 0,7F_{X_s}(d) + 0,3F_{X_f}(d)$, es decir

$$F_D(d) = 0,7\phi\left(\frac{d-53}{5}\right) + 0,3\phi\left(\frac{d-49}{4}\right)$$

(iii) $\mathbb{P}(\{\text{la pila está fallada}\}|D \geq 51) = \dfrac{\mathbb{P}(\text{la pila está fallada y } D \geq 51)}{\mathbb{P}(D \geq 51)} = \dfrac{\mathbb{P}(X_f \geq 51)}{\mathbb{P}(D \geq 51)} = \dfrac{1-\mathbb{P}(X_f < 51)}{1-\mathbb{P}(D < 51)}$, luego

$$\frac{1-\mathbb{P}\left(Z_f \leq \frac{51-49}{4}\right)}{1-\left[0,7\mathbb{P}_{X_s}\left(X_s \leq \frac{51-53}{5}\right) + 0,3\mathbb{P}_{X_f}\left(X_f \leq \frac{51-49}{4}\right)\right]} = \frac{1-\phi\left(\frac{1}{2}\right)}{1-\left[0{,}7\phi(-0,4) + 0{,}3\phi(0,5)\right]}.$$

5. **Definición.** Una variable aleatoria X tiene la propiedad de falta de memoria si para todo par de números reales s y t se verifica $\mathbb{P}(X > s+t | X > t) = \mathbb{P}(X > s)$ con $s, t \geq 0$.

   Probar que una variable aleatoria X posee la propiedad de falta de memoria si y sólo si tiene distribución exponencial de parámetro $\lambda = -\log(\mathbb{P}(X > 1))$.

Solución:

$(\Rightarrow):$ Sea $G(x) = \mathbb{P}(X > x)$, G es monóotona no creciente, continua a derecha y satisface

$$\lim_{x\to+\infty} G(x) = 0 \text{ y } \lim_{x\to-\infty} G(x) = 1.$$

Donde, bajo G, la propiedad de falta de memoria equivale a $G(s+t) = G(s)G(t)$.

En primer lugar, probaremos que $0 < G(1) < 1$.

Como $G(n) = [G(1)]^n$, si $G(1) = 1 \Rightarrow \lim_{n\to\infty} G(n) = 1$ lo cual es un absurdo. Por otro lado, como $G\left(\frac{1}{n}\right) = [G(1)]^{1/n}$, si $G(1) = 0 \Rightarrow \lim_{n\to\infty} G(1/n) = 0$ lo cual también es un absurdo.

De esta manera, se deduce que existe $\lambda > 0$ tal que $G1) = e^{-\lambda}$.

Sea $q \in \mathbb{Q}^+$ (racional positivo), entonces $q = \frac{m}{n}$ con $m, n \in \mathbb{N}$ y

$$G(q) = G\left(\frac{m}{n}\right) = \left[G\left(\frac{1}{n}\right)\right]^m = [G(1)]^{m/n} = [G(1)]^q.$$

Sea $t \in \mathbb{R}^+$ (real positivo), existen dos sucesiones racionales positivas $(q_k)_{k\in\mathbb{N}}$ y $(q'_k)_{k\in\mathbb{N}}$ tales que $q_k \downarrow t$ y $q'_k \uparrow t$, entonces

$$G(q'_k) \geq G(t) \geq G(q_k)\ \forall k \in \mathbb{N}\ |[G(1)]^{q'_k} \geq G(t) \geq [G(1)]^{q_k}\ \forall k \in \mathbb{N} \Rightarrow$$

$$\lim_{k\to\infty} [G(1)]^{q'_k} \geq G(t) \geq \lim_{k\to\infty} [G(1)]^{q_k}\ \forall k \in \mathbb{N} \Rightarrow G(t) = [G(1)]^t = e^{-\lambda t}\ \forall t > 0$$

$$\Rightarrow F_X(x) = 1 - e^{-\lambda x}\ \forall x > 0 \Rightarrow f_X(x) = \lambda e^{-\lambda x}\mathbb{1}_{(0,+\infty)}(x).$$

Por último,

$$\mathbb{P}(X > 1) = \int_1^\infty \lambda e^{-\lambda x} dx = -e^{-\lambda x}\big|_1^\infty = e^{-\lambda} \Rightarrow \mathbb{P}(X > 1) = e^{-\lambda} \Rightarrow \lambda = -\ln(\mathbb{P}(X > 1))$$

Es decir, $X \sim \mathcal{E}(-\ln(\mathbb{P}(X > 1)))$.

$(\Leftarrow):$ Sea $X \sim \mathcal{E}(\lambda)$, con $\lambda = -\ln(\mathbb{P}(X > 1))$. Luego

$$\mathbb{P}(X > s+t | X > t) = \frac{\mathbb{P}(X > s+t \text{ y } X > t)}{\mathbb{P}(X > t)} = \frac{\mathbb{P}(X > s+t)}{\mathbb{P}(X > t)} = \frac{1 - F_X(s+t)}{1 - F_X(t)} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda t}} = e^{-\lambda s} = \mathbb{P}(X > s).$$

Es decir, X posee la propiedad de falta de memoria.

6. Sean $\lambda > 0$ y X una variable aleatoria con distribución $\mathcal{E}(\lambda)$. Probar que $Y = [X] + 1$ tiene distribución geométrica de parámetro $p = 1 - e^{-\lambda}$, donde $[x]$ denota la parte entera de x.

Solución:

Es claro que $R_X = (0, +\infty)$ y $R_Y = \mathbb{N}$ de manera tal que si $k \leq x < k+1$, entonces $[x] = k \Rightarrow Y = k+1$ (con $k \geq 0$). Luego

$$\mathbb{P}(Y = k+1) = \mathbb{P}([X] = k) = \mathbb{P}(k \leq X \leq k+1) = \int_k^{k+1} \lambda e^{-\lambda x} dx = -e^{-\lambda(k+1)} + e^{-\lambda k} \Rightarrow \mathbb{P}(Y = k+1) = e^{-\lambda k}(1 - e^{-\lambda})$$

$$\Rightarrow \mathcal{P}(Y = k) = e^{-\lambda(k-1)}(1 - e^{-\lambda}) = \underbrace{(1 - (1 - e^{-\lambda}))}_{=(1-p)}{}^{k-1} \underbrace{(1 - e^{-\lambda})}_{=p} = p(1-p)^{k-1} \Rightarrow Y \sim \mathcal{G}(p).$$

7. Definimos la función Gamma $\Gamma : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ por la fórmula

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx.$$

(i) Probar que Γ está bien definida.

(ii) Mostrar que $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$ para todo $\alpha > 1$. Deducir que $\Gamma(n) = (n-1)!$ para todo $n \in \mathbb{N}$.

(iii) Probar que $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$.

Sugerencia: Hallar la función de densidad de Z^2 para Z una variable aleatoria con distribución normal estándar. ¿Es la densidad obtenida la de una distribución conocida? ¿Cuál?

Solución:

(i) $\displaystyle\int_0^\infty x^{\alpha-1}e^{-x}dx = \int_0^1 x^{\alpha-1}e^{-x}dx + \int_1^\infty x^{\alpha-1}e^{-x}dx.$

$\underline{0 \le x \le 1:}$ Como $x > 0$, es claro que $x^{\alpha-1}e^{-x} \le x^{\alpha-1}$ para $\alpha > 0$ y, usando que $\displaystyle\lim_{a\to 0^+}\int_a^1 x^{\alpha-1} = \frac{1}{\alpha}$ se tiene que

$$\int_0^1 x^{\alpha-1}e^{-x}dx \le \lim_{a\to 0^+}\int_a^1 x^{\alpha-1} = \frac{1}{\alpha}.$$

$\underline{x \ge 1:}$ Usando L'Hospital, es claro que para $\alpha > 0$ se tiene que $\displaystyle\lim_{x\to\infty}\frac{x^{\alpha-1}}{e^{x/2}} = 0.$

Entonces, $\dfrac{x^{\alpha-1}}{e^{x/2}} \le 1 \Rightarrow x^{\alpha-1} \le e^{x/2} \Rightarrow x^{\alpha-1}e^{-x} \le e^{-x/2}.$

Luego,

$$\lim_{r\to+\infty}\int_1^r x^{\alpha-1}e^{-x}dx \le \lim_{r\to+\infty}\int_1^r e^{-x/2}dx = -2\lim_{r\to+\infty} e^{-x/2}\Big|_1^r = 2e^{-1/2}.$$

Por lo tanto, para todo $x \in (0,+\infty)$ y $\alpha > 0$, se tiene que $\displaystyle\int_0^\infty x^{\alpha-1}e^{-x}dx < \infty.$

(ii) Integrando por partes,

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}dx = \underbrace{\frac{x^\alpha e^{-x}}{\alpha}\Big|_0^\infty}_{=0} + \frac{1}{\alpha}\int_0^\infty x^\alpha e^{-x}dx \Rightarrow \alpha\Gamma(\alpha) = \Gamma(\alpha+1).$$

Para deducir $\Gamma(n) = (n-1)!$ para todo $n \in \mathbb{N}$, basta con notar que

- $\displaystyle\Gamma(1) = \lim_{k\to\infty}\int_0^k e^x dx = \lim_{k\to\infty}(e^0 - e^{-k}) = 1.$
- $\Gamma(n+1) = n\Gamma(n) = n(n-1)\Gamma(n-2) = \ldots = n(n-1)(n-2)\ldots 2\Gamma(1) = n(n-1)(n-2)\ldots 2 = n!$

(iii) $\displaystyle\Gamma\left(\frac{1}{2}\right) = \int_0^\infty \underbrace{x^{\frac{1}{2}-1}e^{-x}}_{z^2=x}dx = 2\int_0^\infty e^{-z^2}dz = \int_{-\infty}^{+\infty} e^{-z^2}dz = \sqrt{\pi}.$

8. Sean $n \in \mathbb{N}$ y Z una variable aleatoria con distribución $\Gamma(n,\lambda)$. Probar que para todo $z > 0$ se tiene

$$F_Z(x) = \mathbb{P}(X_z \ge n)$$

donde X_z es una variable aleatoria con distribución $\mathcal{P}(z\lambda)$.

Solución:

$$Z \sim \Gamma(n,\lambda) \Rightarrow f_Z(z) = \frac{\lambda^n}{\Gamma(n)}z^{n-1}e^{-\lambda z}\mathbb{1}_{(0,+\infty)}(z) \quad \text{y} \quad X_z \sim \mathcal{P}(z\lambda) \Rightarrow \mathbb{P}_{X_z}(k) = \frac{(z\lambda)^k}{k!}e^{-z\lambda} \text{ con } k \in \mathbb{N}_0$$

$$\mathbb{P}(X_z \ge n) = 1 - \mathcal{P}(X_z < n) = 1 - \sum_{k=0}^{n-1}\frac{(z\lambda)^k}{k!}e^{-z\lambda} \Rightarrow \frac{\partial\mathbb{P}(X_z \ge n)}{\partial z} = -\left[-\lambda e^{-z\lambda}\sum_{k=0}^{n-1}\frac{(z\lambda)^k}{k!} + e^{-z\lambda}\sum_{k=0}^{n-1}\frac{(z\lambda)^{k-1}}{k!}k\lambda\right] \Rightarrow$$

$$\frac{\partial\mathbb{P}(X_z \ge n)}{\partial z} = e^{-z\lambda}\lambda\left[\sum_{k=0}^{n-1}\frac{(z\lambda)^k}{k!} - \sum_{k=1}^{n-1}\frac{(z\lambda)^{k-1}}{(k-1)!}\right] = e^{-z\lambda}\lambda\frac{(z\lambda)^{n-1}}{(n-1)!} = \frac{\lambda^n}{\Gamma(n)}z^{n-1}e^{-z\lambda}$$

es decir, $\mathbb{P}(X_z \ge n) = F_Z(z)$.

9. Sea X una variable aleatoria continua con función de distribución acumulada F. Mostrar que $F(X)$ tiene distribución $\mathcal{U}([0,1])$.

Solución:

$$\mathbb{P}(F(X) \le t) = \mathbb{P}(F_X(x) \le t) = \mathbb{P}(X \le F_X^{-1}(t)) = F_X\left(F_X^{-1}(t)\right) = t.$$
$$\Rightarrow \mathbb{P}(F(X) \le t) = t \Rightarrow F(X) \sim \mathcal{U}[(0,1)].$$

10. Sea U una variable con distribución $\mathcal{U}(0,1)$. Encontrar una función g tal que $g(U)$ tenga distribución

(i) $\mathcal{E}(1)$.

(ii) Doble exponencial de parámetro uno, es decir, con función de densidad $f(x) = \frac{1}{2}e^{-|x|}$.

(iii) $Bi\left(5, \frac{1}{3}\right)$.

(iv) Una distribución discreta con rango $R_X = (x_n)_{n\in\mathbb{N}}$ y respectivas probabilidades puntuales $(p_n)_{n\in\mathbb{N}}$.

Solución:

Sea $U \sim \mathcal{U}(0,1)$, entonces $F_U(u) = \begin{cases} 0 & u < 0 \\ u & 0 \le u < 1 \\ 1 & u \ge 1 \end{cases}$

(i) Buscamos $g : \mathbb{R} \to \mathbb{R}$ tal que $X = g(U) \sim \mathcal{E}(1)$. Si g es decreciente,

$$F_X(x) = \mathbb{P}(X \le x) = \mathbb{P}(g(U) \le x) = \mathbb{P}(U \ge g^{-1}(x)) = 1 - \mathbb{P}(U \le g^{-1}(x)) = 1 - F_U(g^{-1}(x)) \Rightarrow$$

$$\frac{\partial F_X(x)}{\partial x} = \frac{1 - \mathbb{P}(U \le g^{-1}(x))}{\partial x} \Rightarrow f_X(x) = -f_U(g^{-1}(x))\frac{\partial(g^{-1}(x))}{\partial x}$$

luego, si $x \in (0, +\infty)$, $f_X(x) = e^{-x}$ y

$$e^{-x} = -\frac{\partial(g^{-1}(x))}{\partial x}\mathbb{1}_{(0,1)}(g^{-1}(x)) \Rightarrow e^{-x} = -\frac{\partial(g^{-1}(x))}{\partial x} \quad \text{y} \quad g^{-1}(x) \in (0,1)$$

$$\Rightarrow u = g^{-1}(x) = e^{-x} \Rightarrow g(u) = -\ln(u) \Rightarrow X = -\ln(U) \sim \mathcal{E}(1).$$

(ii) $f_X(x) = \begin{cases} \frac{1}{2}e^{-x} & x \ge 0 \\ \frac{1}{2}e^{x} & x < 0 \end{cases} \Rightarrow \begin{cases} \text{para } x \ge 0 & \text{se propone una } g_1(u) \text{ decreciente.} \\ \text{para } x < 0 & \text{se propone una } g_2(u) \text{ creciente.} \end{cases}$

Si $x \ge 0$, entonces sea g_1 decreciente tal que

$$F_X(x) = \mathbb{P}(X \le x) = \mathbb{P}(g_1(U) \le x) = \mathbb{P}(U \ge g_1^{-1}(x)) = 1 - \mathbb{P}(U \le g_1^{-1}(x)) = 1 - F_U(g_1^{-1}(x)) \Rightarrow$$

$$\frac{1}{2}e^{-x} = -f_U(g_1^{-1}(x))\frac{\partial(g_1^{-1}(x))}{\partial x} = -\mathbb{1}_{[0,1]}(g_1^{-1}(x))\frac{\partial(g_1^{-1}(x))}{\partial x} \Rightarrow -\frac{\partial(g_1^{-1}(x))}{\partial x} = \frac{1}{2}e^{-x} \quad \text{y} \quad g_1^{-1}(x) \in \left(0, \frac{1}{2}\right)$$

$$\Rightarrow u = g_1^{-1}(x) = \frac{1}{2}e^{-x} \Rightarrow g_1(u) = -\ln(2u).$$

Si $x < 0$, entonces sea g_2 creciente tal que

$$F_X(x) = \mathbb{P}(X \le x) = \mathbb{P}(g_2(U) \le x) = \mathbb{P}(U \le g_2^{-1}(x)) = \mathbb{P}(U \le g_2^{-1}(x)) = F_U(g_2^{-1}(x)) \Rightarrow$$

$$\frac{1}{2}e^{x} = f_U(g_2^{-1}(x))\frac{\partial(g_2^{-1}(x))}{\partial x} = \mathbb{1}_{[0,1]}(g_2^{-1}(x))\frac{\partial(g_2^{-1}(x))}{\partial x} \Rightarrow$$

Como se busca que $g_2^{-1} \in \left(\frac{1}{2}, 1\right)$ para $x \in (-\infty, 0)$, se tiene que

$$\frac{\partial(g_2^{-1}(x))}{\partial x} = \frac{1}{2}e^{x} \quad \text{y} \quad g_2^{-1}(x) \in \left(\frac{1}{2}, 1\right) \Rightarrow u = g_2^{-1}(x) = \frac{1}{2}e^{x} + \frac{1}{2} \Rightarrow g_2(u) = \ln(2u-1).$$

Luego, $g(u) = \begin{cases} -\ln(2u) & u \in \left(0, \frac{1}{2}\right] \\ \ln(2u-1) & u \in \left(\frac{1}{2}, 1\right] \end{cases}$.

11. Don Zoilo tiene dos vacas, Aurora y Belinda. La cantidad de leche en litros que da Aurora en un día es una variable aleatoria X con distribución $\mathcal{E}(0,2)$. Belinda, en cambio, da 5 litros el 20 % de las veces que es ordeñada y el resto no da nada. Don Zoilo ordeña a Belinda solamente los días en que Aurora da menos de 6 litros.

    (i) ¿Cuál es la probabilidad de que Aurora dé más de 6 litros en exactamente dos días de la próxima semana? Cabe aclarar que los fines de semana Don Zoilo no ordeña a sus vacas.

    (ii) ¿Cuál es la probabilidad de que Don Zoilo obtenga más de 8 en un día?

(iii) Con la leche que obtiene de Aurora, Don Zoilo fabrica manteca. La cantidad de manteca en kilos que obtiene con X litros de leche es $W = g(X)$ siendo

$$g(X) = \begin{cases} \sqrt[3]{X} & X \leq 8 \\ \dfrac{1}{7}(X-1)^2 - 5 & 8 < X \leq 15 \\ 2X - 7 & 15 < X \end{cases}$$

Hallar la función de densidad de la cantidad de manteca fabricada.

Solución:

Sean los eventos, X_i = litros de leche que da Aurora en el día i e Y_i = litros de leche que da Belinda en el día i,

(i) $\mathbb{P}\left(\left\{\begin{matrix}\text{Aurora dé más de 6 litros} \\ \text{en exactamente 2 días de la semana}\end{matrix}\right\}\right) = \left(\begin{matrix}5\\2\end{matrix}\right)(\mathbb{P}(X_i \geq 6))^2 \left(\begin{matrix}3\\3\end{matrix}\right)(\mathbb{P}(X_i < 6))^3.$

$$\mathbb{P}(X_i \geq 6) = 1 - \mathbb{P}(X_i < 6) = \int_0^6 0,2e^{-0,2x}dx = -e^{-0,2x}\big|_0^6 = 1 - e^{-1,2} \Rightarrow$$

$$\mathbb{P}\left(\left\{\begin{matrix}\text{Aurora dé más de 6 litros} \\ \text{en exactamente 2 días de la semana}\end{matrix}\right\}\right) = 10(e^{-1,2})^2(1-e^{-1,2})^3.$$

(ii) Buscamos $\mathbb{P}(X_i + Y_i \geq 8)$, entonces

$$\mathbb{P}(X_i + Y_i \geq 8) = \mathbb{P}\left(X_i + Y_i \geq 8 | \left\{\begin{matrix}\text{Belinda}\\\text{es ordeñada}\end{matrix}\right\}\right)\mathbb{P}\left(\left\{\begin{matrix}\text{Belinda}\\\text{es ordeñada}\end{matrix}\right\}\right) +$$
$$+\mathbb{P}\left(X_i + Y_i \geq 8 | \left\{\begin{matrix}\text{Belinda}\\\text{no es ordeñada}\end{matrix}\right\}\right)\mathbb{P}\left(\left\{\begin{matrix}\text{Belinda}\\\text{no es ordeñada}\end{matrix}\right\}\right)$$

luego,

$$\mathbb{P}\left(\left\{\begin{matrix}\text{Belinda}\\\text{es ordeñada}\end{matrix}\right\}\right) = \mathbb{P}(X_i < 6) = 1 - e^{-1,2} \quad \text{y} \quad \mathbb{P}\left(\left\{\begin{matrix}\text{Belinda}\\\text{no es ordeñada}\end{matrix}\right\}\right) \underbrace{=}_{Y_i=0} \mathbb{P}(X_i \geq 6) = e^{-1,2}$$

$$\mathbb{P}\left(X_i + Y_i \geq 8 | \left\{\begin{matrix}\text{Belinda}\\\text{no es ordeñada}\end{matrix}\right\}\right)\mathbb{P}\left(\left\{\begin{matrix}\text{Belinda}\\\text{no es ordeñada}\end{matrix}\right\}\right) = \mathbb{P}\left(X_i + Y_i \geq 8 \bigcap \left\{\begin{matrix}\text{Belinda}\\\text{no es ordeñada}\end{matrix}\right\}\right)$$
$$= \mathbb{P}(X_i + Y_i \geq 8 \text{ y } X_i \geq 6 \text{ y } Y_i = 0) = \mathbb{P}(X_i \geq 8)$$

$$\mathbb{P}\left(X_i + Y_i \geq 8 | \left\{\begin{matrix}\text{Belinda}\\\text{es ordeñada}\end{matrix}\right\}\right)\mathbb{P}\left(\left\{\begin{matrix}\text{Belinda}\\\text{es ordeñada}\end{matrix}\right\}\right) = \mathbb{P}\left(X_i + Y_i \geq 8 \bigcap \left\{\begin{matrix}\text{Belinda}\\\text{es ordeñada}\end{matrix}\right\}\right)$$

$$\mathbb{P}\left(X_i + Y_i \geq 8 \bigcap \left\{\begin{matrix}\text{Belinda}\\\text{es ordeñada}\end{matrix}\right\}\right) = \mathbb{P}\left(X_i + Y_i \geq 8 \bigcap \left\{\begin{matrix}\text{Belinda es}\\\text{ordeñada y da leche}\end{matrix}\right\}\right) +$$
$$+\mathbb{P}\left(X_i + Y_i \geq 8 \bigcap \left\{\begin{matrix}\text{Belinda es ordeñada}\\\text{y no da leche}\end{matrix}\right\}\right)$$

$$\Rightarrow \mathbb{P}\left(X_i + Y_i \geq 8 \bigcap \left\{\begin{matrix}\text{Belinda}\\\text{es ordeñada}\end{matrix}\right\}\right) = \mathbb{P}(X_i + 5 \geq 8)0,2 + \mathbb{P}(X_i \geq 8)0,8$$

Por lo tanto,

$$\mathbb{P}(X_i + Y_i \geq 8) = \mathbb{P}(X_i \geq 8) + 0,2\mathbb{P}(X_i \geq 3) + 0,8\mathbb{P}(X_i \geq 8) = 1,8\mathbb{P}(X_i \geq 8) + 0,2\mathbb{P}(X_i \geq 3).$$

(iii) Sean los eventos X = litros de leche de Aurora y W = kilos de manteca tal que $W = g(X)$ con

$$g(X) = \begin{cases} \sqrt[3]{X} & X \leq 8 \\ \frac{1}{7}(X-1)^2 - 5 & 8 < X \leq 15 \\ 2X - 7 & 15 < X \end{cases}$$

- <u>$w < 2$:</u>
  $F_W(x) = \mathbb{P}(W \leq w) = \mathbb{P}(\sqrt[3]{X} \leq w) = \mathbb{P}(X \leq w^3) = F_X(w^3)$, entonces

$$\frac{\partial F_W(w)}{\partial w} = \frac{\partial F_X(w^3)}{\partial w} \Rightarrow f_W(w) = 3w^2 f_X(w^3) \Rightarrow f_W(w) = 3w^2\left(0,2e^{-0,2w^3}\right).$$

- $\underline{2 \leq w < 23:}$
$F_W(x) = \mathbb{P}(W \leq w) = \mathbb{P}\left(\frac{1}{7}(X-1)^2 - 5 \leq w\right) = \mathbb{P}((X-1)^2 \leq 7(w+5)) = \mathbb{P}(|X-1| \leq \sqrt{7(w+5)})$ y puesto que $8 < X$

$$F_W(x) = \mathbb{P}(X \leq \sqrt{7(w+5)}+1) = F_X(\sqrt{7(w+5)}+1) \Rightarrow \frac{\partial F_W(w)}{\partial w} = \frac{\partial F_X(\sqrt{7(w+5)}+1)}{\partial w} \Rightarrow$$

$$f_W(w) = f_X(\sqrt{7(w+5)}+1)\frac{7}{2}(7w+35)^{-1/2} \Rightarrow f_W(w) = \frac{7}{2}(7w+35)^{-1/2}\left(0,2e^{-0,2-0,2\sqrt{7(w+5)}}\right).$$

- $\underline{w \geq 23:}$
$F_W(x) = \mathbb{P}(W \leq w) = \mathbb{P}(2X-7 \leq w) = \mathbb{P}\left(X \leq \dfrac{w+7}{2}\right) = F_X\left(\dfrac{w+7}{2}\right)$, entonces

$$\frac{\partial F_W(w)}{\partial w} = \frac{\partial F_X\left(\frac{w+7}{2}\right)}{\partial w} \Rightarrow f_W(w) = \frac{1}{2}f_X\left(\frac{w+7}{2}\right) = \frac{1}{2}\left(0,2e^{-0,2\frac{w+7}{2}}\right).$$

Práctica 5

Vectores aleatorios

1. (i) Demostrar que la función

$$F(x,y) = \begin{cases} 1 - e^{-x-y} & \text{si } x \geq 0, y \geq 0 \\ 0 & \text{caso contrario} \end{cases}$$

no es la función de distribución acumulada de ningún vector aleatorio.

(ii) Mostrar que

$$F(x,y) = \begin{cases} (1 - e^{-x})(1 - e^{-y}) & \text{si } x \geq 0, y \geq 0 \\ 0 & \text{caso contrario} \end{cases}$$

sí lo es.

Solución:

(i) Sea B el rectángulo con vértices en $(1,1)$, $(1,0)$, $(0,1)$ y $(0,0)$. Luego,

$$\mathbb{P}(X \in B) = F_X(1,1) - F_X(1,0) - F_X(0,1) + F_X(0,1) = (1 - e^{-2}) - (1 - e^{-1}) - (1 - e^{-1}) + (1 - 1) < 0.$$

por lo tanto, no es la función de distribución acumulada de ningún vector aleatorio.

(ii) Es claro que $F(x,y) = (1-e^{-x})\mathbb{1}_{(0,+\infty)}(x)(1-e^{-y})\mathbb{1}_{(0,+\infty)}(y)$ de manera tal que resulta ser una función distribución de dos variables aleatorias independientes X e Y con $X \sim \mathcal{E}(1)$ e $Y \sim \mathcal{E}(1)$.

2. Una urna contiene 4 bolitas blancas y 6 negras. Se extraen 3 bolitas sin reposición y se definen las siguientes variables aleatorias:

$$X = \begin{cases} 1 & \text{si el número de bolitas blancas extraídas es par} \\ 0 & \text{si el número de bolitas blancas extraídas es impar} \end{cases}$$

$$Y = \text{Número de bolillas negras extraídas}$$

(i) Hallar $\mathbb{P}_{xy}$ y F_{XY}.

(ii) Hallar $\mathbb{P}_X$ y $\mathbb{P}_Y$. Determinar si X e Y son independientes.

Solución:

(i) como $R_X = \{0,1\}$ y $R_Y = \{0,1,2,3\}$ se tiene que

$$\mathbb{P}(0,0) = \mathbb{P}(X = 0 \text{ y } Y = 0) = \mathbb{P}(B = 3 \text{ y } N = 0) = \frac{1}{30}$$

$$\mathbb{P}(1,0) = \mathbb{P}(X = 1 \text{ y } Y = 0) = \mathbb{P}(B = \textit{par} \text{ y } N = 0) = 0$$

$$\mathbb{P}(0,1) = \mathbb{P}(X = 0 \text{ y } Y = 1) = \mathbb{P}(B = \textit{impar} \text{ y } N = 1) = 0$$

$$\mathbb{P}(1,1) = \mathbb{P}(X = 1 \text{ y } Y = 1) = \mathbb{P}(B = \textit{par} \text{ y } N = 1) = \mathbb{P}(B = 2 \text{ y } N = 1) = \frac{1}{10}$$

$$\mathbb{P}(0,2) = \mathbb{P}(X = 0 \text{ y } Y = 2) = \mathbb{P}(B = \textit{impar} \text{ y } N = 2) = \mathbb{P}(B = 1 \text{ y } N = 2) = \frac{1}{2}$$

$$\mathbb{P}(1,2) = \mathbb{P}(X = 1 \text{ y } Y = 2) = \mathbb{P}(B = \textit{par} \text{ y } N = 2) = 0$$

$$\mathbb{P}(0,3) = \mathbb{P}(X = 0 \text{ y } Y = 3) = \mathbb{P}(B = \textit{impar} \text{ y } N = 3) = 0$$

$$\mathbb{P}(1,3) = \mathbb{P}(X = 1 \text{ y } Y = 3) = \mathbb{P}(B = par \text{ y } N = 3) = \frac{1}{6}$$

entonces,

$$F_{XY}(x,y) = \begin{cases} \frac{1}{30} & x < 1 \text{ e } y < 2 \text{ ó } x \in \mathbb{R} \text{ e } y < 1 \\ \frac{1}{30} & x \in \mathbb{R} \text{ e } y < 2 \\ \frac{16}{30} & x < 1 \text{ y } y \in \mathbb{R} \\ \frac{5}{6} & x \in \mathbb{R} \text{ y } y < 3 \\ 1 & \text{en otro caso} \end{cases}$$

(ii) X e Y no son independientes, puesto que

$$\mathbb{P}_{XY}(X = 0 \text{ e } Y = 0)\frac{1}{30} \neq \mathbb{P}_X(X = 0)\mathbb{P}_Y(Y = 0) = \frac{16}{30} \times \frac{1}{30}.$$

3. Se arroja un dado equilibrado 2 veces. Para $i = 1, 2$ sea X_i el número obtenido en la tirada i.
   (i) Sea $Y = X_1 + X_2$. Hallar $\mathbb{P}_Y$.
   (ii) Sea la variable aleatoria

$$Z = \begin{cases} 1 & \text{si } Y \text{ es par} \\ 0 & \text{si } Y \text{ es impar} \end{cases}$$

   Determinar si X_1 y Z son independientes.

Solución:

(i) Como $R_{X_1} = R_{X_2} = \{1, \ldots, 6\}$, entonces $R_Y = \{2, \ldots, 12\}$

$$\mathbb{P}_Y(Y = 2) = \mathcal{P}(X_1 = 1 \text{ y } X_2 = 1) = \frac{1}{6}\frac{1}{6} = \frac{1}{36} = \mathbb{P}_Y(Y = 12)$$

$$\mathbb{P}_Y(Y = 3) = \mathcal{P}(X_1 = 1 \text{ y } X_2 = 2) + \mathcal{P}(X_1 = 2 \text{ y } X_2 = 1) = 2\frac{1}{6}\frac{1}{6} = \frac{2}{36} = \mathbb{P}_Y(Y = 11)$$

$$\mathbb{P}_Y(Y = 4) = \mathcal{P}(X_1 = 1 \text{ y } X_2 = 3) + \mathcal{P}(X_1 = 2 \text{ y } X_2 = 2) + \mathcal{P}(X_1 = 3 \text{ y } X_2 = 1) = 3\frac{1}{6}\frac{1}{6} = \frac{3}{36} = \mathbb{P}_Y(Y = 10)$$

$$\mathbb{P}_Y(Y = 5) = \mathcal{P}(X_1 = 1 \text{ y } X_2 = 4) + \mathcal{P}(X_1 = 2 \text{ y } X_2 = 3) + \mathcal{P}(X_1 = 3 \text{ y } X_2 = 2) + \mathcal{P}(X_1 = 4 \text{ y } X_2 = 1) \Rightarrow$$

$$\mathbb{P}_Y(Y = 5) = 4\frac{1}{6}\frac{1}{6} = \frac{4}{36} = \mathbb{P}_Y(Y = 9)$$

$$\begin{aligned} \mathbb{P}_Y(Y = 6) = {} & \mathcal{P}(X_1 = 1 \text{ y } X_2 = 5) + \mathcal{P}(X_1 = 2 \text{ y } X_2 = 4) + \mathcal{P}(X_1 = 3 \text{ y } X_2 = 3) + \\ & + \mathcal{P}(X_1 = 4 \text{ y } X_2 = 2) + \mathcal{P}(X_1 = 5 \text{ y } X_2 = 1) = 5\frac{1}{6}\frac{1}{6} = \frac{5}{36} = \mathbb{P}_Y(Y = 8) \\ \mathbb{P}_Y(Y = 7) = {} & \mathcal{P}(X_1 = 1 \text{ y } X_2 = 6) + \mathcal{P}(X_1 = 2 \text{ y } X_2 = 5) + \mathcal{P}(X_1 = 3 \text{ y } X_2 = 4) + \mathcal{P}(X_1 = 4 \text{ y } X_2 = 3) \\ & + \mathcal{P}(X_1 = 5 \text{ y } X_2 = 2) + \mathcal{P}(X_1 = 6 \text{ y } X_2 = 1) = 6\frac{1}{6}\frac{1}{6} = \frac{1}{6} = \mathbb{P}_Y(Y = 8) \end{aligned}$$

(ii)

$Z \diagdown X_1$	1	2	3	4	5	6
0	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$
1	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$

Por lo tanto X_1 y Z son independientes.

4. Sean X e Y variables aleatorias independientes. Probar las siguientes afirmaciones:
   (i) $X \sim Bi(n,p)$ e $Y \sim Bi(m,p)$, entonces $X + Y \sim Bi(m + n, p)$.
   (ii) $X \sim \mathcal{P}(\lambda)$ e $Y \sim \mathcal{P}(\mu)$, entonces $X + Y \sim \mathcal{P}(\lambda + \mu)$.
   (iii) $X \sim \mathcal{G}(p)$ e $Y \sim \mathcal{G}(p)$, entonces $X + Y \sim BN(2, p)$.

Solución:

(i) $\mathbb{P}(X+Y=k)=\sum_{i=0}^{k}\mathbb{P}(X=i,Y=k-i)$ y puesto que X e Y son variables aleatorias independientes, se tiene que

$$\mathbb{P}(X+Y=k)=\sum_{i=0}^{k}\mathbb{P}(X=i)\mathbb{P}(Y=k-i)=\sum_{i=0}^{k}\begin{pmatrix} n \\ i \end{pmatrix}p^i(1-p)^{n-i}\begin{pmatrix} m \\ k-i \end{pmatrix}p^{k-i}(1-p)^{n-k+i}\Rightarrow$$

$$\mathbb{P}(X+Y=k)=p^k(1-p)^{n+m-k}\sum_{i=0}^{k}\begin{pmatrix} n \\ i \end{pmatrix}\begin{pmatrix} m \\ k-i \end{pmatrix}=p^k(1-p)^{n+m-k}\begin{pmatrix} n+m \\ k \end{pmatrix}$$

por lo tanto, $X+Y\sim Bi(m+n,p)$.

(ii) $\mathbb{P}(X+Y=k)=\sum_{i=0}^{k}\mathbb{P}(X=i,Y=k-i)$ y puesto que X e Y son variables aleatorias independientes, se tiene que

$$\mathbb{P}(X+Y=k)=\sum_{i=0}^{k}\mathbb{P}(X=i)\mathbb{P}(Y=k-i)=\sum_{i=0}^{k}\frac{\lambda^i}{i!}e^{-\lambda}\frac{\mu^{k-i}}{(k-i)!}e^{-\mu}=e^{-(\lambda+\mu)}\frac{1}{k!}\sum_{i=0}^{k}\frac{k!}{i!(k-i)!}\lambda^i\mu^{k-i}$$

$$\Rightarrow\mathbb{P}(X+Y=k)=e^{-(\lambda+\mu)}\frac{1}{k!}\sum_{i=0}^{k}\begin{pmatrix} k \\ i \end{pmatrix}\lambda^i\mu^{k-i}=e^{-(\lambda+\mu)}\frac{1}{k!}(\lambda+\mu)^k$$

por lo tanto, $X+Y\sim\mathcal{P}(\lambda+\mu)$.

(iii) $\mathbb{P}(X+Y=k)=\sum_{i=0}^{k-1}\mathbb{P}(X=i,Y=k-i)$ y puesto que X e Y son variables aleatorias independientes, se tiene que

$$\mathbb{P}(X+Y=k)=\sum_{i=0}^{k-1}\mathbb{P}(X=i)\mathbb{P}(Y=k-i)=\sum_{i=0}^{k-1}p(1-p)^{i-1}p(1-p)^{k-i-1}=\sum_{i=0}^{k-1}p^2(1-p)^{k-2}\Rightarrow$$

$$\mathbb{P}(X+Y=k)=p^2(1-p)^{k-2}\sum_{i=0}^{k-1}1=p^2(1-p)^{k-2}(k-1)=p^2(1-p)^{k-2}\begin{pmatrix} k-1 \\ 2-1 \end{pmatrix}$$

por lo tanto, $X+Y\sim BN(2,p)$.

5. Se sabe que en la provincia de Salta la proporción de hombres de ojos azules es 20 %, de ojos verdes es 5 %, de ojos negros es 10 % y otro color de ojos es 65 %. Josefina decide viajar de la capital salteña a una ciudad a 200 km. donde se realizará un congreso médico sobre alcoholismo. Para ello debe tomar dos colectivos en los que viajan sólo salteños. Para llevar a cabo una prueba decide tomar una copa de jerez con cada hombre de ojos verdes o azules que encuentre en su viaje. Como su belleza es irresistible, todos los hombres aceptan su invitación. En el primer colectivo viajan 10 hombres de los cuales ninguno transborda al siguiente. En el segundo hay 8 hombres.

   (i) Calcular la probabilidad de que en la primera parte del trayecto haya tomado menos de 4 copas.

   (ii) Calcular la probabilidad de que tome más de 3 copas en total.

Solución:

Sean las variables aleatorias $X_1=$ cantidad de hombres del primer colectivo que beben con Josefina y $X_2=$ cantidad de hombres del segundo colectivo que beben con Josefina. Por lo tanto, X_1 y X_2 son independientes tales que

$$\mathbb{P}\left(\begin{Bmatrix}\text{Beba un hombre}\\ \text{del primer colectivo}\end{Bmatrix}\right)=\mathbb{P}\left(\begin{Bmatrix}\text{Beba un hombre}\\ \text{del segundo colectivo}\end{Bmatrix}\right)=0,20+0,05=0,25$$

$$\Rightarrow X_1\sim Bi(10;0,25)\text{ y }X_2\sim Bi(8;0,25)\Rightarrow X+Y\sim Bi(18;0,25)$$

(i) $\mathbb{P}_{X_1}(X_1<4)=\sum_{k=0}^{4}\begin{pmatrix} 10 \\ k \end{pmatrix}(0,25)^k(0,75)^{10-k}$.

