



INGENIERIA INDUSTRIAL
UNIVERSIDAD DE CHILE

FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
UNIVERSIDAD DE CHILE

Probabilidades

Guía de Estudio

Iván Vidal Romero

AGOSTO, 2026

Índice general

1. Teoría Básica de Conjuntos	1
1.1. Operaciones entre eventos	1
1.1.1. Preguntas	2
1.1.2. Preguntas de Evaluación	4
2. Espacios de Probabilidad	5
2.1. Espacios muestrales y eventos	5
2.1.1. Preguntas	6
2.1.2. Preguntas de Evaluación	7
3. Probabilidad Condicional e Independencia	13
3.1. Probabilidad condicional e independencia	13
3.1.1. Preguntas	14
3.1.2. Preguntas de Evaluación	14
3.2. Probabilidades totales y Teorema de Bayes	19
3.2.1. Preguntas	19
3.2.2. Preguntas de Evaluación	22
4. Principio de Conteo	37
4.1. Técnicas de conteo	37
4.1.1. Preguntas	38
4.1.2. Preguntas de Evaluación	41
5. Variables Aleatorias Discretas	54
5.1. Función de masa, distribución y momentos	54
5.1.1. Preguntas	55
5.1.2. Preguntas de Evaluación	57
5.2. Distribuciones discretas	65
5.2.1. Preguntas	65
5.2.2. Preguntas de Evaluación	70
6. Variables Aleatorias Continuas	90
6.1. Función de distribución y densidad	90
6.1.1. Preguntas	91
6.1.2. Preguntas de Evaluación	92
6.2. Transformaciones y distribuciones continuas	100
6.2.1. Preguntas	101
6.2.2. Preguntas de Evaluación	103

7. Distribución Conjunta	116
7.1. Distribuciones conjuntas	116
7.1.1. Preguntas	117
7.1.2. Preguntas de Evaluación	124
8. Propiedades de la Esperanza	156
8.1. Esperanza condicional, sumas y transformadas	156
8.1.1. Preguntas	157
8.1.2. Preguntas de Evaluación	161
9. Convergencia y Teoremas Límite	176
9.1. Desigualdades, convergencia y teoremas límite	176
9.1.1. Preguntas	177
9.1.2. Preguntas de Evaluación	180

Capítulo 1

Teoría Básica de Conjuntos

1.1. Operaciones entre eventos

Resumen

Conjuntos y cardinalidad.

Un conjunto no considera orden ni repeticiones. Se escribe $x \in A$ cuando x pertenece a A , $A \subseteq B$ cuando todo elemento de A pertenece a B , y $|A|$ para la cardinalidad de A . Además:

- **Conjunto vacío.** $\emptyset$ no contiene elementos y $\emptyset \subseteq A$.
- **Conjunto universo.** Ω contiene todos los elementos posibles en el contexto considerado.
- **Conjunto potencia.** $\mathcal{P}(A) = \{B : B \subseteq A\}$; si A es finito, $|\mathcal{P}(A)| = 2^{|A|}$.

Operaciones.

Para $A, B \subseteq \Omega$,

$$A \cup B = \{x : x \in A \text{ o } x \in B\},$$

$$A \cap B = \{x : x \in A \text{ y } x \in B\},$$

$$A \setminus B = \{x : x \in A, x \notin B\},$$

$$A^c = \Omega \setminus A.$$

Para una familia $\{A_i\}_{i \in I}$,

$$x \in \bigcup_{i \in I} A_i \iff x \in A_i \text{ para algún } i,$$

$$x \in \bigcap_{i \in I} A_i \iff x \in A_i \text{ para todo } i.$$

Disjunción y particiones.

Los conjuntos A_i son disjuntos a pares si $A_i \cap A_j = \emptyset$ para $i \neq j$. Forman una partición de Ω si, además, $\bigcup_i A_i = \Omega$.

Leyes de conjuntos.

Para $A, B, C \subseteq \Omega$:

- **Identidad.** $A \cup \emptyset = A$ y $A \cap \Omega = A$.
- **Dominación.** $A \cup \Omega = \Omega$ y $A \cap \emptyset = \emptyset$.
- **Idempotencia.** $A \cup A = A$ y $A \cap A = A$.
- **Complemento.** $A \cup A^c = \Omega$, $A \cap A^c = \emptyset$ y $(A^c)^c = A$.
- **Conmutatividad.** $A \cup B = B \cup A$ y $A \cap B = B \cap A$.
- **Asociatividad.** $(A \cup B) \cup C = A \cup (B \cup C)$, y análogamente para $\cap$.
- **Distributividad.** $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ y $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.
- **Absorción.** $A \cup (A \cap B) = A$ y $A \cap (A \cup B) = A$.
- **Leyes de De Morgan.** $(\bigcup_i A_i)^c = \bigcap_i A_i^c$ y $(\bigcap_i A_i)^c = \bigcup_i A_i^c$.

Producto y cardinalidad.

El producto cartesiano es $A \times B = \{(a, b) : a \in A, b \in B\}$. Para conjuntos finitos,

$$|A \times B| = |A| |B|,$$

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

La segunda identidad es el caso de dos conjuntos del principio de inclusión-exclusión.

1.1.1. Preguntas

Pregunta 1.

Suponga que este semestre tomó 3 ramos: Probabilidades, Modelamiento y Optimización y Algoritmos y Estructuras de Datos, que podría eliminar. Probabilidades la puede eliminar con probabilidad p_1 ; Modelamiento y Optimización con probabilidad p_2 y Algoritmo y Estructuras de Datos p_3 . También sabe que la probabilidad de eliminar simultáneamente Proba y Opti es $p_1 \cdot p_2$; Opti y Algoritmos es $p_2 \cdot p_3$; Probabilidades y Algoritmos es $p_1 \cdot p_3$. Por último, la probabilidad de eliminar los 3 es $p_1 \cdot p_2 \cdot p_3$, donde $p_1, p_2, p_3 \in [0, 1]$. Calcule la probabilidad de que elimine:

- (a) Alguno de los tres ramos.
- (b) Sólo Probabilidades.
- (c) Sólo Probabilidades y Modelamiento y Optimización.
- (d) Probabilidades, sabiendo que elimina al menos uno.
- (e) Algoritmos, sabiendo que elimina exactamente uno.
- (f) Evalúe numéricamente los resultados anteriores, sabiendo que $p_1 = p_2 = p_3 = \frac{1}{2}$.

Solución

Introduzcamos los eventos asociados a los tres ramos:

- A : eliminar Probabilidades.
- C : eliminar Algoritmos y Estructuras de Datos.
- B : eliminar Modelamiento y Optimización.

(a). Por inclusión-exclusión,

$$\begin{aligned}\mathbb{P}(A \cup B \cup C) &= \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) + \mathbb{P}(A \cap B \cap C) \\ &= p_1 + p_2 + p_3 - p_1p_2 - p_1p_3 - p_2p_3 + p_1p_2p_3 \\ &= 1 - (1 - p_1)(1 - p_2)(1 - p_3).\end{aligned}$$

(b). La probabilidad de eliminar sólo Probabilidades es

$$\begin{aligned}\mathbb{P}(A \cap B^c \cap C^c) &= \mathbb{P}(A) - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) + \mathbb{P}(A \cap B \cap C) \\ &= p_1(1 - p_2)(1 - p_3).\end{aligned}$$

(c). La probabilidad de eliminar sólo Probabilidades y Modelamiento y Optimización es

$$\begin{aligned}\mathbb{P}(A \cap B \cap C^c) &= \mathbb{P}(A \cap B) - \mathbb{P}(A \cap B \cap C) \\ &= p_1p_2(1 - p_3).\end{aligned}$$

(d). Como $A \subseteq A \cup B \cup C$,

$$\begin{aligned}\mathbb{P}(A \mid A \cup B \cup C) &= \frac{\mathbb{P}(A \cap (A \cup B \cup C))}{\mathbb{P}(A \cup B \cup C)} \\ &= \frac{p_1}{1 - (1 - p_1)(1 - p_2)(1 - p_3)}.\end{aligned}$$

Esta expresión está definida cuando $1 - (1 - p_1)(1 - p_2)(1 - p_3) > 0$; si dicha probabilidad es cero, la condicional no está definida.

(e). Si D es el evento de eliminar exactamente un ramo, primero observamos que

$$\begin{aligned}\mathbb{P}(C \cap A^c \cap B^c) &= \mathbb{P}(C) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) + \mathbb{P}(A \cap B \cap C) \\ &= (1 - p_1)(1 - p_2)p_3.\end{aligned}$$

Como los tres casos que forman D son disjuntos,

$$\mathbb{P}(C \mid D) = \frac{(1-p_1)(1-p_2)p_3}{p_1(1-p_2)(1-p_3) + (1-p_1)p_2(1-p_3) + (1-p_1)(1-p_2)p_3}.$$

La fórmula requiere $\mathbb{P}(D) > 0$; en los casos degenerados con $\mathbb{P}(D) = 0$, esta probabilidad condicional no está definida.

(f). Para $p_1 = p_2 = p_3 = 1/2$, los resultados de los incisos (a)–(e) son, respectivamente,

$$\frac{7}{8}, \quad \frac{1}{8}, \quad \frac{1}{8}, \quad \frac{4}{7}, \quad \frac{1}{3}.$$

Pregunta 2.

La probabilidad de que una persona haya visto durante su infancia Hannah Montana es 0.65 y de haber visto Dragon Ball Z es 0.75. También se sabe que la probabilidad de que una persona haya visto ambas es 0.5. Con esto:

- Determine la probabilidad de que haya visto alguno de los dos programas.
- Determine la probabilidad de que solo haya visto Dragon Ball Z.
- Sea A : Haber visto Hannah Montana y B : Haber visto Dragon Ball Z con las probabilidades dadas por enunciado. ¿Cuál es el valor mínimo que debe tener $\mathbb{P}(A \cap B)$ para que se cumpla que $\mathbb{P}(A \cup B) \leq 1$.

Solución

Usaremos los eventos A y B definidos en el inciso (c).