(ii) $\mathbb{P}_{X_1+X_2}(X_1+X_2>3)=1-\mathbb{P}_{X_1+X_2}(X_1+X_2\leq 3)=1-\sum_{k=0}^{3}\begin{pmatrix} 18 \\ k \end{pmatrix}(0,25)^k(0,75)^{18-k}$.

6. Manu llega a la final de un torneo de básquet con su equipo. La probabilidad de que su equipo gane un partido contra el otro equipo finalista es p. Los organizadores proponen jugar un número impar de partidos y declarar ganador del torneo a aquel equipo que gane la mitad más uno de los partidos jugados. Siendo capitán de su equipo, los organizadores le preguntan a Manu cuántos partidos quiere jugar: debe decir si prefiere jugar $2k-1$ ó $2k+1$ partidos, donde k un número natural fijado de antemano por los organizadores. ¿Para qué valores de p le conviene a Manu jugar $2k+1$ partidos?

7. El 10 % de la población fuma cigarrillos negros, el 35 % fuma cigarrillos rubios, el 3 % fuma pipa y el resto no fuma. Se realiza una encuesta a 35 personas y con los resultados obtenidos se definen las variables aleatorias

$$\begin{aligned} Y_1 &= \text{número de personas que no fuman} \\ Y_2 &= \text{número de personas que fuman cigarrillos rubios} \\ Y_3 &= \text{número de personas que fuman cigarrillos negros} \\ Y_4 &= \text{número de personas que fuman pipa} \end{aligned}$$

   (i) Hallar la probabilidad puntual del vector aleatorio (Y_1, Y_2, Y_3, Y_4).
   (ii) Hallar la probabilidad puntual del vector aleatorio (Y_1, Y_2+Y_3, Y_4).
   (iii) Hallar la probabilidad puntual de la variable aleatoria Y_2+Y_3. ¿Se obtiene información adicional a la contenida en la distribución de Y_2+Y_3 si se calcula además la distribución del vector aleatorio (Y_2+Y_3, Y_1+Y_4)? ¿Por qué?

Solución:

Como se encuestan a 35 personas, se tiene que

$$Y_1 \sim Bi(35; 0,52) Y_2 \sim Bi(35; 0,35) \qquad Y_1 \sim Bi(35; 0,10) Y_1 \sim Bi(35; 0,03)$$

Observando que no son variables aleatorias independientes puesto que, por ejemplo, $Y_1 = 35 - (Y_2+Y_3+Y_4)$.

(i) $\widetilde{Y} = (Y_1, Y_2, Y_3, Y_4) \sim \mathcal{M}(35; 0,52; 0,35; 0,10; 0,03)$ entonces

$$\mathbb{P}_{\widetilde{Y}}(y_1, y_2, y_3, y_4) = \begin{cases} \dfrac{35!}{y_1!y_2!y_3!y_4!}(0,52)^{y_1}(0,35)^{y_2}(0,10)^{y_3}(0,03)^{y_4} & y_1+y_2+y_3+y_4 = 35 \\ 0 & y_1+y_2+y_3+y_4 \neq 35 \end{cases}$$

(ii) $\mathbb{P}(y_1, y_2+y_3, y_4) = \mathbb{P}(y_1, y_2, m-y_2, y_4)$ tal que $y_1+y_2+m-y_2+y_4 = 35$ con $0 \leq y_2 \leq 35$ ó $\mathbb{P}(y_1, y_2+y_3, y_4) = 0$ si $y_1+y_2+m-y_2+y_4 \neq 35$ con $0 \leq y_2 \leq 35$.

Luego, para $y_1+y_2+m-y_2+y_4 = 35$ con $0 \leq y_2 \leq 35$

$$\mathbb{P}(y_1, y_2+y_3, y_4) = \sum_{y_2=0}^{35} \frac{35!}{y_1!y_2!(35-y_1-y_4)!y_4!}(0,52)^{y_1}(0,35)^{y_2}(0,10)^{35-y_1-y_4}(0,03)^{y_4}$$

(iii) Las marginales de una distribución Multinomial son distribuciones Binomiales, por lo tanto $Y_2+Y_3 \sim Bi(35; 0,45)$.

Por otra parte, puesto que $(y_1+y_4)+(y_2+y_3) = y_1+y_2+y_3+y_4 = 35$ no se estaría aportando información adicional si se calcula el vector aleatorio (Y_2+Y_3, Y_1+Y_4).

8. (i) El número de personas que ingresa por día a un cierto banco es una variable aleatoria de Poisson de parámetro λ. Cada persona que entra tiene probabilidad p de ser hombre y $1-p$ de ser mujer. Mostrar que el número hombres y de mujeres que ingresan por día a dicho banco son variables aleatorias independientes con distribución Poisson de parámetros λp y $\lambda(1-p)$ respectivamente.
   (ii) Recíprocamente, si el número de hombres y de mujeres que ingresan al banco por día son variables aleatorias de Poisson independientes, ¿qué puede decir sobre la distribución del número total de personas que ingresan al banco por día?
   (iii) El número de hombres y mujeres que ingresan por día a un cierto banco son variables aleatorias independientes con distribución Poisson de parámetros 2 y 2,5 respectivamente. Sabiendo que en un determinado día ingresaron únicamente dos personas al banco, calcular la probabilidad de que hayan sido un hombre y una mujer.

Solución:

Sea N = cantidad de personas que ingresaron, por lo tanto $N \sim \mathcal{P}(\lambda)$.

Notar que si X = cantidad de hombres que ingresaron e Y = cantidad de mujeres que ingresaron, entonces $N = X+Y$.

(i) Por probabilidad total se tiene que

$$\mathbb{P}(X=i,Y=j)=\mathbb{P}(X=i,Y=j|N=i+j)\mathbb{P}(N=i+j)+\underbrace{\mathbb{P}(X=i,Y=j|N\neq i+j)}_{=0}\mathbb{P}(N\neq i+j)\Rightarrow$$

$$\mathbb{P}(X=i,Y=j)=\mathbb{P}(X=i,Y=j)=\mathbb{P}(X=i,Y=j|N=i+j)\mathbb{P}(N=i+j)$$

Por otra parte, $\mathbb{P}(N=i+j)=e^{-\lambda}\dfrac{\lambda^{i+j}}{(i+j)!}$ y es claro que la probabilidad de que entren i hombres de un total de $i+j$ personas es $\begin{pmatrix} i+j \\ i \end{pmatrix} p^i(1-p)^j$.

Por lo tanto,

$$\mathbb{P}(X=i,Y=j)=\begin{pmatrix} i+j \\ i \end{pmatrix} p^i(1-p)^j e^{-\lambda}\frac{\lambda^{i+j}}{(i+j)!}=e^{-\lambda}\frac{(\lambda p)^i}{i!}\frac{[\lambda(1-p)]^j}{j!}=\frac{e^{-\lambda p}(\lambda p)^i}{i!}\frac{e^{-\lambda(1-p)}[\lambda(1-p)]^j}{j!}$$

Luego,

$$\mathbb{P}(X=i)=\sum_{j=0}^{\infty}\frac{e^{-\lambda p}(\lambda p)^i}{i!}\frac{e^{-\lambda(1-p)}[\lambda(1-p)]^j}{j!}=\frac{e^{-\lambda p}(\lambda p)^i}{i!}e^{-\lambda(1-p)}\underbrace{\sum_{j=0}^{\infty}\frac{[\lambda(1-p)]^j}{j!}}_{e^{\lambda(1-p)}}\Rightarrow$$

$$\mathbb{P}(X=i)=\frac{e^{-\lambda p}(\lambda p)^i}{i!}.$$

$$\mathbb{P}(Y=j)=\sum_{i=0}^{\infty}\frac{e^{-\lambda p}(\lambda p)^i}{i!}\frac{e^{-\lambda(1-p)}[\lambda(1-p)]^j}{j!}=\frac{e^{-\lambda(1-p)}[\lambda(1-p)]^j}{j!}e^{-\lambda p}\underbrace{\sum_{j=0}^{\infty}\frac{(\lambda p)^i}{i!}i!}_{e^{\lambda p}}\Rightarrow$$

$$\mathbb{P}(Y=j)=\frac{e^{-\lambda(1-p)}[\lambda(1-p)]^j}{j!}.$$

(ii) Si $X\sim\mathcal{P}(\lambda_1)$ e $Y\sim\mathcal{P}(\lambda_2)$, entonces $N=X+Y\sim\mathcal{P}(\lambda_1+\lambda_2)$.

(iii) Sea $X\sim\mathcal{P}(2)$ e $Y\sim\mathcal{P}(2,5)$, entonces $N=X+Y\sim\mathcal{P}(4,5)$.

$$\mathbb{P}(X=1 \text{ y } Y=1|N=2)=\frac{\mathbb{P}(X=1 \text{ y } Y=1 \text{ y } N=2)}{\mathbb{P}(N=2)}=\frac{\mathbb{P}(X=1)\mathbb{P}(Y=1)}{\mathbb{P}(N=2)}=\frac{2e^{-2}}{1!}\frac{(0,25)e^{-0,25}}{1!}\frac{2!}{(4,5)^2e^{-4,5}}.$$

9. Sea (X,Y) un vector aleatorio absolutamente continuo con función de densidad conjunta

$$f_{XY}(x,y)=\begin{cases}k(3y+x) & 0\le x\le y\le 1\\ 0 & \text{caso contrario}\end{cases}$$

Hallar k, f_X, f_Y, F_X, F_Y, $\mathbb{P}\left(X<\dfrac{Y}{3}\right)$ y $\mathbb{P}\left(X^2-Y^2=\dfrac{1}{3}\right)$.

Solución:

$$1=\int_{\mathbb{R}^2}f_{XY}(x,y)dxdt\Rightarrow\int_0^1\int_x^1 k(3y+x)dydx\Rightarrow\frac{1}{k}=\int_0^1\left(\frac{3y^2}{2}+xy\Big|_x^1\right)dx\Rightarrow$$

$$\frac{1}{k}=\int_0^1\left(\frac{3}{2}+x-\frac{5x^2}{2}\right)dx=\frac{7}{6}\Rightarrow k=\frac{6}{7}\Rightarrow f_{XY}(x,y)=\begin{cases}\frac{6}{7}(3y+x) & 0\le x\le y\le 1\\ 0 & \text{caso contrario}\end{cases}$$

$$f_X(x)=\int_{\mathbb{R}}f_{XY}(x,y)dy=\frac{6}{7}\int_x^1(3y+x)dy=\frac{6}{7}\left(\frac{3y^2}{2}+xy\right)\Big|_x^1=\frac{6}{7}\left(\frac{3}{2}+x-\frac{5x^2}{2}\right).$$

$$f_Y(x)=\int_{\mathbb{R}}f_{XY}(x,y)dx=\frac{6}{7}\int_0^y(3y+x)dx=\frac{6}{7}\left(3yx+\frac{x^2}{2}\right)\Big|_0^y=\frac{6}{7}\left(3y^2+\frac{y^2}{2}\right)=3y^2.$$

$$F_X(x)=\begin{cases}0 & x<0\\ \frac{9}{7}x+\frac{3}{7}x^2-\frac{5}{7}x^3 & 0\le x<1\\ 1 & x\ge 1\end{cases}\quad \text{y } F_Y(y)=\begin{cases}0 & y<0\\ y^3 & 0\le y<1\\ 1 & y\ge 1\end{cases}$$

$$\mathbb{P}\left(X<\frac{Y}{3}\right)=\int_0^1\left(\int_0^{y/3}\frac{6}{7}(3y+x)dx\right)dy=\int_0^1\left(3yx+\frac{x^2}{2}\right)\Big|_0^{y/3}dx=\frac{6}{7}\int_0^1\frac{19}{18}y^2dy=\frac{19}{42}.$$

Como X^2-Y^2 es una variable aleatoria continua, entonces $\mathbb{P}\left(X^2-Y^2=\dfrac{1}{3}\right)=0$.

10. Dada una probabilidad $\mathbb{P}$ definida sobre $(\mathbb{R}^d, \mathcal{B}_d)$ definimos su soporte como el conjunto

$$sop(\mathbb{P}) = \{x \in \mathbb{R}^d : \mathcal{P}(Q_\varepsilon(x)) > 0 \text{ para todo } \varepsilon > 0\}$$

$$\text{donde } Q_\varepsilon(x) = \left\{ y \in \mathbb{R}^d : \max_{1\leq k\leq d} |x_k - y_k| < \varepsilon \right\}$$

(i) Sea $X = (X_1, X_2)$ es un vector aleatorio en $\mathbb{R}^2$. Mostrar que si $sop(P_X) \neq sop(P_{X_1}) \times sop(P_{X_1})$ entonces X_1 y X_2 no son independientes.

(ii) Mostrar con un ejemplo que no vale necesariamente la recíproca.

(iii) Probar que si X es un vector aleatorio en $\mathbb{R}^2$ con distribución uniforme en un rectángulo entonces sus coordenadas son variables aleatorias independientes. ¿Qué sucede si X tiene distribución uniforme sobre un pentágono?

11. Tres integrantes de un equipo de salto juegan un campeonato. Se asigna al equipo la mejor de las tres distancias obtenidas. El atleta A salta con distribución uniforme en el intervalo $[7,8]$ mientras que el atleta B lo hace con una distribución absolutamente continua con función de densidad

$$f(x) = \left(\frac{9-x}{2}\right) \mathbb{1}_{[7,9]}(x)$$

y el atleta C lo hace con otra distribución absolutamente continua cuya densidad es

$$f(x) = \left(\frac{x-7}{2}\right) \mathbb{1}_{[7,9]}(x)$$

(i) Hallar la distribución de la distancia asignada al equipo.

(ii) Encontrar la probabilidad de que la distancia sea mayor que $8,2$.

Solución:

(i) Sea $D = \max\{A, B, C\}$, y $i \in [7,9)$

$$\mathbb{P}_D(D \leq k) = \mathbb{P}(A \leq k \text{ y } B \leq k \text{ y } C \leq k) = \mathbb{P}(A \leq k)\mathbb{P}(B \leq k)\mathbb{P}(C \leq k)$$

$$\left.\begin{array}{r} \mathbb{P}(A \leq k) = \dfrac{1}{2}\displaystyle\int_7^k 1dx = \dfrac{k-7}{2} \\ \mathbb{P}(B \leq k) = \dfrac{1}{2}\displaystyle\int_7^k (9-x)dx = \dfrac{1}{2}\left[9k - \dfrac{k^2}{2} + \dfrac{49}{2}\right] \\ \mathbb{P}(C \leq k) = \dfrac{1}{2}\displaystyle\int_7^k (x-7)dx = \dfrac{1}{2}\left[\dfrac{k^2}{2} - 7k + \dfrac{49}{2}\right] \end{array}\right\} \Rightarrow F_D(k) = \begin{cases} 0 & k \leq 7 \\ \frac{1}{8}(k-7)\left(9k - \frac{k^2}{2} + \frac{49}{2}\right)\left(\frac{k^2}{2} - 7k + \frac{49}{2}\right) & 7 \leq k < 9 \\ 1 & k \geq 9 \end{cases}$$

(ii) $\mathbb{P}(D > 8,2) = 1 - \mathbb{P}(D \leq 8,2) = 1 - \dfrac{1}{8}(8,2-7)\left(9 \times 8,2 - \dfrac{(8,2)^2}{2} + \dfrac{49}{2}\right)\left(\dfrac{(8,2)^2}{2} - 7 \times 8,2 + \dfrac{49}{2}\right).$

12. (i) Dadas $X_1, \ldots, X_n$ variables aleatorias independientes e idénticamente distribuidas con función de distribución acumulada F, se definen sus estadísticos de orden $X^{(1)}, \ldots, X^{(n)}$ como aquellas variables aleatorias que se obtienen ordenando las X_i de manera creciente. En particular, tenemos que

$$X^{(1)} = \min_{1\leq i\leq n} X_i \text{ y } X^{(n)} = \max_{1\leq i\leq n} X_i$$

Hallar para cada $k = 1, \ldots, n$ la función de distribución acumulada de $X^{(k)}$ en términos de F.

(ii) Probar que

$$f_{X^{(k)}}(s) = \frac{n!}{(n-k)!(k-1)!} f(s)(F(s))^{k-1}(1-F(s))^{n-k}$$

(iii) Obtener la distribución de $X^{(1)}$ cuando F viene dada por la densidad

$$f(x) = e^{-(x-\theta)}\mathbb{1}_{(\theta,+\infty)}(x) \text{ para } \theta \in \mathbb{R}.$$

(iv) Obtener la distribución de $X^{(n)}$ cuando F es

$$F(x) = \frac{x}{\theta}\mathbb{1}_{(0,\theta)}(x) + \mathbb{1}_{(\theta,+\infty)}(x) \text{ para } \theta \in \mathbb{R}_{>0}.$$

(v) Probar que si las X_i tienen distribución uniforme en el intervalo $[0,1]$ entonces para cada $k=1,\dots,n$ la variable aleatoria $X^{(k)}$ tiene distribución $\beta(k,n-k+1)$.

Solución:

(i) $F_{X^{(k)}}(s) = \mathbb{P}(X^{(k)} \leq s) = \mathbb{P}(\text{haya al menos } k \text{ variables } X_i \leq s) = \sum_{j=k}^{n} \mathbb{P}(\text{haya exactamente } j \text{ variables } X_i \leq s)$, entonces

$$F_{X^{(k)}}(s) \underbrace{=}_{X_i \text{son iid}} \sum_{j=k}^{n} \binom{n}{j} (F(s))^j (1-F(s))^{n-j}$$

donde se puede deducir que

$$F_{X^{(k)}}(s) = F_{X^{k-1}}(s) - \binom{n}{k-1} (F(s))^{k-1}[1-F(s)]^{n-(k+1)}$$

(ii) Usaremos inducción en k.

- <u>$k=1:$</u>
  $F_{X^{(1)}} = \mathbb{P}(X^{(1)} \leq s) = \mathbb{P}(\text{hay por lo menos 1 variable } X_i \leq s)$, entonces

  $$F_{X^{(1)}} = \mathbb{P}(X^{(1)} \leq s) = 1 - \mathbb{P}(X_1 > s, \dots, X_n > s)$$

  como $X_1,\dots,X_n$ son independientes,

  $$F_{X^{(1)}} = \mathbb{P}(X^{(1)} \leq s) = 1 - \mathbb{P}(X_1 > s)\dots\mathbb{P}(X_n > s) = 1 - [(1-\mathbb{P}(X_1 \leq s)\dots\dots(1-\mathbb{P}(X_n \leq s)]$$

  y puesto que $X_1,\dots,X_n$ son identicamente distribuidas,

  $$F_{X^{(1)}} = 1 - [1-\mathbb{P}(X_1 \leq s)]^n \Rightarrow F_{X^{(1)}} = 1 - [1-\underbrace{F_{X_1}(s)}_{F(s)}]^n \Rightarrow \frac{\partial F_{X^{(1)}}}{\partial s} = f_{X^1}(s) = n[1-F(s)]^{n-1} f(s)$$

- <u>$k-1 \Rightarrow k:$</u>
  Queremos ver que

  $$f_{X^{(k)}}(s) = \frac{n!}{(n-k)!(k-1)!} f(s)(F(s))^{k-1}(1-F(s))^{n-k}$$

  Luego, usando

  $$F_{X^{(k)}}(s) = F_{X^{k-1}}(s) - \binom{n}{k-1} (F(s))^{k-1}[1-F(s)]^{n-(k+1)}$$

  se tiene que

  $$\begin{aligned}\frac{\partial F_{X^{(k)}}}{\partial s} = f_{X^k}(s) = f_{X^{(k-1)}}(s) - \Bigg[&(k-1)\binom{n}{k-1} f(s)(F(s))^{k-2}[1-F(s)]^{n-k+1} - \\ &-(n-k+1)\binom{n}{k-1} f(s)(F(s))^{k-1}[1-F(s)]^{n-k}\Bigg]\end{aligned}$$

  que resulta en

  $$f_{X^{(k)}}(s) = \frac{n!}{(n-k)!(k-1)!} f(s)(F(s))^{k-1}(1-F(s))^{n-k}$$

(iii) $\mathbb{P}(X^{(1)} \leq s) = \mathbb{P}(\text{hay por lo menos 1 variable } X_i \leq s)$, entonces

$$\mathbb{P}(X^{(1)} \leq s) = 1 - \mathbb{P}(X_1 > s, \dots, X_n > s) = 1 - (1 - \mathbb{P}(X_1 \leq s, \dots, X_n \leq s))$$

como $X_1,\dots,X_n$ son independientes,

$$\mathbb{P}(X^{(1)} \leq s) = 1 - [(1-\mathbb{P}(X_1 \leq s)\dots\dots(1-\mathbb{P}(X_n \leq s)] \underbrace{=}_{X_i \text{ iid}} 1 - [1-\mathbb{P}(X_1 \leq s)]^n \Rightarrow$$

$$\mathbb{P}(X^{(1)} \leq s) = \begin{cases} 0 & s < \theta \\ 1 - \left[1 - e^{-(s-\theta)}\right]^n & s \geq \theta \end{cases}$$

(iv) $\mathbb{P}(X^{(n)} \leq s) = \mathbb{P}(\text{hay por lo menos } n \text{ variable } X_i \leq s)$, entonces

$$\mathbb{P}(X^{(n)} \leq s) = \mathbb{P}(X_1 \leq s, \ldots, X_n \leq s)$$

como $X_1, \ldots, X_n$ son independientes,

$$p(X^{(n)} \leq s) = \mathbb{P}(X_1 \leq s) \ldots \mathbb{P}(X_n \leq s) \underbrace{=}_{X_i \text{ iid}} [\mathbb{P}(X_1 \leq s)]^n$$

$$\Rightarrow p(X^{(n)} \leq s) = \begin{cases} 0 & s < 0 \\ \left(\frac{s}{\theta}\right)^n & 0 \leq s < \theta \\ 1 & s \geq \theta \end{cases}$$

(v) Sabemos ya que la función de densidad de $X^{(k)}$ viene dada por

$$f_{X^{(k)}}(s) = \frac{n!}{(n-k)!(k-1)!} f(s)(F(s))^{k-1}(1-F(s))^{n-k}$$

como $X_1, \ldots, X_n \sim \mathcal{U}([0,1])$, se tiene que para $x \in [0,1]$

$$f_{X^{(k)}}(s) = \frac{n!}{(n-k)!(k-1)!} x^{k-1}(1-x)^{n-k} = \frac{\Gamma(n+1)}{\Gamma(k)\Gamma(n-k+1)} x^{k-1}(1-x)^{n-k+1-1}$$

entonces $X^{(k)} \sim \beta(k, n-k+1)$.

13. Sean $X_1, \ldots, X_n$ variables aleatorias independientes con distribución exponencial de parámetros $\alpha_1, \ldots, \alpha_n$ respectivamente.

(i) Mostrar que la distribución de $X^{(1)}$ es exponencial. ¿De qué parámetro?

(ii) Probar que

$$\mathbb{P}\left(X_k = \min_{1 \leq i \leq n} X_i\right) = \frac{\alpha_k}{\alpha_1 + \ldots + \alpha_n}$$

(iii) Calcular $\mathbb{P}\left(\min_{1 \leq i \leq n} Y_i \leq 2\right)$, donde cada $i = 1, \ldots, n$ se define $Y_i = [X_i] + 1$.

Solución:

(i) Por el ejercicio anterior, sabemos que $F_{X^{(1)}}(s) = \mathbb{P}(X^{(1)} \leq s) = \mathbb{P}(\text{hay por lo menos 1 variable } X_i \leq s)$, entonces

$$F_{X^{(1)}}(s) = \mathbb{P}(X^{(1)} \leq s) = 1 - \mathbb{P}(X_1 > s, \ldots, X_n > s)$$

como $X_1, \ldots, X_n$ son independientes,

$$F_{X^{(1)}}(s) = 1 - \mathbb{P}(X_1 > s) \ldots \mathbb{P}(X_n > s) = 1 - [(1 - \mathbb{P}(X_1 \leq s) \ldots\ldots (1 - \mathbb{P}(X_n \leq s)] \Rightarrow$$

$$F_{X^{(1)}}(s) = 1 - \{[1 - (1 - e^{\alpha_1 s})] \ldots\ldots [1 - (1 - e^{\alpha_n s})]\} = 1 - e^{\alpha_1 s} \ldots e^{\alpha_n s} = 1 - e^{(\alpha_1 + \ldots + \alpha_n)s}$$

$$\Rightarrow X^{(1)} \sim \mathcal{E}(\alpha_1 + \ldots + \alpha_n).$$

14. Dos servidores A y B procesan trabajos a medida que van llegando. El tiempo que tarda el servidor A en procesar un trabajo es una variable aleatoria $X \sim \mathcal{E}(\lambda_1)$ mientras que el tiempo que tarda el servidor B es una variable aleatoria $Y \sim \mathcal{E}(\lambda_2)$. Ambos servidores actúan en forma independiente.

(i) Dos trabajos llegan simultáneamente y es atendido uno por A y otro por B. ¿Cuál es la probabilidad de que el servidor A termine con su trabajo antes que B?

(ii) Supongamos que tres trabajos llegan simultáneamente. Uno es atendido por A, otro por B y el tercero queda esperando a que uno de los servidores se libere. Hallar la probabilidad de que el último trabajo en ser atendido sea el último en ser completado si $\lambda_1 = \lambda_2$.

Sugerencia: Hallar la distribución de $X^{(2)} - X^{(1)}$. ¿Pertenece a una familia conocida?

15. (i) Sea X una variable aleatoria con distribución $\Gamma(\alpha, \lambda)$. Probar que $cX \sim \Gamma\left(\alpha, \frac{\lambda}{c}\right)$ para todo $c > 0$.

(ii) Sean X e Y variables aleatorias independientes tales que $X \sim \Gamma(\alpha_1, \lambda)$ e $Y \sim \Gamma(\alpha_2, \lambda)$. Probar que $X + Y$ y $\frac{X}{X+Y}$ son independientes con distribución $\Gamma(\alpha_1 + \alpha_2, \lambda)$ y $\beta(\alpha_1, \alpha_2)$, respectivamente.

(iii) Teniendo en cuenta el ejercicio 7 de la Práctica 4, deducir que si $Z_1, \ldots, Z_n$ son variables aleatorias independientes con distribución normal estándar entonces $Z_1^2 + \ldots + Z_n^2$ tiene distribución $\Gamma\left(\frac{n}{2}, \frac{1}{2}\right)$.

16. Sean X e Y variables aleatorias independientes con distribución $N(0, \sigma^2)$. Sea (ρ, θ) la expresión de (X, Y) en coordenadas polares, es decir $(X, Y) = (\rho\cos(\theta), \rho\sin(\theta))$ con $\rho \geq 0$ y $0 \leq \theta < 2\pi$.

(i) Probar que ρ y θ son variables aleatorias independientes y hallar su distribución. ¿Es alguna de ellas una distribución conocida?¿Cuál?

(ii) Hallar la distribución de ρ^2. ¿Es alguna distribución conocida? ¿Cuál?

(iii) Hallar la probabilidad de que el par (X, Y) caiga en el círculo de centro en el origen y radio σ. Observar que esta probabilidad no depende de σ.

17. Sean X_1, X_2 y X_3 variables aleatorias independientes con distribución $\mathcal{U}([-1, 1])$.

(i) Sea $U = \dfrac{X_1 + X_2}{2}$. Verificar que

$$F_u(u) = (u+1)\mathbb{1}_{(-1,0)}(u) + (1-u)\mathbb{1}_{(0,1)}(u).$$

(ii) Sea $Z = \dfrac{X_1 + X_2 + X_3}{3}$. Usando el item anterior verificar que

$$f_Z(z) = \begin{cases} \frac{27}{16}(1-|z|)^2 & \frac{1}{3} < |z| < 1 \\ \frac{9}{3} - \frac{27}{8}z^2 & |z| \leq \frac{1}{3} \\ 0 & |z| \geq 1 \end{cases}$$

(iii) Verificar que f_U es continua y que f_Z es derivable.

18. (i) Probar que si $X \sim N(\mu, \sigma^2)$ entonces $aX + b \sim N(b + a\mu, a^2\sigma^2)$ para todo $a, b \in \mathbb{R}$.

(ii) Sean X e Y variables aleatorias independientes con distribución normal estándar y $v, w \in \mathbb{R}^2$ dos vectores ortogonales de norma uno. Probar que $V = v \cdot (X, Y)$ y $W = w \cdot (X, y)$ son variables aleatorias independientes con distribución normal. ¿De qué parámetros?

(iii) Deducir que si X e Y son variables aleatorias independientes con distribución $N(\mu, \sigma^2)$ entonces $X+Y$ y $X-Y$ son independientes.

(iv) Deducir que si $X \sim N(\mu_1, \sigma_1^2)$ e $Y \sim N(\mu_2, \sigma_2^2)$ son variables aleatorias independientes entonces $aX + bY + c \sim N(a\mu_1 + b\mu_2 + c, a^2\sigma_1^2 + b^2\sigma_2^2)$ para todo $a, b, c \in \mathbb{R}^6$.

19. Sean Z, X_1 y X_2 variables aleatorias independientes tales que $Z \sim N(0, 1)$, $X_1 \sim \chi^2(n)$ y $X_2 \sim \chi^2(m)$.

(i) Probar que $U = \dfrac{Z}{\sqrt{X/n}}$ tiene distribución t de Student con n grados de libertad cuya densidad viene dada por

$$f_U(u) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{\pi n}}\left(1 + \frac{u^2}{n}\right)^{-\frac{n+1}{2}}$$

(ii) (a) Deducir en particular que si Z_1 y Z_2 son variables aleatorias independientes con distribución normal estándar entonces el cociente $\dfrac{Z_1}{|Z_2|}$ tiene distribución $C(0, 1)$.

(b) Probar que $\dfrac{Z_1}{Z_2}$ también tiene distribución $C(0, 1)$.
Sugerencia: Mostrar que si Z y W son variables aleatorias independientes tales que Z tiene distribución simétrica respecto del cero entonces los vectores aleatorios (Z, W) y $(-Z, W)$ tienen la misma distribución.

(c) Concluir que si X es una variable aleatoria con distribución $C(0, 1)$ entonces $\dfrac{1}{X}$ también lo es.

(iii) Probar que $V = \dfrac{X_1/n}{X_2/m}$ tiene distribución F de Snedecor con n y m grados de libertad cuya densidad viene dada por

$$f_V(v) = \frac{\Gamma\left(\frac{n+n}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{m}{2}\right)}\left(\frac{n}{m}\right)^{n/2} v^{(n/2)-1}\left(1 + \frac{n}{m}v\right)^{-\frac{n+m}{2}} \mathbb{1}_{(0,\infty)}(v)$$

20. Sean $X_1, \ldots, X_n$ variables aleatorias absolutamente continuas independientes e idénticamente distribuidas con función de densidad f y consideremos el vector aleatorio $\overline{X} = (X^{(1)}, \ldots, X^{(n)})$ conformado por sus estadísticos de orden. Mostrar que $\overline{X}$ es absolutamente continuo y que su función de densidad es

$$f_{\overline{X}}(x) = n! \prod_{i=1}^{n} f(x_i) \mathbb{1}_{\{x: 0 < x_1 < \ldots < x_n < T\}}(X)$$

21. Sea N un proceso de Poisson. Dados $T > 0$ y $n \in \mathbb{N}$, mostrar que para cada $0 \leq a < b \leq T$ la variable aleatoria $N_{(a,b]}$ condicionada al evento $\{N_{(0,T]} = n\}$ tiene distribución $Bi(n, p_{b-a,T})$, donde $p_{b-a,T} := \dfrac{b-a}{T}$. Es decir, para todo $0 \leq k \leq n$ se tiene

$$\mathbb{P}\left(N_{(a,b]} = k | N_{(0,T]} = n\right) = \binom{n}{k} p_{b-a,T}^{k} (1 - p_{b-a,T})^{n-k}$$

¿Se anima a conjeturar, basándose en este resultado, cuál debería ser la distribución conjunta condicionada al evento $\{N_{(0,T]} = n\}$ de los n puntos del proceso de Poisson sobre el intervalo $[0, T]$?

Práctica 6

Convergencia en distribución

1. Sea X una variable aleatoria a valores en el intervalo $[0,1]$. Hallar el límite en distribución de $(X^n)_{n\in\mathbb{N}}$ ¿Se trata de alguna distribuciíon conocida?

Solución:

Sea $Y = X^n$, entonces

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(X \leq \sqrt[n]{y}) = F_X(\sqrt[n]{y}) \Rightarrow$$

$$\lim_{n\to\infty} F_Y(y) = \lim_{n\to\infty} F_X(\sqrt[n]{y}) = \lim_{n\to\infty} F_X(y^{1/n}) = \begin{cases} F_X(1) & y \neq 0 \\ 0 & y = 0 \end{cases}$$

2. Sean X e Y variables aleatorias. Para cada $n \in \mathbb{N}$ definimos la variable aleatoria $Z_n = \frac{1}{n}X + \left(1 - \frac{1}{n}\right)Y$. Hallar el límite en distribución de $(Z_n)_{n\in\mathbb{N}}$.

3. Sean $(X_n)_{n\in\mathbb{N}}$ y X variables aleatorias discretas a valores en $\mathbb{N}\cup\{0\}$. Mostrar que $X_n \xrightarrow[n\to\infty]{D} X$ si y sólo si $\lim\limits_{n\to\infty} p_n(k) = p(k)$ para todo $k \in \mathbb{N}_0$.

Solución:

$(\Rightarrow)$: Como $X_n \xrightarrow[n\to\infty]{D} X$, entonces $\lim\limits_{n\to\infty} F_{X_n}(X) = F_X(x)$ para todo x punto de continuidad de F_X.

Luego, para $k = 0, 1, 2, \ldots$ se tiene que

$$\begin{aligned}
\lim_{n\to\infty} p_n(k) &= \lim_{n\to\infty} \left[F_{X_n}\left(k + \frac{1}{2}\right) - F_{X_n}\left(k - \frac{1}{2}\right)\right], \quad \text{puesto que las } X_n \text{ son discretas.} \\
&= \lim_{n\to\infty} F_{X_n}\left(k + \frac{1}{2}\right) - \lim_{n\to\infty} F_{X_n}\left(k - \frac{1}{2}\right), \quad \text{usando que } X_n \xrightarrow[n\to\infty]{D} X \\
&= F_X\left(k + \frac{1}{2}\right) - F_X\left(k - \frac{1}{2}\right) = p(k) \quad \text{para todo } k \in \mathbb{N}_0
\end{aligned}$$

$(\Leftarrow)$: Si $\lim\limits_{n\to\infty} p_n(k) = p(k)$ para todo $k \in \mathbb{N}_0$, entonces para todo $x \in \mathbb{R}$,

$$\lim_{n\to\infty} F_{X_n}(x) = \lim_{n\to\infty} \mathbb{P}(X_n \leq x) = \lim_{n\to\infty} \sum_{k=0}^{[x]} p_n(k)$$

donde $[x]$ es la parte entera de x.

Dado que para cada x fijo, el conjunto $\{1, \ldots, [x]\}$ es finito se tiene que

$$\lim_{n\to\infty} \left(\sum_{k=0}^{[x]} p_n(k)\right) = \sum_{k=0}^{[x]} \left(\lim_{n\to\infty} p_n(k)\right) \Rightarrow$$

$$\lim_{n\to\infty} F_{X_n}(x) = \sum_{k=0}^{[x]} \left(\lim_{n\to\infty} p_n(k)\right) = \sum_{k=0}^{[x]} p(k) = \mathbb{P}(X \leq x) = F_X(x) \Rightarrow X_n \xrightarrow[n\to\infty]{D} X.$$

4. Sean $(p_n)_{n\in\mathbb{N}}$ y $(\lambda_n)_{n\in\mathbb{N}}$ dos sucesiones convergentes tales que vale $0 < p_n < 1$ y $\lambda > 0$ para todo $n \in \mathbb{N}$. Sean $p = \lim\limits_{n\to\infty} p_n$ y $\lambda = \lim\limits_{n\to\infty} \lambda_n$.

   (i) Si $X_n \sim Bi(k, p_n)$ para cada $n \in \mathbb{N}$ entonces $X_n \xrightarrow[n\to\infty]{D} X$ con $X \sim Bi(k,p)$.

   (ii) Si $Y_n \sim \mathcal{G}(p_n)$ para cada $n \in \mathbb{N}$ entonces $Y_n \xrightarrow[n\to\infty]{D} Y$ con $Y \sim \mathcal{G}(p)$.

   (ii) Si $Z_n \sim \mathcal{P}(\lambda_n)$ para cada $n \in \mathbb{N}$ entonces $Z_n \xrightarrow[n\to\infty]{D} Z$ con $Z \sim \mathcal{P}(\lambda)$.

Solución:

(i) $\lim\limits_{n\to\infty} F_{X_n}(x) = \lim\limits_{n\to\infty} \sum\limits_{i=0}^{x} \binom{k}{i} p_n^i (1-p_n)^{k-1}$ como las sumas son finitas y $\lim\limits_{n\to\infty} p_n < \infty$

$$\lim_{n\to\infty} F_{X_n}(x) = \sum_{i=0}^{x} \binom{k}{i} \left(\lim_{n\to\infty} p_n^i\right)\left(\lim_{n\to\infty}(1-p_n)^{k-1}\right) = \sum_{i=0}^{x} \binom{k}{i} p^i(1-p)^{k-1} = F_X(x)$$

con $X \sim Bi(k,p)$.

(ii) $\lim\limits_{n\to\infty} F_{Y_n}(y) = \lim\limits_{n\to\infty} \sum\limits_{i=0}^{y} p_n(1-p_n)^{i-1} = \lim\limits_{n\to\infty} [1-(1-p_n)^y] = 1 - \lim\limits_{n\to\infty}(1-p_n)^y = 1-(1-p)^y = F_Y(y)$
con $Y \sim \mathcal{G}(p)$.

(iii) $\lim\limits_{n\to+\infty} F_{Z_n}(z) = \lim\limits_{n\to+\infty} \left(\sum\limits_{i=0}^{z} \frac{\lambda_n^i}{i!} e^{-\lambda_n}\right)$ como $\lim\limits_{n\to\infty} \lambda_n < \infty$ y la suma es fnita

$$\lim_{n\to+\infty} F_{Z_n}(z) = \lim_{n\to+\infty} e^{-\lambda_n}\left(\sum_{i=0}^{z} \lim_{n\to+\infty} \frac{\lambda_n^i}{i!}\right) = e^{-\lambda}\sum_{i=0}^{z}\frac{\lambda^i}{i!} = F_Z(z), \quad \text{con } Z \sim \mathcal{P}(\lambda)$$

5. Sea X_n una variable aleatoria con distribución binomial de parámetros n y p_n. Probar que si p_n tiende a cero cuando n tiende a infinito de manera tal que $\lim\limits_{n\to+\infty} np_n = \lambda > 0$ entonces $X_n \xrightarrow[n\to\infty]{D} X \sim \mathcal{P}(\lambda)$.