(a). La probabilidad de que haya visto alguno de los dos programas se obtiene por inclusión–exclusión:

$$\begin{aligned}\mathbb{P}(A \cup B) &= \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) \\ &= 0,65 + 0,75 - 0,5 \\ &= 0,9.\end{aligned}$$

Es decir, la probabilidad de haber visto alguna de las dos es 0,9.

(b). Haber visto solo Dragon Ball Z corresponde al evento $B \cap A^c$. Por diferencia,

$$\begin{aligned}\mathbb{P}(B \cap A^c) &= \mathbb{P}(B) - \mathbb{P}(A \cap B) \\ &= 0,75 - 0,5 \\ &= 0,25.\end{aligned}$$

(c). Aunque el enunciado ya fija $\mathbb{P}(A \cap B) = 0,5$, este inciso se interpreta como la búsqueda del menor valor compatible únicamente con las marginales $\mathbb{P}(A) = 0,65$ y $\mathbb{P}(B) = 0,75$. Bajo esa interpretación, toda probabilidad está acotada superiormente por uno. Por tanto,

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) \leq 1$$

Al despejar la probabilidad de la intersección se obtiene

$$\begin{aligned}\mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) &\leq 1 \\ 0,65 + 0,75 - \mathbb{P}(A \cap B) &\leq 1 \\ 1,4 - \mathbb{P}(A \cap B) &\leq 1 \\ \mathbb{P}(A \cap B) &\geq 0,4.\end{aligned}$$

Por lo tanto, el valor mínimo de la intersección es 0,4. En particular, si $\mathbb{P}(A \cap B) = 0,4$, entonces $\mathbb{P}(A \cup B) = 1$.

1.1.2. Preguntas de Evaluación

Pregunta 3. (Control 1 - Otoño 2023, P2)

La municipalidad de Santiago ha anunciado los próximos eventos que se realizarán en el Movistar Arena y dentro de la parrilla de artistas se encuentra Taylor Swift, Coldplay y Paramore. La profesora Natalia es fanática de estos artistas, por lo que le pide a usted, que la ayude a estudiar la probabilidad de asistir a estos conciertos. Para esto, considere los eventos (T) Asistir al concierto de Taylor Swift; (C) Asistir al concierto de Coldplay; (P) Asistir al concierto de Paramore.

La profesora es muy muy fanática de estos grupos, por lo que sí o sí irá a al menos un concierto. Se sabe además que la probabilidad de que asista a Coldplay es de un 40 %, la probabilidad de que asista a Coldplay y Taylor Swift es de un 30 %, y que la probabilidad de que asista solo a Paramore es de un 20 %.

- Escriba en función de los eventos anteriores, el evento “Asistir **exactamente** a 2 de los 3 conciertos”. No tiene que desarrollar ni simplificar la expresión, pero sí dar una breve justificación en palabras.
- Calcule la probabilidad de que asista al concierto de Taylor Swift.

Solución

(a). La unión $(T \cap C) \cup (T \cap P) \cup (C \cap P)$ contiene los resultados en que se asiste a dos o tres conciertos. Para excluir el último caso, el evento solicitado puede escribirse como

$$((T \cap C) \cup (T \cap P) \cup (C \cap P)) \cap (T \cap C \cap P)^c.$$

(b). Como la profesora asistirá al menos a un concierto, el complemento de T se descompone en dos eventos disjuntos: asistir solo a Paramore y asistir a Coldplay sin asistir a Taylor Swift. El segundo tiene probabilidad

$$\mathbb{P}(C \setminus T) = \mathbb{P}(C) - \mathbb{P}(C \cap T) = 0,4 - 0,3 = 0,1.$$

Por lo tanto,

$$\mathbb{P}(T) = 1 - \mathbb{P}(\text{solo } P) - \mathbb{P}(C \setminus T) = 1 - 0,2 - 0,1 = 0,7.$$

Capítulo 2

Espacios de Probabilidad

2.1. Espacios muestrales y eventos

Resumen

Espacio muestral y eventos.

El espacio muestral Ω contiene todos los resultados posibles del experimento. Una σ -álgebra $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ reúne los eventos a los que se asigna probabilidad y satisface:

- **Espacio completo.** $\Omega \in \mathcal{F}$.
- **Cierre por complementos.** $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$.
- **Cierre por uniones numerables.** $A_i \in \mathcal{F}$ para todo $i \Rightarrow \bigcup_{i \geq 1} A_i \in \mathcal{F}$.

En consecuencia, $\emptyset \in \mathcal{F}$ y $\mathcal{F}$ también es cerrada bajo intersecciones numerables.

Espacio de probabilidad.

Un espacio de probabilidad es una tripleta $(\Omega, \mathcal{F}, \mathbb{P})$, donde $\mathbb{P}: \mathcal{F} \rightarrow [0, 1]$ satisface los axiomas:

- **No negatividad.** $\mathbb{P}(A) \geq 0$ para todo $A \in \mathcal{F}$.
- **Normalización.** $\mathbb{P}(\Omega) = 1$.
- **Aditividad numerable.** Si $A_1, A_2, \dots$ son disjuntos a pares,

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

Propiedades básicas.

Para $A, B \in \mathcal{F}$:

- **Probabilidad del vacío.** $\mathbb{P}(\emptyset) = 0$.
- **Cotas.** $0 \leq \mathbb{P}(A) \leq 1$.

- **Complemento.** $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.
- **Monotonía.** Si $A \subseteq B$, entonces $\mathbb{P}(A) \leq \mathbb{P}(B)$.
- **Diferencia.** $\mathbb{P}(A \setminus B) = \mathbb{P}(A) - \mathbb{P}(A \cap B)$.
- **Unión disjunta.** Si $A \cap B = \emptyset$, entonces $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$.
- **Subaditividad.** $\mathbb{P}(A \cup B) \leq \mathbb{P}(A) + \mathbb{P}(B)$.

Inclusión-exclusión.

La intersección debe descontarse porque aparece en ambos términos:

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

Para $A_1, \dots, A_n$ y $[n] = \{1, \dots, n\}$,

$$\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) = \sum_{\emptyset \neq J \subseteq [n]} (-1)^{|J|+1} \mathbb{P}\left(\bigcap_{j \in J} A_j\right).$$

Continuidad de la probabilidad.

Para sucesiones de eventos:

- **Continuidad desde abajo.** Si $A_n \uparrow A$, entonces $\mathbb{P}(A_n) \rightarrow \mathbb{P}(A)$.
- **Continuidad desde arriba.** Si $A_n \downarrow A$, entonces $\mathbb{P}(A_n) \rightarrow \mathbb{P}(A)$.

Ley discreta uniforme.

Si Ω es finito y sus resultados son equiprobables, entonces, para $A \subseteq \Omega$,

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

Los casos favorables y posibles deben contarse dentro del mismo experimento.

2.1.1. Preguntas

Pregunta 4.

- a) En un experimento, un dado es lanzado hasta que un 6 aparece, momento en que se detiene el experimento. Se define E_n como el evento en donde se necesitan n lanzamientos del dado para que termine el experimento. Calcule el tamaño de E_n y por último, ¿Qué significa $(\bigcup_{n=1}^{\infty} E_n)^c$ en términos de conjuntos?
- b) Considere el siguiente experimento: Reordenar, de manera aleatoria, los dígitos 1, 2 y 3. Para $i, j \in \{1, 2, 3\}$ distintos calcule la probabilidad del evento $A_i = \{\text{El } i\text{-ésimo dígito es } i\} \subseteq \Omega$, luego calcule $\mathbb{P}(A_i \cap A_j)$ y finalmente la probabilidad de que al menos un dígito caiga en su posición.

Solución

(a). Si cada resultado se representa mediante la historia finita observada hasta el primer 6, entonces

$$E_n = \{(i_1, \dots, i_{n-1}, 6) : i_j \in \{1, 2, 3, 4, 5\}\}, \quad |E_n| = 5^{n-1}.$$

En este espacio de historias detenidas,

$$\Omega_{\text{det}} = \bigcup_{n=1}^{\infty} E_n, \quad \left(\bigcup_{n=1}^{\infty} E_n \right)^c = \emptyset.$$

También puede modelarse el experimento en $\Omega_{\infty} = \{1, \dots, 6\}^{\mathbb{N}}$. En ese espacio, el complemento contiene las secuencias en las que nunca aparece un 6. Este evento no es vacío, pero tiene probabilidad $\lim_{m \rightarrow \infty} (5/6)^m = 0$.

(b). Las seis permutaciones de $(1, 2, 3)$ son equiprobables. Fijar el dígito i en la posición i deja dos maneras de ordenar los restantes; por ello, $|A_i| = 2$ y

$$\mathbb{P}(A_i) = \frac{2}{6} = \frac{1}{3}, \quad i = 1, 2, 3.$$

Si $i \neq j$, fijar ambos dígitos determina también la tercera posición. En consecuencia,

$$\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_1 \cap A_2 \cap A_3) = \frac{1}{6}.$$

Aplicando inclusión-exclusión,

$$\begin{aligned} \mathbb{P}(A_1 \cup A_2 \cup A_3) &= \sum_{i=1}^3 \mathbb{P}(A_i) - \sum_{1 \leq i < j \leq 3} \mathbb{P}(A_i \cap A_j) + \mathbb{P}(A_1 \cap A_2 \cap A_3) \\ &= 3 \left(\frac{1}{3} \right) - 3 \left(\frac{1}{6} \right) + \frac{1}{6} = \frac{2}{3}. \end{aligned}$$

Esta identidad se obtiene al escribir $(A_1 \cup A_2) \cap A_3 = (A_1 \cap A_3) \cup (A_2 \cap A_3)$ y aplicar dos veces la fórmula de la unión de dos eventos.