Solución:

$$\lim_{n\to\infty} F_{X_n}(x) = \lim_{n\to\infty}\left(\sum_{i=0}^{x}\binom{n}{i} p_n^i(1-p_n)^{n-i}\right) = \lim_{n\to\infty}\left[\sum_{i=0}^{x}\binom{n}{i}\left(\frac{p_n}{1-p_n}\right)^i(1-p_n)^n\right] \Rightarrow$$

$$\lim_{n\to\infty} F_{X_n}(x) = \lim_{n\to\infty}(1-p_n)^n\left[\sum_{i=0}^{x}\frac{n!}{(n-1)!}\frac{1}{i!}\left(\frac{p_n}{1-p_n}\right)^i\right]$$

recordando que $\lim\limits_{n\to\infty}\dfrac{n!}{(n-i)!n^i} = 1$, entonces

$$\lim_{n\to\infty} F_{X_n}(x) = \lim_{n\to\infty}(1-p_n)^n\left[\sum_{i=0}^{x}\frac{1}{i!}\left(\frac{np_n}{1-p_n}\right)^i\right]$$

por otra parte, dado que $\lim\limits_{n\to\infty} p_n = 0$ y $\lim\limits_{n\to\infty} np_n = \lambda$,

$$\lim_{n\to\infty} F_{X_n}(x) = \lim_{n\to\infty}(1-p_n)^n\left[\sum_{i=0}^{x}\frac{1}{i!}\lambda^i\right] = \sum_{i=0}^{x}\frac{\lambda^i}{i!}\left[\lim_{n\to\infty}(1-p_n)^n\right]$$

además,

$$\lim_{n\to\infty}(1-p_n)^n = \lim_{n\to\infty}\left[\left(1-\frac{1}{\frac{n}{np_n}}\right)^{-\frac{n}{np_n}}\right]^{-np_n} = e^{-\lambda}$$

entonces,

$$\lim_{n\to\infty} F_{X_n}(x) = \sum_{i=0}^{x}\frac{\lambda^i}{i!}\left[\lim_{n\to\infty}(1-p_n)^n\right] = \sum_{i=0}^{x}\frac{\lambda^i}{i!}e^{-\lambda} = F_X(x), \quad \text{con } X \sim \mathcal{P}(\lambda)$$

6. De un bolillero que contiene en su interior B bolillas blancas y N bolillas negras se extraen sucesivamente y sin reposición n de ellas. Sea $X_{B,N}$ la cantidad de bolillas blancas obtenidas

(i) ¿Cuál es la distribución de $X_{B,N}$?

(ii) Probar que si B y N tienden a infinito de modo tal que $\frac{B}{N+B} \to p$ entonces $X_{B,N} \xrightarrow[n\to\infty]{D} X \sim Bi(n,p)$.

(iii) Establecer la convergencia en distribución del item anterior en términos de distribuciones conocidas.

Solución:

(i) sea $k = \text{mín}\{B,n\}$, entonces $R_{X_{B,N}} = \{0,1,\ldots,k\}$ y, por lo tanto, $\mathbb{P}_{X_{B,N}}(k) = \dfrac{\left(\begin{array}{c} B \\ k \end{array}\right)\left(\begin{array}{c} N \\ n-k \end{array}\right)}{\left(\begin{array}{c} B+N \\ n \end{array}\right)}$; es decir, $X_{B,N} \sim \mathcal{H}(B+N,B,n)$.

(ii) $$F_{X_{B,N}}(x) = \sum_{k=0}^{x} \frac{\left(\begin{array}{c} B \\ k \end{array}\right)\left(\begin{array}{c} N \\ n-k \end{array}\right)}{\left(\begin{array}{c} B+N \\ n \end{array}\right)} = \sum_{k=0}^{x} \frac{B!}{k!(B-k)!}\frac{N!}{(n-k)!(N-n+k)!}\frac{n!(B+N-n)!}{(B+N)!}$$, entonces

$$F_{X_{B,N}}(x) = \sum_{k=0}^{x} \left(\begin{array}{c} n \\ k \end{array}\right) \underbrace{\frac{B!N!}{(B+N)!}\frac{(B+N-n)!}{(B-k)!(N-n+k)!}}_{\substack{\text{hay k términos que equivalen a } \frac{B}{B+N} \\ \text{hay k términos que equivalen a } \frac{N}{B+N}}} \xrightarrow[\substack{\frac{B}{B+N}\to p \\ \frac{N}{B+N}\to(1-p)}]{n\to\infty} \sum_{k=0}^{x} \left(\begin{array}{c} n \\ k \end{array}\right) p^k(1-p)^{n-k} = F_X(x)$$

con $X \sim Bi(n,p)$.

(iii) Sea $X_n \sim \mathcal{H}(B+N, p(B+N), n) \Rightarrow X_n \xrightarrow[n\to\infty]{D} X$ con $X \sim Bi(n,p)$.

7. De un bolillero que contiene en su interior B bolillas blancas y N negras se extraen sin reposición bolillas al azar. Sea $X_{B,N}$ la cantidad de extracciones que fueron necesarias para obtener la primer bolilla blanca. Probar que si B y N tienden a infinito de modo tal que $\frac{B}{N+B} \to p$ entonces $X_{B,N} \xrightarrow[n\to\infty]{D} X \sim \mathcal{G}(p)$.

Solución:

$$\mathbb{P}(X_{B,N} = k) = \frac{N}{B+N}\frac{N-1}{B+N-1}\cdots\cdots\frac{N-(k-2)}{B+N-(k+2)}\frac{B}{B+N-(k+3)} \Rightarrow$$

$$F_{X_{B,N}}(k) = \sum_{k=1}^{x} \underbrace{\frac{N}{B+N}}_{\to 1-p}\underbrace{\frac{N-1}{B+N-1}}_{\to 1-p}\cdots\cdots\underbrace{\frac{B}{B+N-(k+3)}}_{\to 1-p} = \sum_{k=1}^{x}(1-p)^{k-1}p = F_X(x), \text{ con } X \sim \mathcal{G}(p)$$

8. Sean $(\mu_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ y $(\sigma_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ dos sucesiones convergentes tales que $\lim\limits_{n\to+\infty} \mu_n = \mu$ y $\lim\limits_{n\to+\infty} \sigma_n = \sigma$.

(i) Probar que si $X_n \sim N(\mu_n, \sigma_n^2)$ para cada $n \in \mathbb{N}$ entonces $X_n \xrightarrow[n\to\infty]{D} X \sim N(\mu,\sigma^2)$.

(ii) ¿Qué sucede con la convergencia en distribución si $|\mu| = +\infty$ o $|\sigma| = +\infty$?

Solución:

(i) $F_{X_n}(x) = \displaystyle\int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma_n} \exp\left\{-\frac{(x-\mu_n)^2}{2\sigma}\right\} dx$ y usando la continuidad uniforme,

$$\lim_{n\to\infty} F_{X_n}(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \left(\lim_{n\to\infty} \frac{1}{\sigma_n} \exp\left\{-\frac{(x-\mu_n)^2}{2\sigma}\right\}\right) dx = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}}\frac{1}{\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma}\right\} dx = F_X(x)$$

con $X \sim N(\mu,\sigma^2)$.

9. Hallar el límite en distribución de la sucesión $(Z_n)_{n\in\mathbb{N}}$ en cada uno de los siguientes casos:

(i) Para cada $n \in \mathbb{N}$ la variable aleatoria Z_n tiene distribución uniforme en el conjunto $\left\{\dfrac{i}{n} : 1 \le i \le n\right\}$.

(ii) Para cada $n \in \mathbb{N}$ la variable aleatoria nZ_n tiene distribución geométrica de parámetro $\dfrac{\lambda}{n}$.

Solución:

(i) Sea $X \sim \mathcal{U}([0,1])$ tal que,

$$\mathbb{P}\left(Z_n = \frac{i}{n}\right) = \mathbb{P}\left(X \in \left(\frac{i-1}{n}, \frac{i}{n}\right]\right) = \frac{1}{n} \Rightarrow Z_n = \frac{[nX]+1}{n} \sim \mathcal{U}\left\{\frac{i}{n} : 1 \leq i \leq n\right\}$$

luego,

$$\lim_{n\to\infty} \mathbb{P}(Z_n \leq z) = \lim_{n\to\infty} \mathbb{P}\left(X \leq \frac{[nz]}{n}\right) = \lim_{n\to\infty} \frac{[nz]}{n} = z = \mathbb{P}(X \leq z), \text{ con } X \sim \mathcal{U}([0,1]).$$

(ii) Sea $X \sim \mathcal{E}(\alpha)$ y, definiendo, $Z_n = \dfrac{[X+1]}{n}$

$$\mathcal{P}(nZ_n \leq k) = \mathbb{P}([X+1] \leq k) = \mathbb{P}(X \in (-\infty, k)) = \int_{-\infty}^{k} \alpha e^{-\alpha x} \mathbb{1}_{(0,+\infty)}(x) = 1 - e^{-\alpha k}$$

si $\alpha = -\ln\left(1 - \dfrac{\lambda}{n}\right)$, entonces $\mathcal{P}(nZ_n \leq k) = 1 - \left(1 - \dfrac{\lambda}{n}\right)^k$ y, por lo tanto, $nZ_n \sim \mathcal{G}\left(\dfrac{\lambda}{n}\right)$.

$$F_{Z_n}(x) = \mathbb{P}\left(\frac{[X+1]}{n} \leq x\right) = \mathbb{P}([X+1] \leq nx) = 1 - \left(1 - \frac{\lambda}{n}\right)^{nx} \Rightarrow$$

$$\lim_{n\to\infty} F_{Z_n}(x) = \lim_{n\to\infty} 1 - \left(1 - \frac{\lambda}{n}\right)^{nx} = 1 - \lim_{n\to\infty} \left[\left(1 - \frac{\lambda}{n}\right)^{-\frac{n}{\lambda}}\right]^{-\frac{\lambda}{n}} = 1 - e^{-\frac{\lambda}{n}}$$

entonces, $Z_n \xrightarrow[n\to\infty]{D} Z \sim \mathcal{E}\left(\frac{\lambda}{n}\right)$.

10. Sea $(U_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias independientes con distribución $\mathcal{U}([a,b])$. Para cada $n \in \mathbb{N}$ definimos las variables aleatoria

$$Y_n = \min\{U_1, \ldots, U_n\} \quad \text{y} \quad Z_n = \max\{U_1, \ldots, U_n\}$$

   (i) Hallar los límites en distribución de las sucesiones $(Y_n)_{n\in\mathbb{N}}$ y $(Z_n)_{n\in\mathbb{N}}$.

   (ii) Hallar los límites en distribución de las sucesiones $(V_n)_{n\in\mathbb{N}}$ y $(W_n)_{n\in\mathbb{N}}$, donde para cada $n \in \mathbb{N}$

$$V_n = n(Y_n - a) \quad \text{y} \quad W_n = n(b - Z_n).$$

Solución:

(i) Notemos que

$$F_{Y_n}(u) = 1 - (1 - F_{U_1}(u))^n = \begin{cases} 0 & u < a \\ 1 - \left(1 - \dfrac{x-a}{b-a}\right)^n & a \leq u < b \\ 1 & u \geq b \end{cases} \quad \text{y} \quad F_{Z_n}(u) = (F_{U_1}(u))^n = \begin{cases} 0 & u < a \\ \left(\dfrac{x-a}{b-a}\right)^n & a \leq u < b \\ 1 & u \geq b \end{cases}$$

además, para $u \in [a,b)$ se tiene que $\dfrac{x-a}{b-a} \in [0,1)$, por lo tanto

$$\lim_{n\to\infty} \left(1 - \frac{x-a}{b-a}\right)^n = 0 \quad \text{y} \quad \lim_{n\to\infty} \left(\frac{x-a}{b-a}\right)^n = 0$$

Por lo tanto,

$$\lim_{n\to\infty} F_{Y_n}(u) = \begin{cases} 0 & u < a \\ 1 & u \geq a \end{cases} \quad \text{y} \quad \lim_{n\to\infty} F_{Z_n}(u) = \begin{cases} 0 & u < b \\ 1 & u \geq b \end{cases}$$

(ii) $F_{V_n}(u) = \mathbb{P}(V_n \leq u) \leq \mathbb{P}\left(Y_n \leq \dfrac{u}{n} + a\right) = F_{Y_n}\left(\dfrac{u}{n} + a\right) = 1 - \left[1 - F_{U_1}\left(\dfrac{u}{n} + a\right)\right]^n.$

Luego, $F_{V_n}(u) = \begin{cases} 0 & \dfrac{u}{n} + a < a \\ 1 - \left(1 - \dfrac{u}{n(b-a)}\right)^n & a \leq \dfrac{u}{n} + a \leq b \\ 1 & \dfrac{u}{n} + a > b \end{cases}$

$$\Rightarrow \lim_{n\to\infty} F_{V_n}(u) = \lim_{n\to\infty} 1 - \left(1 - \frac{u}{n(b-a)}\right)^n = \lim_{n\to\infty} \left[\left(1 - \frac{u}{n(b-a)}\right)^{\frac{(b-a)n}{x}}\right]^{\frac{x}{b-a}} = 1 - e^{-\frac{x}{b-a}}, \quad x \in (0, +\infty)$$

$$\Rightarrow V_n \xrightarrow[n\to\infty]{D} X \sim \mathcal{E}\left(\frac{1}{b-a}\right)$$

$$F_{W_n}(u) = \mathbb{P}(W_n \leq u) = \mathbb{P}\left(Z_n \geq b - \frac{u}{n}\right) = 1 - \mathbb{P}\left(Z_n \leq b - \frac{u}{n}\right) = 1 - \left[F_{U_1}\left(b - \frac{u}{n}\right)\right]^n.$$

Luego, $F_{W_n}(u) = \begin{cases} 1 & b - \frac{u}{n} < a \\ 1 - \left(1 - \frac{u}{n(b-a)}\right)^n & a \leq b - \frac{u}{n} \leq b \\ 0 & b - \frac{u}{n} \geq b \end{cases}$

$$\Rightarrow \lim_{n\to\infty} F_{W_n}(u) = \lim_{n\to\infty}\left[1 - \left(1 - \frac{u}{n(b-a)}\right)^n\right] = 1 - e^{-\frac{u}{(b-a)}}, \quad x \in (0, +\infty)$$

$$\Rightarrow W_n \xrightarrow[n\to\infty]{D} X \sim \mathcal{E}\left(\frac{1}{b-a}\right)$$

Práctica 7

Esperanza

1. Sea X el resultado que se obtiene al arrojar un dado equilibrado una vez. Si antes de arrojar el dado se ofrece la opción de elegir entre recibir $\$\frac{2}{7}$ ó $h(X) = \frac{1}{7}\$$, decidir cuál de las dos opciones es preferible, en el sentido de cuál tiene un mayor valor esperado.

Solución:
Como $R_X = \{1, 2, \ldots, 6\}$ tal que $\mathbb{P}_X(k) = \frac{1}{6}$ con $k \in \mathbb{R}_X$, se tiene que

$$\mathbb{E}(h(X)) = \frac{1}{6}\sum_{i=0}^{6} \frac{1}{i} = \frac{49}{120}$$

En particular, es preferible elegir $h(X)$ puesto que tiene mayor valor esperado.

2. En un comercio de artículos para el hogar hay 6 televisores en stock. El número de clientes que entran a comprar un televisor por semana es una variable aleatoria X con distribución $\mathcal{P}(5)$. Cada cliente que entra a comprar un televisor lo comprará si todavía hay stock. ¿Cuál es el número esperado de televisores a ser vendidos durante la próxima semana?

Solución:
Sea $Y = \begin{cases} X & X < 6 \\ 6 & X \geq 6 \end{cases}$, luego

$$\mathbb{E}(Y) = \sum_{i=0}^{5} i\frac{5^i}{i!}e^{-5} + 6\sum_{i\geq 6}^{\infty} \frac{5^i}{i!}e^{-5} = \sum_{i=0}^{5} i\frac{5^i}{i!}e^{-5} + 6\left(1 - \sum_{i=0}^{5}\frac{5^i}{i!}e^{-5}\right) \Rightarrow$$

$$\mathbb{E}(Y) = e^{-5}\left[5 + 5^2 + \frac{5^3}{2} + \frac{5^4}{6} + \frac{5^5}{24}\right] + 6 \cdot e^{-5}\left[e^5 - 1 - 5 - \frac{5^2}{2} - \frac{5^3}{6} - \frac{5^4}{24} - \frac{5^5}{120}\right]$$

3. Un juego consiste en arrojar un dado equilibrado hasta obtener un número mayor o igual que 4. Si X es el número de veces que se arroja el dado a lo largo del juego, el puntaje que se obtiene por jugar es $(4 - X)$ si $1 \leq X \leq 3$ y no se obtiene puntaje en caso contrario. ¿Cuál es el puntaje esperado de este juego?

Solución:
Notando que $\mathbb{P}(\{\text{obtener un número} \geq 4\}) = \frac{1}{2}$, entonces si X es el número de veces que se arroja el dado a lo largo del juego es claro que $X \sim \mathcal{G}\left(\frac{1}{2}\right)$.

$$\mathbb{E}(Y) = 0 + \sum_{i=0}^{3} i\mathbb{P}_X(4 - i) = \left(\frac{1}{2}\right)^3 + 2\left(\frac{1}{2}\right)^2 + 3\left(\frac{1}{2}\right)$$

4. Hallar la esperanza y varianza de las siguientes variables aleatorias:

   (i) $Bi(n, p)$. *Sugerencia:* ¿Qué distribución tiene la suma de n variables aleatorias independientes $Be(p)$?

   (ii) $\mathcal{G}(p)$. *Sugerencia:* Recordar que una serie de potencias es derivable dentro de su radio de convergencia.

   (iii) $BN(r, p)$. *Sugerencia:* ¿Qué distribución tiene la suma de r variables aleatorias independientes $\mathcal{G}(p)$?

(iv) $\mathbb{P}(\lambda)$. *Sugerencia:* Para hallar $\mathbb{V}(X)$ calcular primero $\mathbb{E}(X(X-1))$.

(v) $\Gamma(\alpha,\lambda)$. *Sugerencia:* Si $X \sim \Gamma(\alpha,\lambda)$ entonces $\int_{\mathbb{R}} f_X(x)dx = 1$ para cualquier α y λ positivos.

(vi) $\mathcal{E}(\lambda)$.

(vii) $\chi^2(n)$.

(viii) $N(\mu,\sigma^2)$.

(ix) $\mathcal{U}([a,b])$.

(x) $\beta(a,b)$. *Sugerencia:* Si $X \sim \beta(a,b)$ entonces $\int_{\mathbb{R}} f_X(x)dx = 1$ para cualquier a y b positivos.

Solución:

(i) Si $X_i \sim Be(p)$, entonces $\mathbb{E}(X_i) = p$ y $\mathbb{E}(X_i^2) = 1^2p = p$. Por lo tanto, $\mathbb{E}(X_i) = p$ y $\mathbb{V}(X_i) = p - p^2 = p(1-p)$.

Luego, si $X_1,\dots,X_n$ son iid tales que $X_i \sim Be(p)$ se tiene que $X = \sum_{i=1}^{n} X_i \sim Bi(n,p)$, y, por lo tanto,

$$\mathbb{E}(X) = \mathbb{E}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n}\mathbb{E}(X_i) = \sum_{i=1}^{n} p = np.$$

$$\mathbb{V}(X) = \mathbb{V}\left(\sum_{i=1}^{n} X_i\right) \underbrace{=}_{X_i,\, iid} \sum_{i=1}^{n}\mathbb{V}(X_i) = \sum_{i=1}^{n} p(1-p) = np(1-p).$$

(ii) $\mathbb{E}(X) = \sum_{k\in\mathbb{N}} k\mathbb{P}_X(k) = \sum_{k\in\mathbb{N}} kp(1-p)^{k-1} = -p\sum_{k\in\mathbb{N}} \frac{\partial[(1-p)^k]}{\partial p}$ y, como $p<1$

$$\mathbb{E}(X) = -p\frac{\partial}{\partial p}\left[\frac{1}{1-(1-p)}\right] = -p\frac{\partial}{\partial p}\left(\frac{1}{p}\right) = \frac{1}{p}.$$

Para calcular la varianza, es necesario saber cuánto vale $\mathbb{E}(X^2)$. Luego,

$$\mathbb{E}(X^2) = \sum_{k\in\mathbb{N}} k^2p(1-p)^{k-1} = \sum_{k\in\mathbb{N}} k(k-1)p(1-p)^{k-1} + \sum_{k\in\mathbb{N}} kp(1-p)^{k-1} = p(1-p)\sum_{k\in\mathbb{N}} k(k-1)(1-p)^{k-2} + p\sum_{k\in\mathbb{N}} kp(1-p)^{k-1} \Rightarrow$$

$$\mathbb{E}(X^2) = p(1-p)\sum_{k=2}^{\infty} k(k-1)(1-p)^{k-2} + p\sum_{k\in\mathbb{N}} kp(1-p)^{k-1} = p(1-p)\frac{\partial^2}{\partial p^2}\left[\sum_{k\in\mathbb{N}_0}(1-p)^k\right] - p\frac{\partial}{\partial p}\left[\sum_{k\in\mathbb{N}_0}(1-p)^k\right] \Rightarrow$$

$$\mathbb{E}(X^2) \underbrace{=}_{p<1} p(1-p)\frac{\partial^2}{\partial p^2}\left[\frac{1}{1-(1-p)}\right] - p\frac{\partial}{\partial p}\left[\frac{1}{1-(1-p)}\right] = p(1-p)\frac{\partial^2[p^{-1}]}{\partial p^2} - p\frac{\partial[p^{-1}]}{\partial p} = p(1-p)\frac{2}{p^3} - p\left(-\frac{1}{p^2}\right)$$

Luego, $\mathbb{V}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \frac{1}{p^2} - \frac{1}{p}$.

(iii) Si $X_i,\dots,X_n$ iid tales que $X_i \sim \mathcal{G}(p)$, entonces $X = \sum_{i=1}^{n} X_i \sim\ BN(r,p)$. Luego,

$$\mathbb{E}(X) = \mathbb{E}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n}\mathbb{E}(X_i) \underbrace{=}_{X_i,\, iid} \sum_{i=1}^{n}\mathbb{E}(X_1) = \sum_{i=1}^{n}\frac{1}{p} = \frac{r}{p}.$$

$$\mathbb{V}(X)) = \mathbb{V}\left(\sum_{i=1}^{n} X_i\right) \underbrace{=}_{X_i,\, ind.} \sum_{i=1}^{n}\mathbb{V}(X_i) \underbrace{=}_{X_i,\, iid} \sum_{i=1}^{n}\mathbb{V}(X_1) = r\left(\frac{1}{p^2} - \frac{1}{p}\right).$$

(iv) $\mathbb{E}(X) = \sum_{k\in\mathbb{N}_0} k\frac{\lambda^k}{k!}e^{-\lambda} = \lambda e^{-\lambda}\sum_{k\in\mathbb{N}}\frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda}\underbrace{\sum_{i\in\mathbb{N}_0}\frac{\lambda^i}{i!}}_{e^{\lambda}} = \lambda.$

$$\mathbb{E}(X^2) = \sum_{k\in\mathbb{N}_0} k^2\frac{\lambda^k}{k!}e^{-\lambda} = \sum_{k\in\mathbb{N}_0}[k(k-1)+k]\frac{\lambda^k}{k!}e^{-\lambda} = e^{-\lambda}\left[\sum_{k\in\mathbb{N}_0}\lambda^2\frac{\lambda^{k-2}}{(k-2)!} + \sum_{k\in\mathbb{N}_0}\lambda\frac{\lambda^{k-1}}{(k-1)!}\right] = e^{-\lambda}(\lambda^2e^{\lambda} + \lambda e^{\lambda}) = \lambda^2 + \lambda$$

Luego, $\mathbb{V}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$.

(v) $\mathbb{E}(X) = \int_{\mathbb{R}} x \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} \mathbb{1}_{(0,+\infty)}(x)dx = \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)} \frac{\lambda^\alpha}{\lambda^{\alpha+1}} \underbrace{\int_{\mathbb{R}} \frac{\lambda^{\alpha+1}}{\Gamma(\alpha+1)} x^{\alpha+1-1} e^{-\lambda x} \mathbb{1}_{(0,+\infty)}(x)dx}_{=1}$, luego

$$\mathbb{E}(X) = \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)} \frac{\lambda^\alpha}{\lambda^{\alpha+1}} = \frac{\alpha\Gamma(\alpha)}{\Gamma(\alpha)} \frac{\lambda^\alpha}{\lambda\lambda^\alpha} = \frac{\alpha}{\lambda}.$$

$$\mathbb{E}(X^2) = \int_{\mathbb{R}} x^2 \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} \mathbb{1}_{(0,+\infty)}(x)dx = \frac{\Gamma(\alpha+2)}{\Gamma(\alpha)} \frac{\lambda^\alpha}{\lambda^{\alpha+2}} \underbrace{\int_{\mathbb{R}} \frac{\lambda^{\alpha+2}}{\Gamma(\alpha+2)} x^{\alpha+2-1} e^{-\lambda x} \mathbb{1}_{(0,+\infty)}(x)dx}_{=1} \Rightarrow$$

$$\mathbb{E}(X^2) = \frac{(\alpha+1)\alpha\Gamma(\alpha)}{\Gamma(\alpha)} \frac{\lambda^\alpha}{\lambda^2\lambda^\alpha} = \frac{(\alpha+1)\alpha}{\lambda^2}.$$

Luego, $\mathbb{V}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \frac{(\alpha+1)\alpha}{\lambda^2} - \frac{\alpha^2}{\lambda^2} = \frac{\alpha}{\lambda^2}$.

(vi) Como $X \sim \mathcal{E}(\lambda) = \Gamma(1, \lambda)$ se tiene que $\mathbb{E}(X) = \frac{1}{\lambda}$ y $\mathbb{V}(X) = \frac{1}{\lambda^2}$.

(vii) Como $X \sim \chi(n) = \Gamma\left(\frac{n}{2}, \frac{1}{2}\right)$ se tiene que $\mathbb{E}(X) = \frac{n/2}{1/2} = n$ y $\mathbb{V}(X) = \frac{n/2}{1/4} = 2n$.

(viii) Sea $Z = \frac{(X-\mu)}{\sigma}$, entonces $\mathbb{E}(X) = \mu + \sigma\mathbb{E}(Z)$ y $\mathbb{V}(X) = \sigma^2\mathbb{V}(Z)$.

$$\mathbb{E}(Z) = \int_{\mathbb{R}} x f_Z(x)dx = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} x e^{-\frac{x^2}{2}} dx = -\frac{1}{\sqrt{2\pi}} \underbrace{e^{-\frac{x^2}{2}} \Big|_{-\infty}^{+\infty}}_{=0} = 0$$

$$\mathbb{E}(Z^2) = \int_{\mathbb{R}} \frac{x^2}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \underbrace{-x\, e^{-\frac{x^2}{2}} \Big|_{-\infty}^{+\infty}}_{=0} + \underbrace{\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx}_{=1} = 1$$

Luego,$\mathbb{E}(X) = \mu + \sigma\mathbb{E}(Z) = 0\mu$ y $\mathbb{V}(X) = \sigma^2\mathbb{V}(Z) = \sigma^2$.

(ix) $\mathbb{E}(X) = \int_a^b \frac{x}{b-a} dx = \frac{1}{b-a} \frac{x^2}{2}\Big|_a^b = \frac{b-a}{2}$ y $\mathbb{E}(X^2) = \int_a^b \frac{x^2}{b-a} dx = \frac{1}{b-a} \frac{x^3}{3}\Big|_a^b = \frac{1}{3} \frac{b^3-a^3}{b-a}$.

Luego, $\mathbb{V}(X) = \frac{1}{3} \frac{b^3-a^3}{b-a} - \frac{1}{4}(b+a)^2 = \frac{(b-a)^2}{12}$.

(x) $\mathbb{E}(X) = \int_{\mathbb{R}} x \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1} \mathbb{1}_{(0,1)}(x)dx = \frac{\Gamma(a+b)}{\Gamma(a+b+1)} \frac{\Gamma(a+1)}{\Gamma(a)} \underbrace{\int_0^1 \frac{\Gamma(a+b+1)}{\Gamma(a+1)\Gamma(b)} x^{a+1} x^{b-1} dx}_{=1}$, entonces

$\mathbb{E}(X^2) = \frac{a}{(a+b)}$.

$\mathbb{E}(X^2) = \int_{\mathbb{R}} x^2 \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1} \mathbb{1}_{(0,1)}(x)dx = \frac{\Gamma(a+b)}{\Gamma(a+b+2)} \frac{\Gamma(a+2)}{\Gamma(a)} \underbrace{\int_0^1 \frac{\Gamma(a+b+2)}{\Gamma(a+2)\Gamma(b)} x^{a+1} x^{b-1} dx}_{=1}$, entonces

$\mathbb{E}(X^2) = \frac{(a+1)a}{(a+b+1)(a+b)}$.

Luego, $\mathbb{V}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \frac{(a+1)a}{(a+b+1)(a+b)} - \frac{a^2}{(a+b)^2} = \frac{ab}{(a+b)^2(a+b+1)}$.

5. Sean $X_1, \ldots, X_n$ variables aleatorias positivas independientes e idénticamentes distribuídas. Probar que

$$\mathbb{E}\left(\frac{X_1 + \ldots + X_k}{X_1 + \ldots + X_n}\right) = \frac{k}{n}$$

Sugerencia: Probar que $\frac{X_i}{X_1 + \ldots + X_n} \sim \frac{X_j}{X_1 + \ldots + X_n}$ para todo par $1 \leq i, j \leq n$.

Solución:

Sea $T_i = \dfrac{X_i}{X_1 + \ldots + X_n}$, luego

$$\mathbb{P}(T_i \le k) = \mathbb{P}(X_i \le k(X_1 + \ldots + X_n)) \underbrace{=}_{X_i, iid} \mathbb{P}(X_j \le k(X_1 + \ldots + X_n)) = \mathbb{P}(T_j \le k) \Rightarrow F_{T_j}(k) = F_{T_j}(k) \Rightarrow$$

$$\frac{X_i}{X_1 + \ldots + X_n} \sim \frac{X_j}{X_1 + \ldots + X_n}$$

$$\sum_{i=1}^{n} T_i = 1 \Rightarrow 1 = \mathbb{E}\left(\sum_{i=1}^{n} T_i\right) = \sum_{i=1}^{n} \mathbb{E}(T_i) \underbrace{=}_{T_i, iid} \sum_{i=1}^{n} \mathbb{E}(T_1) = n\mathbb{E}(T_1) \Rightarrow \mathbb{E}(T_i) = \frac{1}{n} \quad \text{para cada } i \in \{1, \ldots, n\}$$

luego,

$$\mathbb{E}\left(\frac{X_1 + \ldots + X_k}{X_1 + \ldots + X_n}\right) = \mathbb{E}(T_1 + \ldots + T_k) = \sum_{i=1}^{k} \mathbb{E}(T_i) = \sum_{i=1}^{k} \frac{1}{n} = \frac{k}{n}.$$

6. Se distribuyen al azar N bolillas indistinguibles en m urnas. Sean X el número de urnas vacías, Y el número de urnas que contienen exactamente una bolilla y Z el número de urnas que contienen dos o más bolillas.

   (i) Hallar $\mathbb{E}(X)$.
   Sugerencia: Sea
   $$X_i = \begin{cases} 1 & \text{si la i-ésima urna está vacía} \\ 0 & \text{en caso contrario} \end{cases}$$
   Verificar que $X = \sum_{i=1}^{m} X_i$.

   (ii) Hallar $\mathbb{E}(Y)$ y $\mathbb{E}(Z)$.

   (iii) Un centro cultural dispone de m cuentas de correo electrónico para comunicarse con el público. Durante un día en particular, N personas envían sus inquietudes vía e-mail al centro cultural, eligiendo una cuenta al azar para hacerlo. Hallar la esperanza del número de cuentas de correo que no son usadas durante dicho día.

Solución:

(i) Puesto que $X_i \sim Be(p)$ donde p es la probabilidad de que la urna i esté vacía; ie, $p = \left(\dfrac{m-1}{m}\right)^N$.

$$\mathbb{E}(X) = \mathbb{E}\left(\sum_{i=1}^{m} X_i\right) = \sum_{i=1}^{m} \mathbb{E}(X_i) = m\left(\frac{m-1}{m}\right)^N.$$

(ii) Sea $Y_i = \begin{cases} 1 & \text{si la i-ésima urna contiene exactamente una bolilla} \\ 0 & \text{en caso contrario} \end{cases}$.

Entonces $Y_i \sim Be(p)$ donde p es la probabilidad de que la urna i contenga exactamente una bolilla; ie, $p = \dfrac{N(m-1)^{N-1}}{m^N}$.

$$\mathbb{E}(Y) = \mathbb{E}\left(\sum_{i=1}^{m} Y_i\right) = \sum_{i=1}^{m} \mathbb{E}(Y_i) = m\frac{N(m-1)^{N-1}}{m^N}.$$

Por otro lado, $Z = m - X - Y$, entonces $\mathbb{E}(Z) = m - \mathbb{E}(X) - \mathbb{E}(Y) = m - m\left(\dfrac{m-1}{m}\right)^N - N\left(\dfrac{m-1}{m}\right)^{N-1}$.

(iii) El problema es equivalente a calcular la esperanza del número de urnas vacías que hay cuando se distribuyen N bolillas indistinguibles al azar en m urnas.

Luego, $\mathbb{E}(X) = m\left(\dfrac{m-1}{m}\right)^N$.

7. **Muestreo estratificado.**

   Se quiere saber cuántos habitantes viven en una cierta ciudad. Se sabe que dicha ciudad tiene n manzanas, de las cuales n_j tienen x_j habitantes cada una ($n_1 + n_2 + \ldots = n$). Sea $m = \sum_j \dfrac{n_j x_j}{n}$ el número medio de habitantes por manzana y sea $a^2 = \dfrac{1}{n}\sum_j n_j x_j^2 - m^2$. Para averiguarlo se sortean al azar r manzanas para encuestarlas. Un encuestador es enviado a cada una de las manzanas sorteadas para contar la cantidad de habitantes que viven en ellas. Definamos las variables aleatorias

$$\begin{aligned} Z_i &= \text{cantidad de habitantes que viven en la i-ésima manzana sorteada} \\ Y &= \text{cantidad de personas encuestadas} \end{aligned}$$

(i) Hallar $\mathbb{E}(Z_i)$ y $\mathbb{V}(Z_i)$.

(ii) Mostrar que

$$\begin{aligned} \mathbb{E}(Y) &= mr \\ \mathbb{V}(Y) &= \frac{a^2 r(n-r)}{n-1} \end{aligned}$$

Solución:

(i) Es claro que $\mathbb{E}(Z_i) = m$ puesto que m representa el número medio de habitantes por manzana. Luego,

$$\mathbb{E}(Z_j^2) = \sum_j x_j^2 \mathbb{P}(Z_j = x_j) = \sum_j x_i^2 \frac{n_j}{n} \Rightarrow \mathbb{V}(Z_j) = \mathbb{E}(Z_j^2) - \mathbb{E}^2(Z_j) = \frac{1}{n}\sum_j x_j^2 n_j - m^2.$$

(ii) Es claro que $Y = \sum_{i=1}^{r} Z_i$, entonces $\mathbb{E}(Y) = \sum_{i=1}^{r} \mathbb{E}(Z_i) = rm$.

Por otra parte, $Cov(Z_i, Z_j) = \mathbb{E}(Z_i, Z_j) - m^2$. Luego,

$$\mathbb{E}(Z_i, Z_j) = \sum_{k=1}^{r}\sum_{l=1}^{r} x_k x_l \mathbb{P}(Z_i = x_j | Z_j = x_k)\mathbb{P}(Z_j = x_k), \quad \text{donde } \mathbb{P}(Z_i = x_j | Z_j = x_k)\mathbb{P}(Z_j = x_k) = \begin{cases} \dfrac{n_k n_l}{n(n-1)} & x_k \neq x_l \\ \dfrac{n_l(n_l - 1)}{n(n-1)} & x_k = x_l \end{cases}$$

$$\mathbb{E}(Z_i, Z_j) = \sum_{k=1}^{r}\sum_{l=1}^{r} x_k x_l \frac{n_k n_l}{n(n-1)} - \sum_{k=1}^{r} \frac{x_k^2 n_k^2}{n(n-1)} + \sum_{k=1}^{r} \frac{x_k^2 n_k (n_k - 1)}{n(n-1)} = \frac{1}{n(n-1)}\left[nm^2 - \sum_{k=1}^{r} x_k^2 n_k\right]$$

$$\Rightarrow \mathbb{E}(Z_i, Z_j) = \frac{1}{n(n-1)}[nm^2 - n(a^2 + m^2)] = \frac{(n-1)m^2 - a^2}{n-1} \Rightarrow Cov(Z_i, Z_j) = -\frac{a^2}{n-1}$$

Luego, $\mathbb{V}(Y) = \sum_{i=1}^{r} \mathbb{V}(Z_i) + \sum_{\substack{j=1 \\ j\neq i}}^{r} Cov(Z_i, Z_j) = ra^2 + r(r-1)\left(-\dfrac{a^2}{n-1}\right)\dfrac{a^2 r(n-r)}{n-1}$.

8. Sean $X_1, \ldots, X_n$ variables aleatorias independientes tales que

$$\mathbb{P}(X_k \leq t) = \begin{cases} 0 & t < 0 \\ t^k & 0 \leq t \leq 1 \quad k = 1, \ldots, n \\ 1 & t > 1 \end{cases}$$

(i) Sea $Y = \max(X_1, \ldots, X_n)$. Hallar F_y, f_Y y $\mathbb{E}(Y)$.

(ii) Hallar $\mathbb{E}\left(\prod_{i=1}^{n} X_i\right)$.

Solución:

(i) $F_Y(t) = \mathbb{P}(Y \leq t) = \mathbb{P}(X_1 \leq t, X_2 \leq t, \ldots, X_n \leq t) \underbrace{=}_{X_i,\, iid} \mathbb{P}(X_1 \leq t)^n = \begin{cases} 0 & t < 0 \\ \prod_{i=1}^{n} t^i & 0 \leq t < 1 \\ 1 & t \geq 1 \end{cases}$.

Por lo tanto, $F_Y(y) = y^{\frac{n(n+1)}{2}} \mathbb{1}_{(0,1)}(y) + 1\mathbb{1}_{[1,\infty)}(y)$, entonces $f_Y(y) = \dfrac{n(n+1)}{2} y^{\frac{n(n+1)-2}{2}} \mathbb{1}_{(0,1)}(y)$. Es decir, $Y \sim \beta\left(\dfrac{n(n+1)}{2}, 1\right)$ y, entonces $\mathbb{E}(Y) = \dfrac{n(n+1)}{n(n+1)+2}$.

(ii) Como $X_1, \dots, X_n$ son independientes, $\mathbb{E}\left(\prod_{i=1}^{n} X_i\right) = \prod_{i=1}^{n} \mathbb{E}(X_i)$ y

$$\mathbb{E}(X_k) = \int_{\mathbb{R}} x_k f_{X_k}(x_k) dx_k = \int_0^1 x_k k x^{k-1} dx_k = k \left. \frac{x_k^{k+1}}{k+1} \right|_0^1 = \frac{k}{k+1}$$

$$\Rightarrow \mathbb{E}\left(\prod_{k=1}^{n} X_k\right) = \prod_{k=1}^{n} \frac{k}{k+1} = \frac{1}{2}\frac{2}{3} \dots \frac{n-1}{n}\frac{n}{n+1} = \frac{1}{n+1}.$$

9. Una urna contiene 4 bolitas blancas y 6 negras. Se extraen 3 bolitas sin reposición y se definen las siguientes variables aleatorias:

$$X = \begin{cases} 1 & \text{si el número de bolitas blancas extraídas es par} \\ 0 & \text{si el número de bolitas blancas extraídas es impar} \end{cases}$$

$$Y = \text{cantidad de bolillas negras extraídas}$$

(i) Hallar $\mathbb{E}(X)$, $\mathbb{E}(Y)$, $\mathbb{V}(X)$, $\mathbb{V}(Y)$, $Cov(X,Y)$, $\mathbb{V}(X-Y)$ y $\rho(X,Y)$.

(ii) Calcular $\mathbb{E}(Z)$ y $\mathbb{V}(Z)$ donde $Z = -4X + 1$.

Solución:

(i) $\mathbb{E}(X) = 1 \cdot \mathbb{P}\left(\left\{\begin{matrix}\text{el número de bolitas} \\ \text{blancas extraídas es par}\end{matrix}\right\}\right) + 0 \cdot \mathbb{P}\left(\left\{\begin{matrix}\text{el número de bolitas} \\ \text{blancas extraídas es impar}\end{matrix}\right\}\right)$, entonces

$$\mathbb{E}(X) = \mathbb{P}\left(\left\{\begin{matrix}\text{extraer 0 blancas} \\ \text{y 3 negras}\end{matrix}\right\}\right) + \mathbb{P}\left(\left\{\begin{matrix}\text{extraer 2 blancas} \\ \text{y 1 negras}\end{matrix}\right\}\right) = \binom{3}{0}\frac{6}{10}\frac{5}{9}\frac{4}{8} + \binom{3}{2}\frac{4}{10}\frac{3}{9}\frac{6}{8} = \frac{7}{15}.$$

$$\mathbb{E}(Y) = \sum_{i=0}^{3} i\mathbb{P}\left(\left\{\begin{matrix}\text{el número de bolillas} \\ \text{negras extraídas} = \text{i}\end{matrix}\right\}\right) = 1\binom{3}{1}\frac{6}{10}\frac{4}{9}\frac{3}{8} + 2\binom{3}{2}\frac{6}{10}\frac{5}{9}\frac{4}{8} + 3\binom{3}{3}\frac{6}{10}\frac{4}{9}\frac{3}{8} = \frac{9}{5}.$$

Por otra parte, $\mathbb{E}(X^2) = 1^2 \frac{7}{15} \Rightarrow \mathbb{V}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \frac{56}{225}$ y

$$\mathbb{E}(Y^2) = 1^2\binom{3}{1}\frac{6}{10}\frac{4}{9}\frac{3}{8} + 2^2\binom{3}{2}\frac{6}{10}\frac{5}{9}\frac{4}{8} + 3^2\binom{3}{3}\frac{6}{10}\frac{4}{9}\frac{3}{8} = \frac{19}{5} \Rightarrow \mathbb{V}(Y) = \mathbb{E}(Y^2) - [\mathbb{E}(Y)]^2 = \frac{14}{25}$$

$$\begin{aligned} \mathbb{E}(X \cdot Y) = {} & 0 \cdot 0\mathbb{P}(X=0 \text{ y } Y=0) + 0 \cdot 1\mathbb{P}(X=0 \text{ y } Y=1) + 0 \cdot 2\mathbb{P}(X=0 \text{ y } Y=2) + 0 \cdot 3\mathbb{P}(X=0 \text{ y } Y=3) \\ & + 1 \cdot 0\mathbb{P}(X=1 \text{ y } Y=0) + 1 \cdot 1\mathbb{P}(X=1 \text{ y } Y=1) + 1 \cdot 2\mathbb{P}(X=1 \text{ y } Y=2) + 1 \cdot 3\mathbb{P}(X=1 \text{ y } Y=3) \end{aligned}$$

$$\Rightarrow \mathbb{E}(X \cdot Y) = 1 \cdot \mathbb{P}\left(\left\{\begin{matrix}\text{extraer 0 blancas} \\ \text{y 3 negra}\end{matrix}\right\}\right) + 3 \cdot \mathbb{P}\left(\left\{\begin{matrix}\text{extraer 2 blancas} \\ \text{y 1 negra}\end{matrix}\right\}\right) = 3\binom{3}{0}\frac{6}{10}\frac{5}{9}\frac{4}{8} + 1\binom{3}{2}\frac{4}{10}\frac{3}{9}\frac{6}{8}$$

luego, $Cov(X,Y) = \mathbb{E}(X \cdot Y) - \mathbb{E}(X)\mathbb{E}(Y) = \frac{4}{5} - \frac{7}{15}\frac{9}{5} = -\frac{1}{25}$.

$$\mathbb{V}(X-Y) = Cov(X-Y, X-Y) = \mathbb{V}(X) - 2Cov(X,Y) + \mathbb{V}(Y) = \frac{56}{225} + \frac{2}{25} + \frac{14}{25} = \frac{8}{9}.$$

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{\mathbb{V}(X)\mathbb{V}(Y)}} = -\frac{3}{28}.$$

10. Sea (X,Y) un vector aleatorio con función de densidad conjunta

$$f_{XY}(x,y) = \begin{cases} \frac{6}{7}(3y+x) & 0 \leq x \leq y \leq 1 \\ 0 & \text{en caso contrario} \end{cases}$$

Hallar $\mathbb{E}(X)$, $\mathbb{E}(Y)$, $\mathbb{V}(X)$, $\mathbb{V}(Y)$, $Cov(X,Y)$, $\mathbb{V}(X-Y)$ y $\rho(X,Y)$

Solución:

$\mathbb{E}(X)=\int_{\mathbb{R}} x f_{XY}(x,y)dxdy=\int_0^1 x\frac{6}{7}\left[\int_x^1(3y+x)dy\right]dx=\int_0^1 x\frac{6}{7}\left[\left(\frac{3y^2}{2}+xy\right)\Big|_x^1\right]dx$, luego

$$\mathbb{E}(X)=\frac{6}{7}\int_0^1\left(\frac{3x}{2}+x^2-\frac{5x^3}{2}\right)dx=\frac{6}{7}\left(\frac{3x^2}{4}+\frac{x^3}{3}-\frac{5x^4}{8}\right)\Big|_0^1=\frac{11}{28}.$$

$\mathbb{E}(Y)=\int_{\mathbb{R}} y f_{XY}(x,y)dxdy=\int_0^1 \frac{6}{7}\left[\int_x^1(3y^2+xy)dy\right]dx=\frac{6}{7}\int_0^1\left[\left(y^3+\frac{xy^2}{2}\right)\Big|_x^1\right]dx$, luego

$$\mathbb{E}(Y)=\frac{6}{7}\int_0^1\left(1+\frac{x}{2}-\frac{3x^3}{2}\right)dx=\frac{6}{7}\left(x+\frac{x^2}{4}-\frac{3x^4}{8}\right)\Big|_0^1=\frac{3}{4}.$$

$\mathbb{E}(X^2)=\int_{\mathbb{R}} x^2 f_{XY}(x,y)dxdy=\int_0^1 x^2\frac{6}{7}\left[\int_x^1(3y+x)dy\right]dx=\int_0^1 x^2\frac{6}{7}\left[\left(\frac{3y^2}{2}+xy\right)\Big|_x^1\right]dx$, luego

$$\mathbb{E}(X^2)=\frac{6}{7}\int_0^1\left(\frac{3x^2}{2}+x^3-\frac{5x^4}{2}\right)dx=\frac{6}{7}\left(\frac{x^3}{2}+\frac{x^4}{4}-\frac{x^5}{2}\right)\Big|_0^1=\frac{3}{14}.$$

$\mathbb{E}(Y^2)=\int_{\mathbb{R}} y^2 f_{XY}(x,y)dxdy=\int_0^1 \frac{6}{7}\left[\int_x^1(3y^3+xy^2)dy\right]dx=\int_0^1 \frac{6}{7}\left[\left(\frac{3y^4}{4}+x\frac{y^3}{3}\right)\Big|_x^1\right]dx$, luego

$$\mathbb{E}(Y^2)=\frac{6}{7}\int_0^1\left(\frac{3}{4}+\frac{x}{3}-\frac{13x^4}{12}\right)dx=\frac{6}{7}\left(\frac{3x}{4}+\frac{x^2}{6}-\frac{13x^5}{60}\right)\Big|_0^1=\frac{3}{5}.$$

$\mathbb{V}(X)=\frac{3}{14}-\left(\frac{11}{28}\right)^2$ y $\mathbb{V}(Y)=\frac{3}{5}-\left(\frac{3}{4}\right)^2$.

$$\mathbb{E}(X\cdot Y)=\int_0^1\int_x^1 xy\frac{6}{7}(3y+x)dydx=\int_0^1\frac{6}{7}x\left(y^3+\frac{xy^2}{2}\right)\Big|_x^1 dx=\frac{6}{7}\int_0^1\left(x+\frac{x^2}{2}-\frac{3x^4}{2}\right)dx=\frac{6}{7}\left(\frac{x^2}{2}+\frac{x^3}{6}-\frac{3x^5}{10}\right)\Big|_0^1=\frac{11}{35}.$$

Luego, $Cov(X,Y)=\mathbb{E}(X\cdot Y)-\mathbb{E}(X)\mathbb{E}(Y)=\frac{11}{35}-\frac{11}{28}\frac{3}{4}$.

Por otra parte, $\mathbb{V}(X-Y)=\mathbb{V}(X)-2Cov(X,Y)+\mathbb{V}(Y)$ y $\rho(X,Y)=\frac{Cov(X,Y)}{\sqrt{\mathbb{V}(X)\mathbb{V}(Y)}}$ (cuentas).

11. Sea $(X_1,\ldots,X_k)$ un vector aleatorio con distribución $\mathcal{M}(p_1,\ldots,p_k,n)$ para $n>2$.
    (i) $\mathbb{E}(X_i)$, $\mathbb{V}(X_i)$, $Cov(X_i,X_j)$. Interpretar.
    (ii) Hallar el mejor predictor lineal de X_1 basado en X_2+X_3 y el error de predicción.

Solución:

(i) $\mathbb{E}(X_i)=\sum_{\substack{x_1,\ldots,x_k|\\ x_1+\ldots+x_k=n}}\frac{x_i n!}{x_1!\ldots x_i!\ldots x_k!}p_1^{x_1}\ldots p_i^{x_i}\ldots p_k^{x_k}=p_i\sum_{\substack{x_1,\ldots,x_k|\\ x_1+\ldots+x_k=n}}\frac{n(n-1)!}{x_1!\ldots(x_i-1)!\ldots x_k!}p_1^{x_1}\ldots p_i^{x_i-1}\ldots p_k^{x_k}$,

luego,

$$\mathbb{E}(X_i)=p_i n\sum_{\substack{x_1,\ldots,x_k\\ x_1+\ldots+x_k=n-1}}\frac{n(n-1)!}{x_1!\ldots(x_i-1)!\ldots x_k!}p_1^{x_1}\ldots p_i^{x_i-1}\ldots p_k^{x_k}$$

y como $\sum_{\substack{x_1,\ldots,x_k\\ x_1+\ldots+x_k=n-1}}\frac{n(n-1)!}{x_1!\ldots(x_i-1)!\ldots x_k!}p_1^{x_1}\ldots p_i^{x_i-1}\ldots p_k^{x_k}=1$, se tiene que $\mathbb{E}(X_i)=np_i$.

$$\mathbb{E}(X_i^2)=\sum_{\substack{x_1,\ldots,x_k|\\ x_1+\ldots+x_k=n}}\frac{x_i^2 n!}{x_1!\ldots x_i!\ldots x_k!}p_1^{x_1}\ldots p_i^{x_i}\ldots p_k^{x_k}=\sum_{\substack{x_1,\ldots,x_k|\\ x_1+\ldots+x_k=n}}\frac{[x_i(x_1-1)+x_i]n!}{x_1!\ldots x_i!\ldots x_k!}p_1^{x_1}\ldots p_i^{x_i}\ldots p_k^{x_k}\Rightarrow$$

$$\mathbb{E}(X_i^2)=\sum_{\substack{x_1,\ldots,x_k|\\ x_1+\ldots+x_k=n}}\frac{x_i(x_i-1)n!}{x_1!\ldots x_i!\ldots x_k!}p_1^{x_1}\ldots p_i^{x_i}\ldots p_k^{x_k}+\mathbb{E}(X_i)=p_i^2\sum_{\substack{x_1,\ldots,x_k|\\ x_1+\ldots+x_k=n}}\frac{n(n-1)(n-2)!}{x_1!\ldots(x_i-2)!\ldots x_k!}p_1^{x_1}\ldots p_i^{x_i-2}\ldots p_k^{x_k}+\mathbb{E}(X_i)$$

$$\Rightarrow\mathbb{E}(X_i^2)=n(n-1)p_i^2+np_i\Rightarrow\mathbb{V}(X_i)=n(n-1)p_i^2+np_i-p_i^2n^2=np_i(1-p_i).$$

Por último, $Cov(X_i,X_j)=n(n-1)p_ip_j-n^2p_ip_j=-np_ip_j$.

(ii) El mejor predictor lineal es $h(X_1) = a(X_2 + X_3) + b$ donde $a = \dfrac{Cov(X_2 + X_3, X_1)}{\mathbb{V}(X_2 + X_3)}$ y $b = \mathbb{E}_(X_1) - a\mathbb{E}(X_2 + X_3)$ con un error de pedicción $ECM(h) = \mathbb{V}(X_1)[1 - \rho^2(X_1, X_2 + X_3)]$.

$$a = \frac{Cov(X_2 + X_3, X_1)}{\mathbb{V}(X_2 + X_3)} = \frac{Cov(X_2, X_1) + Cov(X_3, X_1)}{\mathbb{V}(X_2) + 2Cov(X_2, X_3)X_3) + \mathbb{V}(X_3)} = \frac{-np_1p_2 - np_1p_3}{np_2(1 - p_2) + 2np_2p_3 + np_3(1 - p_3)}$$

$$a = \frac{-p_1(p_2 + p_3)}{-(p_2 - p_3)^2 + p_2 + p_3} = \frac{p_1}{p_2 - p_3 - 1}.$$

$$b = np_1 + \frac{p_1}{1 + p_3 - p_2}(np_2 + np_3) = \frac{np_1(1 + p_3)}{1 + p_3 - p_2}.$$

$$\rho^2(X_1, X_2 + X_3) = \frac{Cov^2(X_1, X_2 + X_3)}{\mathbb{V}(X_1\mathbb{V}(X_2 + X_3)} = \frac{n^2p_1^2(p_2 + p_3)^2}{np_1(1 - p_1)n[p_2(1 - p_2) + 2p_2p_3 + p_3(1 - p_3)]}.$$

12. Dada una urna con N bolillas de las cuales D son blancas y $N - D$ son negras, se extraen n sin reposición. Sean

$$X = \text{número de bolillas blancas extraídas}$$

$$X_i = \begin{cases} 1 & \text{si la i-ésima bolilla extraída es blanca} \\ 0 & \text{si la i-ésima bolilla extraída es negra} \end{cases}$$

Observar que $X \sim \mathcal{H}(N, D, n)$.

(i) Mostrar que $\mathbb{P}(X_i = 1) = \dfrac{D}{N}$ para todo $i = 1, \ldots, n$ y que para $i \neq j$ se tiene

$$\mathbb{P}(X_i = 1, X_j = 1) = \frac{D(D-1)}{N(N-1)}.$$

Determinar la distribución conjunta del vector (X_i, X_j).

(ii) Calcular $\mathbb{E}(X_i)$, $\mathbb{V}(X_i)$.

(iii) Calcular $Cov(X_i, X_j)$ para $i \neq j$.

(iv) Hallar $\mathbb{E}(X)$. Verificar que

$$\mathbb{V}(X) = \frac{(N - n)nD(N - D)}{(N - 1)N^2}$$

Solución:

(i) $\mathbb{P}(X_1 = 1) = \mathbb{P}(\{\text{la primer bolilla extraída es blanca}\}) = \dfrac{D}{N}$.

$$\begin{aligned}
\mathbb{P}(X_2) &= \mathbb{P}(\{\text{la segunda bolilla extraída es blanca}\}) = \\
&= \mathbb{P}\left(\left\{\begin{matrix}\text{la segunda bolilla}\\ \text{extraída es blanca}\end{matrix}\right\}\middle|\left\{\begin{matrix}\text{la primer bolilla}\\ \text{extraída es blanca}\end{matrix}\right\}\right)\mathbb{P}\left(\left\{\begin{matrix}\text{la primer bolilla}\\ \text{extraída es blanca}\end{matrix}\right\}\right) + \\
&+ \mathbb{P}\left(\left\{\begin{matrix}\text{la segunda bolilla}\\ \text{extraída es blanca}\end{matrix}\right\}\middle|\left\{\begin{matrix}\text{la primer bolilla}\\ \text{extraída no es blanca}\end{matrix}\right\}\right)\mathbb{P}\left(\left\{\begin{matrix}\text{la primer bolilla}\\ \text{extraída no es blanca}\end{matrix}\right\}\right)
\end{aligned}$$

$$\Rightarrow \mathbb{P}(X_2) = \frac{D-1}{N-1}\frac{D}{N} + \frac{D}{N-1}\frac{N-D}{N} = \frac{D^2 - D + ND - D^2}{N(N-1)} = \frac{D}{N.}$$

$$\begin{aligned}
\mathbb{P}(X_i = 1) &= \sum_{j=0}^{i-1} \mathbb{P}\left(\left\{\begin{matrix}\text{la i-ésima bolilla}\\ \text{extraída es blanca}\end{matrix}\right\}\middle|\left\{\begin{matrix}\text{se extrajeron j blancas y}\\ \text{(i-1)-j negras}\end{matrix}\right\}\right)\mathbb{P}\left(\left\{\begin{matrix}\text{se extrajeron j blancas y}\\ \text{(i-1)-j negras}\end{matrix}\right\}\right) \\
&= \sum_{j=0}^{i-1} \frac{D-j}{N-j-(i-1-j)} \frac{\binom{D}{j}\binom{N-D}{i-1-j}}{\binom{N}{i-1}} \\
&= \sum_{j=0}^{i-1} \frac{D-j}{N-i+1}\frac{D!}{j!(D-j)!}\frac{(N-D)!}{(i-1-j)!(N-D-i+1+j)!}\frac{(i-1)!(N-i+1)!}{N!} \\
&= \frac{1}{\binom{N}{D}}\sum_{j=0}^{i-1}\binom{i-1}{j}\binom{N-i}{D-j-1} = \frac{1}{\binom{N}{D}}\binom{i-1+N-i}{D-1} = \frac{D}{N}
\end{aligned}$$

(ii) Como $X_i \sim Be\left(\frac{D}{N}\right)$, entonces $\mathbb{E}(X_i) = \frac{D}{N}$ y $\mathbb{V}(X_i) = \frac{D}{N}\left(1 - \frac{D}{N}\right)$.

(iii) $\mathbb{E}(X \cdot Y) = \sum_{(x_i,x_j) \in R_{X_i X_j}} (x_i, x_j)\mathbb{P}(X_i = x_i, X_j = x_j) = \mathbb{P}(X_i = 1, X_j = 1) = \frac{D(D-1)}{N(N-1)}$, luego

$$Cov(X_i, X_j) = \begin{cases} \frac{D}{N}(1 - \frac{D}{N}) & i = j \\ \frac{D(D-1)}{N(N-1)} - \left(\frac{D}{N}\right)^2 & i \neq j \end{cases}$$

(iv) Como $X \sim \mathcal{H}(N, D, n)$, entonces $\mathbb{E}(X) = n\frac{D}{N}$.

$$\mathbb{V}(X) = Cov(X, X) = Cov\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{n} X_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} Cov(X_i, X_j) = \sum_{i=1}^{n} Cov(X_i, X_i) + \sum_{i=1}^{n}\sum_{\substack{j=1 \\ j\neq i}}^{n} Cov(X_i, X_j)$$

$$\Rightarrow \mathbb{V}(X) = n\frac{D}{N}(1 - \frac{D}{N}) + (n^2 - n)\left[\frac{D(D-1)}{N(N-1)} - \left(\frac{D}{N}\right)^2\right] = \frac{nD}{N} - \frac{nD^2}{N^2} + \frac{D(n^2-n)(D-N)}{N^2(N-1)} = \frac{(N-n)nD(N-D)}{N^2(N-1)}.$$

13. **El problema del coleccionista de cupones**
Un hombre colecciona cupones de un álbum compuesto por n cupones distintos. El hombre adquiere sus cupones comprando uno por día en el kiosko de la esquina de su casa y, cada vez que adquiere uno, éste tiene igual probabilidad de ser cualquiera de los n que componen el álbum.

(i) Hallar la esperanza del número de cupones diferentes que hay en un conjunto de k cupones.

(ii) Hallar el número esperado de cupones que es necesario juntar para completar el álbum.

Solución:

Se define c como la clase de copón, entonces $c = 1, \ldots, n$ y para todo c, se tiene que

$$\mathbb{P}(\{\text{cupón comprado sea de la clase c}\}) = \frac{1}{n}.$$

(i) Sea X el número de cupones diferentes que hay entre k cupones comprados.

$$\underbrace{n_1}_{clase\ 1} \ldots \underbrace{n_i}_{clase\ i} \underbrace{n_n}_{clase\ n} \text{ tal que } \sum_{i=1}^{n} n_i = k.$$

Si algún $n_i = 0$, entonces no hay cupones de la clase i. Por lo tanto, $X = \sharp\{i \in \{1, \ldots, n\} | n_i \neq 0\}$.

Sea $X_i = \begin{cases} 1 & n_i \neq 0 \\ 0 & n_i = 0 \end{cases}$, entonces $\mathbb{P}(X_i = 1) = 1 - \left(\frac{n-1}{n}\right)^k$ y como $X = \sum_{i=1}^{n} X_i$, se tiene que

$$\mathbb{E}(X) = \sum_{i=1}^{n} \mathbb{E}(X_i) \sum_{i=1}^{n} \left[1 - \left(\frac{n-1}{n}\right)^k\right] = n\left[1 - \left(\frac{n-1}{n}\right)^k\right].$$

(ii) Sean Y_i el número de cupones adicionales que se necesitan para tener un cupón distinto a los de la clase i $(0 \leq i \leq n-1)$ e $Y = \sum_{i=0}^{n-1} Y_i$. Por lo tanto, Y resulta ser el número de cupones recolectados justo para completar el álbum.

$$\mathbb{P}(Y_i = k) \underbrace{=}_{k \geq 1} \frac{n-i}{n}\left(\frac{i}{n}^{k-1}\right) \Rightarrow Y_i \sim \mathcal{G}\left(\frac{n-i}{n}\right) \Rightarrow \mathbb{E}(Y_i) = \frac{n}{n-i} \Rightarrow \mathbb{E}(Y) = \sum_{i=0}^{n-1} \mathbb{E}(Y_i) = \sum_{i=0}^{n-1} \frac{n}{n-i}.$$

14. (i) Mostrar con un ejemplo que $Cov(X, Y) = 0$ no implica que X e Y sean independientes.
Sugerencia: Puede considerar un vector aleatorio con densidad uniforme en el pentágono de vértices $(-1, 1)$, $(1, 1)$, $(-1, 0)$, $(0, -1)$ y $(1, 0)$.

(ii) Sean X e Y dos variables aleatorias que toman sólo dos valores cada una. Probar que si $Cov(X, Y) = 0$ entonces X e Y son independientes.
Sugerencia: Considerar primero el caso en que los dos valores que toma cada variable son 0 y 1.

Solución:

(i) Sea el vector $\boldsymbol{X} = (X, Y)$ y R el pentágono de vértices $(-1, 1); (1, 1); (-1, 0); (1, 0); (0, -1)$, tal que $\boldsymbol{X} \sim \mathcal{U}(R)$.

Puesto que el área de R es igual a 3, entonces $f_{(X,Y)}(x, y) = \begin{cases} \dfrac{1}{3} & (x, y) \in R \\ 0 & (x, y) \notin R \end{cases}$.

$$\begin{aligned}\mathbb{E}(X \cdot Y) &= \int_R xyf_{(X,Y)}(x, y)dxdy = \frac{1}{3}\int_{-1}^{1}\int_{0}^{1} xydxdy + \frac{1}{3}\int_{-1}^{0}\int_{-y-1}^{y+1} xydxdy \\ &= \frac{1}{3}\int_{-1}^{1} x\left(\frac{y^2}{2}\bigg|_0^1\right)dx + \frac{1}{3}\int_{-1}^{0} y\left(\frac{x^2}{2}\bigg|_{-y-1}^{y+1}\right)dy = \frac{1}{6}\left(\frac{x^2}{2}\right)\bigg|_{-1}^{1} = 0\end{aligned}$$

$$\mathbb{E}(X) = \frac{1}{3}\int_{-1}^{1}\int_{0}^{1} xdydx + \frac{1}{3}\int_{-1}^{0}\int_{-y-1}^{y+1} xdxdy = \frac{1}{3}\int_{-1}^{1} xdx + \frac{1}{3}\int_{-1}^{0}\left(\frac{x^2}{2}\bigg|_{-y-1}^{y+1}\right)dy = \frac{1}{3}\left(\frac{x^2}{2}\right)\bigg|_{-1}^{1} = 0$$

Luego, $Cov(X, Y) = 0$ pero X e Y no son independientes pues $sop(f_{(X,Y)})$ no es un rectángulo.

(ii) • $\underline{R_X = R_Y = \{0, 1\}:}$

$$\begin{aligned}\mathbb{P}(X = 0)\mathbb{P}(Y = 0) &= (1 - \mathbb{P}(X = 1))(1 - \mathbb{P}(Y = 1)) = 1 + \mathbb{P}(X = 1 \text{ y } Y = 1) - \mathbb{P}(X = 1) - \mathbb{P}(Y = 1) \\ &= 1 + \mathbb{P}(X = 1 \text{ y } Y = 1) - 1 + \mathbb{P}(X = 0) - 1 + \mathbb{P}(Y = 0) \\ &= \mathbb{P}(X = 1 \text{ y } Y = 1) + \mathbb{P}(X = 0 \text{ y } Y = 0) + \mathbb{P}(X = 0 \text{ y } Y = 1) \\ &- 1 + \mathbb{P}(X = 0 \text{ y } Y = 0) + \mathbb{P}(X = 1 \text{ y } Y = 0) = \mathbb{P}(X = 0 \text{ y } Y = 0)\end{aligned}$$

puesto que $\mathbb{P}(X = 1 \text{ y } Y = 1) + \mathbb{P}(X = 0 \text{ y } Y = 0) + \mathbb{P}(X = 0 \text{ y } Y = 1) + \mathbb{P}(X = 1 \text{ y } Y = 0) = 1$.

$$\begin{aligned}\mathbb{P}(X = 0)\mathbb{P}(Y = 1) &= (1 - \mathbb{P}(X = 1))\mathbb{P}(Y = 1) = \mathbb{P}(Y = 1) - \mathbb{P}(X = 1 \text{ y } Y = 1) \\ &= \mathbb{P}(X = 0 \text{ y } Y = 1) + \mathbb{P}(X = 1 \text{ y } Y = 0) - \mathbb{P}(X = 1 \text{ y } Y = 0) = \mathbb{P}(X = 0 \text{ y } Y = 1)\end{aligned}$$

$$\begin{aligned}\mathbb{P}(X = 1)\mathbb{P}(Y = 0) &= \mathbb{P}(X = 1)(1 - \mathbb{P}(Y = 1)) = \mathbb{P}(X = 1) - \mathbb{P}(X = 1 \text{ y } Y = 1) \\ &= \mathbb{P}(X = 1 \text{ y } Y = 0) + \mathbb{P}(X = 1 \text{ y } Y = 0) - \mathbb{P}(X = 1 \text{ y } Y = 0) = \mathbb{P}(X = 1 \text{ y } Y = 0)\end{aligned}$$

Por último, para ver que $\mathbb{P}(X = 1 \text{ y } Y = 1) = \mathbb{P}(X = 1)\mathbb{P}(Y = 1)$ usamos que $Cov(X, Y) = 0$.

$$\mathbb{E}(X \cdot Y) = \sum_{x \in R_X}\sum_{y \in R_Y} xy\mathbb{P}(X = x \text{ y } Y = y) = \mathbb{P}(X = 1 \text{ y } Y = 1)$$

$$\mathbb{E}(X) = \sum_{x \in R_X} x\mathbb{P}(X = x \text{ y } Y = y) = \mathbb{P}(X = 1 \text{ y } Y = 0) + \mathbb{P}(X = 1 \text{ y } Y = 1) = \mathbb{P}(X = 1)$$

$$\mathbb{E}(Y) = \sum_{y \in R_Y} y\mathbb{P}(X = x \text{ y } Y = y) = \mathbb{P}(X = 0 \text{ y } Y = 1) + \mathbb{P}(X = 1 \text{ y } Y = 1) = \mathbb{P}(Y = 1)$$

Luego, como $0 = Cov(X, Y) = \mathbb{E}(X \cdot Y) - \mathbb{E}(X)\mathbb{E}(Y)$ entonces $\mathbb{P}(X = 1 \text{ y } Y = 1) = \mathbb{P}(X = 1)\mathbb{P}(Y = 1)$.

• $\underline{R_X = \{x_1, x_2\} \text{ y } R_Y = \{y_1, y_2\}:}$
Como X e Y son dos variables aleatorias que toman sólo dos valores cada una suponemos que $x_1 \neq x_2$ e $y_1 \neq y_2$.
Luego, definiendo las variables $\widetilde{X} = \dfrac{X - x_1}{x_2 - x_1}$ e $\widetilde{Y} = \dfrac{Y - y_1}{y_2 - y_1}$ se tiene que $R_{\widetilde{X}} = R_{\widetilde{Y}} = \{0, 1\}$.
Como $Cov(X, Y) = 0$ por hipótesis, $0 = Cov((x_2 - x_1)\widetilde{X} + x_1, (y_2 - y_1)\widetilde{Y} + y_1)$, entonces $0 = (x_2 - x_1)(y_2 - y_1)Cov(\widetilde{X}, \widetilde{Y}) \Rightarrow Cov(\widetilde{X}, \widetilde{Y}) = 0$, entonces $\widetilde{X}$ e $\widetilde{Y}$ son independientes y, por lo tanto, X e Y también lo son.

15. Probar que si X es una variable aleatoria con distribución simétrica respecto de $\theta \in \mathbb{R}$ y $\mathbb{E}(|X|) < \infty$, entonces $\mathbb{E}(X) = \theta$.

Solución:

Sea $Y = X - \theta$ de manera tal que Y es una variable aleatoria con distribución simétrica respecto de cero. Es decir, $f_Y(y) = f_Y(-y)$. Luego,

$$\mathbb{E}(Y) = \int_{\mathbb{R}} yf_Y(y)dy = \int_{-\infty}^{0} yf_Y(y)dy + \int_{0}^{+\infty} yf_Y(y)dy = \int_{0}^{+\infty}(-y)\underbrace{f_Y(-y)}_{=f_Y(y)}(-dy) + \int_{0}^{+\infty} yf_Y(y)dy = 0$$

Por lo tanto, $0 = \mathbb{E}(Y) = \mathbb{E}(X - \theta) = \mathcal{E}(X) - \theta \Rightarrow \mathbb{E}(X) = \theta$.

16. **Esquema de Polya**
De un bolillero que contiene B bolillas blancas y R rojas se extrae una bolilla al azar y se la devuelve al bolillero junto con otras c bolillas del mismo color. Se repite este procedimiento sucesivamente comenzando en cada nuevo paso con la composición del bolillero resultante del paso anterior. Sean

$$X_i = \begin{cases} 1 & \text{si la i-ésima bolilla extraída es roja} \\ 0 & \text{si la i-ésima bolilla extraída es blanca} \end{cases}$$

(i) ¿Qué distribución tienen las variables aleatorias X_i?

(ii) $\mathbb{E}(X_i)$, $\mathbb{V}(X_i)$, $Cov(X_i, X_j)$ y $\rho(x_i, X_j)$ para $i < j$. ¿Son independientes las variables X_i? Contestar a partir las cantidades calculadas.

(iii) Sea $S_j = \displaystyle\sum_{i=1}^{j} X_i$, el número de bolillas rojas extraídas luego de j extracciones. Hallar $\mathbb{E}(S_j)$, $\mathbb{V}(S_j)$, $Cov(S_i, S_j)$, $\rho(S_i, S_j)$ para $i < j$.

Solución:

(i) $X_i \sim Be(p)$ con $p = \mathbb{P}(\{\text{la i-ésima bolilla extraída es roja}\}) = \dfrac{R}{R+B}$.

(ii) $\mathbb{E}(X_i) = \dfrac{R}{R+B}$ y $\mathbb{V}(X_i) = \dfrac{R}{R+B}\left(1 - \dfrac{R}{R+B}\right) = \dfrac{RB}{(R+B)^2}$.

$$Cov(X_i, X_j) = \mathbb{E}(X_i \cdot X_j) - \mathbb{E}(X_i)\mathbb{E}(X_j) = \left(\frac{R}{R+B}\right)^2 - \frac{R}{R+B}\frac{R}{R+B} = 0.$$

$$\rho(X_i, X_j) = \frac{Cov(X_i, X_j)}{\sqrt{\mathbb{V}(X_i)\mathbb{V}(X_j)}} \underbrace{=}_{Cov(X_i,X_j)=0} 0$$

Como X_i y X_j sólo toman dos valores y $Cov(X_i, X_j) = 0$ se tiene que X_i y X_j son independientes.

(iii) Como $S_j = \displaystyle\sum_{i=1}^{j} X_i$ con $X_i \sim Be\left(\dfrac{R}{R+B}\right)$ se tiene que $S_j \sim Bi\left(j, \dfrac{R}{R+B}\right)$. Entonces, $\mathbb{E}(S_j) = j\dfrac{R}{R+B}$ y $\mathbb{V}(S_j) = j\dfrac{RB}{(R+B)^2}$.

$$\begin{aligned}
\mathbb{E}(S_i, S_j) &= \sum_{k=0}^{i}\sum_{t=0}^{j} kt \binom{i}{k}\binom{j}{t}\left(\frac{R}{R+B}\right)^k\left(1 - \frac{R}{R+B}\right)^{i+j-k-t} \\
&= \left[\sum_{k=0}^{i}\binom{i}{k}\left(\frac{R}{R+B}\right)^k\left(1 - \frac{R}{R+B}\right)^{i-k}\right]\left[\sum_{t=0}^{j}\binom{j}{t}\left(\frac{R}{R+B}\right)^t\left(1 - \frac{R}{R+B}\right)^{j-t}\right] \\
&= \mathbb{E}(S_i)\mathbb{E}(S_j)
\end{aligned}$$

$$\Rightarrow Cov(S_i, S_j) = 0 \Rightarrow \rho(S_i, S_j) = 0.$$

17. Sea $\boldsymbol{X} = (X_1, X_2)$ un vector aleatorio de marginales X_1 y X_2 independientes y con distribución $N(0, 1)$. Dada $A \in \mathbb{R}^{2\times 2}$ inversible y $\boldsymbol{c} \in \mathbb{R}^2$, sea $\boldsymbol{Y} = (Y_1, Y_2)$ el vector aleatorio definido por $\boldsymbol{Y} = A \cdot \boldsymbol{X} + \boldsymbol{c}$.

(i) Calcular $\mathbb{E}(Y_i)$ y $\mathbb{V}(Y_i)$ para todo i. Observar que, por el ejercicio 18 de la Práctica 5, cada Y_i tiene distribución normal.

(ii) Hallar $Cov(Y_1, Y_2)$ y $\rho(Y_1, Y_2)$ a partir del cálculo de $Cov(X_1, X_2)$. Observar que $|\rho(Y_1, Y_2)| < 1$.

(iii) Verificar que si para $\boldsymbol{Y}$ definimos su vector de medias $\mu \in \mathbb{R}^2$ por $\mu_i = \mathbb{E}(Y_i)$ y su matriz de covarianzas $\Sigma \in \mathbb{R}^{2\times 2}$ como $\Sigma_{ij} = Cov(Y_i, Y_j)$ entonces $\mu = \boldsymbol{c}$ y $\Sigma = A \cdot A^t$.

(iv) Concluir a partir del item anterior que Σ es inversible y que tanto ella como su inversa son matrices simétricas y definidas positivas.

(v) Mostrar que la función de densidad conjunta de $\boldsymbol{Y}$ viene dada por

$$f_{Y_1Y_2}(y_1, y_2) = \frac{1}{2\pi\sigma_{Y_1}\sigma_{Y_1}} e^{\left\{-\frac{1}{2(1-\rho)^2}\left[\left(\frac{y_1-\mu_{Y_1}}{\sigma_{Y_1}}\right)^2 + \left(\frac{y_2-\mu_{Y_2}}{\sigma_{Y_2}}\right)^2 - 2\rho\left(\frac{y_1-\mu_{Y_1}}{\sigma_{Y_1}}\right)\left(\frac{y_2-\mu_{Y_2}}{\sigma_{Y_2}}\right)\right]\right\}} \quad (1)$$

donde notamos $\mu_{Y_i} = \mathbb{E}(X_i)$, $\sigma^2_{Y_i} = \mathbb{V}(Y_i)$ y $\rho = \rho(Y_1, Y_2)$. Verificar que la densidad también puede escribirse en notación matricial como

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = \frac{1}{2\pi[\det(\Sigma)]^{\frac{1}{2}}} exp\left\{-\frac{1}{2}(\boldsymbol{y}-\mu)^t\Sigma^{-1}(\boldsymbol{y}-\mu)\right\}$$

¿Qué tipo de curvas de nivel tiene $f_{\boldsymbol{Y}}$?

(vi) Deducir que Y_1 e Y_2 son independientes si y sólo si $Cov(Y_1, Y_2) = 0$.

(vii) Sea $\boldsymbol{Z}$ un vector aleatorio continuo con densidad conjunta como en (1). Probar que existen $\boldsymbol{c} \in \mathbb{R}^2$ y $A \in \mathbb{R}^{2\times 2}$ con $\det(A) \neq 0$ tales que $\boldsymbol{Z} = A \cdot \boldsymbol{X} + \boldsymbol{c}$ para cierto vector aleatorio $\boldsymbol{X}$ de marginales independientes con distribución $N(0, 1)$.

(viii) Sean $\boldsymbol{Z}$ un vector aleatorio con distribución $N_2(\mu, \Sigma)$, $\boldsymbol{d} \in \mathbb{R}^2$ y $B \in \mathbb{R}^{2\times 2}$ con $\det(B) \neq 0$. Probar que $B \cdot \boldsymbol{Z} + \boldsymbol{d} \sim N_2(B \cdot \mu + \boldsymbol{d}, B \cdot \Sigma \cdot B^t)$.

(ix) Probar que si $\boldsymbol{Z}$ es un vector con distribución normal multivariada 2-dimensional entonces $\boldsymbol{v} \cdot \boldsymbol{Z}$ tiene distribución normal para todo $\boldsymbol{v} \in \mathbb{R}^2\setminus\{0\}$.

Solución:

(i) Como $\boldsymbol{Y} = A \cdot \boldsymbol{X} + \boldsymbol{c}$, entonces $\begin{cases} Y_1 = a_{11}X_1 + a_{12}X_2 + c_1 \\ Y_2 = a_{21}X_1 + a_{22}X_2 + c_2 \end{cases}$.