2.1.2. Preguntas de Evaluación

Pregunta 5. (Control 1 - Otoño 2022, P1)

Usted y dos compañeros realizarán un experimento que tiene como conjunto de posibles resultados a $\Omega = \{a, b\}$. Usted le pregunta a sus compañeros qué probabilidad creen que tienen los distintos eventos en Ω . Ellos le entregan las funciones $\mathbb{P}_1 : \mathcal{P}(\Omega) \rightarrow \mathbb{R}$ y $\mathbb{P}_2 : \mathcal{P}(\Omega) \rightarrow \mathbb{R}$, dadas por

$$\begin{aligned}\mathbb{P}_1(\emptyset) &= 0, \mathbb{P}_1(\{a\}) = 0.6, \mathbb{P}_1(\{b\}) = 0.7, \mathbb{P}_1(\Omega) = 1; \\ \mathbb{P}_2(\emptyset) &= 0, \mathbb{P}_2(\{a\}) = 5, \mathbb{P}_2(\{b\}) = 10, \mathbb{P}_2(\Omega) = 15.\end{aligned}$$

- (a) Verifique si $\mathbb{P}_1$ y $\mathbb{P}_2$ son funciones de probabilidad o no, y explique por qué. En caso de que alguna no sea función de probabilidad, muestre una propiedad o axioma que no se cumpla, pero que toda función de probabilidad debería cumplir.

- (b) Explique por qué existe un único $\alpha > 0$ tal que $\mathbb{P} := \alpha \cdot \mathbb{P}_2$ es una función de probabilidad. Calcule dicho α .

Nota: Si su respuesta en la parte (a) es que $\mathbb{P}_2$ sí es una función de probabilidad, entonces $\alpha = 1$. Si no, entonces $\alpha \neq 1$.

Para las partes (c) y (d) tome la función de probabilidad $\mathbb{P}$ encontrada en la parte (b) como la función de probabilidad asociada al experimento. Considere la variable aleatoria $X : \Omega \rightarrow \mathbb{R}$ dada por

$$X(a) = 9, X(b) = 21.$$

- (c) Calcule el rango de X . Calcule la PMF y la CDF de X y grafíquelas.
(d) Calcule $\mathbb{E}[X]$ y $\text{Var}(X)$.

Solución

(a). Ninguna de las dos funciones es una medida de probabilidad. Como $\{a\}$ y $\{b\}$ son disjuntos y su unión es Ω , la aditividad exigiría

$$\mathbb{P}_1(\{a\}) + \mathbb{P}_1(\{b\}) = \mathbb{P}_1(\Omega).$$

Sin embargo, $0,6 + 0,7 = 1,3 \neq 1$. Por otra parte, $\mathbb{P}_2(\Omega) = 15$, lo que contradice el axioma de normalización.

(b). Para que $\mathbb{P} = \alpha \mathbb{P}_2$ sea una medida de probabilidad se requiere

$$\mathbb{P}(\Omega) = \alpha \mathbb{P}_2(\Omega) = 15\alpha = 1.$$

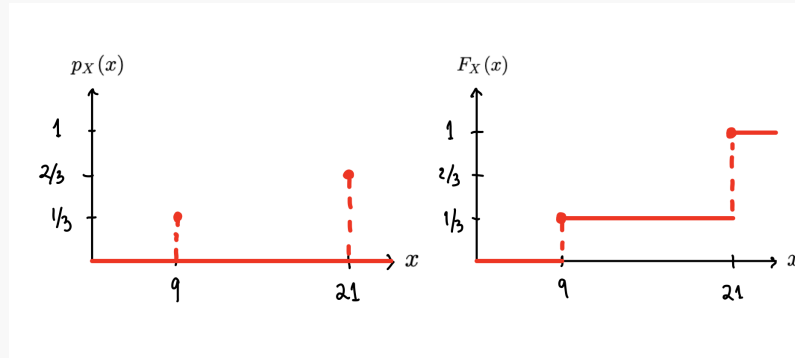
En consecuencia, necesariamente $\alpha = 1/15$, lo que demuestra también la unicidad. Con este valor, $\mathbb{P}$ es no negativa y aditiva porque $\mathbb{P}_2$ posee ambas propiedades, y queda normalizada por el cálculo anterior.

(c). El rango de X es $\{9, 21\}$. Como $\mathbb{P}(\{a\}) = 1/3$ y $\mathbb{P}(\{b\}) = 2/3$, su PMF es

$$p_X(x) = \begin{cases} 1/3, & x = 9, \\ 2/3, & x = 21, \\ 0, & \text{en otro caso,} \end{cases}$$

y su CDF es

$$F_X(x) = \begin{cases} 0, & x < 9, \\ 1/3, & 9 \leq x < 21, \\ 1, & x \geq 21. \end{cases}$$

Figura 2.1: PMF (izquierda) y CDF (derecha) de X .

(d). La esperanza y el segundo momento son

$$\mathbb{E}[X] = 9 \left(\frac{1}{3} \right) + 21 \left(\frac{2}{3} \right) = 17,$$

$$\mathbb{E}[X^2] = 9^2 \left(\frac{1}{3} \right) + 21^2 \left(\frac{2}{3} \right) = 321.$$

Por la identidad $\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$,

$$\text{Var}(X) = 321 - 17^2 = 32.$$

Pregunta 6. (Control 1 - Primavera 2023, P1)

(i). Los sistemas eléctricos en la vida real suelen estar conformados por varios componentes, los cuales pueden ir conectados entre sí **en paralelo** o **en serie**. Cuando la conexión es **en paralelo** (ver Figura 2.2), el sistema funcionará cuando al menos uno de los componentes funcione; mientras que si la conexión es **en serie** (ver Figura 2.3), el sistema funcionará solo si todos los componentes están operativos.

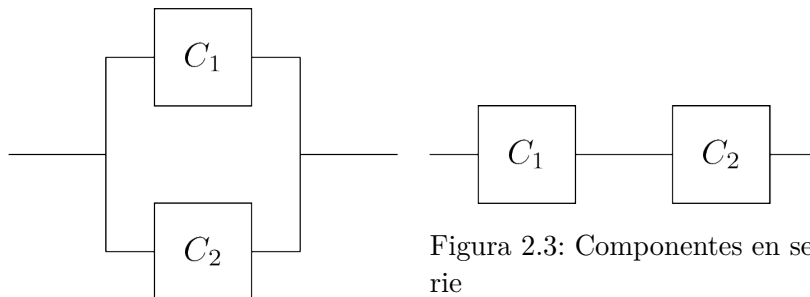
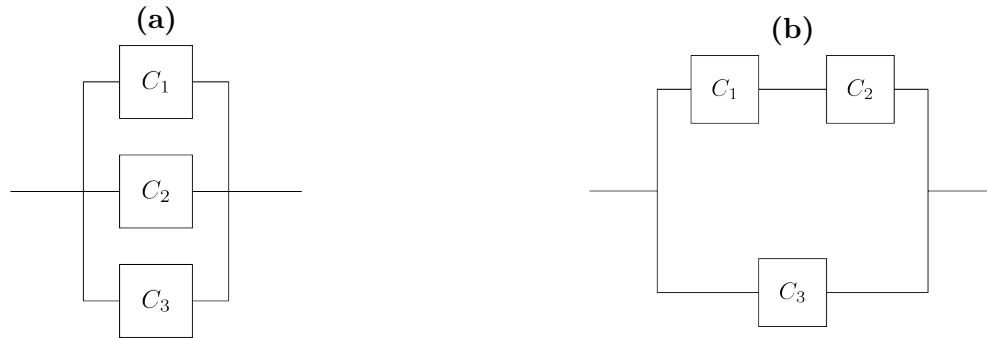


Figura 2.2: Componentes en paralelo

Figura 2.3: Componentes en serie

Para los siguientes sistemas, partes a) y b) a continuación, calcule la probabilidad de que **el sistema funcione**. Asuma que cada componente C_k está operativo con una probabilidad p_k independiente de los demás. Sea claro en indicar el evento que se pregunta en cada caso, y entregue todo el desarrollo con todos los pasos de cómo llega a la respuesta.



(ii). Cuatro especialidades de la facultad: Computación (C), Geología (G), Industrias (I) y Matemática (M) compiten en el torneo de *la mona*, donde solo uno de ellos será el ganador. Los equipos C y G tienen la misma probabilidad de ganar, el equipo I tiene el doble de probabilidad de ganar que M , mientras que la probabilidad de que el equipo C o el equipo I gane el torneo es 0,6. Encuentre las probabilidades de cada equipo de ganar el torneo.

Solución

Sea F el evento de que el circuito funcione.

(i)(a). Los tres componentes están en paralelo, de modo que el circuito falla solamente si fallan los tres. Usando independencia,

$$\begin{aligned}\mathbb{P}(F) &= 1 - \mathbb{P}(C_1^c \cap C_2^c \cap C_3^c) \\ &= 1 - (1 - p_1)(1 - p_2)(1 - p_3).\end{aligned}$$

(i)(b). Los componentes C_1 y C_2 forman una rama en serie, que funciona cuando ocurre $C_1 \cap C_2$, y esta rama está en paralelo con C_3 . El circuito falla cuando falla la primera rama y también falla C_3 . Por independencia,

$$\begin{aligned}\mathbb{P}(F) &= 1 - \mathbb{P}((C_1 \cap C_2)^c \cap C_3^c) \\ &= 1 - (1 - \mathbb{P}(C_1 \cap C_2))\mathbb{P}(C_3^c) \\ &= 1 - (1 - p_1 p_2)(1 - p_3).\end{aligned}$$

(ii). Como solo una especialidad puede ganar, los cuatro eventos son disjuntos. Si $x = \mathbb{P}(C) = \mathbb{P}(G)$ y $y = \mathbb{P}(M)$, entonces $\mathbb{P}(I) = 2y$. Las probabilidades suman uno y, además, $\mathbb{P}(C \cup I) = 0,6$, por lo que

$$2x + 3y = 1, \quad x + 2y = 0,6.$$

Al resolver este sistema se obtiene

$$\mathbb{P}(C) = 0,2, \quad \mathbb{P}(G) = 0,2, \quad \mathbb{P}(I) = 0,4, \quad \mathbb{P}(M) = 0,2.$$

Pregunta 7. (Control 1 - Otoño 2024, P1)

En el mes de abril, el cometa 12P/Pons-Brooks será visible desde la Tierra, y los observatorios ESO, Mauna Kea y ALMA han firmado un acuerdo en el cual se comprometen a compartir las mediciones que realicen. Los observatorios se encuentran distribuidos en distintos lugares del mundo: ESO se encuentra en Múnich, Mauna Kea en Hawái, y ALMA en San Pedro de Atacama (los dos primeros lugares están en el hemisferio norte, mientras que el último está en el hemisferio sur). Para llevar a cabo la medición por parte de cada observatorio, es primordial que el cielo esté despejado en la zona de medición. En Múnich, esto ocurre el 30 % del tiempo; en Hawái, el 70 % del tiempo; y en San Pedro de Atacama, el 50 % del tiempo. Además, se sabe que el 80 % del tiempo en al menos una de las ubicaciones está nublado. Considere que el clima entre San Pedro de Atacama y Múnich son independientes. Por otra parte, un 25 % de las veces está despejado en San Pedro de Atacama y nublado en Hawái. Se sabe que para obtener una medición **precisa** de la trayectoria del cometa 12P/Pons-Brooks,

se deben tener al menos dos mediciones realizadas, una en cada hemisferio.

Calcule la probabilidad de obtener una medición **precisa**.

Solución

Introduzcamos los eventos de cielo despejado:

- S : cielo despejado en San Pedro de Atacama. ■ M : cielo despejado en Múnich.
- H : cielo despejado en Hawái.

Los datos no definen ninguna medida de probabilidad. En efecto, por complemento,

$$\mathbb{P}(S \cap H \cap M) = 1 - \mathbb{P}(S^c \cup H^c \cup M^c) = 0,2.$$

Pero la independencia de S y M implica

$$\mathbb{P}(S \cap M) = 0,5 \cdot 0,3 = 0,15,$$

lo que contradice $S \cap H \cap M \subseteq S \cap M$. Por tanto, la probabilidad solicitada no puede calcularse mientras no se corrija alguno de los datos.

Pregunta 8. (Control 1 - Primavera 2024, P1)

Un estudiante de la Facultad de Ciencias Físicas y Matemáticas reside en Lampa y hoy debe rendir un importante control en la universidad. Para asegurarse de llegar a tiempo, el estudiante evalúa dos alternativas de transporte. La primera opción es optar por un Uber que lo transportaría directamente desde su hogar en Lampa hasta la facultad. Aunque esta ruta es la más rápida, enfrenta un desafío significativo, la disponibilidad de Ubers en la zona puede ser limitada, y existe la posibilidad real de no encontrar un vehículo disponible. La probabilidad de que encuentre un Uber disponible es p_1 .

La segunda alternativa implica tomar un bus rural hasta la estación Los Libertadores, desde donde continuaría su viaje en metro hasta la universidad. Esta opción, aunque es más prolongada, también podría complicarse por retrasos en el bus o problemas en el metro, afectando su llegada a tiempo. Sea p_2 la probabilidad de que el bus rural no experimente contratiempos, mientras que p_3 representa la probabilidad de que el bus rural llegue a tiempo al paradero y que el metro presente algún inconveniente. Por último, la probabilidad, de que ninguno de los tres métodos de transporte falle es p_4 . La ocurrencia de los fenómenos mencionados en los medios de transporte no tienen por qué ser independientes.

- (a) Exprese el evento en el que el estudiante llegue a tiempo a su evaluación, a partir de eventos que remitan a un solo medio de transporte.
- (b) Calcule la probabilidad de que el estudiante llegue a tiempo a la universidad.
- (c) Defina una VA igual al número de transportes que “funcionan adecuadamente” de los tres medios de transporte mencionados. Funcionar adecuadamente para: un Uber, es que hay disponibilidad en la zona; el bus, es que no presenta retrasos; el metro, es que no presenta inconvenientes. Exprese la CDF de dicha variable aleatoria evaluada en 1, como una probabilidad, $\mathbb{P}(\cdot)$, de un evento que se construya a partir de tres eventos, cada uno de estos definidos sobre un solo medio de transporte.
- (d) Respecto a la VA definida en la parte (c), ¿Hay algún(os) valor(es) de su rango tales que su PMF corresponde(n) a alguna de las probabilidades entregadas al inicio (p_1, p_2, p_3, p_4)? Justifique claramente su respuesta.
- (e) Respecto a la VA definida en la parte (c), diga la relación entre la CDF evaluada en 1 y la CDF evaluada en 1,5. Específicamente, cuál es mayor (o igual), o menor (o igual), o

si son iguales, o si no se puede afirmar nada. Justifique claramente su respuesta. No es necesario que haga cálculos.

Solución

Introduzcamos los eventos asociados a cada medio de transporte:

- U : hay un Uber disponible.
- M : el metro no presenta inconvenientes.
- B : el bus no presenta retrasos.

(a). El estudiante llega a tiempo si encuentra un Uber o si tanto el bus como el metro funcionan adecuadamente. El evento solicitado es

$$U \cup (B \cap M).$$

(b). Por inclusión-exclusión,

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

Los datos indican $\mathbb{P}(U) = p_1$, $\mathbb{P}(B) = p_2$, $\mathbb{P}(B \cap M^c) = p_3$ y $\mathbb{P}(U \cap B \cap M) = p_4$. Como $B = (B \cap M) \dot{\cup} (B \cap M^c)$,

$$\mathbb{P}(B \cap M) = p_2 - p_3.$$

Al sustituir,

$$\mathbb{P}(U \cup (B \cap M)) = p_1 + p_2 - p_3 - p_4.$$

(c). Una definición directa de la variable es $X = \mathbb{1}_{\{U\}} + \mathbb{1}_{\{B\}} + \mathbb{1}_{\{M\}}$. El evento $\{X \leq 1\}$ contiene el caso en que ningún medio funciona y los tres casos en que funciona exactamente uno. En consecuencia,

$$F_X(1) = \mathbb{P}((U^c \cap B^c \cap M^c) \cup (U \cap B^c \cap M^c) \cup (U^c \cap B \cap M^c) \cup (U^c \cap B^c \cap M)).$$

Los cuatro eventos de la unión son disjuntos, por lo que esta probabilidad también puede escribirse como la suma de sus probabilidades.

(d). El rango de X es $\{0, 1, 2, 3\}$. La probabilidad p_4 coincide directamente con la masa en 3, pues

$$p_X(3) = \mathbb{P}(X = 3) = \mathbb{P}(U \cap B \cap M) = p_4.$$

En cambio, p_1 y p_2 describen el funcionamiento de un solo medio sin especificar los otros dos, y p_3 describe dos medios sin especificar el Uber. Por ello, ninguno de estos tres valores es, en general, una masa puntual de X .

(e). Como X solo toma valores enteros, los eventos $\{X \leq 1\}$ y $\{X \leq 1,5\}$ son iguales. Por lo tanto,

$$F_X(1) = F_X(1,5).$$

Pregunta 9. (Control 1 - Otoño 2025, P1)

Tres amigos están perdidos tras hacer trekking y, al caer la noche, deciden intentar hacer fuego para protegerse del frío. Para aumentar sus posibilidades, acuerdan que cada uno lo intente por su cuenta, eligiendo entre el valle o la montaña. El primer amigo propone: “Deberíamos intentar los tres en el valle”. El segundo responde: “Mejor que dos lo intenten en el valle y uno en la montaña; así diversificamos el riesgo”. Considere que cada uno de los amigos logra hacer fuego por su cuenta con probabilidad p en el valle. Si los tres lo intentan en el valle, la probabilidad de que al menos uno falle es r . Mientras que si dos intentan en el valle y uno en la montaña: La probabilidad de que al menos uno de los que están en el valle tenga éxito es s y la probabilidad de que el amigo en la montaña tenga éxito es q , donde $q < p$, considere que este último evento es independiente del resultado en el valle.

Se quiere maximizar la probabilidad de que al menos uno de los amigos logre hacer fuego.

Entregue una condición, en base a los parámetros del problema, para que la opción sugerida por el segundo amigo sea mejor (en el sentido descrito al inicio de esta pregunta) que la estrategia sugerida por el primer amigo.

Solución

Sea V_i el evento de que el amigo i logra hacer fuego en el valle. El resultado requiere interpretar que la ley conjunta de los éxitos en el valle no depende de la estrategia y es intercambiable entre amigos, de modo que $\mathbb{P}(V_i \cup V_j) = s$ para todo par $i \neq j$. Sin esta hipótesis, la probabilidad de éxito cuando los tres están en el valle no queda determinada. Bajo esta interpretación, $\mathbb{P}(V_i) = p$ y, por inclusión-exclusión,

$$\mathbb{P}(V_i \cap V_j) = \mathbb{P}(V_i) + \mathbb{P}(V_j) - \mathbb{P}(V_i \cup V_j) = 2p - s.$$

La probabilidad de que al menos uno falle es r , por lo que $\mathbb{P}(V_1 \cap V_2 \cap V_3) = 1 - r$. Aplicando inclusión-exclusión a los tres eventos de éxito,

$$\begin{aligned} \mathbb{P}(V_1 \cup V_2 \cup V_3) &= \sum_{i=1}^3 \mathbb{P}(V_i) - \sum_{1 \leq i < j \leq 3} \mathbb{P}(V_i \cap V_j) + \mathbb{P}(V_1 \cap V_2 \cap V_3) \\ &= 3p - 3(2p - s) + (1 - r) \\ &= 1 - r - 3p + 3s. \end{aligned}$$

Para la segunda estrategia, sea M el evento de que el amigo ubicado en la montaña logra hacer fuego. Los dos amigos del valle fallan simultáneamente con probabilidad $1 - s$, mientras que el amigo de la montaña falla con probabilidad $1 - q$. Como estos dos resultados son independientes,

$$\begin{aligned} \mathbb{P}(V_1 \cup V_2 \cup M) &= 1 - \mathbb{P}(V_1^c \cap V_2^c \cap M^c) \\ &= 1 - (1 - s)(1 - q). \end{aligned}$$

Por consiguiente, la estrategia del segundo amigo es mejor precisamente cuando

$$1 - (1 - s)(1 - q) > 1 - r - 3p + 3s,$$

o, de manera equivalente,

$$(1 - s)(1 - q) < 3p + r - 3s.$$

Capítulo 3

Probabilidad Condicional e Independencia

3.1. Probabilidad condicional e independencia

Resumen

Probabilidad condicional.