Luego,

$$\begin{cases} \mathbb{E}(Y_1) = a_{11}\mathbb{E}(X_1) + a_{12}\mathbb{E}(X_2) + c_1 = c_1 \\ \mathbb{E}(Y_2) = a_{21}\mathbb{E}(X_1) + a_{22}\mathbb{E}(X_2) + c_2 = c_2 \end{cases}$$

$$\begin{cases} \mathbb{V}(Y_1) = \mathbb{V}(a_{11}X_1 + a_{12}X_2 + c_1) = a_{11}^2\mathbb{V}(X_1) + a_{12}^2\mathbb{V}(X_2) = a_{11}^2 + a_{12}^2 \\ \mathbb{V}(Y_2) = \mathbb{V}(a_{21}X_1 + a_{22}X_2 + c_2) = a_{21}^2\mathbb{V}(X_1) + a_{22}^2\mathbb{V}(X_2) = a_{21}^2 + a_{22}^2 \end{cases}$$

(ii) $Cov(Y_1, Y_2) = Cov(a_{11}X_1 + a_{12}X_2 + c_1, a_{21}X_1 + a_{22}X_2 + c_2)$, luego

$$\begin{aligned} Cov(Y_1, Y_2) = {} & a_{11}a_{21}\underbrace{Cov(X_1, X_1)}_{=1} + a_{11}a_{22}\underbrace{Cov(X_1, X_2)}_{=0} + a_{11}\underbrace{Cov(X_1, c_2)}_{=0} + a_{12}a_{21}\underbrace{Cov(X_1, X_2)}_{=0} \\ & + a_{12}a_{22}\underbrace{Cov(X_2, X_2)}_{=1} + a_{12}\underbrace{Cov(X_2, c_2)}_{=0} + a_{21}\underbrace{Cov(X_1, c_1)}_{=0} + a_{22}\underbrace{Cov(X_2, c_1)}_{=0} + \underbrace{Cov(c_1, c_2)}_{=0} \end{aligned}$$

$$\Rightarrow Cov(Y_1, Y_2) = a_{11}a_{21} + a_{12}a_{22}.$$

$$\rho(Y_1, Y_2) = \frac{a_{11}a_{21} + a_{12}a_{22}}{\sqrt{(a_{11}^2 + a_{12}^2)(a_{21}^2 + a_{22}^2)}} = \frac{(a_{11}, a_{12})(a_{21}, a_{12})}{\|(a_{11}, a_{12})\|\|(a_{21}, a_{22})\|}$$

Sea α el ángulo formado por (a_{11}, a_{12}) y (a_{21}, a_{22}), entonces $\rho(Y_1, Y_2) = \cos(\alpha)$. Además, como A es inversible, se tiene que sus filas no son múltiplos, entonces $|\cos(\alpha)| \neq 1 \Rightarrow |\rho(Y_1, Y_2)| < 1$.

(iii) $Cov(\boldsymbol{Y}, \boldsymbol{Y}) = A \cdot Cov(\boldsymbol{X}, \boldsymbol{X}) \cdot A^t = A \cdot Id_2 \cdot A^t = A \cdot A^t$ y puesto que $\mathbb{E}(Y_i) = c_i$ para $i = 1, 2$, entonces $\mu = \boldsymbol{c}$.

(iv) $\Sigma = A \cdot A^t$,entonces $\det(\Sigma) = [\det(A)]^2 \neq 0$ puesto $\det(A) \neq 0$. Luego, Σ es inversible.

$$\Sigma^t = (A \cdot A^t)^t = A \cdot A^t = \Sigma \Rightarrow \Sigma \text{ es simétrica}$$

y como $\det(\Sigma) > 0$ se tiene que Σ es definida positiva.

$$\left.\begin{array}{r} \Sigma \text{ simétrica} \Rightarrow \Sigma^{-1} \text{ simétrica} \\ \det(\Sigma) > 0 \Rightarrow \det(\Sigma^{-1}) > 0 \end{array}\right\} \Rightarrow \Sigma^{-1} \text{ es definida positiva.}$$

(v) Como $\boldsymbol{Y} = g(\boldsymbol{X}) = A \cdot \boldsymbol{X} + \boldsymbol{c}$, entonces

$$\begin{cases} Y_1 = a_{11}X_1 + a_{12}X_2 + c_1 \\ Y_2 = a_{21}X_1 + a_{22}X_2 + c_2 \end{cases} \Rightarrow \begin{cases} X_1 = \dfrac{1}{\det(A)}[(Y_1 - c_1)a_{22} - (Y_2 - c_2)a_{12}] \\ X_2 = \dfrac{1}{\det(A)}[(Y_2 - c_2)a_{11} - (Y_1 - c_1)a_{21}] \end{cases}$$

$$\mathbb{P}(\boldsymbol{Y} \in (-\infty, y_1] \times (-\infty, y_2]) = \mathbb{P}(\boldsymbol{X} \in g^{-1}((-\infty, y_1] \times (-\infty, y_2])) \underbrace{=}_{\substack{X_1, X_2 \\ ind.}} \mathbb{P}(X_1 \in g^{-1}((-\infty, y_1]))\mathbb{P}(X_2 \in g^{-1}((-\infty, y_2]))$$

$$f_{\boldsymbol{Y}}(y)=\frac{1}{2\pi}exp\left\{-\frac{1}{2(\det(A))^2}\left[(Y_1-c_1;Y_2-c_2)\begin{pmatrix}a_{21}^2+a_{22}^2 & -a_{11}a_{21}-a_{12}a_{22}\\ -a_{11}a_{21}-a_{12}a_{22} & a_{21}^2+a_{22}^2\end{pmatrix}\begin{pmatrix}Y_1-c_1\\ Y_2-c_2\end{pmatrix}\right]\right\}\mathcal{J}_{\{g(\boldsymbol{X})\}}$$

donde $\mathcal{J}_{\{g(\boldsymbol{X})\}}$ es el jacobiano de la transformación, por lo tanto

$$\mathcal{J}_{\{g(\boldsymbol{X})\}}=\frac{1}{(\det(A))^2}\left|\begin{pmatrix}a_{22} & -a_{12}\\ -a_{21} & a_{11}\end{pmatrix}\right|$$

$$f_{\boldsymbol{Y}}(y)=\frac{1}{2\pi\det(A)}exp\left\{-\frac{1}{2(\det(A))^2}\left[(Y_1-c_1;Y_2-c_2)\begin{pmatrix}a_{21}^2+a_{22}^2 & -a_{11}a_{21}-a_{12}a_{22}\\ -a_{11}a_{21}-a_{12}a_{22} & a_{21}^2+a_{22}^2\end{pmatrix}\begin{pmatrix}Y_1-c_1\\ Y_2-c_2\end{pmatrix}\right]\right\}$$

Por otra parte, $\Sigma=A\cdot A^t$ entonces

- $\det(A)=(\det(\Sigma))^{1/2}$.
- $\Sigma^{-1}=(A^t)^{-1}\cdot A^{-1}$, entonces

$$\Sigma^{-1}=\frac{1}{\det(A)}\begin{pmatrix}a_{22} & -a_{21}\\ -a_{12} & a_{11}\end{pmatrix}\cdot\frac{1}{\det(A)}\begin{pmatrix}a_{22} & -a_{12}\\ -a_{21} & a_{11}\end{pmatrix}=\frac{1}{(\det(A))^2}\begin{pmatrix}a_{22}^2+a_{21}^2 & -a_{12}a_{22}-a_{21}a_{11}\\ -a_{12}a_{22}-a_{21}a_{11} & a_{22}^2+a_{21}^2\end{pmatrix}$$

Luego,

$$f_{\boldsymbol{Y}}(y)=\frac{1}{2\pi(\det(\Sigma))^{1/2}}e^{-\frac{1}{2}(\boldsymbol{y}-\mu)^t\Sigma^{-1}(\boldsymbol{y}-\mu)}$$

Para saber qué tipos de curvas de nivel, tenemos que $(\boldsymbol{y}-\mu)^t\Sigma^{-1}(\boldsymbol{y}-\mu)=cte$. Como Σ^{-1} es definida positiva, entonces sus autovalores son positivos y, por lo tanto, sus curvas de nivel son elipses.

(vi) Es claro que si Y_1 e Y_2 son independientes entonces $Cov(Y_1,Y_2)=0$.
Supongamos entonces que $Cov(Y_1,Y_2)=0$, entonces

- $\rho\equiv 0$.
- $0=Cov(Y_1,Y_2)=a_{11}a_{21}+a_{12}a_{22}$.
- $f_{Y_1Y_2}(y_1,y_2)=\dfrac{1}{2\pi\sigma_{Y_1}\sigma_{Y_1}}e^{\left\{-\frac{1}{2}\left[\left(\frac{y_1-\mu_{Y_1}}{\sigma_{Y_1}}\right)^2+\left(\frac{y_2-\mu_{Y_2}}{\sigma_{Y_2}}\right)^2\right]\right\}}$

Por lo tanto,

$$f_{Y_1Y_2}(y_1,y_2)=\frac{1}{2\pi\sigma_{Y_1}\sigma_{Y_1}}e^{\left\{-\frac{1}{2}\left[\left(\frac{y_1-\mu_{Y_1}}{\sigma_{Y_1}}\right)^2+\left(\frac{y_2-\mu_{Y_2}}{\sigma_{Y_2}}\right)^2\right]\right\}}=\frac{1}{\sqrt{2\pi}\sigma_{Y_1}}e^{-\frac{1}{2}\left(\frac{y_1-\mu_{Y_1}}{\sigma_{Y_1}}\right)^2}\frac{1}{\sqrt{2\pi}\sigma_{Y_2}}e^{-\frac{1}{2}\left(\frac{y_2-\mu_{Y_2}}{\sigma_{Y_2}}\right)^2}$$

Es decir, $f_{Y_1Y_2}(y_1,y_2)=f_{Y_1}(y_1)f_{Y_2}(y_2)$ y, por lo tanto, Y_1 e Y_2 son independientes.

(vii) Sea $\boldsymbol{Z}=(Z_1,Z_2)$ un vector aleatorio continuo tal que $Z_1\sim N(\mu_1,\sigma_1^2)$ y $Z_2\sim N(\mu_2,\sigma_2^2)$ con $\sigma:1\neq\sigma_2$.
Sea $\boldsymbol{X}=(X_1,X_2)$ un vector aleatorio continuo tal que $X_1\sim N(0,1)$ y $X_2\sim N(0,1)$ con X_1 y X_2 independientes, entonces definiendo $\boldsymbol{c}=(\mu_1,\mu_2)$ y $A=\begin{pmatrix}\sigma_1 & 0\\ 0 & \sigma_2\end{pmatrix}$ se tiene que

$$A\cdot\boldsymbol{X}+\boldsymbol{c}=(\sigma_1X_1+\mu_1,\sigma_2X_2+\mu_2)=(Z_1,Z_2)=\boldsymbol{Z}$$

18. Sean X y W variables aleatorias independientes tales que $X\sim N(0,1)$ y $W\sim Be\left(\frac{1}{2}\right)$. Definimos

$$Y=\begin{cases}X & W=1\\ -X & W=0\end{cases}$$

(i) Probar que Y es una variable aleatoria con distribución $N(0,1)$.
(ii) ¿Son X e Y independientes?
(iii) ¿Son Y e W independientes?
(iv) Mostrar que $Cov(X,Y)=0$.
(v) Deducir que el vector aleatorio (X,Y) no tiene distribución normal multivariada a pesar de tener marginales con distribución normal. ¿Contradice esto los resultados del ejercicio anterior?¿Por qué?

Solución:

(i) $\mathbb{P}(Y \le y) = \mathbb{P}(X \le y|W = 1)\mathbb{P}(W = 0) + \mathbb{P}(-X \le y|W = 0)\mathbb{P}(W = 1)$, como X y W son variables aleatorias independientes y $\mathbb{P}(W = 0) = \mathbb{P}(W = 1) = \frac{1}{2}$ se tiene que

$$\mathbb{P}(Y \le y) = \frac{1}{2}[\mathbb{P}(X \le y) + \mathbb{P}(X \ge -y)]$$

Puesto que X es simétrica respecto de cero $\mathbb{P}(X \le y) = \mathbb{P}(X \ge -y)$ y, por lo tanto, $\mathbb{P}(Y \le y) = \mathbb{P}(X \le y)$. Entonces Y es una variable aleatoria con distribución $N(0, 1)$.

(ii) X e Y no son independientes dado que Y es función de X.

(iii) W e Y no son independientes dado que Y es función de W.

(iv) $\mathbb{E}(X \cdot Y) = \frac{1}{2}\left[\int_{\mathbb{R}} x^2 f_X(x)dx + \int_{\mathbb{R}} x(-x)f_X(x)dx\right] \equiv 0$

(v) el vector aleatorio (X, Y) no tiene distribución normal multivariada a pesar de tener marginales con distribución normal pues X e Y no son independientes.

19. Sean $U_1, \dots, U_n$ variables aleatorias independientes con distribución $\mathcal{U}([0, 1])$ y consideremos sus estadísticos de orden $U^{(1)}, \dots, U^{(n)}$.

   (i) Hallar $\mathbb{E}(U^{(i)})$. ¿Qué relación guardan las esperanzas entre sí?
   Sugerencia: Recordar la distribución de $U^{(i)}$ calculada en el Ejercicio 13 de la Práctica 5.

   (ii) Calcular $\mathbb{V}(U^{(i)})$.

   (iii) ¿Para qué valor de i se minimiza la varianza? ¿Para cuál se maximiza?

Solución:

(i) Sabemos ya que la función de densidad de $U^{(k)}$ viene dada por

$$f_{U^{(k)}}(s) = \frac{n!}{(n-k)!(k-1)!}f(s)(F(s))^{k-1}(1 - F(s))^{n-k}$$

como $U_1, \dots, U_n \sim \mathcal{U}([0, 1])$, se tiene que para $x \in [0, 1]$

$$f_{U^{(k)}}(x) = \frac{n!}{(n-k)!(k-1)!}x^{k-1}(1-x)^{n-k} = \frac{\Gamma(n+1)}{\Gamma(k)\Gamma(n-k+1)}x^{k-1}(1-x)^{n-k+1-1}$$

entonces $U^{(k)} \sim \beta(k, n-k+1)$ y, por lo tanto, $\mathbb{E}(U^{(k)}) = \dfrac{k}{n+1} = k\mathbb{E}(U^{(1)})$.

(ii) Puesto que $U^{(k)} \sim \beta(k, n-k+1)$, entonces $\mathbb{V}(U^{(k)}) = \dfrac{b(n-k+1)}{(n+1)^2(n+2)}$.

(iii) $\dfrac{\partial \mathbb{V}(U^{(k)})}{\partial k} = \dfrac{2-k+(+1)}{(n+1)^2(n+2)}$ y buscamos k tal que $\dfrac{\partial \mathbb{V}(U^{(k)})}{\partial k} = 0$, por lo tanto

$$\frac{\partial \mathbb{V}(U^{(k)})}{\partial k} = 0 \Leftrightarrow k = \frac{n+1}{2}$$
$$\frac{\partial^2 \mathbb{V}(U^{(k)})}{\partial k^2} = -\frac{2}{(n+1)^2(n+2)}$$

Luego, la varianza es máxima para $k = \begin{cases} \dfrac{n+1}{2} & \text{si } n \text{ es impar} \\ \lfloor \dfrac{n+1}{2} \rfloor \text{ y } \lfloor \dfrac{n}{2} \rfloor & \text{si } n \text{ es par} \end{cases}$

De esta manera se deduce que la varianza será mínima para $k = 1$ y $k = n$.

20. (i) Sea X una variable aleatoria discreta con $R_X \subseteq \mathbb{N}_0$. Probar que

$$\mathbb{E}(X) = \sum_{n \in \mathbb{N}_0} \mathbb{P}(X > n) = \sum_{n \in \mathbb{N}} \mathbb{P}(X \ge n).$$

   (ii) Sea X una variable aleatoria arbitraria. Probar que

$$\sum_{n \in \mathbb{N}} \mathbb{P}(|X| \ge n) \le \mathbb{E}(|X|) \le 1 + \sum_{n \in \mathbb{N}} \mathbb{P}(|X| \ge n).$$

   Conclur que

$$\mathbb{E}(|X|) < \infty \Leftrightarrow \sum_{n \in \mathbb{N}} \mathbb{P}(|X| \ge n) < \infty.$$

   Sugerencia: Considerar la variable aleatoria discreta $[|X|]$.

Práctica 8

Convergencias - Ley de los Grandes Números

1. Una máquina produce artículos de 3 clases: A, B y C en proporciones 25 %, 25 % y 50 % respectivamente. Las longitudes de los artículos A y B siguen distribuciones $\mathcal{U}([0,1])$ y $\mathcal{U}([0,2)]$ respectivamente y las longitudes de los artículos C se distribuyen según la densidad $f(x) = \left(1 - \frac{x}{2}\right)\mathbb{1}_{[0,2]}(x)$. Se eligen n artículos al azar de la producción total y se calcula el promedio de sus longitudes.

   (i) Dar una cota inferior para la probabilidad de que el promedio de las longitudes esté comprendido entre $\frac{15}{24}$ y $\frac{19}{24}$ si el tamaño de la muestra es $n = 100$.

   (ii) ¿Cuál debe ser el tamaño de la muestra para que la probabilidad de que el promedio de las longitudes esté comprendido entre $\frac{15}{24}$ y $\frac{19}{24}$ sea mayor o igual que 0,90?

Solución:
Sea L_A, L_B y L_C las variables aleatorias correspondientes a las longitudes A, B y C respectivamente. Definimos X_i al promedio de las longitudes; i.e., $X_i = 0,25L_A + 0,25L_B + 0,50L_C$.

$$\mathbb{E}(X_i) = 0,25\mathbb{E}(L_A) + 0,25\mathbb{E}(L_B) + 0,50\mathbb{E}(L_C) = \frac{1}{4}\frac{1}{2} + \frac{1}{4} + \frac{1}{2}\int_0^2 x\left(1 - \frac{x}{2}\right)dx = \frac{1}{8} + \frac{1}{4} + \frac{1}{2}\left(\frac{x^2}{2} - \frac{x^3}{6}\right)\Big|_0^2 = \frac{17}{24},$$

$$\mathbb{E}(X_i^2) = \frac{1}{4}\int_0^1 x^2dx + \frac{1}{4}\int_0^2 x^2dx + \frac{1}{2}\int_0^2 x^2\left(1 - \frac{x}{2}\right)dx = \frac{1}{12} + \frac{2}{3} + \frac{1}{3} = \frac{13}{12},$$

$$\Rightarrow \mathbb{V}(X_i) = \frac{13}{12} - \left(\frac{17}{24}\right)^2.$$

(i) Sea $X = \frac{1}{100}\sum_{i=1}^{100} X_i$ tal que $\mathbb{E}(X) = \mathbb{E}(X_i) = \frac{17}{24}$ y $\mathbb{V}(X) = \frac{\mathbb{V}(X_i)}{100} = \frac{1}{100}\left[\frac{13}{12} - \left(\frac{17}{24}\right)^2\right]$.

Por la desigualdad de Chebyshev $\mathbb{P}(|X - \mathbb{E}(X)| < \varepsilon) > 1 - \frac{\mathbb{V}(X)}{\varepsilon}$. Luego,

$$\mathbb{P}(|X - \mathbb{E}(X)| < \varepsilon) = \mathbb{P}\left(\left|X - \frac{17}{24}\right| < \varepsilon\right) \underbrace{=}_{\varepsilon = \frac{1}{12}} \mathbb{P}\left(\left|X - \frac{17}{24}\right| < \frac{1}{12}\right) \Rightarrow$$

$$\mathbb{P}(|X - \mathbb{E}(X)| < \varepsilon) = \mathbb{P}\left(\frac{15}{24} < X < \frac{19}{24}\right) > 1 - 12\mathbb{V}(X) = 1 - \frac{3}{25}\left[\frac{13}{12} - \left(\frac{17}{24}\right)^2\right] = \frac{893}{960}.$$

(ii) Sea $X = \frac{1}{n}\sum_{i=1}^{n} X_i$ tal que $\mathbb{E}(X) = \mathbb{E}(X_i) = \frac{17}{24}$ y $\mathbb{V}(X) = \frac{\mathbb{V}(X_i)}{n} = \frac{1}{n}\frac{335}{576}$.

Por la desigualdad de Tchebychev $\mathbb{P}(|X - \mathbb{E}(X)| < \varepsilon) > 1 - \frac{\mathbb{V}(X)}{\varepsilon}$. Luego,

$$\mathbb{P}\left(\frac{15}{24} < X < \frac{19}{24}\right) > 1 - 12\mathbb{V}(X) \geq 0,90 \Rightarrow$$

$$1 - 12\mathbb{V}(X) \geq 0,90 \Leftrightarrow 1 - \frac{12}{n}\frac{335}{576} = 1 - \frac{1}{n}\frac{335}{48} \geq 0,90 \Leftrightarrow 0,1 \geq \frac{1}{n}\frac{335}{48} \Leftrightarrow n \geq \frac{1675}{24} \approx 69,79$$

Por lo tanto, basta con que $n \geq 70$.

2. Sea p la probabilidad de que una persona elegida al azar apoye la legalización del consumo de marihuana (p es desconocida). Se toma una muestra de 50 personas elegidas al azar, se les pregunta si apoyan o no la legalización y se estima p a partir de la frecuencia relativa f_r que se define por

$$f_r = \frac{\text{N}^{\underline{\text{o}}} \text{ de personas encuestadas que están a favor de la legalización}}{50}$$

Observar que f_r es una variable aleatoria y p es simplemente un número. Cuanto más cerca se encuentre f_r de p, mejor estimador será. Hallar una cota superior para $\mathbb{P}(|f_r - p| > 0,1)$ que no dependa de p.

Solución:

Sea la variable aleatoria $X_i \sim Be(p)$ definida como $X_i = \begin{cases} 1 & \text{si la persona } i \text{ está a favor} \\ 0 & sino \end{cases}$.

Es claro que $X_i, \ldots, X_n$ son iid y $X = \sum_{i=1}^{50} X_i \sim Bi(50, p)$ representa la cantidad de personas encuestadas que están a favor de la legalización.

$$\mathbb{P}(|f_r - p| > 0,1) = \mathbb{P}(|f_r - \mathbb{E}(f_r)| > 0,1) \leq \frac{\mathbb{V}(f_r)}{(0{,}1)^2} = 100\frac{\mathbb{V}(X)}{50^2} = \frac{1}{25} \cdot 50p(1-p) = 2p(1-p).$$

Para encontrar una cota superior, basta con observar que la función $f(x) = -2x^2 + 2x$ alcanza su máximo en $x = \dfrac{1}{2}$.

Por lo tanto, si $p = \dfrac{1}{2}$ se tiene que $\mathbb{P}(|f_r - p| > 0,1) < 2p(1-p) \leq \dfrac{1}{2}$.

3. **Desigualdad de Tchebychev a un lado**

   (i) Sea X una variable aleatoria con $\mathbb{E}(X) = 0$ y $\mathbb{V}(X) = \sigma^2 < \infty$. Probar que para todo $a > 0$ vale

$$\mathbb{P}(X > a) \leq \frac{\sigma^2}{\sigma^2 + a^2}.$$

   Sugerencia: Observar que para todo $b > 0$ vale la desigualdad

$$\mathbb{P}(X > a) = \mathbb{P}(X + b > a + b) \leq \mathbb{P}((X+b)^2 > (a+b)^2) \tag{1}$$

   Aplicar la desigualdad de Markov en (1) y calcular el mínimo en b de la cota hallada.

   (ii) Un conjunto de 200 personas (integrado por 100 mujeres y 100 hombres) se divide aleatoriamente en 100 pares de 2 personas cada uno. Utilizar la desigualdad de Tchebychev a una lado para hallar una cota superior para la probabilidad de que menos de 30 de estos pares estén formados por una mujer y un hombre. Comparar con la cota obtenida a partir de la desigualdad de Tchebychev original.

Solución:

(i) sea $b \in \mathbb{R}$, entonces

$$\mathbb{P}(X > a) = \mathbb{P}(X + b > a + b) \leq \mathbb{P}((X+b)^2 > (a+b)^2) = \mathbb{P}(|X+b| > |a+b|) \underbrace{\leq}_{\substack{\text{Desigualdad}\\ \text{Markov}}} \frac{\mathbb{E}(|X+b|)}{|a+b|} = \frac{\mathbb{E}[(X+b)^2]}{(a+b)^2} \Rightarrow$$

$$\mathbb{P}(X > a) \leq \frac{\mathbb{E}(X^2) + 2b\overbrace{\mathbb{E}(X)}^{=0} + b^2}{(a+b)^2} = \frac{\mathbb{V}(X) + b^2}{(a+b)^2} = \frac{\sigma^2 + b^2}{(a+b)^2}$$

Sea $f(x) = \dfrac{\sigma^2 + x^2}{(a+x)^2}$, cuyo mínimo se alcanza en $x = \dfrac{\sigma^2}{a}$ pues $f'(x) = \dfrac{2(ax - \sigma^2)}{(a+x)^3} = 0 \Leftrightarrow x = \dfrac{\sigma^2}{a}$.

Por lo tanto,

$$\mathbb{P}(X > a) \leq \frac{\sigma^2 + \dfrac{\sigma^4}{a^2}}{\left(a + \dfrac{\sigma^2}{a}\right)^2} = \frac{\sigma^2\left(1 + \frac{\sigma^2}{a^2}\right)}{a^2\left(1 + \frac{\sigma^2}{a^2}\right)^2} = \frac{\sigma^2}{a^2 + \sigma^2}.$$

(ii) Sea $X_i = \begin{cases} 1 & \text{si la pareja i-ésima es mixta} \\ 0 & \text{si no} \end{cases}$, entonces $X_i \sim Be\left(\frac{200}{200}\frac{100}{199}\right)$.

Definimos X como la cantidad de parejas mixtas, entonces $X = \sum_{i=1}^{100} X_i$ y $\mathbb{E}(X) = \sum_{i=1}^{100} \mathbb{E}(X_i) = \frac{(100)^2}{199} \approx 50,25$.

Como las $X_1, \ldots, X_{100}$ no son independientes

$$\mathbb{V}(X) = Cov\left(\sum_{i=1}^{100} X_i, \sum_{j=1}^{100} X_j\right) = \sum_{i=1}^{100}\sum_{j=1}^{100} Cov(X_i, X_j)$$

$$\mathbb{E}(X_i, X_j) = \begin{cases} \mathbb{P}(X_i = 1 \text{ y } X_i = 1) = \mathbb{P}(X_i = 1 | X_j = 1)\mathbb{P}(X_j = 1) = \frac{99}{197}\frac{100}{199} & i \neq j \\ \mathbb{E}(X_i^2) = \mathbb{V}(X_i) + [\mathbb{E}(X_i)]^2 = \frac{100}{199}\frac{99}{199} + \left(\frac{100}{199}\right)^2 & i = j \end{cases}$$

por lo tanto,

$$\begin{aligned} \mathbb{V}(X) &= \sum_{i=1}^{100}\sum_{j=1}^{100} Cov(X_i, X_j) = \sum_{i=1}^{100} Cov(X_i, X_i) + 2\sum_{i<j}^{100} Cov(X_i, X_i) \\ &= 100\frac{100}{199}\frac{99}{199} + 2\binom{100}{2}\left[\frac{99}{197}\frac{100}{199} - \left(\frac{100}{199}\right)^2\right]^2 \\ &\approx 25,126 \end{aligned}$$

Luego,

$$\mathbb{P}(X \leq 30) = \mathbb{P}(X - \mathbb{E}(X) \leq -20,25) = \mathbb{P}(\mathbb{E}(X) - X \geq 20,25) \leq \frac{25,126}{(20,25)^2 + 25,126} \approx 0,058.$$

Mientras que si la usamos desigualdad de Tchebychev original

$$\mathbb{P}(X \leq 30) = \mathbb{P}\left(|X - \mathbb{E}(X)| \geq 20,25\right) \leq \frac{25,126}{(20,25)^2} \approx 0,061.$$

4. Sea $(X_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias independientes tal que $X_1 \equiv 0$ y para $n \geq 2$,

$$\mathbb{P}(X_k = k) = \begin{cases} \frac{1}{n^3} & k = \pm 1, \pm 2, \ldots, \pm n \\ 1 - \frac{2}{n^2} & k = 0 \end{cases}$$

Probar que si $\alpha > \frac{1}{2}$ entonces

$$\sum_{i=1}^{n} \frac{X_i}{n^\alpha} \xrightarrow{P} 0.$$

Sugerencia: $\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$.

Solución:

$$\mathbb{E}(X_i) = 0\mathbb{P}(X_i = 0) + \sum_{i=1}^{n} k\mathbb{P}(X_i = k) + \sum_{i=-n}^{-1} k\mathbb{P}(X_i = k) = \sum_{i=1}^{n} k\frac{1}{n^3} + \sum_{i=1}^{n}(-k)\frac{1}{n^3} \equiv 0.$$

$$\mathbb{V}(X_i) = \mathbb{E}(X_i^2) = \sum_{i=1}^{n} k^2\mathbb{P}(X_i = k) + \sum_{i=-n}^{-1} k^2\mathbb{P}(X_i = k) = \sum_{i=1}^{n} d\frac{k^2}{n^3} + \sum_{i=1}^{n}\frac{(-k)^2}{n^3} = 2\sum_{i=1}^{n} d\frac{k^2}{n^3} = \frac{2}{n^3}\frac{n(n+1)(2n+1)}{6}$$

$$\mathbb{E}\left(\sum_{i=1}^{n}\frac{X_i}{n^\alpha}\right) = \frac{1}{n^\alpha}\sum_{i=1}^{n}\mathbb{E}(X_i) = 0 \quad \text{y} \quad \mathbb{V}\left(\sum_{i=1}^{n}\frac{X_i}{n^\alpha}\right) = \frac{1}{n^{2\alpha}}\sum_{i=1}^{n}\mathbb{V}(X_i) = \frac{1}{n^{2\alpha}}\frac{2}{n^3}\frac{n^2(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{3n^{2\alpha+1}}.$$

Por la desigualdad de Tchebychev se tiene

$$\mathbb{P}\left(\left|\sum_{i=1}^{n}\frac{X_i}{n^\alpha} - \mathbb{E}\left(\sum_{i=1}^{n}\frac{X_i}{n^\alpha}\right)\right| > \varepsilon\right) = \mathbb{P}\left(\left|\sum_{i=1}^{n}\frac{X_i}{n^\alpha}\right| > \varepsilon\right) \leq \frac{1}{\varepsilon^2}\mathbb{V}\left(\sum_{i=1}^{n}\frac{X_i}{n^\alpha}\right) = \frac{1}{\varepsilon^2}\frac{(n+1)(2n+1)}{3n^{2\alpha+1}}$$

Si $\alpha > \frac{1}{2}$ entonces, $\lim_{n\to\infty} \frac{(n+1)(2n+1)}{3n^{2\alpha+1}} = 0 \Rightarrow \forall\varepsilon > 0$ se tiene que $\frac{1}{\varepsilon^2} \lim_{n\to\infty} \frac{(n+1)(2n+1)}{3n^{2\alpha+1}} = 0$. Entonces,

$$\mathbb{P}\left(\left|\sum_{i=1}^{n} \frac{X_i}{n^\alpha}\right| > \varepsilon\right) \leq \frac{1}{\varepsilon^2}\frac{(n+1)(2n+1)}{3n^{2\alpha+1}} \xrightarrow[n\to\infty]{} 0 \Rightarrow \sum_{i=1}^{n} \frac{X_i}{n^\alpha} \xrightarrow{P} 0.$$

5. Sea $(X_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias tal que $\lim_{n\to\infty} \mathbb{V}(X_n) = 0$.

   (i) Probar que si $\mathbb{E}(X_n) = 0$ para todo $n \in \mathbb{N}$ entonces $X_n \xrightarrow[n\to\infty]{P} 0$.

   (ii) Probar que si $\lim_{n\to\infty} \mathbb{E}(X_n) = \mu \in \mathbb{R}$ entonces $X_n \xrightarrow[n\to\infty]{P} \mu$.

Solución:

(i) Por la desigualdad de Tchebychev, para todo $\varepsilon > 0$ se tiene que

$$\lim_{n\to\infty} \mathbb{P}(|X_n - \mathbb{E}(X_n)| \geq \varepsilon) \leq \lim_{n\to\infty} \frac{\mathbb{V}(X_n)}{\varepsilon^2} = \frac{1}{\varepsilon^2} \underbrace{\lim_{n\to\infty} \mathbb{V}(X_n)}_{=0} \Rightarrow X_n \xrightarrow[n\to\infty]{P} 0.$$

(ii) Como $\lim_{n\to\infty} \mathbb{E}(X_n) = \mu \in \mathbb{R}$ entonces para todo $\varepsilon' > 0$ existe $n_0 \in \mathbb{N}$ tal que si $n \geq n_0$ se tiene que $|\mathbb{E}(X_n) - \mu| < \varepsilon'$.

Por otra parte, por la desigualdad de Tchebychev, para todo $\varepsilon > \varepsilon' > 0$ se tiene que

$$\lim_{n\to\infty} \mathbb{P}(|X_n-\mu| \geq \varepsilon) = \lim_{n\to\infty} \mathbb{P}(|X_n \pm \mathbb{E}(X_n)-\mu| \geq \varepsilon) \leq \lim_{n\to\infty} \mathbb{P}(|X_n-\mathbb{E}(X_n)| + \underbrace{|\mathbb{E}(X_n) - \mu|} \geq \varepsilon) = \mathbb{P}(|X_n-\mu| \geq \varepsilon-\varepsilon') \Rightarrow$$

$$\lim_{n\to\infty} \mathbb{P}(|X_n - \mu| \geq \varepsilon) \leq \frac{\mathbb{V}(X_n)}{(\varepsilon - \varepsilon')^2} \xrightarrow[n\to\infty]{\mathbb{V}(X_n)\to 0} 0 \Rightarrow X_n \xrightarrow[n\to\infty]{P} \mu.$$

6. Sea $(X_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias y X otra variable aleatoria, todas ellas definidas sobre el mismo espacio.

   (i) Escribir el conjunto $\{X_n \to X\}$ en términos de eventos de la forma $\{|X_n - X| \leq \alpha\}$ con $\alpha > 0$ y verificar que es un evento perteneciente a la σ-álgebra $\mathcal{F}$.

   (ii) Escribir el conjunto $\{X_n \nrightarrow X\}$ en términos de numerables eventos de la forma $\{|X_n - X| > \alpha\}$ con $\alpha > 0$.

   (iii) Verificar $\{X_n \to X\} = \{X_n - X \to 0\}$.

   (iv) Sea $L^+ = \limsup_{n\to\infty} X_n$; i.e., para cada $\omega \in \Omega$ se define $L^+(\omega) = \limsup_{n\to\infty} X_n(\omega)$.

      a. Para cada $\alpha > 0$ escribir el conjunto $\{L^+ \geq \alpha\}$ en términos de numerables eventos de la forma $\{X_n > r$ para infinitos valores de $n\}$ con $r \in \mathbb{R}$ y verificar que es un evento perteneciente a la σ-álgebra $\mathcal{F}$. Deducir que L^+ es una variable aleatoria.

      b. Para cada $\alpha > 0$ escribir el conjunto $\{L^+ > \alpha\}$ en términos de numerables eventos de la forma $\{X_n > r$ para infinitos valores de $n\}$ con $r \in \mathbb{R}$ y verificar que es un evento perteneciente a la σ-álgebra $\mathcal{F}$.

   (v) Demostrar afirmaciones análogas a las del ítem anterior para $L^- = \liminf_{n\to\infty} X_n$.

Solución:

Decimos que $X_n \xrightarrow[n\to\infty]{c.s.} X$ si y sólo si $\mathbb{P}\left(\left\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\right\}\right) = 1$. De donde se puede observar entonces que

$$X_n \xrightarrow[n\to\infty]{c.s.} X \Leftrightarrow \mathbb{P}(\{X_n \to X\}) = 1 \Leftrightarrow \forall\varepsilon > 0,\ \lim_{N\to\infty} \mathbb{P}\left(\bigcup_{n=N}^{\infty} \{|X_n - X| > \varepsilon\}\right) = 0.$$

(i) Sea el conjutno $A = \{X_n \to X\}$ y la idea es escribir al conjunto en términos de elementos de la σ-álgebra $\mathcal{F}$.

En primer lugar, notemos que como $|X_n - X|$ es una variable aleatoria, entonces $\{|X_n - X| \leq \varepsilon\}$ es la preimagen del boreliano $[-\varepsilon, \varepsilon]$ y, entonces, $(X_n - X)^{-1}([-\varepsilon, \varepsilon]) \in \mathcal{F}$.

Luego,

$$\begin{aligned}
&\omega \in A \Leftrightarrow \forall\varepsilon > 0,\ \exists n_0 \in \mathbb{N} \text{ tal que } |X_n(\omega) - X(\omega)| \leq \varepsilon\ \forall n \geq n_0 \\
&\Leftrightarrow \omega \in \bigcap_{\varepsilon>0}\left[\bigcup_{n_0\in\mathbb{N}}\left(\bigcap_{n\geq n_0} \{|X_n - X| \leq \varepsilon\}\right)\right] \\
&\Leftrightarrow \omega \in \bigcap_{k\in\mathbb{N}} \bigcup_{n_0\in\mathbb{N}} \bigcap_{n\geq n_0} \left\{|X_n - X| \leq \frac{1}{k}\right\}.
\end{aligned}$$

$$\Rightarrow A = \{X_n \to X\} = \bigcap_{k\in\mathbb{N}} \bigcup_{n_0\in\mathbb{N}} \bigcap_{n\geq n_0} \left\{|X_n - X| \leq \frac{1}{k}\right\}$$

(ii) Tomando complemento en el resultado del ítem anterior,

$$A^c = \{X_n \nrightarrow X\} = \bigcup_{k\in\mathbb{N}} \bigcap_{n_0\in\mathbb{N}} \bigcup_{n\geq n_0} \left\{|X_n - X| > \frac{1}{k}\right\}$$

(iii) Se deduce usando la misma definición de A usada en el ítem (i).

(iv) Recordar que $L^+(\omega) = \limsup\limits_{n\to\infty} X_n(\omega) = \lim\limits_{n\to\infty}\left(\sup\limits_{m\geq n} X_m(\omega)\right) = \inf\limits_{n\in\mathbb{N}}\left(\sup\limits_{m\geq n} X_m(\omega)\right)$.

a. $L^+(\omega)$ es el mayor punto de acumulación de la sucesión $(X_n(\omega))_{n\in\mathbb{N}}$; i.e., el límite más grande de alguna subsucesión. Por lo tanto, para todo $k \in \mathbb{N}$ definimos el conjunto

$$\left\{X_n > \alpha - \frac{1}{k}\text{para infinitos } n\right\}$$

entonces

$$\bigcap_{k\in\mathbb{N}} \underbrace{\left\{X_n > \alpha - \frac{1}{k}\text{para infinitos } n\right\}}_{\limsup\limits_{n\to\infty}\left\{X_n > \alpha - \frac{1}{k}\right\}} = \{L^+ \geq \alpha\}$$

De esta manera, se deduce que L^+ es una variable aleatoria por ser intersección numerable de eventos de la σ-álgebra $\mathcal{F}$.

b. Como L^+ es una variable aleatoria, entonces el cojunto $\{L^+ > \alpha\}$ es medible y

$$\omega \in \{L^+ > \alpha\} \Leftrightarrow \exists k \in \mathbb{N} \text{ tal que } X_n(\omega) > \alpha\frac{1}{k} \text{ para infinitos } n \Rightarrow \{L^+ > \alpha\} = \bigcup_{k\in\mathbb{N}} \underbrace{\left\{X_n > \alpha + \frac{1}{k}\text{para infinitos } n\right\}}_{\limsup\limits_{n\to\infty}\left\{X_n > \alpha + \frac{1}{k}\right\}}$$

(v) Basta con observar que $\liminf\limits_{n\to\infty} X_n = -\liminf\limits_{n\to\infty}(-X_n)$.