Para $A, B \in \mathcal{F}$ con $\mathbb{P}(B) > 0$,

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

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

Condicionar restringe el espacio muestral a B y vuelve a normalizar la probabilidad dentro de ese evento.

Regla de la cadena.

Siempre que las probabilidades condicionales estén bien definidas,

$$\mathbb{P}\left(\bigcap_{i=1}^n A_i\right) = \mathbb{P}(A_1) \prod_{i=2}^n \mathbb{P}\left(A_i \mid \bigcap_{j<i} A_j\right).$$

Propiedades condicionales.

Fijado $C \in \mathcal{F}$ con $\mathbb{P}(C) > 0$, $\mathbb{P}(\cdot | C)$ es una medida de probabilidad. Para $A, B \in \mathcal{F}$:

- **Vacío y cotas.** $\mathbb{P}(\emptyset | C) = 0$ y $0 \leq \mathbb{P}(A | C) \leq 1$.
- **Complemento.** $\mathbb{P}(A^c | C) = 1 - \mathbb{P}(A | C)$.
- **Monotonía.** Si $A \subseteq B$, entonces $\mathbb{P}(A | C) \leq \mathbb{P}(B | C)$.
- **Diferencia.** $\mathbb{P}(A \setminus B | C) = \mathbb{P}(A | C) - \mathbb{P}(A \cap B | C)$.
- **Inclusión-exclusión.**

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

Independencia.

Los eventos A y B son independientes si

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

Si $\mathbb{P}(B) > 0$, esto equivale a $\mathbb{P}(A | B) = \mathbb{P}(A)$.

Para $A_1, \dots, A_n$, escriba $[n] = \{1, \dots, n\}$. La independencia mutua exige que, para todo $J \subseteq [n]$ con $|J| \geq 2$,

$$\mathbb{P}\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} \mathbb{P}(A_j).$$

Propiedades que conviene distinguir:

- **Simetría.** Si A es independiente de B , entonces B es independiente de A .
- **Complementos.** La independencia se conserva al complementar uno o ambos eventos.
- **Disjunción.** Dos eventos disjuntos con probabilidades positivas no son independientes.
- **Independencia a pares.** Exige la factorización solo para cada par y no implica independencia mutua.

Independencia condicional.

Para $\mathbb{P}(C) > 0$, los eventos A y B son condicionalmente independientes dado C si

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

La independencia condicional no implica independencia sin condicionar; tampoco vale, en general, la implicación recíproca.

3.1.1. Preguntas

Pregunta 10.

Se dispone de un revólver con capacidad para $n \in \mathbb{N}$ balas y se insertan $k \in \{0, \dots, n\}$ al azar. Dos personas juegan a la ruleta rusa, girando la nuez después de cada disparo.

- (a) Calcule la probabilidad de que el jugador que comienza disparando muera.
- (b) Evalúe en $n = 6$ y $k = 1$.
- (c) ¿Cómo deben ser n y k para que el juego sea equilibrado?

Solución

(a). Supongamos $1 \leq k \leq n$ y que cada giro selecciona una recámara de manera uniforme e independiente de los giros anteriores. Si M_r denota que la primera bala se dispara en el turno r , entonces

$$\mathbb{P}(M_r) = \left(\frac{n-k}{n}\right)^{r-1} \frac{k}{n}.$$

El primer jugador dispara en los turnos impares. Así,

$$\mathbb{P}(\text{muere el primer jugador}) = \frac{k}{n} \sum_{i=0}^{\infty} \left(\frac{n-k}{n}\right)^{2i} = \frac{n}{2n-k}.$$

Si $k = 0$, no se dispara ninguna bala y esta probabilidad es 0.

(b). Para $n = 6$ y $k = 1$,

$$\mathbb{P}(\text{muere el primer jugador}) = \frac{6}{11}.$$

(c). Para $1 \leq k \leq n$, las probabilidades de morir del primer y segundo jugador son $n/(2n-k)$ y $(n-k)/(2n-k)$, respectivamente. La primera es estrictamente mayor si $k > 0$, de modo que no existe un juego equilibrado con al menos una bala. El caso $k = 0$ es degenerado: nadie puede morir.

3.1.2. Preguntas de Evaluación

Pregunta 11. (Control 1 - Primavera 2022, P1)

Trinidad y Dominique tienen un emprendimiento en entretención de cumpleaños para niños y niñas entre 3 y 8 años. Para este fin de semana, irán al cumpleaños de Pedro, quien celebrará su cumpleaños número 4, al cual invitará a sus amigos del jardín. Trinidad y Dominique tienen tres actividades preparadas: *Cuenta cuentos*, *Baila a tu ritmo*, y *Moldea con plasticina*. Sin embargo, por motivos de tiempo, ellas solo podrán hacer dos de las tres actividades en el cumpleaños de Pedro. Usted debe ayudar a las emprendedoras a decidir qué actividades realizar en el cumpleaños de Pedro, con el objetivo de dejar *contentos* a quienes estarán presente. Considere que un asistente al cumpleaños queda *contento* si al menos una de las dos actividades a realizar son de su gusto o agrado. Motivo de lo anterior, Dominique le hace unas preguntas a la profesora del jardín, quien responde lo siguiente:

“A todos los niños y niñas del curso les gusta al menos una de las tres actividades (*contar cuentos*, *bailar*, o *moldear*). Podría decir que a un 60 % le gusta *bailar*, mientras que a un 70 % le gustan los *cuentos*. Además, a un 35 % no le gusta jugar con *plasticinas*, mientras que a un 30 % le gustan los *cuentos* y no le gusta jugar con *plasticina*. Por cierto, a un 55 % no le gusta *bailar* o no le gustan los *cuentos*. Por último, solo a un 25 % le agradan todas las (i.e., las 3) actividades.”

- (a) Encuentre las dos actividades que se deberían llevar a cabo en el cumpleaños para maximizar la probabilidad de que los asistentes estén *contentos*.

Suponga que ahora, además de las tres actividades mencionadas, hay otra posible actividad a realizar en el cumpleaños: *Títeres*. Respecto a esto la profesora del jardín dice (además de lo dicho anteriormente):

“Podría afirmar que al 55 % no le gustan los *títeres*, mientras que al 22 % no le gustan los *títeres* ni *bailar*.”

- (b) Determine si es que es cierto que “a un niño le gustan los *títeres*” es independiente de que “a un niño le gusta bailar”.

Solución

(a). Introduzcamos los eventos asociados a las actividades:

- C : gustan los cuentos.
- M : gusta moldear con plastilina.
- B : gusta bailar.

De los complementos y las intersecciones entregadas se obtiene

$$\begin{aligned}\mathbb{P}(M) &= 1 - \mathbb{P}(M^c) = 0,65, \\ \mathbb{P}(C \cap M) &= \mathbb{P}(C) - \mathbb{P}(C \cap M^c) = 0,70 - 0,30 = 0,40, \\ \mathbb{P}(B \cap C) &= 1 - \mathbb{P}(B^c \cup C^c) = 1 - 0,55 = 0,45.\end{aligned}$$

Como todos prefieren al menos una actividad, $\mathbb{P}(B \cup C \cup M) = 1$. Por inclusión-exclusión,

$$\begin{aligned}1 &= \mathbb{P}(B) + \mathbb{P}(C) + \mathbb{P}(M) - \mathbb{P}(B \cap C) - \mathbb{P}(C \cap M) - \mathbb{P}(B \cap M) + \mathbb{P}(B \cap C \cap M) \\ &= 0,60 + 0,70 + 0,65 - 0,45 - 0,40 - \mathbb{P}(B \cap M) + 0,25,\end{aligned}$$

de donde $\mathbb{P}(B \cap M) = 0,35$. Las coberturas de cada pareja son

$$\begin{aligned}\mathbb{P}(B \cup C) &= 0,60 + 0,70 - 0,45 = 0,85, \\ \mathbb{P}(B \cup M) &= 0,60 + 0,65 - 0,35 = 0,90, \\ \mathbb{P}(C \cup M) &= 0,70 + 0,65 - 0,40 = 0,95.\end{aligned}$$

La mayor cobertura se obtiene con cuenta cuentos y moldear con plastilina.

(b). Sea T el evento de que gusten los títeres. Se tiene

$$\mathbb{P}(T^c \cap B^c) = 0,22 = 0,55 \cdot 0,4 = \mathbb{P}(T^c)\mathbb{P}(B^c).$$

Así, T^c y B^c son independientes y, en consecuencia, también lo son sus complementos T y B .

Pregunta 12. (Control 1 - Primavera 2022, P3)

1. Suponga que próximamente se llevará a cabo un festival al que usted tiene muchas ganas de asistir, pues tendría la posibilidad de ver a su artista favorito. Considere que:
 - Su artista favorito tiene planeado presentarse uno de los tres días (Viernes, Sábado o Domingo), con probabilidad de 50 %, 30 %, y 20 %, respectivamente.
 - Si llueve el día en que el artista tiene planeado presentarse, entonces se cancela su presentación sin posibilidad de reagendar.
 - La probabilidad que llueva el día viernes es de 20 %.
 - En caso de que llueva un día, la probabilidad que llueva al día siguiente es de 60 %. En caso de que en un día no llueva, la probabilidad de que llueva al día siguiente es de 25 %.

Dado que llueve el día Sábado, ¿cuál es la probabilidad de que usted haya visto a su artista favorito el viernes?

2. Sean los eventos A , B y C pertenecientes al mismo espacio muestral Ω . Demuestre que si A y B son condicionalmente independiente dado C , entonces se tiene que $\mathbb{P}(A \mid B \cap C) = \mathbb{P}(A \mid C)$. No se complique con casos donde las probabilidades pueden ser cero.

Solución

(1). Introduzcamos los eventos relevantes:

- V : el artista se presenta el viernes.
- L_S : llueve el sábado.
- L_V : llueve el viernes.

Haber visto al artista el viernes corresponde a $V \cap L_V^c$. El enunciado no relaciona la elección del día con el clima; bajo la hipótesis implícita de que ambos mecanismos son independientes,

$$\begin{aligned}\mathbb{P}(V \cap L_V^c \mid L_S) &= \frac{\mathbb{P}(V)\mathbb{P}(L_V^c \cap L_S)}{\mathbb{P}(L_S)} \\ &= \frac{\mathbb{P}(V)\mathbb{P}(L_S \mid L_V^c)\mathbb{P}(L_V^c)}{\mathbb{P}(L_S)}.\end{aligned}$$

La probabilidad del denominador se obtiene por la ley de probabilidades totales:

$$\begin{aligned}\mathbb{P}(L_S) &= \mathbb{P}(L_S \mid L_V)\mathbb{P}(L_V) + \mathbb{P}(L_S \mid L_V^c)\mathbb{P}(L_V^c) \\ &= 0,60(0,20) + 0,25(0,80) = 0,32.\end{aligned}$$

Por tanto,

$$\mathbb{P}(V \cap L_V^c \mid L_S) = \frac{0,50(0,25)(0,80)}{0,32} = 0,3125.$$

(2). Suponemos $\mathbb{P}(C) > 0$ y $\mathbb{P}(B \cap C) > 0$, como permite el enunciado. La independencia condicional de A y B dado C significa que

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

Por otra parte, la regla del producto condicional entrega

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

La condición $\mathbb{P}(B \cap C) > 0$ implica $\mathbb{P}(B \mid C) > 0$. Al igualar las dos expresiones y cancelar este factor se obtiene

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

Pregunta 13. (Control 1 - Primavera 2025, P2)

Para el paseo de Industrial se organizará una Yincana en la que los estudiantes compiten **en pareja**. La Yincana consta de tres pruebas que se realizan según la siguiente secuencia: (1) Carrera de sacos, (2) Correr con un huevo en la cuchara y (3) Encontrar el dulce en la harina. La probabilidad de superar cada prueba en un intento es, respectivamente, $p \in (0, 1)$, $q \in (0, 1)$ y $r \in (0, 1)$.

La pareja de Benjamín y Catalina adopta la siguiente estrategia. Benjamín inicia y avanza prueba por prueba: si supera la prueba (1) pasa a la (2), si supera la (2) pasa a la (3), y si supera la (3) ganan. Si Benjamín falla en alguna prueba, no podrá seguir compitiendo, y Catalina toma el relevo en dicha prueba. En este caso, para que la pareja gane, Catalina debe superar todas las pruebas, desde la prueba en la que toma el relevo hasta la (3); si falla en cualquiera de ellas, la pareja pierde. Todos los intentos son independientes entre pruebas e independientes entre integrantes.

- (a) Calcule la probabilidad de que la pareja gane la Yincana. En su desarrollo, escriba el evento del cual se pregunta, en función de eventos que describan el resultado de cada participante en cada una de las pruebas.
- (b) Encuentre la probabilidad de que Catalina haya intentado alguna prueba dado que la pareja ganó.

Solución

Introduzcamos los eventos asociados al desempeño de la pareja:

- B_j : Benjamín supera la prueba j .
- W : la pareja gana la Yincana.
- C_j : Catalina supera la prueba j .

(a). La pareja gana en uno de cuatro escenarios disjuntos: Benjamín completa las tres pruebas, o falla por primera vez en una de ellas y Catalina completa desde ese punto. Por tanto,

$$\begin{aligned} W &= (B_1 \cap B_2 \cap B_3) \cup (B_1^c \cap C_1 \cap C_2 \cap C_3) \\ &\quad \cup (B_1 \cap B_2^c \cap C_2 \cap C_3) \cup (B_1 \cap B_2 \cap B_3^c \cap C_3). \end{aligned}$$

Como los escenarios son disjuntos, sus probabilidades se suman. La independencia entre intentos permite multiplicar dentro de cada escenario:

$$\begin{aligned} \mathbb{P}(W) &= pqr + (1-p)pqr + p(1-q)qr + pq(1-r)r \\ &= pqr[1 + (1-p) + (1-q) + (1-r)] \\ &= pqr(4 - p - q - r). \end{aligned}$$

(b). Sea A el evento de que Catalina intentó alguna prueba. Condicionalmente a que la pareja ganó, A ocurre precisamente en los tres escenarios de relevo. Por ello,

$$\begin{aligned} \mathbb{P}(A \cap W) &= \mathbb{P}(W) - \mathbb{P}(B_1 \cap B_2 \cap B_3) \\ &= pqr(4 - p - q - r) - pqr \\ &= pqr(3 - p - q - r). \end{aligned}$$

Aplicando la definición de probabilidad condicional y el resultado del inciso anterior,

$$\mathbb{P}(A | W) = \frac{\mathbb{P}(A \cap W)}{\mathbb{P}(W)} = \frac{3 - p - q - r}{4 - p - q - r}.$$

Pregunta 14. (Examen - Primavera 2023, P1)

1. Don Eusebio posee un campo con 32 vacas. En las últimas semanas, se han producido varios robos de vacas en otros campos. Motivo de esto, don Eusebio ha decidido poner un dispositivo GPS a sus vacas para así poder rastrearlas en caso de sufrir un robo. Dado el alto precio de los dispositivos GPS, don Eusebio está evaluando a cuántas vacas debería poner dicho dispositivo. Es sabido, que la banda que se dedica a robar vacas, roba el 25 % de las vacas que hay en el campo donde cometen el delito. Asuma que la banda roba exactamente el 25 % de las vacas. Además, cada vaca tiene la misma probabilidad de ser robada. La banda no puede identificar cuáles vacas tienen un GPS y cuáles no.
 - (a) Don Eusebio desea saber cuál es el mínimo número de vacas que debiese tener un GPS para que, en caso de sufrir un robo por la banda mencionada, pueda rastrear a la banda con al menos un 95 % de probabilidad (en otras palabras, la banda se roba al menos una vaca que tiene un GPS). Formule el problema de optimización que debe resolver don Eusebio.
 - (b) Suponga ahora que hay otro proveedor de GPS para animales, el cual vende a un

menor precio que el proveedor anterior. Debido a esto, don Eusebio decide ponerle GPS a todas sus vacas. Durante los últimos días, algunos clientes (del proveedor de GPS con menor precio) se han quejado que a veces el GPS viene con fallas, por lo que no se puede rastrear. Suponga que hay un 60 % de probabilidad que un GPS esté con falla. Se sabe además que las fallas entre un GPS y otro son independientes. Encuentre una expresión para la probabilidad de que, en caso de sufrir un robo por la banda descrita anteriormente, se pueda rastrear al menos una de las vacas.

2. Dé un ejemplo con un **dado** balanceado de seis caras, en que se tengan dos eventos A y B , tales que $\mathbb{P}(B | A) > \mathbb{P}(B)$. Debe ser claro en definir los eventos.
3. Explique en palabras, sin fórmulas, para qué sirve la Ley de Probabilidades Totales. En particular, indique en qué caso sería adecuado usar dicha ley.

Solución

(1)(a). Si n vacas tienen GPS, sea X el número de esas vacas entre las ocho que roba la banda. Entonces

$$X \sim \text{HGeom}(N = 32, K = n, m = 8).$$

Don Eusebio puede rastrear a la banda si $X \geq 1$. Como

$$\mathbb{P}(X = 0) = \frac{\binom{32-n}{8}}{\binom{32}{8}},$$

el número mínimo de dispositivos se obtiene resolviendo

$$\begin{aligned} & \min_{n \in \{0, 1, \dots, 32\}} n \\ & \text{sujeto a } \frac{\binom{32-n}{8}}{\binom{32}{8}} \leq 0,05. \end{aligned}$$

La restricción equivale a $\mathbb{P}(X \geq 1) \geq 0,95$.

(1)(b). Ahora las ocho vacas robadas tienen GPS. Cada dispositivo funciona con probabilidad 0,4, de manera independiente. Si Y es el número de GPS funcionales entre esas vacas, entonces

$$Y \sim \text{Binomial}(8, 0,4).$$

Por complemento,

$$\mathbb{P}(Y \geq 1) = 1 - \mathbb{P}(Y = 0) = 1 - 0,6^8.$$

(2). Considere un dado balanceado, $A = \{2, 4, 6\}$ y $B = \{2\}$. Condicionado a A , los tres resultados pares son equiprobables, de modo que

$$\mathbb{P}(B | A) = \frac{1}{3} > \frac{1}{6} = \mathbb{P}(B).$$

(3). La ley de probabilidades totales permite calcular la probabilidad de un evento separando el espacio muestral en escenarios disjuntos que cubren todos los casos. En cada escenario se calcula la probabilidad del evento y luego se combinan esos valores, ponderados por la probabilidad del escenario. Es especialmente útil cuando el evento es difícil de estudiar directamente, pero resulta sencillo una vez que se conoce cuál de los escenarios ocurrió.

3.2. Probabilidades totales y Teorema de Bayes

Resumen

Particiones.

Los eventos $A_1, \dots, A_n \in \mathcal{F}$ forman una partición de Ω si son disjuntos a pares y cubren todo el espacio:

$$A_i \cap A_j = \emptyset \text{ si } i \neq j, \quad \bigcup_{i=1}^n A_i = \Omega.$$

Por ello, cualquier evento B se descompone como la unión disjunta $B = \bigcup_{i=1}^n (B \cap A_i)$.

Ley de probabilidades totales.

Si $A_1, \dots, A_n$ forman una partición de Ω y $\mathbb{P}(A_i) > 0$, entonces

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

Más generalmente, si $\mathbb{P}(C) > 0$,

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

siempre que los términos condicionales estén bien definidos.

Teorema de Bayes.

Para $\mathbb{P}(A), \mathbb{P}(B) > 0$,

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

Bajo la partición anterior y para $\mathbb{P}(B) > 0$,

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

Los términos tienen roles distintos:

- **Probabilidad a priori.** $\mathbb{P}(A_k)$ representa la información disponible antes de observar B .
- **Verosimilitud.** $\mathbb{P}(B | A_k)$ mide cuán compatible es la observación con el escenario A_k .
- **Evidencia.** $\mathbb{P}(B)$ normaliza las probabilidades de todos los escenarios.
- **Probabilidad a posteriori.** $\mathbb{P}(A_k | B)$ actualiza la probabilidad del escenario después de observar B .

Evidencia compuesta.