7. Sea $(X_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias independientes con distribución $\mathcal{E}(1)$ y para cada $n \in \mathbb{N}$ definamos la variable aleatoria

$$Y_n = \frac{X_n}{\log(n+1)}.$$

(i) Probar que $Y_n \xrightarrow[n\to\infty]{P} 0$.

(ii) Probar que $\mathbb{P}(L^+ = 1) = 1$, donde $L^+ = \limsup\limits_{n\to\infty} Y_n$.

(iii) Probar que $\mathbb{P}(L^- = 0) = 1$, donde $L^- = \liminf\limits_{n\to\infty} Y_n$.

(iv) Deducir de los items anteriores que la sucesión $(Y_n)_{n\in\mathbb{N}}$ no tiene límite casi seguro.

8. Sea $(X_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias independientes con distribución $N(0,1)$ y para cada $n \in \mathbb{N}$ definamos la variable aleatoria

$$Y_n = \frac{X_n}{\sqrt{\log n}}$$

(i) Probar que $Y_n \xrightarrow[n\to\infty]{P} 0$.

(ii) Probar que $\mathbb{P}(L^+ = \sqrt{2}) = 1$, donde $L^+ = \limsup\limits_{n\to\infty} Y_n$.

Sugerencia: Probar primero que si $X \sim N(0,1)$ entonces para todo $x > 0$ valen las desigualdades

$$\frac{f_X(x)}{1+\frac{1}{x}} \leq \mathbb{P}(X > x) \leq \frac{f_X(x)}{x}.$$

(iii) Probar que $\mathbb{P}(L^- = -\sqrt{2}) = 1$, donde $L^- = \liminf\limits_{n\to\infty} Y_n$.

(iv) Concluir a partir de los items anteriores que la sucesión $(Y_n)_{n\in\mathbb{N}}$ no tiene límite casi seguro.

(v) Probar que $\mathbb{P}\left(\left|\sum_{k=1}^{n} X_k\right| < 2\sqrt{n\log n},\ \text{para todo n suficientemente grande}\right) = 1$. Deducir a partir de este resultado la ley fuerte de los grandes números para variables aleatorias con distribución normal.

9. Se elige al azar un número X en el intervalo $[0,1]$.

Observación: Si un número admite dos desarrollos decimales se optará por el finito.

(i) Dados $k\in\mathbb{N}$ y una secuencia ordenada de k dígitos $(a_1,\dots,a_k)\in\{0,\dots,9\}^k$, calcular la probabilidad para cada $n\in\mathbb{N}$ de que dicha secuencia coincida con la de los dígitos del desarrollo decimal de X entre los lugares n y $n+k-1$.

(ii) Dados $k\in\mathbb{N}$ y una secuencia ordenada de k dígitos $(a_1,\dots,a_k)\in\{0,\dots,9\}^k$, calcular la probabilidad de que dicha secuencia aparezca infinitas veces en el desarrollo decimal de X.

(iii) Calcular $\mathbb{P}(\text{Ocurren infinitos } A_n)$, donde para cada $n\in\mathbb{N}$ se define el evento A_n como

$$A_n = \{\text{El 9 aparece } n \text{ veces consecutivas en los } 2n \text{ primeros lugares del desarrollo decimal de } X\}.$$

(iv) ¿Cuál es la probabilidad de que X sea racional?

10. Se tira infinitas veces una moneda de manera independiente y con probabilidad p de obtener cara en cada lanzamiento.

(i) Dado $k\in\mathbb{N}$ calcular la probabilidad de obtener infinitas rachas de k caras consecutivas.

(ii) Sea A_n el evento de obtener una racha de caras consecutivas de longitud no menor que n entre los lanzamientos 2^n y $2^{n+1}-1$. Probar que

$$p(\text{Ocurren infinitos } A_n) = \begin{cases} 0 & p<\frac{1}{2} \\ 1 & p\geq\frac{1}{2}\end{cases}$$

Sugerencia: SI $p\geq\dfrac{1}{2}$ entonces $\displaystyle\sum_{n\in\mathbb{N}}\left[1-(1-p^n)^{\lfloor\frac{2^n}{n}\rfloor}\right] = +\infty$.

11. Probar que una sucesión de variables aleatorias $(X_n)_{n\in\mathbb{N}}$ converge en probabilidad a una variable aleatoria X si y sólo si toda subsucesión de $(X_n)_{n\in\mathbb{N}}$ contiene otra subsucesión que converge casi seguramente a X.

12. Una colección $(X_i)_{i\in\mathcal{I}}$ de variables aleatorias se dice *acotada en probabilidad* o *tight* si dado $\varepsilon>0$ existe un compacto K_ε tal que

$$\sup_{i\in\mathcal{I}}\mathbb{P}(X_i\notin K_\varepsilon)<\varepsilon.$$

(i) Probar que toda colección finita de variables aleatorias es acotada en probabilidad.

(ii) Mostrar que una familia infinita de variables aleatorias no es necesariamente acotada en probabilidad.

(iii) Sea $(X_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias y X_0 una variable aleatoria, $X_n \xrightarrow[n\to\infty]{P} X_0$. Probar que la familia $(X_n)_{n\in\mathbb{N}_0}$ está acotada en probabilidad.

13. Sean $(X_n)_{n\in\mathbb{N}}$ e $(Y_n)_{n\in\mathbb{N}}$ dos sucesiones de variables aleatorias.

(i) Probar que si $X_n \xrightarrow[n\to\infty]{P} 0$ e $(Y_n)_{n\in\mathbb{N}}$ está acotada en probabilidad entonces $X_nY_n \xrightarrow[n\to\infty]{P} 0$.

(ii) Probar que si $X_n \xrightarrow[n\to\infty]{P} X$ e $Y_n \xrightarrow[n\to\infty]{P} Y$ entonces $X_n+Y_n \xrightarrow[n\to\infty]{P} X+Y$ y que $X_nY_n \xrightarrow[n\to\infty]{P} XY$.

(iii) Probar que si $(X_n)_{n\in\mathbb{N}}$ es una sucesión acotada entonces para todo $p\geq 1$ vale $X_n \xrightarrow[n\to\infty]{P} 0 \Leftrightarrow X_n \xrightarrow[n\to\infty]{L^p} 0$.

(iv) Mostrar con un ejemplo que la equivalencia del ítem anterior puede no valer si la sucesión $(X_n)_{n\in\mathbb{N}}$ no es acotada. ¿Es alguna de las implicaciones cierta siempre? ¿Cuál?

14. Sea $(\boldsymbol{X}_n)_{n\in\mathbb{N}}$ una sucesión de vectores aleatorios sobre $\mathbb{R}^n$ y $\boldsymbol{X}_0$ otro vector aleatorio sobre $\mathbb{R}^n$.

(i) Probar que si $\boldsymbol{X}_n \xrightarrow[n\to\infty]{P} \boldsymbol{X}_0$ y $g:\mathbb{R}^n\to\mathbb{R}^m$ es una función continua entonces $g(\boldsymbol{X}_n) \xrightarrow[n\to\infty]{P} g(\boldsymbol{X}_0)$.

Sugerencia: Tener presente que $(\boldsymbol{X}_n)_{n\in\mathbb{N}_0}$ es una sucesión acotada en probabilidad y que toda función continua es uniformemente continua sobre compactos.

(ii) Probar que si $\boldsymbol{X}_n \xrightarrow[n\to\infty]{cs} \boldsymbol{X}_0$ y $g:\mathbb{R}^n\to\mathbb{R}^m$ es una función continua entonces $g(\boldsymbol{X}_n) \xrightarrow[n\to\infty]{P} g(\boldsymbol{X}_0)$.

(iii) Probar que si $\boldsymbol{X}_n \xrightarrow[n\to\infty]{P} \boldsymbol{X}_0$ y $g : \mathbb{R}^n \to \mathbb{R}$ es una función continua y acotada entonces $g(\boldsymbol{X}_n) \xrightarrow[n\to\infty]{L^p} g(\boldsymbol{X}_0)$ para todo $p \geq 1$.

15. Sea $(X_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias.

 (i) Supongamos que $X_n \xrightarrow[n\to\infty]{P} k$ para cierta constante $k \in \mathbb{R}$ no nula. Se define la sucesión $(Z_n)_{n\in\mathbb{N}}$ como

$$Z_n = \begin{cases} \dfrac{1}{X_n} & X_n \neq 0 \\ 0 & X_n = 0 \end{cases}$$

 Probar que Z_n es una variable aleatoria para todo $n \in \mathbb{N}$ y que $Z_n \xrightarrow[n\to\infty]{P} \dfrac{1}{k}$. Adoptaremos la notación $\dfrac{1}{X_n} \xrightarrow[n\to\infty]{P} \dfrac{1}{k}$ para referirnos a este hecho.

 (ii) Supongamos que $X_n \xrightarrow[n\to\infty]{P} X$ con $\mathbb{P}(X = 0) = 0$. Se define la sucesión $(Z_n)_{n\in\mathbb{N}}$ como en el ítem anterior. Probar que $Z_n \xrightarrow[n\to\infty]{P} Z$, donde Z está defnidia por la fórmula

$$Z_n = \begin{cases} \dfrac{1}{X} & X \neq 0 \\ 0 & X = 0 \end{cases}$$

 Adoptaremos la notación $\dfrac{1}{X_n} \xrightarrow[n\to\infty]{P} \dfrac{1}{X}$ para referirnos a este hecho.

 Sugerencia: Observar que $\dfrac{X_n}{X} \xrightarrow[n\to\infty]{P} 1$ y aplicar el resultado establecido en el ítem anterior.

16. Sean f y g funciones continuas definidas sobre el intervalo $[0,1]$ tales que $0 < f(x) \leq kg(x)$ para todo $x \in [0,1]$, donde k es una cierta constante positiva.

 (i) Sea $\boldsymbol{X} = (X_1, \ldots, X_n)$ un punto elegido al azar en el cubo $[0,1]^n$. Definimos las sucesiones de variables aleatorias $(Y_n)_{n\in\mathbb{N}}$ y $(Z_n)_{n\in\mathbb{N}}$ por las fórmulas

$$Y_n = \sum_{i=1}^{n} \frac{f(X_i)}{n} \quad \text{y} \quad Z_n = \sum_{i=1}^{n} \frac{g(X_i)}{n}$$

 Probar que

$$Y_n \xrightarrow[n\to\infty]{P} \int_0^1 f(x)dx \quad \text{y} \quad Z_n = \xrightarrow[n\to\infty]{P} \int_0^1 g(x)dx$$

 (ii) Probar que

$$\lim_{n\to\infty} \int_0^1 \cdots \int_0^1 \frac{f(x_1) + \ldots + f(x_n)}{g(x_1) + \ldots + g(x_n)} dx_1 \ldots dx_n = \frac{\displaystyle\int_0^1 f(x)dx}{\displaystyle\int_0^1 g(x)dx}$$

17. Sea $(X_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias independientes con distribución $\mathcal{U}([0,1])$. Hallar el límite casi seguro de la sucesión $(Y_n)_{n\in\mathbb{N}}$, donde para cada $n \in \mathbb{N}$ la variable aleatoria Y_n se define como

$$Y_n = \sqrt[n]{\prod_{i=1}^{n} X_i}.$$

18. Sea $(X_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias independientes e idénticamente distribuidas tal que $\mathbb{E}(X) < 1, \mathbb{E}(X^2) = 2$ y $\mathbb{E}(X^4) < +\infty$. Probar que

$$\frac{\displaystyle\sum_{i=1}^{n} X_i}{\sqrt{n}\displaystyle\sum_{i=1}^{n} X_i^2} \xrightarrow[n\to\infty]{P} \frac{1}{\sqrt{2}}.$$

19. Sea $(X_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias independientes con distribución $N(0,1)$. Dados $a, b \in \mathbb{R}$ definimos para cada $n \in \mathbb{N}$ la variable aleatoria

$$Z_n = e^{aS_n - b_n}$$

donde $S_n = \sum_{k=1}^{n} X_k$. Probar que $Z_n \xrightarrow[n\to\infty]{cs} 0 \Leftrightarrow b > 0$ pero para $r \geq 1$ se tiene $Z_n \xrightarrow[n\to\infty]{L^p} 0 \Leftrightarrow r < \dfrac{2b}{a^2}$.

20. Sea $(X_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias i.i.d. con $\mathbb{E}(|X_1|) = +\infty$.

   (i) Probar que para todo $k \in \mathbb{N}$ se verifica $\sum_{n\in\mathbb{N}} \mathbb{P}(|X_n| > kn) = +\infty$.

   (ii) Probar que $\limsup_{n\to\infty} \dfrac{|X_n|}{n} = +\infty$.

   (iii) Probar que $\limsup_{n\to\infty} \dfrac{|S_n|}{n} = +\infty$, donde $S_n = \sum_{k=1}^{n} X_k$.

   (iv) Concluir del ítem anterior que si $(X_n)_{n\in\mathbb{N}}$ es una sucesión de variables aleatorias independientes e idénticamente distribuidas tal que $\dfrac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow[n\to\infty]{cs} \mu$ para cierto $\mu \in \mathbb{R}$ entonces $\mathbb{E}(|X_1|) = +\infty$.

Práctica 9

Distribuciones Condicionales y Esperanza Condicional

1. (i) Sean X e Y variables aleatorias independientes con $X \sim Bi(n,p)$ e $Y \sim Bi(m,p)$, respectivamente. Probar que $X|X+Y=k \sim \mathcal{H}(m+n,n,k)$ y hallar $\mathbb{E}(X|X+Y=k)$.

   (ii) Sean X e Y dos variables aleatorias definidas en un mismo espacio con $X+Y \sim \mathcal{P}(\lambda)$ para $\lambda > 0$. Probar que $X|X+Y=k \sim Bi(k,p)$ para cierto $0 < p < 1$ si y sólo si X e Y son independientes y tienen distribución $X \sim \mathcal{P}(\lambda p)$ e $Y \sim \mathcal{P}(\lambda(1-p))$, respectivamente.

<u>Solución:</u>

(i) Como X e Y son variables aleatorias independientes se tiene que $X+Y \sim Bi(n+m,p)$ y

$$\mathbb{P}(X=x|X+Y=k) = \frac{\mathbb{P}(X=x \text{ y } X+Y=k)}{\mathbb{P}(X+Y=k)} = \frac{\mathbb{P}(X=x \text{ y } Y=k-x)}{\mathbb{P}(X+Y=k)} \Rightarrow$$

$$\mathbb{P}(X=x|X+Y=k) = \frac{\binom{n}{x} p^x(1-p)^{n-x} \binom{m}{k-x} p^{k-x}(1-p)^{m-k}}{\binom{m+n}{k} p^k(1-p)^{n+m-k}} = \frac{\binom{n}{x}\binom{m}{k-x}}{\binom{m+n}{k}}.$$

$$\Rightarrow X|X+Y=k \sim \mathcal{H}(m+n,n,k) \Rightarrow \mathbb{E}(X|X+Y=k) = k\frac{n}{m+n}.$$

(ii)

<u>$(\Rightarrow)$:</u> $\mathbb{P}(X=x|X+Y=k) = \dfrac{\mathbb{P}(X=x \text{ y } X+Y=k)}{\mathbb{P}(X+Y=k)}$, entonces

$$\begin{aligned}
\mathbb{P}(X=x \text{ y } X+Y=k) &= \mathbb{P}(X=x|X+Y=k)\mathbb{P}(X+Y=k) \\
&= \binom{k}{x} p^x(1-p)^{k-x}\frac{\lambda^k}{k!}e^{-\lambda} \\
&= \frac{1}{x!}\frac{1}{(k-x)!}p^x(1-p)^{k-x}\lambda^{k\pm x}e^{-\lambda\pm\lambda p} \\
&= \underbrace{\left[\frac{(\lambda p)^x}{x!}e^{-\lambda p}\right]}_{\mathbb{P}_X(x)} \underbrace{\left[\frac{(\lambda(1-p))^{k-x}}{(k-x)!}e^{-\lambda(1-p)}\right]}_{\mathbb{P}_Y(k-x)}
\end{aligned}$$

Como quedó factorizada como producto de dos probabiliades correspondientes a dos variables aleatorias de distribución $\mathcal{P}(\lambda p)$ e $\mathcal{P}(\lambda(1-p))$ se sigue que X e Y con independientes tales que $X \sim \mathcal{P}(\lambda p)$ e $Y \sim \mathcal{P}(\lambda(1-p))$.

<u>$(\Leftarrow)$:</u>

$$\begin{aligned}
\mathbb{P}(X=x|X+Y=k) &= \frac{\mathbb{P}(X=x \text{ y } Y=k-x)}{\mathbb{P}(X+Y=k)} = \frac{\mathbb{P}(X=x)\mathbb{P}(Y=k-x)}{\mathbb{P}(X+Y=k)} \\
&= \frac{(\lambda p)^x}{x!}e^{-\lambda p}\frac{(\lambda(1-p))^{k-x}}{(k-x)!}e^{-\lambda(1-p)}\frac{k!}{\lambda^k e^{-\lambda}} \\
&= p^x(1-p)^{k-x}\frac{k!}{x!(k-x)!} = \binom{k}{k} p^x(1-p)^{k-x}
\end{aligned}$$

$$\Rightarrow X|X+Y=k \sim Bi(k,p).$$

2. Un señor se va a pescar un fin de semana. La cantidad de peces que pican en el lapso de una hora sigue una distribución $\mathcal{P}(\lambda)$. Además cada pez tiene probabilidad q de zafarse y $1-q=p$ de ser atrapado. Sean $X=$ cantidad de peces que pican e $Y=$ cantidad de peces atrapados.
   (i) Encontrar el valor esperado del número de peces atrapados sabiendo que picaron k peces.
   (ii) Encontrar el valor esperado del número de peces atrapados.
   (iii) Mostrar que Y tiene distribución $\mathcal{P}(\lambda p)$.

Solución:

Sea $Y_i = \begin{cases} 1 & \text{si el pez } i \text{ fue artapado} \\ 0 & \text{si no} \end{cases}$ tal que $\{Y_i\}$ resultan independientes e identicamente distribuídas $Y_i \sim Be(p)$.

(i) $Y|X=k = \sum_{i=1}^{k} Y_i \sim Bi(k,p)$, por lo tanto $\mathbb{E}(Y|X=k)=kp$.

(ii) $\mathbb{E}(Y)=\mathbb{E}[\mathbb{E}(Y|X)]=\mathbb{E}(pX)=p\mathbb{E}(X)=p\lambda$.

(iii) $\mathbb{P}(Y=y|X=x)=\dfrac{\mathbb{P}(Y=y \text{ y } X=x)}{\mathbb{P}(X=x)}$, entonces

$$\begin{aligned}\mathbb{P}(Y=y \text{ y } X=x) &= \mathbb{P}(Y=y|X=x)\mathbb{P}(X=x)\\ &= \binom{x}{y} p^y(1-p)^{x-y}\frac{\lambda^x}{x!}e^{-\lambda} = \frac{p^y}{y!}\frac{(1-p)^{x-y}}{(x-y)!}\lambda^x e^{-\lambda}\end{aligned}$$

entonces

$$\mathbb{P}(Y=y)=\frac{\lambda^y p^y}{y!}e^{-\lambda}\sum_{x\in\mathbb{N}_0}\frac{(1-p)^{x-y}}{(x-y)!}\lambda^{x-y}=\frac{e^{-\lambda}(p\lambda)^y}{y!}\sum_{x\in\mathbb{N}_0}\frac{[\lambda(1-p)]^{x-y}}{(x-y)!}=\frac{e^{-\lambda}(p\lambda)^y}{y!}\sum_{j=y}^{\infty}\frac{[\lambda(1-p)]^j}{j!}$$

$$\Rightarrow \mathbb{P}(Y=y)=\frac{e^{-\lambda}(p\lambda)^y}{y!}e^{\lambda(1-p)}=\frac{e^{-\lambda p}(p\lambda)^y}{y!}\Rightarrow Y\sim\mathcal{P}(\lambda p).$$

3. Se tienen dos urnas numeradas 0 y 1 con bolas blancas y negras. La urna 0 tiene 2 bolas blancas y 8 negras, mientras que la urna 1 tiene 7 bolas blancas y 3 negras.
   (i) Elegimos una urna al azar. Luego sacamos de ella, con reposición, 5 bolas. Definimos las variables aleatorias

   $X=$ cantidad de bolas blancas obtenidas en las 5 extracciones.
   $Y=$ urna elegida.

   Calcular $\mathbb{E}(X|Y=0), \mathbb{E}(X|Y=1)$ y $\mathbb{E}(X)$. Compararlas entre sí.
   (ii) Ahora cambiamos el procedimiento. Elegimos una urna al azar, y de ella extraemos 1 bola, que luego reponemos en la urna correspondiente. Luego repetimos el experimento 5 veces, de manera independiente. Sea $Z=$ cantidad de bolas blancas obtenidas en las 5 extracciones. Hallar $\mathbb{E}(Z)$ y compararla con la $\mathbb{E}(X)$ hallada en el ítem anterior.

Solución:

(i) Para poder calcular $\mathbb{E}(X|Y=0)$ y $\mathbb{E}(X|Y=1)$, basta con notar que

$$X|Y=0\sim Bi\left(5,\frac{1}{5}\right)\Rightarrow\mathbb{E}(X|Y=0)=5\frac{1}{5}=1.$$

$$X|Y=1\sim Bi\left(5,\frac{7}{10}\right)\Rightarrow\mathbb{E}(X|Y=1)=5\frac{7}{10}=\frac{35}{10}.$$

Por otra parte, $\mathbb{E}(X)=\sum_{i=0}^{5} i\mathbb{P}(X=i)$. Por lo tanto, calculamos $\mathbb{P}(X=i)$.

$$\begin{aligned}\mathbb{P}(X=x) &= \mathbb{P}(X=x|Y=0)\mathbb{P}(Y=0)+\mathbb{P}(X=x|Y=1)\mathbb{P}(Y=1)\\ &= \frac{1}{2}\binom{5}{x}\left(\frac{1}{5}\right)^x\left(\frac{4}{5}\right)^{5-x}+\frac{1}{2}\binom{5}{x}\left(\frac{7}{10}\right)^x\left(\frac{3}{10}\right)^{5-x}\\ &= \frac{1}{2}\binom{5}{x}\left[\left(\frac{1}{4}\right)^x\left(\frac{4}{5}\right)^5+\left(\frac{7}{3}\right)^x\left(\frac{3}{10}\right)^5\right]\end{aligned}$$

luego,

$$\mathbb{E}(X) = \sum_{i=0}^{5} i\mathbb{P}(X = i) = \frac{1}{2}\left(\frac{4}{5}\right)^5 \sum_{i=0}^{5} i \binom{5}{i} \left(\frac{1}{4}\right)^i + \frac{1}{2}\left(\frac{3}{10}\right)^5 \sum_{i=0}^{5} i \binom{5}{i} \left(\frac{7}{3}\right)^i$$

(ii) $\mathbb{P}(Z = z) = \mathbb{P}\left(\left\{\begin{matrix}\text{extraer k}\\ \text{blancas}\end{matrix}\right\}\right) = \sum_{j=0}^{z} \mathbb{P}\left(\left\{\begin{matrix}\text{j blancas de}\\ \text{la urna 0}\end{matrix}\right\}\right) \mathbb{P}\left(\left\{\begin{matrix}\text{z-j blancas de}\\ \text{la urna 1}\end{matrix}\right\}\right)$, luego

$$\mathbb{P}(Z = z) = \sum_{j=0}^{z} \mathbb{P}(X = j|Y = 0)\mathbb{P}(X = z - j|Y = 1) = \sum_{j=0}^{z} \binom{5}{j} \left(\frac{1}{5}\right)^j \left(\frac{4}{5}\right)^{5-j} \binom{5}{z-j} \left(\frac{7}{10}\right)^{z-j} \left(\frac{3}{10}\right)^{5-z+j}$$

por lo tanto, $\mathbb{E}(Z) = \sum_{k=0}^{5} k\mathbb{P}(Z = z)$.

4. Sean $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$. Probar que

$$Cov[X - \mathbb{E}(X|Y), \mathbb{E}(X|Y)] = 0.$$

Interpretar geométricamente.

Solución:

$$\begin{aligned} Cov[X - \mathbb{E}(X|Y), \mathbb{E}(X|Y)] &= Cov[X, \mathbb{E}(X|Y)] - Cov[\mathbb{E}(X|Y), \mathbb{E}(X|Y)] \\ &= \mathbb{E}[X\mathbb{E}(X|Y)] - \mathbb{E}(X) \underbrace{\mathbb{E}[\mathbb{E}(X|Y)]}_{=\mathbb{E}(X)} - \mathbb{E}[\mathbb{E}(X|Y)\mathbb{E}(X|Y)] + \underbrace{\mathbb{E}[\mathbb{E}(X|Y)]\mathbb{E}[\mathbb{E}(X|Y)]}_{=(\mathbb{E}(X))^2} \\ &= \mathbb{E}[\mathbb{E}(X\mathbb{E}(X|Y)|Y)] - \mathbb{E}[\mathbb{E}(X|Y)\mathbb{E}(X|Y)] = \mathbb{E}[\mathbb{E}(X|Y)\mathbb{E}(X|Y)] - \mathbb{E}[\mathbb{E}(X|Y)\mathbb{E}(X|Y)] = 0. \end{aligned}$$

La funci ón $\mathbb{E}(X|Y)$ es la proyección de X sobre el espacio de funciones de Y en L^2.

5. Sean $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$. Se define la varianza condicional de X dada Y como la variable aleatoria

$$\mathbb{V}(X|Y) := \mathbb{E}\left[(X - \mathbb{E}(X|Y))^2|Y\right].$$

Probar que

$$\mathbb{V}(X) = \mathbb{V}[\mathbb{E}(X|Y)] + \mathbb{E}[\mathbb{V}(X|Y)].$$

Solución:

$$\begin{aligned} \mathbb{V}[\mathbb{E}(X|Y)] + \mathbb{E}[\mathbb{V}(X|Y)] &= \mathbb{E}[\mathbb{E}(X|Y)\mathbb{E}(X|Y)] - \mathbb{E}(\mathbb{E}(X|Y))\mathbb{E}(\mathbb{E}(X|Y)) + \mathbb{E}[\mathbb{E}(X - \mathbb{E}(X|Y))^2|Y] \\ &= \mathbb{E}[\mathbb{E}(X|Y)\mathbb{E}(X|Y)] - \mathbb{E}(\mathbb{E}(X|Y))\mathbb{E}(\mathbb{E}(X|Y)) + \mathbb{E}[\mathbb{E}(XX|Y)] - \\ &\quad (2\mathbb{E}[\mathbb{E}(X\mathbb{E}(X|Y)|Y)] - \mathbb{E}[\mathbb{E}(\mathbb{E}(X|Y)\mathbb{E}(X|Y)|Y)]) \\ &= \mathbb{E}[\mathbb{E}(X|Y)\mathbb{E}(X|Y)] - \mathbb{E}(X)\mathbb{E}(X) + \mathbb{E}(XX) - 2\underbrace{\mathbb{E}[X\mathbb{E}(X|Y)]}_{\mathbb{E}[\mathbb{E}(X|Y)\mathbb{E}(X|Y)]} + \mathbb{E}[\mathbb{E}(X|Y)\mathbb{E}(X|Y)] \\ &= \mathbb{E}(XX) - \mathbb{E}(X)\mathbb{E}(X) = \mathbb{V}(X) \end{aligned}$$

6. Sean $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y X e Y variables aleatorias en $L^2(\Omega, \mathcal{F}, \mathbb{P})$.

   (i) Supongamos que (X, Y) es un vector discreto con probabilidad conjunta $\mathbb{P}_{XY}$. Sea $g : R_Y \to \mathbb{R}$ medible Borel tal que $\mathbb{E}(X|Y) = g(Y)$ y definamos $\sigma^2 : R_Y \to \mathbb{R}$ por la fórmula

$$\sigma^2(y) = \sum_{x \in R_X} [x - g(y)]^2 \mathbb{P}_{X|Y=y}(x).$$

   Es decir, para cada $y \in R_Y$, $\sigma^2(y)$ denota la varianza condicional de X dado el evento $\{Y = y\}$. Probar que $\mathbb{V}(X|Y) = \sigma^2(Y)$.

   (ii) Supongamos que (X, Y) es un vector absolutamente continuo con función de densidad conjunta f_{XY}. Si $sop(Y) = \left\{y \in \mathbb{R} : \int_{\mathbb{R}} f_{XY}(x, y)dx > 0\right\}$ y $g : \mathbb{R} \to \mathbb{R}$ medible Borel es tal que $\mathbb{E}(X|Y) = g(Y)$, definamos $\sigma^2 : sop(Y) \to \mathbb{R}$ por la fórmula

$$\sigma^2(y) = \int_{\mathbb{R}} [x - g(y)]^2 f_{X|Y=y}(x)dx.$$

   Probar que $\mathbb{V}(X|Y) = \sigma^2(Y)$.

7. (i) Sean X e Y variables aleatorias tales que para cada $y \in \mathbb{R}$ se tiene que $\mathbb{P}(Y \leq y|X)$ es una constante $F(y)$ que no depende de X. Probar que $F_Y(y) = F(y)$ para todo $y \in \mathbb{R}$ y que X e Y son independientes.

   (ii) Contando con el resultado del ítem anterior, rehacer el ejercicio 1b.

Solución:

(i) $F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{E}\left(\mathbb{1}_{(-\infty,y)}(y)\right) = \mathbb{E}\left[\mathbb{E}\left(\mathbb{1}_{(-\infty,y)}(y)|X\right)\right] = \mathbb{E}\left[\mathbb{E}\left(\mathbb{1}_{(-\infty,y)}(y|x)\right)\right] = \mathbb{P}(Y \leq y|X) = F(y)$.

(ii) $\mathbb{P}(Y \leq y|X = x) = \dfrac{\mathbb{P}(Y \leq y \text{ y } X = x)}{\mathbb{P}(X = x)} = F_Y(y)$, entonces

$$\mathbb{P}(Y \leq y \text{ y } X = x) = F_Y(y)\mathbb{P}(X = x)$$

con $F_Y(y)$ no depende de X y $\mathbb{P}(X = x)$ no depende de y. Luego,

$$\mathbb{P}(Y \leq y|X = x) = \mathbb{P}(Y \leq y)\mathbb{P}(X = x) \Rightarrow X \text{ e } Y \text{ son independientes.}$$

8. Sean X e Y variables aleatorias definidas en un mismo espacio con Y discreta. Probar que $\mathbb{E}(X|Y) = g(Y)$ donde $g : R_Y \to \mathbb{R}$ se define para cada $y \in R_Y$ mediante la fórmula

$$g(y) = \frac{1}{\mathbb{P}(Y = y)}\mathbb{E}\left(X\mathbb{1}_{\{Y=y\}}\right).$$

Solución:

$$\mathbb{E}(X|Y = y) = \sum_{x \in R_X} \frac{x\mathbb{P}(X = x, Y = y)}{\mathbb{P}(Y = y)} = \frac{1}{\mathbb{P}(Y = y)} \sum_{x \in R_X} x\mathbb{P}(X = x, Y = y) = \frac{1}{\mathbb{P}(Y = y)}\mathbb{E}(X\mathbb{1}_{Y=y}) = g(y)$$

$$\mathbb{E}(X|Y = y) = \int_{R_X} \frac{x\mathbb{P}(X = x, Y = y)}{\mathbb{P}(Y = y)}dx = \frac{1}{\mathbb{P}(Y = y)} \int_{R_X} x\mathbb{P}(X = x, Y = y)dx = \frac{1}{\mathbb{P}(Y = y)}\mathbb{E}(X\mathbb{1}_{Y=y}) = g(y)$$

9. Sean $(X_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias i.i.d. con esperanza y varianza finitas y N una variable aleatoria a valores en $\mathbb{N}$ con esperanza y varianza finitas e independiente de $(X_n)_{n\in\mathbb{N}}$.

   (i) Probar que $\mathbb{E}\left[\sum_{i=1}^{N} X_i\right] = \mathbb{E}(N)\mathbb{E}(X_1)$.

   (ii) Probar que
   $$\mathbb{V}\left(\sum_{i=1}^{N} X_i\right) = \mathbb{E}(N)\mathbb{V}(X_1) + \mathbb{E}^2(X_1)\mathbb{V}(N).$$

   (iii) El número de reclamos recibido por una compañía de seguros en una semana es una variable aleatoria Poisson de parámetro λ. El monto pagado por cada uno de esos reclamos es una variable aleatoria con media μ y varianza σ^2. Hallar la esperanza y la varianza del total pagado por la compañía en concepto de reclamos en una semana. ¿Qué hipótesis se están haciendo? ¿Son razonables?

Solución:

(i) Sabemos que para cada k fijo, como $(X_n)_{n\in\mathbb{N}}$ es una sucesión de variables aleatorias i.i.d. con esperanza finitas, se tiene que $\mathbb{E}\left(\sum_{i=1}^{k} X_i\right) = k\mathbb{E}(X_1)$. Luego,

$$k\mathbb{E}(X_1) = \mathbb{E}\left(\sum_{i=1}^{k} X_i \middle| N = k\right) \Rightarrow \mathbb{E}\left(\sum_{i=1}^{N} X_i \middle| N\right) = N\mathbb{E}(X_1)$$

Por lo tanto,

$$\mathbb{E}\left(\sum_{i=1}^{N} X_i\right) = \mathbb{E}\left[\mathbb{E}\left(\sum_{i=1}^{N} X_i \middle| N\right)\right] = \mathbb{E}(N\mathbb{E}(x_1)) \underbrace{=}_{\substack{X \text{ y } N \\ ind}} \mathbb{E}(N)\mathbb{E}(X_1)$$

(ii) $\mathbb{V}\left(\sum_{i=1}^{N} X_i\right) = \mathbb{E}\left[\left(\sum_{i=1}^{N} X_i\right)^2\right] - \left[\mathbb{E}\left(\sum_{i=1}^{N} X_i\right)\right]^2$. Luego,

$$\left[\mathbb{E}\left(\sum_{i=1}^{N} X_i\right)\right]^2 = \mathbb{E}^2(N)\mathbb{E}^2(X_1) \quad \text{y} \quad \mathbb{E}\left[\left(\sum_{i=1}^{N} X_i\right)^2\right] = \mathbb{E}\left[\mathbb{E}\left[\left(\sum_{i=1}^{N} X_i\right)^2 \middle| N = n\right]\right]$$

$$\mathbb{E}\left[\left(\sum_{i=1}^{N} X_i\right)^2 \middle| N = n\right] = \mathbb{V}\left(\sum_{i=1}^{n} X_i\right) + \left[\mathbb{E}\left(\sum_{i=1}^{n} X_i\right)\right]^2 = n\mathbb{V}(X_1) + [n\mathbb{E}(X_1)]^2$$

luego,

$$\mathbb{E}\left[\left(\sum_{i=1}^{N} X_i\right)^2\right] = \sum_{i\in\mathbb{N}} \left[n\mathbb{V}(X_1) + (n\mathbb{E}(X_1))^2\right] \mathbb{P}(N = n) = \mathbb{E}(N)\mathbb{V}(X_1) + \mathbb{E}(N^2)\mathbb{E}^2(X_1)$$

entonces,

$$\mathbb{V}\left(\sum_{i=1}^{N} X_i\right) = \mathbb{E}(N)\mathbb{V}(X_1) + \mathbb{E}(N^2)\mathbb{E}^2(X_1) - \mathbb{E}^2(N)\mathbb{E}^2(X_1) = \mathbb{E}(N)\mathbb{V}(X_1) + \mathbb{E}^2(X_1)\mathbb{V}(N)$$

(iii) Sea $N =$ número de reclamos recibidos por una compañía de seguros en una semana ($N \sim \mathcal{P}(\lambda)$) y $X_i =$ monto pagado en el reclamo i ($\mathbb{E}(X_i) = \mu$ y $\mathbb{V}(X_i) = \sigma^2$).

Luego, la esperanza y la varianza del total pagado por la compañía en concepto de reclamos en una semana viene dada por

$$\mathbb{E}\left(\sum_{i=1}^{N} X_i\right) = \mathbb{E}(N)\mathbb{E}(X_1) = \lambda\mu$$

$$\mathbb{V}\left(\sum_{i=1}^{N} X_i\right) = \lambda\sigma^2 + \lambda\mu^2 = \lambda(\sigma^2 + \mu^2)$$

notemos que el monto que se paga por cada reclamo es independiente de la cantidad de reclamos.

10. Se tiene un bolillero con 100 bolillas, siendo 60 de ellas de color verde y el resto blancas. Se procede a realizar extracciones sucesivas del bolillero obedeciendo la siguiente regla: para cada $n \in \mathbb{N}$, en la n-ésima extracción se sacan del bolillero X_n bolillas, que luego son devueltas a dicho bolillero para proceder con la siguiente extracción. Para cada $n \in \mathbb{N}$ la cantidad de bolillas X_n extraídas en la n-ésima instancia es

$$X_n = \begin{cases} 10 & n = 1 \\ \text{cantidad de bolillas verdes obtenidas en la } (n-1) - \text{ésima extracción} & n \geq 2 \end{cases}$$

(i) Calcular $\mathbb{E}(X_n)$ para todo $n \in \mathbb{N}$.

(ii) Probar que $X_n \xrightarrow[n\to\infty]{cs} 0$.

(iii) Sea S_n la cantidad de bolillas verdes obtenidas en las primeras n extracciones. Calcular $\lim\limits_{n\to\infty} \mathbb{E}(S_n)$.

(iv) Calcular $Cov(X_2, X_3)$.

11. Sea $(X_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias discretas independientes. Para cada $n \in \mathbb{N}$ definimos $S_n = \sum_{i=1}^{n} X_i$.

Observar que S_n es una variable discreta para todo $n \in \mathbb{N}$.

(i) Dados números naturales $k < n < m$, mostrar que S_k y S_m son condicionalmente independientes dado S_n, i.e., para todo $x_k \in R_{S_k}$ y $x_m \in R_{S_m}$ se verifica

$$\mathbb{P}(S_k = x_k, S_m = x_m | S_n) = \mathbb{P}(S_k = x_k | S_n)\mathbb{P}(S_m = x_m | S_n).$$

Sugerencia: Mostrar que $\mathbb{P}(S_k = x_k, S_m = x_m | S_n) = \mathbb{P}(S_k = x_k | S_n)\mathbb{P}(S_m = x_m | S_n)$ para todo $x_n \in S_n$.

(ii) Mostrar que si $\mathbb{E}(X_i) = 0$ para todo $i \in \mathbb{N}$ entonces dados $n < m$ se verifica $\mathbb{E}(S_m | S_n) = S_n$.

Solución:

(i)

$$\begin{aligned}
\mathbb{P}(S_k = x_k \text{ y } S_m = x_m | S_n = x_n) &= \mathbb{P}\left(\sum_{i=1}^{k} X_i = x_k \text{ y } \sum_{i=1}^{m} X_i = x_m \middle| \sum_{i=1}^{n} X_i = x_n \right) \\
&= \frac{\mathbb{P}\left(\sum_{i=1}^{k} X_i = x_k \text{ y } \sum_{i=1}^{m} X_i = x_m \text{ y } \sum_{i=1}^{n} X_i = x_n \right)}{\mathbb{P}\left(\sum_{i=1}^{n} X_i = x_n \right)} \\
&\underbrace{=}_{\substack{(X_n)_{n\in\mathbb{N}} \\ independientes}} \frac{\mathbb{P}\left(X_{k+1} + \ldots + X_n = x_n - x_k \text{ y } X_{n+1} + \ldots + X_m = x_m - x_n\right)}{\mathbb{P}\left(X_1 + \ldots + X_n = x_n\right)} \\
&= \frac{\mathbb{P}\left(X_{k+1} + \ldots + X_n = x_n - x_k\right)\mathbb{P}\left(X_{n+1} + \ldots + X_m = x_m - x_n\right)}{\mathbb{P}\left(X_1 + \ldots + X_n = x_n\right)} \\
&= \frac{\mathbb{P}\left(\sum_{i=1}^{k} X_i = x_k \text{ y } \sum_{i=1}^{n} X_i = x_n \right)}{\mathbb{P}\left(X_1 + \ldots + X_n = x_n\right)} \frac{\mathbb{P}\left(\sum_{i=1}^{m} X_i = x_m \text{ y } \sum_{i=1}^{n} X_i = x_n \right)}{\mathbb{P}\left(X_1 + \ldots + X_n = x_n\right)} \\
&= \mathbb{P}(S_k = x_k | S_n = x_n)\mathbb{P}(S_m = x_m | S_n = x_n).
\end{aligned}$$

(ii)

$$\begin{aligned}
\mathbb{E}(S_m | S_n = x_n) &= \mathbb{E}\left(\sum_{i=1}^{m} X_i \middle| \sum_{i=1}^{n} X_i = x_n \right) \\
&\underbrace{=}_{n<m} \mathbb{E}\left(x_n + X_{n+1} + \ldots + X_m\right) \\
&= x_n + \sum_{i=n+1}^{m} \underbrace{\mathbb{E}(X_i)}_{=0} = x_n
\end{aligned}$$

Por lo tanto, $\mathbb{E}(S_m|S_n = x_n) = x_n \Rightarrow \mathbb{E}(S_m|S_n) = S_n$.

12. La llegada de un tren a la estación se produce con distribución uniforme entre las 10 am y las $10:20$ am. La partida del mismo se produce con distribución uniforme entre su llegada y las 11 am. Sean X la hora de llegada de dicho tren e Y la hora de su partida.
    (i) Calcular $\mathbb{E}(Y)$ y $cov(X, Y)$.
    (ii) Hallar la función de densidad de Y.
    (iii) Calcular la probabilidad de que el tren haya llegado a la estación entre las $10:10$ am y las $10:15$ am sabiendo que partió después de las $10:25$ am.

Solución:

Es claro que $X \sim \mathcal{U}([0, 20])$ e $Y \sim \mathcal{U}([x, 60])$ con X e Y dependientes.

(i) $\mathbb{E}(Y|X = x) = \displaystyle\int_x^{60} \frac{y}{60 - x} dy = \frac{x + 60}{2}$, por lo tanto $\mathbb{E}(Y||X) = \dfrac{X + 60}{2}$. Luego,

$$\mathbb{E}(Y) = \mathbb{E}\left(\frac{X + 60}{2}\right) = \frac{\mathbb{E}(X) + 60}{2} = \frac{10 + 60}{2} = 35.$$

$$\mathbb{E}(X \cdot Y) = \mathbb{E}\left(\mathbb{E}(X \cdot Y|X)\right) = \mathbb{E}(X\mathbb{E}(Y|X)) = \mathbb{E}\left(X\frac{X + 60}{2}\right) = \frac{\mathbb{E}(X^2)}{2} + 30\mathbb{E}(X) = \frac{\mathbb{V}(X) + \mathbb{E}^2(X)}{2} + 300 = \frac{1100}{3}$$

$$Cov(X, Y) = \mathbb{E}(X \cdot Y) - \mathbb{E}(X)\mathbb{E}(Y) = \frac{50}{3}$$

(ii)

$$\begin{aligned}
f_Y(y) &= \int_{\mathbb{R}} f_{XY}(x,y)dx = \int_{\mathbb{R}} f_{Y|X=x}(y) f_X(x)dx \\
&= \int_{\mathbb{R}} \frac{1}{60-x} \mathbb{1}_{[x,60]}(y) \mathbb{1}_{[0,20]}(x)dx \\
&\underbrace{=}_{\substack{0\le x\le 20 \\ x\le y\le 60}} \int_0^{20} \frac{1}{60-x} \mathbb{1}_{[20,60]}(y)dx + \int_0^{y} \frac{1}{60-x} \mathbb{1}_{[0,20]}(y)dx \\
&= \ln\left(\frac{3}{2}\right) \mathbb{1}_{[20,60]}(y) + \ln\left(\frac{60-y}{60}\right) \mathbb{1}_{[0,20]}(y)
\end{aligned}$$

(iii)

$$\begin{aligned}
\mathbb{P}(10 \le X \le 15 | Y > 25) &= \frac{\mathbb{P}(10 \le X \le 15 \text{ y } Y > 25)}{\mathbb{P}(Y > 25)} \\
&= \frac{\displaystyle\int_{10}^{15} \left(\int_{25}^{+\infty} \frac{1}{60-x} \mathbb{1}_{[x,60]}(y)dy \right) \mathbb{1}_{[0,20]}(x)dx}{\displaystyle\int_{25}^{+\infty} \ln\left(\frac{3}{2}\right) \mathbb{1}_{[20,60]}(y)dy} \\
&= \frac{\displaystyle\int_{10}^{15} \frac{1}{60-x}(60-25)dx}{35 \ln\left(\frac{3}{2}\right)} = \frac{\ln\left(\frac{10}{9}\right)}{\ln\left(\frac{3}{2}\right)}
\end{aligned}$$

13. (i) Sean X e Y dos variables aleatorias definidas en un mismo espacio con $X+Y \sim \Gamma(2,\lambda)$ para $\lambda > 0$. Probar que $X|X+Y=z \sim \mathcal{U}([0,z])$ para todo $z > 0$ si y sólo si X e Y son independientes y tienen distribución $\mathcal{E}(\lambda)$.
    (ii) Sean X y Z variables aleatorias que satisfacen $Z \sim \Gamma(2,1)$ y $X|Z=z \sim \mathcal{U}([0,z])$ para todo $z > 0$.
        a. Calcular $\mathbb{P}(Z \ge 2 | X \le 1)$.
        b. Calcular $\mathbb{P}(Z - X \ge 2 | X \le 1)$.

Solución:

(i)$\underline{(\Rightarrow):}$

$$f(X=x \text{ y } Z=z) = f_{X|Z=z}(X=x|Z=z) f_Z(z) = \frac{1}{z}\mathbb{1}_{[0,z]} \lambda^2 z e^{-\lambda z} \mathbb{1}_{(0,+\infty)} = \lambda^2 e^{-\lambda z} \mathbb{1}_{[0,z]}(x) \mathbb{1}_{(0,+\infty)}(z)$$

$$\Rightarrow f(X=x \text{ y } Z=z) = \lambda^2 e^{-\lambda(z \pm x)} \mathbb{1}_{(0,\infty)}(x) \mathbb{1}_{(0,+\infty)}(z)$$

$$\Rightarrow f(X=x \text{ y } Y=z-x) = \lambda e^{-\lambda x} \mathbb{1}_{(0,\infty)}(x) \lambda e^{-\lambda(z-x)} \mathbb{1}_{(0,+\infty)}(z-x)$$

Por lo tanto, X e Y son independientes y tienen distribución $\mathcal{E}(\lambda)$.

$\underline{(\Leftarrow):}$

$$f_{X|Z=z}(X=x|Z=z) = \frac{f_{XY}(X=x \text{ y } Z=z)}{f_Z(z)} = \frac{f_{XY}(X=x \text{ y } Y=z-x)}{f_Z(z)} = \frac{\lambda e^{-\lambda x} \mathbb{1}_{(0,\infty)}(x) \lambda e^{-\lambda(z-x)} \mathbb{1}_{(0,\infty)}(z-x)}{\lambda^2 z e^{-\lambda z} \mathbb{1}_{(0,+\infty)}(z)}$$

$$\Rightarrow f_{X|Z=z}(X=x|Z=z) = \frac{1}{z} \frac{\mathbb{1}_{(0,\infty)}(x) \mathbb{1}_{(0,\infty)}(z-x)}{\mathbb{1}_{(0,+\infty)}(z)} = \frac{1}{z}\mathbb{1}_{(0,z)}(x) \Rightarrow X|X+Y=z \sim \mathcal{U}([0,z])$$

14. Sean X e Y variables aleatorias independientes con distribución $N(0,1)$. Probar que su cociente $\dfrac{X}{Y}$ tiene distribución de Cauchy sin apelar al teorema de cambio de variables.

15. Dados $\alpha > 0$ y $\lambda > 0$ sea (X,Y) un vector aleatorio absolutamente continuo con densidad conjunta

$$f_{XY}(x,y) = \frac{\lambda}{\Gamma(\alpha)} \left(\frac{\lambda}{x}\right)^{\alpha} y^{\alpha-1} e^{-\lambda\left(x+\frac{y}{x}\right)} \mathbb{1}_{(0,+\infty)\times(0,+\infty)}(x,y).$$

    (i) Hallar la distribucion de X y la distribucið condicional de Y dada X. Indicar si son distribuciones conocidas.

(ii) Calcular $\mathbb{E}(Y|X)$ y $\mathbb{V}(Y|X)$. Deducir los valores de $\mathbb{E}(Y)$, $\mathbb{V}(Y)$ y $Cov(X,Y)$.

(iii) Probar que el cociente $\dfrac{Y}{Y}$ tiene distribución $\Gamma(\alpha,\lambda)$ y es independiente de X sin apelar al teorema de cambio de variables.

Solución:

(i) $f_X(x) = \dfrac{\lambda}{\Gamma(\alpha)}\left(\dfrac{\lambda}{x}\right)^{\alpha} e^{-\lambda x}\displaystyle\int_0^{\infty} y^{\alpha-1}e^{-\frac{y\lambda}{x}}\mathbb{1}_{(0,+\infty)}(x)dy$, usando $t=\dfrac{\lambda y}{x}$, entonces $dt = dy\dfrac{\lambda}{x}$

$$f_X(x) = \frac{\lambda}{\Gamma(\alpha)}\left(\frac{\lambda}{x}\right)^{\alpha} e^{-\lambda x}\mathbb{1}_{(0,+\infty)}(x)\underbrace{\int_0^{\infty}\left(\frac{x}{\lambda}\right)^{\alpha-1}\left(\frac{x}{\lambda}\right)t^{\alpha-1}e^{-t}dt}_{=\Gamma(\alpha)} = \lambda e^{-\lambda x}\mathbb{1}_{(0,+\infty)}(x) \Rightarrow X\sim\mathcal{E}(\lambda)$$

$$\begin{aligned} f_{Y|X=x}(y) &= \frac{f_{XY}(x,y)}{f_X(x)} = \frac{\dfrac{\lambda}{\Gamma(\alpha)}\left(\dfrac{\lambda}{x}\right)^{\alpha} y^{\alpha-1}e^{-\lambda\left(x+\dfrac{y}{x}\right)}\mathbb{1}_{(0,+\infty)\times(0,+\infty)}(x,y)}{\lambda e^{-\lambda x}\mathbb{1}_{(0,+\infty)}(x)} \\ &= \frac{1}{\Gamma(\alpha)}\left(\frac{\lambda}{x}\right)^{\alpha} y^{\alpha-1}e^{-\lambda\dfrac{y}{x}}\mathbb{1}_{(0,+\infty)}(y) \\ &\Rightarrow Y|X=x\sim\Gamma\left(\alpha,\frac{\lambda}{x}\right) \end{aligned}$$

(ii) $\mathbb{E}(Y|X=x) = \dfrac{\alpha x}{\lambda} \Rightarrow \mathbb{E}(Y|X) = \dfrac{\alpha X}{\lambda}$ y $\mathbb{V}(Y|X=x) = \alpha\left(\dfrac{x}{\lambda}\right)^2 = \dfrac{\alpha X^2}{\lambda^2}$.

$$\begin{aligned} \mathbb{E}(Y) &= \mathbb{E}[\mathbb{E}(Y|X)] = \frac{\alpha\mathbb{E}(X)}{\lambda} = \frac{\alpha}{\lambda^2} \\ \mathbb{V}(Y) &= \mathbb{E}[\mathbb{V}(Y|X)] + \mathbb{V}[\mathbb{E}(Y|X)] = \mathbb{E}\left(d\frac{\alpha X^2}{\lambda^2}\right) + \mathbb{V}\left(\frac{\alpha X}{\lambda}\right) \\ &= \frac{\alpha}{\lambda^2}\mathbb{E}(X^2) + \left(\frac{\alpha}{\lambda}\right)^2\mathbb{V}(X) \\ &= \frac{\alpha}{\lambda^2}\frac{2}{\lambda^2} + \left(\frac{\alpha}{\lambda}\right)^2\frac{1}{\lambda^2} = \frac{\alpha^2+2\alpha}{\lambda^4} \end{aligned}$$

$$\begin{aligned} \mathbb{E}(X\cdot Y) &= \int_0^{\infty}\int_0^{\infty} xy\frac{\lambda}{\Gamma(\alpha)}\left(\frac{\lambda}{x}\right)^{\alpha} y^{\alpha-1}e^{-\lambda\left(x+\dfrac{y}{x}\right)}dxdy \\ &= \int_0^{\infty}\frac{\lambda}{\Gamma(\alpha)}\left(\frac{\lambda}{x}\right)^{\alpha} xe^{-\lambda x}\left(\int_0^{\infty} y^{\alpha}e^{-\lambda\dfrac{y}{x}}dy\right)dx \\ &= \int_0^{\infty}\frac{\lambda}{\Gamma(\alpha)}\left(\frac{\lambda}{x}\right)^{\alpha} xe^{-\lambda x}\left(\frac{x}{\lambda}\right)^{\alpha+1}\Gamma(\alpha+1)dx \\ &= \int_0^{\infty}\alpha x^2 2e^{-\lambda x} = \frac{2\alpha}{\lambda^3} \end{aligned}$$

$$\Rightarrow Cov(X,Y) = \mathbb{E}(X\cdot Y) - \mathbb{E}(X)\mathbb{E}(Y) = \frac{2\alpha}{\lambda^3} - \frac{\alpha}{\lambda^3} = \frac{\alpha}{\lambda^3}$$

16. Un vector aleatorio (X,Y) tiene distribución normal bivariada si tiene función de densidad conjunta

$$f_{XY}(x,y) = \frac{e^{-\frac{1}{2(1-\rho)^2}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2+\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2-2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right)\right]}}{2\pi\sigma_Y\sigma_X\sqrt{1-\rho^2}}.$$

(i) Mostrar que la distribución condicional de X dada Y es $N\left[\mu_X+\rho\dfrac{\sigma_X}{\sigma_Y}(Y-\mu_Y), \sigma_X^2(1-\rho^2)\right]$.

(ii) Calcular $\mathbb{E}(Y|X)$ y $\mathbb{V}(Y|X)$.

(iii) Hallar el predictor óptimo lineal de X basado en Y y calcular el error de predicción.

Solución:

(i) si $z_x = \dfrac{x-\mu_X}{\sigma_X}$ y $z_y = \dfrac{y-\mu_Y}{\sigma_Y}$ entonces

$$\begin{aligned}
f_Y(y) &= \frac{e^{-\frac{z_y^2}{2(1-\rho)^2}}}{2\pi\sigma_Y\sigma_X\sqrt{1-\rho^2}}\int_{\mathbb{R}}\frac{1}{\sigma_X}e^{-\frac{z_x^2}{2(1-\rho)^2}+\frac{\rho z_x z_y}{(1-\rho)^2}}dz_x \\
&= \frac{e^{-\frac{z_y^2}{(1-\rho)^2}}}{\pi\sigma_Y}2\sqrt{\pi}\sqrt{2(1-\rho^2)}e^{\frac{(1-\rho^2)}{2}\left[\frac{z_y^2\rho^2}{2(1-\rho^2)}+\frac{2}{1-\rho^2}\right]} \\
&= \sqrt{\frac{2}{\pi\sigma_Y^2}}e^{-\frac{z_y^2-\rho^2 z_y^2-1}{2(1-\rho^2)}} = \sqrt{\frac{2}{\pi\sigma_Y^2}}e^{-\frac{z_y^2}{2}+\frac{1}{2(1-\rho^2)}}
\end{aligned}$$

$$\begin{aligned}
f_{X|Y=y}(x) &= \frac{e^{-\frac{1}{2(1-\rho)^2}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2+\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2-2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right)\right]}}{2\pi\sigma_Y\sigma_X\sqrt{1-\rho^2}}\frac{\sqrt{\pi\sigma_Y^2}}{\sqrt{2}e^{-\frac{z_y^2}{2}+\frac{1}{2(1-\rho^2)}}} \\
&= \frac{1}{2\sqrt{2\pi}\sigma_X\sqrt{1-\rho^2}}e^{-\frac{z_x^2}{2(1-\rho^2)}}e^{-\frac{\rho^2 z_y^2}{2(1-\rho^2)}}e^{\frac{\rho z_x z_y}{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}} \\
&= \frac{e^{-\frac{1}{2(1-\rho^2)}}}{2\sqrt{2\pi}\sigma_X\sqrt{1-\rho^2}}e^{-\frac{(z_x-\rho z_y)^2}{2(1-\rho^2)}} \\
&= \frac{e^{-\frac{1}{2(1-\rho^2)}}}{2\sqrt{2\pi}\sigma_X\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}\left(\frac{x-\mu_X-\rho\frac{\sigma_X}{\sigma_Y}(y-\mu_Y)}{\sigma_X}\right)^2} \\
\Rightarrow X|Y=y &\sim N\left[\mu_X+\rho\frac{\sigma_X}{\sigma_Y}(Y-\mu_Y),\sigma_X^2(1-\rho^2)\right]
\end{aligned}$$

(ii) $\mathbb{E}(X|Y) = \mu_X + \rho\dfrac{\sigma_X}{\sigma_Y}(Y-\mu_Y)$ y $\mathbb{V}(X|Y) = \sigma_X^2(1-\rho^2)$.

(iii) Sabemos que el predictor óptimo lineal de X basado en Y es de la pinta,

$$p_{lineal} = aY + b. \text{ con } Y = \mathbb{E}(X|Y) \text{ y } b = \mathbb{E}(X) - a\mathbb{E}(Y) \Rightarrow$$

$$p_{lineal} = \rho\frac{\sigma_X}{\sigma_Y}Y + \frac{\mu_X\sigma_Y - \mu_Y\sigma_X\rho}{\sigma_Y}$$

$$ECM = \mathbb{E}\left[(X-\mathbb{E}(X|Y)^2\right] = \mathbb{E}\left[\mathbb{E}\left[(X-\mathbb{E}(X|Y))^2|Y\right]\right] = \mathbb{E}[\mathbb{V}(X|Y)] = \sigma_X(1-\rho)^2$$

17. Sea (X,Y) un vector aleatorio bidimensional tal que $f_{X|Y=y}(x) = \dfrac{3x^2}{y^3}\mathbb{1}_{(0,y)}(x)$ y $f_Y(y) = 5y^4\mathbb{1}_{(0,1)}(y)$.

    (i) Calcular la función de densidad del cociente $\dfrac{X}{Y}$ y probar que es independiente de Y sin apelar al teorema de cambio de variables.

    (ii) Hallar la densidad condicional $f_{Y|X=x}$ para cada $x \in (0,1)$.

18. Sean X e Y variables aleatorias tales que $Y \sim \Gamma(r,\theta)$ con $r \in \mathbb{N}$ y $X|Y=y \sim \mathcal{P}(\lambda y)$ para todo $y > 0$.

    (i) Probar que $X + r \sim BN\left(r, \dfrac{\theta}{\lambda+\theta}\right)$.

    (ii) Probar que $Y|X=k \sim \Gamma(r+k,\lambda+\theta)$ para todo $k \in \mathbb{N}_0$ y calcular $\mathbb{E}\left(\dfrac{Y}{r+X}\right)$.

19. Sea (X,Y) un vector aleatorio a valores en $\{0,1,\dots,n\}\times(0,1)$ tal que para todo $k=0,\dots,n$ y $0 \le t \le 1$

$$\mathbb{P}(X=k, Y\le t) = \int_0^t \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)}\left(\begin{array}{c} n \\ k \end{array}\right) y^{p+k-1}(1-y)^{q+n-k-1}dy.$$

para ciertos $p,q > 0$.

   (i) Hallar la distribución condicional de Y dado X y calcular $\mathbb{E}\left(\dfrac{Y}{p+X}\right)$.

(ii) Hallar la distribución condicional de X dado Y y calcular $\mathbb{E}(X \cdot Y)$.
Sugerencia: Hallar también la distribución de Y y verificar si se trata de una distribución conocida.

Solución:

(i)

$$\begin{aligned}
\mathbb{P}(Y \leq t | X = k) &= \frac{\mathbb{P}(Y \leq t \text{ y } X = k)}{\mathbb{P}(X = k)} \\
&= \frac{\displaystyle\int_0^t \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} \binom{n}{k} y^{p+k-1}(1-y)^{q+n-k-1} dy}{\displaystyle\int_0^1 \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} \binom{n}{k} y^{p+k-1}(1-y)^{q+n-k-1} dy} \\
&= \frac{\displaystyle\int_0^t y^{p+k-1}(1-y)^{q+n-k-1} dy}{\displaystyle\frac{\Gamma(p+k)\Gamma(q+n-k)}{\Gamma(p+k+q+n-k)} \underbrace{\int_0^1 \frac{\Gamma(p+k+q+n-k)}{\Gamma(p+k)\Gamma(q+n-k)} y^{p+k-1}(1-y)^{q+n-k-1} dy}_{=1}} \\
&= \int_0^t \frac{\Gamma(p+k+q+n-k)}{\Gamma(p+k)\Gamma(q+n-k)} y^{p+k-1}(1-y)^{q+n-k-1} dy \\
&\Rightarrow Y | X = k \sim \beta(p+k, q+n-k).
\end{aligned}$$

$$\mathbb{E}\left(\frac{Y}{p+X}\right) = \mathbb{E}\left[\mathbb{E}\left(\frac{Y}{p+X}\middle| X\right)\right] = \mathbb{E}\left[\frac{1}{p+X}\mathbb{E}(X|Y)\right] = \mathbb{E}\left[\frac{1}{p+X}\frac{p+X}{p+X+q+n-X}\right] = \frac{1}{p+q+n}.$$

(ii)

$$\begin{aligned}
\mathbb{P}(Y \leq y) &= \sum_{k=0}^{n} \int_0^t \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} \binom{n}{k} y^{p+k-1}(1-y)^{q+n-k-1} dy \\
&= \int_0^t \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} y^{p-1}(1-y)^{q-1} \underbrace{\left[\sum_{k=0}^{n} \binom{n}{k} y^k (1-y)^{n-k}\right]}_{(y+1-y)^n} dy \\
&= \int_0^t \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} y^{p-1}(1-y)^{q-1} dy \\
&\Rightarrow Y \sim \beta(p, q).
\end{aligned}$$

$$\mathbb{P}(X = k | Y \leq y) = \frac{f_{XY}(X = k, Y \leq t)}{f_Y(y)} = \frac{\frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} \binom{n}{k} y^{p+k-1}(1-y)^{q+n-k-1}}{\frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} y^{p-1}(1-y)^{q-1}} = \binom{n}{k} y^k (1-y)^{n-k}$$

$$\Rightarrow X | Y = y \sim Bi(n, y).$$

$$\mathbb{E}(X \cdot Y) = \mathbb{E}[\mathbb{E}(X \cdot Y|Y)] = \mathbb{E}[\mathbb{E}(Y)\mathbb{E}(X|Y)] = \mathbb{E}\left(\frac{p}{p+q} nY\right) = \frac{np\mathbb{E}(Y)}{p+q} = n\left(\frac{p}{p+q}\right)^2.$$

20. Sea X una variable aleatoria absolutamente continua y $g : \mathbb{R} \to \mathbb{R}$ medible Borel tal que $\mathbb{E}(|g(X)|) < +\infty$.

(i) Utilizando argumentos intuitivos, ver que

$$\mathbb{E}(g(X)|X^2 = t) = \frac{g(\sqrt{t}) f_X(\sqrt{t}) + g(-\sqrt{t}) f_X(-\sqrt{t})}{f_X(\sqrt{t}) + f_X(-\sqrt{t})}$$

(ii) Utilizando la definición de esperanza condicional, probar que

$$\mathbb{E}(g(X)|X^2) = \frac{g(\sqrt{X^2}) f_X(\sqrt{X^2}) + g(-\sqrt{X^2}) f_X(-\sqrt{X^2})}{f_X(\sqrt{X^2}) + f_X(-\sqrt{X^2})}$$

(iii) Observar que si $X \in L^2$ vale que $\mathbb{E}(X^2|X) = X^2$ mientras que $\mathbb{E}(X|X^2) \neq X$. ¿Es razonable esto?

21. Sean X e Y variables aleatorias tales que $Y \sim \mathcal{U}([2, 3])$ y para cada $y \in [2, 3]$ se verifica

x	-1	0	1
$\mathbb{P}_{X\|Y=y}(x)$	$\dfrac{3-y}{2}$	$y-2$	$\dfrac{3-y}{2}$

(i) Hallar $\mathbb{P}_X$.

(ii) Calcular la función de distribución $F_{Y|X=x}(y)$ para cada $x \in R_X$. ¿Es razonable esto?

Solución:

(i)

$$\mathbb{P}_X(-1) = \int_2^3 \frac{3-y}{2} dy = \frac{1}{4}, \qquad \mathbb{P}_X(0) = \int_2^3 (y-2)dy = \frac{1}{2} \quad \text{y} \quad \mathbb{P}_X(1) = \int_2^3 \frac{3-y}{2} dy = \frac{1}{4}$$

(ii)

$$\begin{aligned}
\mathbb{P}(Y=y|X=-1) &= \frac{\mathbb{P}(Y=y \text{ y } X=-1)}{\mathbb{P}(X=-1)} = \frac{\mathbb{P}(X=-1|Y=y)f_Y(y)}{\mathbb{P}(X=-1)} = 4\frac{(3-y)}{2}\mathbb{1}_{[2,3]}(y) \\
\mathbb{P}(Y=y|X=0) &= \frac{\mathbb{P}(Y=y \text{ y } X=0)}{\mathbb{P}(X=0)} = \frac{\mathbb{P}(X=0|Y=y)f_Y(y)}{\mathbb{P}(X=0)} = 2(y-2)\mathbb{1}_{[2,3]}(y) \\
\mathbb{P}(Y=y|X=1) &= \mathbb{P}(Y=y|X=-1)
\end{aligned}$$

22. Sean X e Y variables aleatorias tales que X es discreta con probabilidad puntual

x	1	2
$\mathbb{P}_X(x)$	$\dfrac{1}{4}$	$\dfrac{3}{4}$

y la distribución condicional de Y dada X es $\mathcal{E}(X)$.

(i) Calcular $F_Y(y)$ y $f_Y(y)$.

(ii) Calcular $\mathbb{P}(X=x, Y \leq y)$ para $y > 0$ y $x \in R_X$.

(iii) Para cada $x \in R_X$ sea $g_x : \mathbb{R} \to \mathbb{R}$ una aplicación medible Borel tal que $\mathbb{P}(X=x|Y) = g_x(Y)$. Para cada $x \in R_X$ expresar $\mathbb{P}(X=x, Y \leq y)$ en términos de g_x.

(iv) Para cada $x \in R_X$ calcular explícitamente $\mathbb{P}(X=x|Y)$.

Solución:

(i)

$$\begin{aligned}
F_Y(y) &= \mathbb{P}(Y \leq y) = \mathbb{P}(Y \leq y|X=1)\mathbb{P}(X=1) + \mathbb{P}(Y \leq y) = \mathbb{P}(Y \leq y|X=2)\mathbb{P}(X=2) \\
&= \frac{1}{4}\int_0^y e^{-t}dt + \frac{3}{4}\int_0^y 2e^{-2t}dt = \frac{1-e^{-y}}{4} + \frac{3(1-e^{-2y})}{4}, \qquad \text{si } y \in (0,+\infty) \\
\Rightarrow f_Y(y) &= \left(\frac{e^{-y}}{4} + \frac{3e^{-2y}}{2}\right)\mathbb{1}_{(0,+\infty)(y)}
\end{aligned}$$

(ii) $\mathbb{P}(X=x|Y \leq y) = \dfrac{\mathbb{P}(X=x\text{e } Y \leq y)}{\mathbb{P}(Y \leq y)} = \dfrac{\mathbb{P}(Y \leq y|X=x)\mathbb{P}(X=x)}{\mathbb{P}(Y \leq y)}$, entonces

$$\mathbb{P}(X=x|Y \leq y) = \frac{\displaystyle\int_0^y xe^{-xy}dy\mathbb{P}(X=x)}{\dfrac{1-e^{-y}}{4} + \dfrac{3(1-e^{-2y})}{4}} = \frac{4\left(1-\dfrac{e^{-xy}}{x}\right)\mathbb{P}(X=x)}{4-3e^{-2y}-e^{-y}}$$

(iii) $\mathbb{P}(X=x|Y \leq y) = \displaystyle\int_0^y g_x(t)f_Y(t)dt.$

(iv)

$$\mathbb{P}(X=1|Y=y)=\frac{\mathbb{P}(\mathbb{P}(Y=y|X=1)\mathbb{P}(X=1))}{f_Y(y)}=\frac{e^{-y}}{e^{-y}+6e^{-2y}}\mathbb{1}_{(0,+\infty)}(y)$$
$$\mathbb{P}(X=2|Y=y)=\frac{\mathbb{P}(\mathbb{P}(Y=y|X=2)\mathbb{P}(X=2))}{f_Y(y)}=\frac{6e^{-2y}}{e^{-y}+6e^{-2y}}\mathbb{1}_{(0,+\infty)}(y)$$

23. Sea $(X_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias independientes e idénticamente distribuidas con segundo momento finito.

 (i) Hallar $\mathbb{E}(X_1|X_1+\ldots+X_n)$ para cada $n\in\mathbb{N}$.
 Sugerencia: Probar que $\mathbb{E}(X_1|X_1+\ldots+X_n)=\mathbb{E}(X_k|X_1+\ldots+X_n)$ para todo $k=2,\ldots,n$.

 (ii) Probar que $\mathbb{E}(X_1|X_1+\ldots+X_n)\xrightarrow[n\to\infty]{cs}\mathbb{E}(X_1)$.

Solución:

(i) Como $\mathbb{E}(X_1|X_1+\ldots+X_n)=\mathbb{E}(X_k|X_1+\ldots+X_n)$ para todo $k=2,\ldots,n$, entonces

$$\mathbb{E}(X_1+\ldots+X_n|X_1+\ldots+X_n)=\mathbb{E}(X_1|X_1+\ldots+X_n)+\ldots+\mathbb{E}(X_n|X_1+\ldots+X_n)=n\mathbb{E}(X_1|X_1+\ldots+X_n)$$

$$\mathbb{E}(X_1|X_1+\ldots+X_n=y)=\frac{\mathbb{E}(X_1+\ldots+X_n|X_1+\ldots+X_n=y)}{n}=\frac{y}{n}\Rightarrow\mathbb{E}(X_1|X_1+\ldots+X_n)=\frac{X_1+\ldots+X_n}{n}$$

(ii) Puesto que $(X_n)_{n\in\mathbb{N}}$ es una sucesión de variables aleatorias iid con segundo momento finito, por la Ley fuerte de los grandes números $\mathbb{E}(X_1|X_1+\ldots+X_n)=\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}X_n\xrightarrow[n\to\infty]{cs}\mathbb{E}(X_1)$.

Práctica 10

Funciones características y Teorema del Límite Central

1. Hallar la función característica de las siguientes distribuciones:

   (i) $\mathbb{P}(\lambda)$.

   (ii) $Bi(n,p)$.

   (iii) $\mathcal{U}(a,b)$.

   (iv) $\mathcal{E}(\lambda)$.

   (v) $\Gamma(n,\lambda)$

Solución:

(i) $\phi_X(t) = \mathbb{E}(e^{itX}) = \displaystyle\sum_{k\in\mathbb{N}_0} e^{itk}\frac{\lambda^k}{k!}e^{-\lambda} = e^{-\lambda}\sum_{k\in\mathbb{N}_0}\frac{\left(\lambda e^{it}\right)^k}{k!} = e^{-\lambda}e^{\lambda[\cos(t)+i\sin(t)]} = e^{\lambda\left(e^{it}-1\right)}$.

(ii) Sea $X_i \sim Be(p)$, entonces $\phi_{X_i}(t) = pe^{it} + (1-p)$. Como $X = X_1 + \ldots + X_n$, con $X_i \sim Be(p)$ independientes para todo $1 \leq i \leq n$ se tiene que

$$\phi_X(t) = [\phi_{X_i}(t)]^n = \left(pe^{it} + 1 - p\right)^n$$

(iii) $\phi_X(t) = \displaystyle\int_{\mathbb{R}} e^{tx} f_X(x)dx = \frac{1}{b-a}\int_a^b e^{tx}dx = \begin{cases} \dfrac{e^{tb}-e^{ta}}{t(b-a)} & t \neq 0 \\ 1 & t = 0 \end{cases}$.

(iv) $\phi_X(t) = \displaystyle\int_{\mathbb{R}} e^{tx} f_X(x)dx = \lambda\int_0^{\infty} e^{x(it-\lambda)}dx = \frac{\lambda}{it-\lambda}$.

(v) Sea $X \sim \Gamma(n,\lambda)$, entonces existen $X_i \sim \mathcal{E}(\lambda)$ independientes para todo $1 \leq i \leq n$ tal que $X = X_1 + \ldots + X_n$, entonces

$$\phi_X(t) = [\phi_{X_i}(t)]^n = \left(\frac{\lambda}{it-\lambda}\right)^n$$

2. Sean X e Y variables aleatorias independientes. Compruebe utilizando funciones características que

   (i) $X \sim \mathcal{P}(\lambda_1), Y \sim \mathcal{P}(\lambda_2) \Rightarrow X + Y \sim \mathbb{P}(\lambda_1 + \lambda_2)$.

   (ii) $X \sim Bi(n,p), Y \sim Bi(m,p) \Rightarrow X + Y \sim Bi(n+m,p)$.

Solución:

(i) $\phi_{X+Y}(t) = \phi_X(t)\phi_Y(t) = e^{\lambda_1(e^{it}-1)}e^{\lambda_2(e^{it}-1)} = e^{(e^{it}-1)(\lambda_1+\lambda_2)}$, entonces $X + Y \sim \mathbb{P}(\lambda_1+\lambda_2)$.

(ii) $\phi_{X+Y}(t) = \phi_X(t)\phi_Y(t) = \left(pe^{it}+1-p\right)^n\left(pe^{it}+1-p\right)^m = \left(pe^{it}+1-p\right)^{n+m}$, entonces $X+Y \sim Bi(n+m,p)$.

3. Probar que una variable aleatoria X es simétrica respecto del origen si y sólo si $\phi_X(t) \in \mathbb{R}$ para todo t.

Solución:

$(\Rightarrow)$:

$$\mathbb{E}(e^{itX}) = \int_{\mathbb{R}} e^{itx}f_X(x)dx = \int_{\mathbb{R}} e^{itx}f_X(-x)dx = \int_{\mathbb{R}} e^{it(-x)}\underbrace{f_X(x)}_{=f_X(-x)}dx = \mathbb{E}(e^{\overline{itX}}) = \overline{\mathbb{E}(e^{itX})} \Rightarrow \phi_X(t) \in \mathbb{R}\ \forall t \in \mathbb{R}$$

$(\Leftarrow):$ Suponemos $\phi_X(t) \in \mathbb{R}$ para todo t, entonces $\mathbb{E}(e^{itX}) = \mathbb{E}(e^{-itX})$ luego

$$\int_{\mathbb{R}} \sin(tx) f_X(x) dx = 0 \ \forall t \in \mathbb{R}$$

$$\int_{\mathbb{R}} \cos(tx) f_X(x) dx \underbrace{=}_{\forall t \in \mathbb{R}} \int_{\mathbb{R}} \cos(-tx) f_X(x) dx \underbrace{=}_{\forall t \in \mathbb{R}} \int_{\mathbb{R}} \cos(tx) f_X(-x) dx$$

$$\Rightarrow \int_{\mathbb{R}} \cos(tx)[f_X(x) - f_X(-x)] dx = 0 \ \forall t \in \mathbb{R} \Rightarrow f_X(x) = f_X(-x) \ \forall t \in \mathbb{R}$$

4. (i) Calcular la función característica asociada a una variable aleatoria con distribución Cauchy $C(0,1)$. *Sugerencia:* Utilizar la igualdad
$$\int_{-\infty}^{+\infty} \frac{\cos(xt)}{1+x^2} dx = \pi e^{-|t|}$$
   (ii) Verificar que si X es una variable aleatoria con distribución Cauchy $C(0,1)$ entonces $\phi_{2X} = \phi_X^2$. Observar que esto dice que ϕ_{X+Y} puede coincidir con $\phi_X \cdot \phi_Y$ aunque X e Y no sean independientes.
   (iii) Sea $(X_i)_{i\in\mathbb{N}}$ una sucesión de variables aleatorias independientes e idénticamente distribuidas con distribución Cauchy $C(0,1)$. Hallar para cada $n \in \mathbb{N}$ la distribución del promedio $X_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. ¿Contradicen sus resultados algún teorema conocido?

Solución:

(i) Como $X \sim C(0,1)$ es una variable aleatoria simétrica respecto del origen, entonces

$$\phi_X(t) = \int_{\mathbb{R}} \frac{e^{itx}}{\pi(1+x^2)} dx \underbrace{=}_{\phi_X(t)\in\mathbb{R}} \int_{\mathbb{R}} \frac{\cos(tx)}{\pi(1+x^2)} dx = e^{-|t|}$$

(ii) $\phi_{2X}(t) = \phi_X(2t) = e^{-2|t|} = \left(e^{-|t|}\right)^2 = [\phi_X(t)]^2$.

(iii) $\phi_{X_n}(t) = \left[\phi_{\frac{X_1}{n}}(t)\right]^n = \left[\phi_{X_1}\left(\frac{t}{n}\right)\right]^n = \left(e^{-\frac{|t|}{n}}\right)^n = e^{-|t|} = \phi_{X_1}(t)$.