Si $B_1, \dots, B_m$ son independientes condicionalmente a cada A_i , su verosimilitud conjunta factoriza:

$$\mathbb{P}\left(\bigcap_{j=1}^m B_j \mid A_i\right) = \prod_{j=1}^m \mathbb{P}(B_j | A_i).$$

Esta factorización debe justificarse; no se desprende de la independencia sin condicionar.

3.2.1. Preguntas

Pregunta 15.

Una urna contiene inicialmente dos bolitas, una blanca y una negra. Se extrae una bolita al azar, se observa su color y se devuelve a la urna junto con otra bolita del color opuesto. Luego se extrae una segunda bolita al azar, que se devuelve con otra del mismo color. Sean los sucesos $B_k =$ “La urna resultante contiene k bolitas blancas”, $A_i =$ “La i -ésima bolita extraída es blanca”, para $i = 1, 2, 3, 4$. Calcule:

- (a) $\mathbb{P}(B_k)$ para $k = 1, 2, 3, 4$.
- (b) $\mathbb{P}(A_2)$ y $\mathbb{P}(A_2 | A_1), \mathbb{P}(A_1 | A_2)$.

Solución

(a). Después de la primera extracción, la urna contiene una bolita blanca y dos negras si ocurre A_1 , y dos blancas y una negra si ocurre A_1^c . Al considerar también el color de la segunda

extracción,

$$B_1 = A_1 \cap A_2^c,$$

$$\mathbb{P}(B_1) = \frac{1}{2} \frac{2}{3} = \frac{1}{3},$$

$$B_2 = (A_1 \cap A_2) \cup (A_1^c \cap A_2^c),$$

$$\mathbb{P}(B_2) = \frac{1}{2} \frac{1}{3} + \frac{1}{2} \frac{1}{3} = \frac{1}{3},$$

$$B_3 = A_1^c \cap A_2,$$

$$\mathbb{P}(B_3) = \frac{1}{2} \frac{2}{3} = \frac{1}{3},$$

$$B_4 = \emptyset,$$

$$\mathbb{P}(B_4) = 0.$$

(b). Por probabilidades totales,

$$\mathbb{P}(A_2) = \mathbb{P}(A_2 \mid A_1)\mathbb{P}(A_1) + \mathbb{P}(A_2 \mid A_1^c)\mathbb{P}(A_1^c) = \frac{1}{2}.$$

Además,

$$\mathbb{P}(A_2 \mid A_1) = \frac{1}{3}, \quad \mathbb{P}(A_1 \mid A_2) = \frac{\mathbb{P}(A_1 \cap A_2)}{\mathbb{P}(A_2)} = \frac{1/6}{1/2} = \frac{1}{3}.$$

Pregunta 16.

Usted es un espectador de la final de la Copa América, en la que se están enfrentando Argentina y Colombia. Los equipos no logran romper el empate en el tiempo reglamentario, por lo que el ganador se definirá en la tanda de penales. Luego de ejecutar los 5 lanzamientos de penales por equipo, no se ha definido un ganador. Motivo de lo anterior, se procede a la fase de *tie break*, en la cual, cada equipo ejecuta un penal por ronda. Las rondas continúan hasta que se tenga una ronda en la cual solo uno de los dos equipos anote y se declare vencedor. Sean A_i el evento en donde Colombia anota en la i -ésima ronda de penales, C_i el evento en donde Argentina anota en la i -ésima ronda, y L_i el evento en donde se realizan exactamente i rondas de lanzamientos de penales; con $i \in \{1, 2, 3, \dots\}$. Note que cada vez que nos referimos a *la primera ronda*, nos referimos a la primera ronda de la fase de *tie break*. Similarmente, cuando nos referimos a la i -ésima ronda, nos referimos a la i -ésima ronda considerando desde la primera ronda de la fase de *tie break*.

Por simplicidad asuma que todos los penales de esta pregunta son en la fase de *tie break*. Además, suponga que Argentina es el equipo que patea en primer lugar.

- a) Escriba el evento *Lanzar exactamente i -rondas de penales* en función de $\{A_j, C_j\}_{j=1}^{\infty}$ para todo $i \in \{1, 2, 3, \dots\}$. Además defina el rango de la cantidad de penales anotados en las i – rondas lanzadas.

A su lado se encuentra Davo, un streamer argentino, quien le comenta que la UBA creo una pagina llamada 301060, la cual indica que la probabilidad de que un jugador convierta su penal es p excepto en el caso de que sea el segundo jugador de la ronda en patear y que además el jugador del equipo rival haya anotado su penal en la misma ronda, pues entonces el jugador siente la presión del ambiente, ya que en caso de no convertir el gol su equipo perderá. La pagina indica que la probabilidad de que convierta el penal un jugador en la situación descrita es q , donde $q < p$.

- b) Calcule la probabilidad de lanzar exactamente i rondas de penales para $i \in \{1, 2, 3, \dots\}$, es decir, calcule $\mathbb{P}(L_i)$.

(hint: notar que las rondas de lanzamiento de penal son independientes entre si.)

- c) Se define G como el evento que gane el equipo que empieza pateando en la tanda de penales, obtenga una expresión para $\mathbb{P}(G \mid L_3)$. Generalice el caso para $\mathbb{P}(G \mid L_i)$

- d) Calcule la probabilidad de ganar del equipo que pateo en primer lugar.
(hint: Utilice probabilidades totales utilizando como partición el número de rondas que dura el *tie break*.)
- e) Finalmente calcule la probabilidad de que gane Argentina, se lancen exactamente 2 rondas de penales, y que solo Julián Álvarez anote su gol.

Solución

(a). El evento *lanzar i rondas de penales* se escribe de la siguiente manera:

$$L_i = ((A_i \cap C_i^c) \cup (A_i^c \cap C_i)) \cap \bigcap_{k=1}^{i-1} ((A_k \cap C_k) \cup (A_k^c \cap C_k^c))$$

para $i \in \{1, 2, 3, \dots\}$, con la convención $\bigcap_{k=1}^0 (\cdot) = \Omega$ cuando $i = 1$.

El primer factor exige que exactamente uno de los equipos anote en la ronda i ; el segundo, que cada ronda anterior termine empatada. Su intersección es precisamente L_i .

Finalmente, el rango de K_i , la cantidad de penales anotados dado que se lanzaron i rondas, está dado por $\{1, 3, 5, \dots, 2i - 1\}$.

(b). Para que se jueguen exactamente i rondas, las primeras $i - 1$ deben terminar empatadas y la última debe terminar con un solo gol. En una ronda hay empate si ambos anotan, con probabilidad pq , o si ambos fallan, con probabilidad $(1 - p)^2$. La ronda es decisiva si Argentina anota y Colombia falla, con probabilidad $p(1 - q)$, o si Argentina falla y Colombia anota, con probabilidad $(1 - p)p$. Como las rondas son independientes,

$$\begin{aligned} \mathbb{P}(L_i) &= \prod_{j=1}^{i-1} \mathbb{P}((A_j \cap C_j) \cup (A_j^c \cap C_j^c)) \\ &\quad \times \mathbb{P}((A_i \cap C_i^c) \cup (A_i^c \cap C_i)) \\ &= (pq + (1 - p)^2)^{i-1} (p(1 - q) + (1 - p)p). \end{aligned}$$

(c). Sea G el evento de que gane Argentina, el equipo que pateo primero. Dado L_i , Argentina gana exactamente cuando en la última ronda ocurre $C_i \cap A_i^c$; las $i - 1$ anteriores deben terminar empatadas. Por tanto,

$$\begin{aligned} \mathbb{P}(G \mid L_i) &= \frac{\mathbb{P}(G \cap L_i)}{\mathbb{P}(L_i)} \\ &= \frac{(pq + (1 - p)^2)^{i-1} p(1 - q)}{(pq + (1 - p)^2)^{i-1} (p(1 - q) + (1 - p)p)} \\ &= \frac{p(1 - q)}{p(1 - q) + (1 - p)p}, \end{aligned}$$

para todo $i \geq 1$, y en particular para $i = 3$.

(d). Sea $r = pq + (1 - p)^2$ la probabilidad de empate en una ronda. Como la probabilidad de una ronda decisiva es $p(1 - q) + (1 - p)p = p(2 - p - q) > 0$, se tiene $r < 1$ y

$$\mathbb{P}(\text{empates para siempre}) = \lim_{m \rightarrow \infty} r^m = 0.$$

Por tanto, $\sum_{i \geq 1} \mathbb{P}(L_i) = 1$ y podemos aplicar probabilidades totales:

$$\begin{aligned} \mathbb{P}(G) &= \sum_{i=1}^{\infty} \mathbb{P}(G \mid L_i) \mathbb{P}(L_i) \\ &= \frac{p(1 - q)}{p(1 - q) + (1 - p)p} \sum_{i=1}^{\infty} \mathbb{P}(L_i) \\ &= \frac{p(1 - q)}{p(1 - q) + (1 - p)p} = \frac{1 - q}{2 - p - q}. \end{aligned}$$

La segunda igualdad usa que la probabilidad condicional del inciso (c) no depende de i , y la tercera, que las probabilidades de la partición suman uno.

(e). Suponiendo que Julián Álvarez ejecuta el penal de Argentina en la segunda ronda, para que se jueguen exactamente dos rondas, gane Argentina y el único gol sea el de Julián Álvarez, en la primera ronda ambos deben fallar y, en la segunda, Julián debe anotar mientras Colombia falla. Por independencia,

$$\mathbb{P}(A_1^c \cap C_1^c \cap C_2 \cap A_2^c) = (1-p)^2 p(1-q).$$

3.2.2. Preguntas de Evaluación

Pregunta 17. (Control 1 - Otoño 2021, P1)

1. Se tienen 2 mochilas. La primera contiene 10 dulces azules, mientras que la segunda contiene 15 dulces azules. Seleccione una de las mochilas de manera aleatoria (i.e. con igual probabilidad escojo cada mochila) e introduzco 6 dulces negros en ella. Luego, revuelvo los dulces en esta mochila y saco un dulce. Si este dulce resulta ser negro, ¿cuál es la probabilidad de haber seleccionado la primera mochila?
2. Se tienen k mochilas con M dulces azules y N dulces negros cada una. Un dulce es seleccionado aleatoriamente desde la primera mochila y es transferido a la segunda, luego se selecciona un dulce al azar desde la segunda mochila y es transferido a la tercera mochila, y así hasta llegar a la k -ésima mochila. Si después de este proceso un dulce es seleccionado de manera aleatoria desde la k -ésima mochila, muestre que la probabilidad de que este dulce sea de color azul es la misma que la probabilidad de que el primer dulce transferido desde la primera a la segunda mochila sea azul (i.e. $p = \frac{M}{M+N}$). *Hint: Podría ser de utilidad analizar dicha probabilidad en la segunda mochila, y luego para una j -ésima mochila cualquiera.*

Solución

(1). Introduzcamos los eventos necesarios:

- N : extraer un dulce negro.
- M_1 : elegir la primera mochila.