5. Sea X una variable aleatoria con media cero y varianza finita tal que para cierta variable aleatoria $Y \sim X$ independiente de X se verifica que $X+Y$ y $X-Y$ son independientes.
   (i) Probar que $\phi_X(t) = \left[\phi_X\left(\frac{t}{2}\right)\right]^3 \phi_X\left(-\frac{t}{2}\right)$ para todo $t \in \mathbb{R}$. Deducir que si $\phi_X(t) \neq 0$ entonces
$$\frac{\phi_X(t)}{\phi_X(-t)} = \left[\frac{\phi_X\left(\frac{t}{2}\right)}{\phi_X\left(-\frac{t}{2}\right)}\right]^2 = \ldots = \left[\frac{\phi_X\left(\frac{t}{2^n}\right)}{\phi_X\left(-\frac{t}{2^n}\right)}\right]^{2^n}$$
   para todo $n \in \mathbb{N}$.
   (ii) Probar que $\phi_X(t) = \phi_X(-t)$ para todo $t \in \mathbb{R}$. Deducir que $\phi_X(t) = \left[\phi_X\left(\frac{t}{2}\right)\right]^4$ para todo $t \in \mathbb{R}$.
   Sugerencia: Estudiar el cociente $\frac{\phi_X(t)}{\phi_X(-t)}$ utilizando el desarrollo de Taylor de orden 2 de ϕ_X en $t_0 = 0$.
   (iii) Concluir que $X \sim N(0, \sigma^2)$.
   (iv) Probar que el resultado sigue valiendo aún si la media de X es distinta de cero.

Solución:

(i) Sean $Z = X+Y$ y $W = X-Y$, tal que

$$\phi_Z(t) = \phi_X(t)\phi_Y(t) \underbrace{=}_{Y\sim X} [\phi_X(t)]^2 \quad \text{y} \quad \phi_W(t) = \phi_X(t)\phi_{-Y}(t) = \phi_X(t)\phi_Y(-t) \underbrace{=}_{Y\sim X} \phi_X(t)\phi_X(-t)$$

$$\phi_X(t) = \phi_{\frac{Z}{2}}(t)\phi_{\frac{W}{2}}(t) = \phi_Z\left(\frac{t}{2}\right)\phi_W\left(\frac{t}{2}\right) = \left[\phi_X\left(\frac{t}{2}\right)\right]^3 \phi_X\left(-\frac{t}{2}\right)$$

Luego, si $\phi_Z(t) \neq 0$

$$\frac{\phi_X(t)}{\phi_X(-t)} = \frac{\left[\phi_X\left(\frac{t}{2}\right)\right]^3 \phi_X\left(-\frac{t}{2}\right)}{\left[\phi_X\left(-\frac{t}{2}\right)\right]^3 \phi_X\left(\frac{t}{2}\right)} = \left[\frac{\phi_X\left(\frac{t}{2}\right)}{\phi_X\left(-\frac{t}{2}\right)}\right]^2$$

como vale para todo $t \in \mathbb{R}$, entonces, en particular $\frac{\phi_X\left(\frac{t}{2}\right)}{\phi_X\left(-\frac{t}{2}\right)} = \left[\frac{\phi_X\left(\frac{t}{4}\right)}{\phi_X\left(-\frac{t}{4}\right)}\right]^2 \Rightarrow \frac{\phi_X(t)}{\phi_X(-t)} = \left[\left(\frac{\phi_X\left(\frac{t}{2^2}\right)}{\phi_X\left(-\frac{t}{2^2}\right)}\right)^2\right]^2$.

De la misma manera, como vale para todo $t \in \mathbb{R}$,en particular

$$\frac{\phi_X\left(\frac{t}{4}\right)}{\phi_X\left(-\frac{t}{4}\right)} = \left[\frac{\phi_X\left(\frac{t}{8}\right)}{\phi_X\left(-\frac{t}{8}\right)}\right]^2 \Rightarrow \frac{\phi_X(t)}{\phi_X(-t)} = \left[\left[\left(\frac{\phi_X\left(\frac{t}{2^3}\right)}{\phi_X\left(-\frac{t}{2^3}\right)}\right)^2\right]^2\right]^2.$$

Por lo tanto, de manera inductiva, se tiene que para todo $t \in \mathbb{R}$

$$\frac{\phi_X(t)}{\phi_X(-t)} = \left[\frac{\phi_X\left(\frac{t}{2}\right)}{\phi_X\left(-\frac{t}{2}\right)}\right]^2 = \ldots = \left[\frac{\phi_X\left(\frac{t}{2^n}\right)}{\phi_X\left(-\frac{t}{2^n}\right)}\right]^{2^n}$$

(ii) Como $\mu_X = 0$, entonces f_X es simétrica respecto al origen y, por lo tanto, $\phi_X(t) = \phi_X(-t)$ para todo $t \in \mathbb{R}$. Luego, para todo $t \in \mathbb{R}$

$$\phi_X(t) = \left[\phi_X\left(\frac{t}{2}\right)\right]^3 \phi_X\left(-\frac{t}{2}\right) = \left[\phi_X\left(\frac{t}{2}\right)\right]^4$$

6. (i) Sea $(Z_i)_{i\in\mathbb{N}}$ una sucesión de variables aleatorias independientes con distribución $N(\mu, \sigma^2)$. Verificar que para todo $n \in \mathbb{N}$ se tiene que $\sqrt{n}(\overline{Z}_n - \mu) \sim N(0, \sigma^2)$.
   (ii) Sea $(X_i)_{i\in\mathbb{N}}$ una sucesión de variables aleatorias independientes idénticamente distribuidas con media μ y varianza σ^2 tal que para algún $n_0 \in \mathbb{N}_{\geq 2}$ se verifica que $\sqrt{n_0}\left(\overline{X}_{n_0} - \mu\right) \sim X_1 - \mu$. Probar que $X_1 \sim N(\mu, \sigma^2)$.

7. Probar que no existen variables aleatorias X y Z independientes e idénticamente distribuidas tales que

$$X - Z \sim \mathcal{U}([-1, 1]).$$

Solución:

$$\phi_{X-Z}(t) = \phi_X(t)\phi_{-X}(t) = \phi_X(t)\phi_X(-t) = \begin{cases} \dfrac{\sin(t)}{t} & t \neq 0 \\ 1 & t = 0 \end{cases}$$

y, por otra parte,

$$\begin{aligned}
\phi_X(t)\phi_X(-t) &= \mathbb{E}(e^{itX})\mathbb{E}(e^{-itX}) \\
&= \left[\int_{\mathbb{R}} \cos(xt) f_X(x)dx + i\int_{\mathbb{R}} \sin(xt) f_X(x)dx\right]\left[\int_{\mathbb{R}} \cos(xt) f_X(x)dx - i\int_{\mathbb{R}} \sin(xt) f_X(x)dx\right] \\
&= \left[\int_{\mathbb{R}} \cos(xt) f_X(x)dx\right]^2 + \left[\int_{\mathbb{R}} \sin(xt) f_X(x)dx\right]^2
\end{aligned}$$

Luego, por un lado, para todo $t \in \mathbb{R}$, se tiene que $\phi_X(t)\phi_X(-t) \geq 0$, pero para la función $\frac{\sin(t)}{t}$ no es siempre positiva para todo $t \in \mathbb{R}$.

8. Probar mediante el estudio de funciones características que la distribución binomial de parámetros n y p converge en distribución cuando $n \to +\infty$ y $np \to \lambda > 0$ a la distribución de Poisson de parámetro λ.

Solución:

Sean X e Y variables aleatorias tales que $X \sim Bi(n, p)$ e $Y \sim \mathcal{P}(\lambda)$, luego

$$\lim_{\substack{n\to\infty \\ np\to\lambda}} \phi_X(t) = \lim_{\substack{n\to\infty \\ np\to\lambda}} \left(pe^{it} + 1 - p\right)^n = \lim_{\substack{n\to\infty \\ np\to\lambda}} \left(1 + np\frac{e^{it} - 1}{n}\right)^n = e^{\lambda(e^{it}-1)} = \phi_Y(t)$$

$$\Rightarrow \phi_X(t) \xrightarrow[\substack{n\to\infty \\ np\to\lambda}]{} \phi_X(t)\ \forall t \in \mathbb{R} \Rightarrow X \xrightarrow[\substack{n\to\infty \\ np\to\lambda}]{D} Y.$$

9. Probar mediante el estudio de funciones características la ley débil de los grandes números para variables aleatorias i.i.d. con esperanza finita sin asumir la finitud del segundo momento.

Solución:

Sean $X_1, \ldots, x_2$ variables aleatorias i.i.d tales que $\mathbb{E}(X_i) = \mu < +\infty$ para todo $1 \leq i \leq n$,

$$\begin{aligned}
\lim_{n\to\infty} \phi_{\overline{X}_n}(t) &= \lim_{n\to\infty} \left[\phi_X\left(\frac{t}{n}\right)\right]^n \\
&= \lim_{n\to\infty} \left[1 + \phi'_X(0)\frac{t}{n} + o\left(\frac{t}{n}\right)\right]^n \\
&= \lim_{n\to\infty} \left[\left(1 + i\mu\frac{t}{n}\right)^{\frac{n}{i\mu t}}\right]^{i\mu t} = e^{i\mu t}
\end{aligned}$$

Como $e^{i\mu t} = \phi_\mu(t)$, se tiene que $\overline{X}_n \xrightarrow[n\to\infty]{D} \mu$ y, por lo tanto, $\overline{X}_n \xrightarrow[n\to\infty]{P} \mu$.

10. Sean $(Y_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias independientes con distribución Cauchy $C(0,1)$ y $(c_n)_{n\in\mathbb{N}}$ una sucesión de números reales. Consideremos la sucesión $(X_n)_{n\in\mathbb{N}}$ de variables aleatorias definidas para cada $n \in \mathbb{N}$ como $X_n = c_n Y_n$. Probar que

   (i) Si $\sum_{n\in\mathbb{N}} |c_n| = c < +\infty$, entonces $\sum_{k=1}^{n} X_k \xrightarrow[n\to\infty]{D} cX$ con $X \sim C(0,1)$.

   (ii) Si $\sum_{n\in\mathbb{N}} |c_n| = +\infty$, entonces no existe ninguna variable aleatoria Z tal que $\sum_{k=1}^{n} X_k \xrightarrow[n\to\infty]{D} Z$.

   (iii) Si $c_n = \frac{1}{n}$ entonces $\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow[n\to\infty]{P} 0$.

   Sugerencia: Puede resultarle útil la desigualdad $\sum_{i=1}^{n} \frac{1}{i} \leq \log(n) + 1$ para todo $n \in \mathbb{N}$.

Solución:

Recordar que si $X \sim C(0,1) \Rightarrow \phi_{cX}(t) = e^{-|ct|}$.

(i) $\lim_{n\to\infty} \phi_{X_1+\ldots+X_n}(t) = \lim_{n\to\infty} \prod_{i=1}^{n} \phi_{X_i}(t) = \lim_{n\to\infty} \prod_{i=1}^{n} e^{-|c_i t|} = \lim_{n\to\infty} exp\left\{-|t|\sum_{i=1}^{n}|c_i|\right\} \xrightarrow[n\to\infty]{} e^{-|t|c} = \phi_X(t).$

Luego, $\sum_{k=1}^{n} X_k \xrightarrow[n\to\infty]{D} X$.

(ii) $\lim_{n\to\infty} \phi_{X_1+\ldots+X_n}(t) = \lim_{n\to\infty} \prod_{i=1}^{n} \phi_{X_i}(t) = \lim_{n\to\infty} \prod_{i=1}^{n} e^{-|c_i t|} = \lim_{n\to\infty} exp\left\{-|t|\sum_{i=1}^{n}|c_i|\right\} \xrightarrow[n\to\infty]{} 0$ pues $\sum_{n\in\mathbb{N}} |c_n| = +\infty$.

Entonces, $\sum_{k=1}^{n} X_k$ no converge en distribución a ninguna variable aleatoria.

11. (i) Sea $(X_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias independientes e idénticamente distribuidas con distribución dada por $\mathbb{P}(X=1) = \mathbb{P}(X=-1) = \frac{1}{2}$ y para cada $n \in \mathbb{N}$ definamos la variable aleatoria

$$Y_n = \sum_{k=1}^{n} \frac{X_k}{2^k}.$$

   Probar que $Y_n \xrightarrow[n\to\infty]{D} \mathcal{U}([-1,1])$.

   Sugerencia: Utilizar la identidad $\sin(2\theta) = 2\cos(\theta)\sin(\theta)$.

   (ii) Sea $(W_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias independientes e idénticamente distribuidas con distribución Bernoulli de parámetro $\frac{1}{2}$ y para cada $n \in \mathbb{N}$ definamos la variable aleatoria

$$Z_n = \sum_{k=1}^{n} \frac{W_k}{2^k}.$$

   Probar que $Z_n \xrightarrow[n\to\infty]{D} \mathcal{U}([0,1])$. Interpretar el resultado obtenido.

Solución:

(i) Notemos que

- $\phi_X(t) = \mathbb{E}(e^{itX}) = \frac{e^{it}}{2} + \frac{e^{-it}}{2} = \cos(t)$.
- Usando la sugerencia, se llega a que $\cos\left(\frac{\theta}{2^k}\right) = \frac{1}{2}\frac{\sin\left(\frac{\theta}{2^{k-1}}\right)}{\sin\left(\frac{\theta}{2^k}\right)}$.

Luego,

$$\phi_{Y_n} = \prod_{k=1}^{n} \phi_{X_k}\left(\frac{t}{2^k}\right) = \prod_{k=1}^{n} \cos\left(\frac{t}{2^k}\right) = \frac{1}{2^n}\prod_{k=1}^{n} \frac{\sin\left(\frac{t}{2^{k-1}}\right)}{\sin\left(\frac{t}{2^k}\right)} = \frac{1}{2^n}\frac{\sin(t)}{\sin\left(\frac{t}{2^n}\right)}$$

$$\Rightarrow \lim_{n\to\infty} \phi_{Y_n} = \lim_{n\to\infty} \frac{1}{2^n}\frac{\sin(t)}{\sin\left(\frac{t}{2^n}\right)} = \frac{\sin(t)}{t}\underbrace{\lim_{n\to\infty}\frac{t}{2^n\sin\left(\frac{\theta}{2^n}\right)}}_{=1} = \frac{\sin(t)}{t} = \phi_Y(t) \text{ con } Y \sim \mathcal{U}([-1,1]).$$

(ii) Notar que $W_i = \frac{X_i}{2} + \frac{1}{2}$, con X_i definida en el ítem anterior. Luego,

- $\phi_{W_i}(t) = e^{\frac{it}{2}}\phi_{X_i}\left(\frac{t}{2}\right)$.
- $Z_n = \sum_{k=1}^{n}\frac{1}{2^k}\left(\frac{X_k}{2}+\frac{1}{2}\right) = \frac{1}{2}Y_n + \sum_{k=1}^{n}\frac{1}{2^{k+1}} = \frac{1}{2}Y_n + \sum_{k=0}^{n}\frac{1}{2^k} - 1 - \frac{1}{2} = \frac{1}{2}Y_n + \frac{1-(1/2)^{n+1}}{1-(1/2)} - \frac{3}{2} = \frac{Y_n - 1}{2} - \frac{1}{2^n}$.

Entonces
$$\lim_{n\to\infty}\phi_{Z_n}(t) = \lim_{n\to\infty} e^{-\frac{it}{2^n}}\phi_{\frac{Y_n-1}{2}}(t) = \lim_{n\to\infty}\phi_{\frac{Y_n-1}{2}}(t)$$

Por otra parte, sea g una función continua definida por $g(y) = \frac{y-1}{2}$ y como $\lim_{n\to\infty}\phi_{Y_n}(t) = \phi_Y(t)$ con $Y \sim \mathcal{U}([-1,1])$, se tiene que $\lim_{n\to\infty}\phi_{g(Y_n)}(t) = \phi_{g(Y)}(t)$.

Por lo tanto, $\lim_{n\to\infty}\phi_{Z_n}(t) = \lim_{n\to\infty}\phi_Z(t)$ con $Z = g(Y)$. Entonces, si $Y \sim \mathcal{U}([-1,1])$ se tiene que $Z = \frac{Y-1}{2} \sim \mathcal{U}([0,1]) \Rightarrow Z_n \xrightarrow[n\to\infty]{D} Z \sim \mathcal{U}([0,1])$.

12. Sea $(X_n, Y_n)_{n\in\mathbb{N}}$ una sucesión de vectores aleatorios tal que para todo n las variables aleatorias X_n e Y_n son independientes y vale que $X_n \xrightarrow[n\to\infty]{D} X$ e $Y_n \xrightarrow[n\to\infty]{D} Y$. Probar que $X_n + Y_n \xrightarrow[n\to\infty]{D} X' + Y'$, donde X' e Y' son independientes y tales que $X' \sim X$ e $Y' \sim Y$.

Solución:

Como $X_n \xrightarrow[n\to\infty]{D} X$ e $Y_n \xrightarrow[n\to\infty]{D} Y$, entonces $\lim_{n\to\infty}\phi_{X_n}(t) = \phi_X(t)$ y $\lim_{n\to\infty}\phi_{Y_n}(t) = \phi_Y(t)$. Luego,

$$\lim_{n\to\infty}\phi_{X_n+Y_n}(t) \underbrace{=}_{\substack{X_n \text{ e } Y_n \\ independientes}} \lim_{n\to\infty}\phi_{X_n}(t)\phi_{Y_n}(t) = \phi_X(t)\phi_Y(t)$$

Si $X_n + Y_n \xrightarrow[n\to\infty]{D} X' + Y'$ con X' e Y' independientes, entonces

$$\lim_{n\to\infty}\phi_{X_n}(t)\phi_{Y_n}(t) = \lim_{n\to\infty}\phi_{X_n+Y_n}(t) = \phi_{X'+Y'}(t) = \phi_{X'}(t)\phi_{Y'}(t)$$

luego,

$$\phi_{X'+Y'}(t) = \phi_{X'}(t)\phi_{Y'}(t) = \phi_X(t)\phi_Y(t) \underbrace{=}_{\substack{si\ X \text{ e } Y \\ independientes}} \phi_{X+Y}(t) \Rightarrow \phi_{X'+Y'}(t) = \phi_{X+Y}(t)$$

entonces, $F_{X'+Y'} = F_{X+Y}$.

Por lo tanto, basta con usar $X' \sim X$ e $Y' \sim Y$ tales que $F_{X'} = F_X$ y $F_{Y'} = F_Y$.

13. Sean $(a_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ una sucesión que satisface $\lim\limits_{n\to+\infty} a_n = +\infty$ y $(X_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias tal que $a_n(X_n - X) \xrightarrow[n\to\infty]{D} Z$ para ciertas variables aleatorias X y Z.

    (i) Probar que $X_n \xrightarrow[n\to\infty]{P} X$.

    (ii) Sea $(X_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias i.i.d. con media μ y varianza σ^2. Mostrar que

$$n^\alpha\left(\overline{X}_n - \mu\right) \xrightarrow[n\to\infty]{P} \begin{cases} 0 & 0 \le \alpha < \dfrac{1}{2} \\ \infty & \alpha > \dfrac{1}{2} \end{cases}$$

donde $n^\alpha\left(\overline{X}_n - \mu\right) \xrightarrow[n\to\infty]{P} \infty$ significa que para todo $M > 0$ se tiene $\lim\limits_{n\to\infty} \mathbb{P}\left(\left|n^\alpha\left(\overline{X}_n - \mu\right)\right| \le M\right) = 0$

¿Qué sucede si $\alpha = \dfrac{1}{2}$?

Solución:

(i) $a_n(X_n - X) \xrightarrow[n\to\infty]{D} Z \Rightarrow (X_n - X) \xrightarrow[n\to\infty]{D} \dfrac{Z}{a_n} \xrightarrow[n\to\infty]{} 0 \Rightarrow (X_n - X) \xrightarrow[n\to\infty]{D} 0 \Rightarrow (X_n - X) \xrightarrow[n\to\infty]{P} 0 \Rightarrow X_n \xrightarrow[n\to\infty]{P} X.$

(ii) Por la Ley Fuerte de los Grandes Números, $\overline{X}_n \xrightarrow[n\to\infty]{cd} \mu$, entonces $(\overline{X}_n - \mu) \xrightarrow[n\to\infty]{D} 0$ y por el Teorema Centra del Límite $\sqrt{n}\dfrac{(\overline{X}_n - \mu)}{\sigma} \xrightarrow[n\to\infty]{D} N(0,1)$. Entonces,

$$\lim_{n\to\infty} n^\alpha(\overline{X}_n - \mu) = \lim_{n\to\infty} \frac{n^\alpha}{\sqrt{n}}\sigma\left[\sqrt{n}\left(\frac{(\overline{X}_n - \mu)}{\sigma}\right)\right] = \begin{cases} \lim_{n\to\infty} \underbrace{n^{\alpha-\frac{1}{2}}\sigma}_{\xrightarrow[\alpha<\frac{1}{2}]{}0} \underbrace{\left[\sqrt{n}\left(\frac{(\overline{X}_n - \mu)}{\sigma}\right)\right]}_{\xrightarrow{D} N(0,1)} = 0 & 0 \le \alpha < \frac{1}{2} \\ \lim_{n\to\infty} \underbrace{n^{\alpha-\frac{1}{2}}\sigma}_{\xrightarrow[\alpha>\frac{1}{2}]{}+\infty} \underbrace{\left[\sqrt{n}\left(\frac{(\overline{X}_n - \mu)}{\sigma}\right)\right]}_{\xrightarrow{D} N(0,1)} = +\infty & \alpha > \frac{1}{2} \end{cases}$$

Si $\alpha = \dfrac{1}{2}$,

$$n^\alpha(\overline{X}_n - \mu) = \sigma\left[\sqrt{n}\left(\frac{(\overline{X}_n - \mu)}{\sigma}\right)\right] \xrightarrow[n\to\infty]{D} N(0, \sigma^2)$$

pero esta convergencia no implica que sea en probabilidad.

14. Sean $(X_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias y X otra variable aleatoria no necesariamente definidas sobre un mismo espacio. Probar que $X_n \xrightarrow[n\to\infty]{D} X$ si y sólo si $g(X_n) \xrightarrow[n\to\infty]{D} g(X)$ para toda $g : \mathbb{R} \to \mathbb{R}$ continua.

Solución:

$\underline{(\Rightarrow):}$ Como $X_n \xrightarrow[n\to\infty]{D} X$, entonces existen $(Y_n)_{n\in\mathbb{N}}$ e Y tales que $Y_n \sim F_{X_n}$, $Y \sim F_X$ y $Y_n \xrightarrow[n\to\infty]{cs} Y$.

Además,

$$\begin{aligned} Y_n \sim X_n &\Leftrightarrow \mathbb{P}(Y_n \in B) = \mathbb{P}(X_n \in B) = \mathbb{P}(g(Y_n) \le t) = \mathbb{P}(g(X_n) \le t) \\ &\Leftrightarrow \mathbb{P}(Y_n \in g^{-1}(-\infty, t]) = \mathbb{P}(X_n \in g^{-1}(-\infty, t]) \\ &\Leftrightarrow \mathbb{P}(g(Y_n) \in (-\infty, t]) = \mathbb{P}(g(X_n) \in (-\infty, t]) \end{aligned}$$

Luego, como $Y_n \xrightarrow[n\to\infty]{cs} Y \Rightarrow Y_n \xrightarrow[n\to\infty]{D} Y$ y, por lo tanto, para g continua, $g(Y_n) \xrightarrow[n\to\infty]{cs} g(Y) \Rightarrow g(Y_n) \xrightarrow[n\to\infty]{D} g(Y)$.

Es decir, si $X_n \xrightarrow[n\to\infty]{D} X$ entonces $g(X_n) \xrightarrow[n\to\infty]{D} g(X)$.

15. Sean $(a_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ una sucesión que satisface $\lim\limits_{n\to+\infty} a_n = +\infty$ y $(X_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias tal que $a_n(X_n - \mu) \xrightarrow[n\to\infty]{D} X$ para un determinado parámetro μ y cierta variable aleatoria X.

(i) Probar que si $g : \mathbb{R} \to \mathbb{R}$ es una función derivable en μ entonces $a_n[g(X_n) - g(\mu)] \xrightarrow[n\to\infty]{D} g'(\mu)X$.
Sugerencia: Obsevar que para todo $x \in \mathbb{R}$ tenemos la escritura

$$g(x) = g(\mu) + [g'(\mu) + T(x)](x - \mu)$$

donde $T : \mathbb{R} \to \mathbb{R}$ es una función continua con $T(\mu) = 0$.

(ii) Probar que si $g : \mathbb{R} \to \mathbb{R}$ es una función con derivada segunda en μ y tal que $g'(\mu) = 0$ entonces

$$a_n^2(g(X_n) - g(\mu)) \xrightarrow[n\to\infty]{D} \frac{g''(\mu)}{2}X^2.$$

Solución:

(i) $a_n[g(X_n) - g(\mu)] = [g'(\mu) + T(X_n)]a_n(X_n - \mu)$. Luego,

- $a_n(X_n - \mu) \xrightarrow[n\to\infty]{D} X$ por hipótesis, por lo tanto, $X_n \xrightarrow[n\to\infty]{P} X$.
- Por el ítem anterior, $X_n \xrightarrow[n\to\infty]{P} X \Rightarrow T(X_n) \xrightarrow[n\to\infty]{P} T(\mu) = 0$ puesto que T es continua.
- De los ítems anteriores, se tiene que $T(X_n)a_n(X_n - \mu) \xrightarrow[n\to\infty]{D} 0$.

$$a_n[g(X_n) - g(\mu)] = [g'(\mu) + T(X_n)]a_n(X_n - \mu) = g'(\mu)a_n(X_n - \mu) + T(X_n)a_n(X_n - \mu) \xrightarrow[n\to\infty]{D} g'(\mu)X + 0 = g'(\mu).$$

(ii) $g(x) = g(\mu) + \underbrace{g'(\mu)}_{=0}(x - \mu) + \frac{1}{2}\left[g''(\mu) + T(x)\right](x - \mu)^2$, donde $T : \mathbb{R} \to \mathbb{R}$ es una función continua con $T(\mu) = 0$.

- $a_n(X_n - \mu) \xrightarrow[n\to\infty]{D} X$ por hipótesis, por lo tanto, $a_n^2(X_n - \mu)^2 \xrightarrow[n\to\infty]{D} X^2$ y $X_n \xrightarrow[n\to\infty]{P} X$.
- Por el ítem anterior, $X_n \xrightarrow[n\to\infty]{P} X \Rightarrow T(X_n) \xrightarrow[n\to\infty]{P} T(\mu) = 0$ puesto que T es continua.
- De los ítems anteriores, se tiene que $T(X_n)a_n^2(X_n - \mu)^2 \xrightarrow[n\to\infty]{D} 0$.

$$a_n^2[g(X_n) - g(\mu)] = a_n^2\frac{1}{2}\left[g''(\mu) + T(x)\right](X_n - \mu)^2 = \frac{g''(\mu)}{2}a_n^2(X_n - \mu)^2 - \frac{T(x)}{2}a_n^2(X_n - \mu)^2 \xrightarrow[n\to\infty]{D} \frac{g''(\mu)}{2}X^2.$$

16. Sean $X_1, \dots, X_{100}$ variables aleatorias independientes e idénticamente distribuidas con densidad dada por

$$f_X(x) = \frac{2}{x^3}\mathbb{1}_{(1,\infty)}(x).$$

Calcular aproximadamente $\mathbb{P}\left(\prod_{i=1}^{100} X_i > e^{55}\right)$.
Sugerencia: Hallar la distribución de $\log(X)$.

Solución:

Sea $Y = \ln(X)$, entonces $F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(X \leq e^y) = F_X(e^y)$. Por lo tanto, $f_Y(y) = e^y f_X(e^y) = 2e^{-2y}\mathbb{1}_{(0,+\infty)}(y)$. Es decir, $Y \sim \mathcal{E}(2)$.

$$\mathbb{P}\left(\prod_{i=1}^{100} X_i > e^{55}\right) = \mathbb{P}\left(\sum_{i=1}^{100} \ln(X_i) > 55\right)$$

Por otra parte, $Y_n = \sum_{i=1}^{n} \ln(X_i) \sim \Gamma(n, 2) \Rightarrow \mathbb{E}(Y_n) = \frac{n}{2}$ y $\mathbb{V}(Y_n) = \frac{n}{4}$. Luego,

$$\mathbb{P}\left(\prod_{i=1}^{100} X_i > e^{55}\right) = \mathbb{P}\left[\frac{1}{\sqrt{25}}\left(\sum_{i=1}^{100} \ln(X_i) - 50\right) > \frac{55 - 50}{\sqrt{25}}\right] \underbrace{\approx}_{\substack{Z\sim N(0,1)\\ TCL}} \mathbb{P}(Z > 1) = 1 - \mathbb{P}(Z \leq 1) \approx 0{,}1587$$

17. Rehacer los ejercicios 1 y 2 de la práctica 8 utilizando el Teorema del Límite Central: pero donde se pide acotar la probabilidad, calcularla de manera aproximada. Comparar con las cotas obtenidas a partir de la desigualdad de Tchebychev.

18. Sea $(X_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias independientes e idénticamente distribuidas con función de distribución acumulada F. Para cada $n \in \mathbb{N}$ y $t \in \mathbb{R}$ se define la variable aleatoria

$$F_n(t) = \frac{\sharp\{i \in \{1,\dots,n\} : X_i \leq t\}}{n}.$$

(i) Probat que para todo $t \in \mathbb{R}$ se satisface $F_n(t) \xrightarrow[n\to\infty]{cs} F(t)$.

(ii) Mostrar que $\sqrt{n}(F_n(t) - F(t)) \xrightarrow[n\to\infty]{D} Z \sim N[0, F(t)(1-F(t))]$.

19. Sea $(X_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias independientes con distribución $\mathcal{U}([0,\theta])$ para cierto $\theta > 0$. Probar que $\sqrt{n}[\ln(2\overline{X}_n) - \ln(\theta)] \xrightarrow[n\to\infty]{D} N\left(0, \frac{1}{3}\right)$.

Solución: Por el Teorema Central del Límite, $\sqrt{n}\dfrac{[\overline{X}_n - \mathbb{E}(\overline{X}_n)]}{\sqrt{\mathbb{V}(\overline{X}_n)}} \xrightarrow[n\to\infty]{D} N(0,1)$.

$$\sqrt{n}\frac{[\overline{X}_n - \mathbb{E}(\overline{X}_n)]}{\sqrt{\mathbb{V}(\overline{X}_n)}} = \sqrt{n}\frac{\left[\overline{X}_n - \dfrac{\theta}{2}\right]}{\dfrac{\theta}{\sqrt{12}}} = \frac{\sqrt{3}}{\theta}\sqrt{n}(2\overline{X} - \theta)$$

$$\Rightarrow \frac{\sqrt{3}}{\theta}\sqrt{n}(2\overline{X} - \theta) \xrightarrow[n\to\infty]{D} N(0,1) \Rightarrow \sqrt{n}(2\overline{X} - \theta) \xrightarrow[n\to\infty]{D} \theta N\left(0, \frac{1}{3}\right)$$

$$\Rightarrow \sqrt{n}[\ln(2\overline{X}) - \ln(\theta)] \xrightarrow[n\to\infty]{D} \theta \left.\frac{\partial \ln(x)}{\partial x}\right|_{x=\theta} N\left(0, \frac{1}{3}\right) = N\left(0, \frac{1}{3}\right)$$

20. Sea $(X_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias independientes e identicamente distribuidas tal que $\mathbb{E}(X_1) = 0, \mathbb{E}(X_1^2) = 2$ y $\mathbb{E}(X_1^4) < \infty$. Si para cada $n \in \mathbb{N}$ definimos las variables aleatorias

$$Y_n = \frac{\sqrt{n}\sum_{i=1}^{n} X_i}{\sum_{i=1}^{n} X_i^2} \qquad Z_n = \frac{\sum_{i=1}^{n} X_i}{\sqrt{\sum_{i=1}^{n} X_i^2}} \qquad W_n = n\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^2$$

hallar el límite en distribución de las sucesiones $(Y_n)_{n\in\mathbb{N}}$, $(Z_n)_{n\in\mathbb{N}}$ y $(W_n)_{n\in\mathbb{N}}$

Solución:

Sea $(X_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias i.i.d., luego si $S_n = \sum_{i=1}^{n} X_i$ se tiene que $\mathbb{E}(S_n) = 0$ y $\mathbb{V}(S_n) = 2n$.

Por el Teorema Central del Límite

$$\frac{S_n - \mathbb{E}(S_n)}{\sqrt{\mathbb{V}(S_n)}} \xrightarrow[n\to\infty]{D} N(0,1) \Rightarrow \frac{S_n}{\sqrt{n}} \xrightarrow[n\to\infty]{D} N(0,2) \Rightarrow \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \xrightarrow[n\to\infty]{D} N(0,2).$$

Por otra parte,

$$\sum_{i=1}^{n}(X_i - \overline{X}_n)^2 = \sum_{i=1}^{n} X_i^2 - 2\overline{X}_n\sum_{i=1}^{n} X_i + \sum_{i=1}^{n}\overline{X}_n^2 = \sum_{i=1}^{n} X_i^2 - 2\overline{X}_n n\overline{X}_n^2 + n\overline{X}_n^2 = \sum_{i=1}^{n} X_i^2 - n\overline{X}_n^2$$

y, por la Ley de los Grandes Números, $\overline{X}_n \xrightarrow[n\to\infty]{P} \mathbb{E}(X_1) = 0 \Rightarrow \overline{X}_n^2 \xrightarrow[n\to\infty]{P} 0$ y $\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}_n^2 \xrightarrow[n\to\infty]{P} \mathbb{E}(X_1^2) = 2$.

$$\text{(i) } Y_n = \frac{\sqrt{n}\sum_{i=1}^{n} X_i}{\sum_{i=1}^{n} X_i^2} = \underbrace{\left(\sqrt{n}\frac{1}{n}\sum_{i=1}^{n} X_i\right)}_{\xrightarrow[n\to\infty]{D} N(0,2)} \underbrace{\frac{1}{\frac{1}{n}\left(\sum_{i=1}^{n} X_i^2 - n\overline{X}_n^2\right)}}_{\xrightarrow[n\to\infty]{P} \frac{1}{\mathbb{E}(X_1^2)}} \xrightarrow[n\to\infty]{D} \frac{1}{\mathbb{E}(X_1^2)}N(0,2) = \frac{1}{2}N(0,2) \sim N\left(0, \frac{1}{2}\right).$$

(ii) $$Z_n = \frac{\sum_{i=1}^{n} X_i}{\sqrt{\sum_{i=1}^{n} X_i^2}} = \underbrace{\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i\right)}_{\xrightarrow[n\to\infty]{D} N(0,2)} \frac{1}{\underbrace{\sqrt{\frac{1}{n}\sum_{i=1}^{n} X_i^2}}_{\xrightarrow[n\to\infty]{P} \frac{1}{\sqrt{\mathbb{E}(X_1^2)}}}} \xrightarrow[n\to\infty]{D} \frac{1}{\sqrt{\mathbb{E}(X_1^2)}} N(0,2) = \frac{1}{\sqrt{2}} N(0,2) \sim N(0,1).$$

(iii) $$W_n = n\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^2 = \left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i\right)^2 \xrightarrow[n\to\infty]{D} [N(0,2)]^2.$$

21. Sea $(X_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias independientes con distribución Poisson de parámetro λ. Hallar el límite en distribución de $\sqrt{n}\left(\overline{X}_n^2 - \lambda^2\right)$.

Solución:

Sea $(X_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias i.i.d., luego si $S_n = \sum_{i=1}^{n} X_i$ se tiene que $\mathbb{E}(S_n) = n\lambda$ y $\mathbb{V}(S_n) = n\lambda$.

Por el Teorema Central del Límite

$$\frac{S_n - \mathbb{E}(S_n)}{\sqrt{\mathbb{V}(S_n)}} \xrightarrow[n\to\infty]{D} N(0,1) \Rightarrow \sqrt{n}\frac{\overline{X}_n - \lambda}{\sqrt{\lambda}} \xrightarrow[n\to\infty]{D} N(0,1) \Rightarrow \sqrt{n}(\overline{X}_n - \lambda) \xrightarrow[n\to\infty]{D} \sqrt{\lambda}N(0,1) \sim N(0,\lambda)$$

Sea $g : \mathbb{R} \to \mathbb{R}$ derivable definida por $g(x) = x^2$, entonces como $\sqrt{n}(\overline{X}_n - \lambda) \xrightarrow[n\to\infty]{D} N(0,\lambda)$ se tiene que

$$\sqrt{n}\left(\overline{X}_n^2 - \lambda^2\right) = \sqrt{n}[g(\overline{X}_n) - g(\lambda)] \xrightarrow[n\to\infty]{D} \left.\frac{\partial g(x)}{\partial x}\right|_{x=\lambda} N(0,\lambda) = 2\lambda N(0,\lambda) \sim N(0,4\lambda^3).$$

22. Sean X_n e Y_m m dos variables aleatorias independientes con distribución Poisson de parámetros n y m, respectivamente.

   (i) Probar que
   $$\frac{(X_n - n) + (Y_m - m)}{\sqrt{X_n + Y_m}} \xrightarrow[n\to\infty]{D} Z \sim N(0,1) \qquad \text{cuando } n, m \to +\infty.$$

   (ii) Probar que
   $$\frac{(X_n - n) - (Y_m - m)}{\sqrt{X_n + Y_m}} \xrightarrow[n\to\infty]{D} Z \sim N(0,1) \qquad \text{cuando } n, m \to +\infty.$$

23. Sea $(X_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias independientes e idénticamente distribuidas con función de densidad
$$f(x) = \frac{1}{\alpha} x^{-\frac{1}{\alpha}-1} \mathbb{1}_{(1,+\infty)}(x)$$
para cierto $0 < \alpha < 1$.

   (i) Hallar el límite de
   $$Y_n = \sqrt[n]{\prod_{k=1}^{n} X_k}.$$
   ¿En qué sentidos puede garantizar esta convergencia? No analizar convergencia en L^p.

   (ii) Probar que
   $$Z_n = \frac{\sqrt{n}\left(\log \sqrt[n]{\prod_{k=1}^{n} X_k} - \alpha\right)}{\sqrt{\log \sqrt[n]{\prod_{k=1}^{n} X_k}}} \xrightarrow[n\to\infty]{D} Z \sim N(0,\alpha).$$

(iii) Sea $(\alpha_n)_{n\in\mathbb{N}}$ una sucesión de números reales tal que $\lim\limits_{n\to\infty} \alpha_n = \alpha$. Probar que

$$W_n = \frac{1}{\alpha_n}\left[\left(\frac{\sqrt{n}(\log R_n - \alpha)}{\sqrt{\log R_n}}\right)^2 + \left(\frac{\sqrt{n}(\log T_n - \alpha)}{\sqrt{\log T_n}}\right)^2\right] \xrightarrow[n\to\infty]{D} W \sim \chi^2(2).$$

donde, para cada $n \in \mathbb{N}$

$$R_n = \sqrt[n]{\prod_{k=1}^{n} X_{2k-1}} \qquad \text{y} \qquad R_n = \sqrt[n]{\prod_{k=1}^{n} X_{2k}}$$