Las mochilas se eligen con la misma probabilidad, mientras que

$$\mathbb{P}(N \mid M_1) = \frac{6}{16}, \quad \mathbb{P}(N \mid M_1^c) = \frac{6}{21}.$$

La ley de probabilidades totales entrega

$$\mathbb{P}(N) = \frac{6}{16} \left(\frac{1}{2} \right) + \frac{6}{21} \left(\frac{1}{2} \right) = \frac{37}{112}.$$

Por el Teorema de Bayes,

$$\begin{aligned} \mathbb{P}(M_1 \mid N) &= \frac{\mathbb{P}(N \mid M_1) \mathbb{P}(M_1)}{\mathbb{P}(N)} \\ &= \frac{3/16}{37/112} = \frac{21}{37} \approx 0,568. \end{aligned}$$

(2). Para $j = 1, \dots, k-1$, sea A_j el evento de que el dulce transferido desde la mochila j hacia la mochila $j+1$ sea azul. Denotemos además por A_k el evento de que el dulce extraído al final desde la mochila k sea azul. En la primera transferencia,

$$\mathbb{P}(A_1) = \frac{M}{M+N}.$$

Supongamos que $\mathbb{P}(A_{j-1}) = M/(M+N)$. Si el dulce recibido por la mochila j fue azul, esta contiene $M+1$ dulces azules y N negros antes de la siguiente extracción; si fue negro, contiene M azules y $N+1$ negros. Por la ley de probabilidades totales,

$$\begin{aligned}\mathbb{P}(A_j) &= \frac{M+1}{M+N+1}\mathbb{P}(A_{j-1}) + \frac{M}{M+N+1}\mathbb{P}(A_{j-1}^c) \\ &= \frac{M+1}{M+N+1} \frac{M}{M+N} + \frac{M}{M+N+1} \frac{N}{M+N} \\ &= \frac{M}{M+N}.\end{aligned}$$

El mismo cálculo se aplica a la extracción final desde la mochila k . Por inducción,

$$\mathbb{P}(A_k) = \frac{M}{M+N},$$

que coincide con la probabilidad de que el primer dulce transferido sea azul.

Pregunta 18. (Control 1 - Primavera 2021, P1)

Don Beltrán tiene una empresa con varios trabajadores los cuales poseen una cuenta de correo electrónico de la empresa. Últimamente, la empresa ha tenido problemas debido a la apertura de emails con virus, ya que el servidor de correos de la empresa no cuenta con ningún sistema de detección de emails con virus. Por este motivo, don Beltrán ha contratado una consultora en tecnologías de la información para abordar el tema. Luego de un mes, la consultora le afirma lo siguiente: “Hemos detectado que el $100v\%$ de los correos recibidos tiene virus. Además, hemos detectado que de todos los correos que contienen virus, el $100\alpha\%$ contiene la palabra w . Por otra parte, de los correos que no tienen virus, solo el $100\beta\%$ contiene la palabra w ”.

- Determine la probabilidad de que un correo que tiene la palabra w , posea virus.
- La política de la empresa es que un correo que contiene la palabra w no se debe abrir. Por el contrario, aquellos correos que no tengan la palabra w se deben abrir. Se sabe que el costo de abrir un correo con virus es de c_A , mientras que no abrir un correo sin virus tiene un costo de c_N . Calcule el costo esperado de un correo.

Solución

Introduzcamos los eventos asociados al correo:

- A : el correo contiene la palabra w .
- V : el correo contiene un virus.

Los datos son $\mathbb{P}(V) = v$, $\mathbb{P}(A | V) = \alpha$ y $\mathbb{P}(A | V^c) = \beta$.

(a). La ley de probabilidades totales aplicada a la partición $\{V, V^c\}$ da

$$\begin{aligned}\mathbb{P}(A) &= \mathbb{P}(A | V)\mathbb{P}(V) + \mathbb{P}(A | V^c)\mathbb{P}(V^c) \\ &= \alpha v + \beta(1 - v).\end{aligned}$$

Al sustituir esta expresión en el Teorema de Bayes se obtiene

$$\mathbb{P}(V | A) = \frac{\mathbb{P}(A | V)\mathbb{P}(V)}{\mathbb{P}(A)} = \frac{\alpha v}{\alpha v + \beta(1 - v)}.$$

Esta probabilidad condicional está definida siempre que $\alpha v + \beta(1 - v) > 0$. Si este valor es cero, el evento de observar la palabra w tiene probabilidad nula.

(b). Se incurre en el costo c_A cuando se abre un correo con virus, es decir, cuando ocurre $A^c \cap V$. El costo c_N corresponde a no abrir un correo sin virus, esto es, al evento $A \cap V^c$. Si C denota el

costo por correo, entonces

$$\begin{aligned}\mathbb{E}[C] &= c_A \mathbb{P}(A^c \cap V) + c_N \mathbb{P}(A \cap V^c) \\ &= c_A \mathbb{P}(A^c | V) \mathbb{P}(V) + c_N \mathbb{P}(A | V^c) \mathbb{P}(V^c) \\ &= c_A(1 - \alpha)v + c_N \beta(1 - v).\end{aligned}$$

Pregunta 19. (Control 1 - Otoño 2022, P2)

En un curso de $n = 100$ alumnos aparte de usted, están buscando cambiar la fecha de la clase auxiliar. Para saber la opinión del curso, pregunta a 3 personas al azar (sin reemplazo).

Para su estudio considere 2 escenarios equiprobables:

1. Hay 75 alumnos que quieren cambiar la fecha y 25 que no.
 2. Hay 25 alumnos que no quieren cambiar la fecha y 75 que sí.
- (a) ¿Cuál es la probabilidad de que las 3 personas encuestadas quieran cambiar la fecha del auxiliar?
- (b) ¿Cuál es la probabilidad de que ocurra el escenario 1 dado que las 3 personas quieren cambiar la fecha?

Considere ahora los siguientes escenarios equiprobables, con $i \in \{0, 1, 2, \dots, 100\}$:

$$E_i = \{\text{escenario donde } i \text{ personas quieren cambiar la fecha} \\ \text{y } 100 - i \text{ no quieren cambiarla}\}.$$

- (c) ¿Cuál es la probabilidad de que las 3 personas encuestadas quieran cambiar la fecha de la clase auxiliar?
- (d) ¿Cuál es la probabilidad de que más de la mitad de las personas quieran cambiar la fecha dado que las 3 personas encuestadas quieren cambiar la fecha?

Solución

Introduzcamos los eventos de la primera parte:

- H : las tres personas encuestadas quieren cambiar la fecha.
- S_1 : hay 75 personas que quieren cambiar la fecha.
- S_2 : hay 25 personas que quieren cambiar la fecha.

(a). Como ambos escenarios son equiprobables, la ley de probabilidades totales entrega

$$\begin{aligned}\mathbb{P}(H) &= \mathbb{P}(H | S_1) \mathbb{P}(S_1) + \mathbb{P}(H | S_2) \mathbb{P}(S_2) \\ &= \frac{1}{2} \frac{\binom{75}{3}}{\binom{100}{3}} + \frac{1}{2} \frac{\binom{25}{3}}{\binom{100}{3}} \\ &\approx 0,216.\end{aligned}$$

(b). Por el Teorema de Bayes,

$$\begin{aligned}\mathbb{P}(S_1 | H) &= \frac{\mathbb{P}(H | S_1) \mathbb{P}(S_1)}{\mathbb{P}(H | S_1) \mathbb{P}(S_1) + \mathbb{P}(H | S_2) \mathbb{P}(S_2)} \\ &= \frac{\binom{75}{3}}{\binom{75}{3} + \binom{25}{3}} \approx 0,967.\end{aligned}$$

(c). Ahora los 101 escenarios E_i , para $i = 0, \dots, 100$, son equiprobables. Condicionado a E_i ,

$$\mathbb{P}(H \mid E_i) = \frac{\binom{i}{3}}{\binom{100}{3}},$$

donde $\binom{i}{3} = 0$ para $i < 3$. Por la ley de probabilidades totales y la identidad $\sum_{i=3}^{100} \binom{i}{3} = \binom{101}{4}$,

$$\begin{aligned} \mathbb{P}(H) &= \frac{1}{101 \binom{100}{3}} \sum_{i=3}^{100} \binom{i}{3} \\ &= \frac{\binom{101}{4}}{101 \binom{100}{3}} = \frac{1}{4}. \end{aligned}$$

(d). Sea $U = \bigcup_{i=51}^{100} E_i$, el evento de que más de la mitad quiera cambiar la fecha. Los escenarios son disjuntos y tienen la misma probabilidad previa, de modo que el Teorema de Bayes entrega

$$\begin{aligned} \mathbb{P}(U \mid H) &= \frac{\sum_{i=51}^{100} \mathbb{P}(H \mid E_i) \mathbb{P}(E_i)}{\sum_{i=0}^{100} \mathbb{P}(H \mid E_i) \mathbb{P}(E_i)} \\ &= \frac{\sum_{i=51}^{100} \binom{i}{3}}{\sum_{i=3}^{100} \binom{i}{3}} \\ &= \frac{\binom{101}{4} - \binom{51}{4}}{\binom{101}{4}} \approx 0,939. \end{aligned}$$

Pregunta 20. (Control 1 - Primavera 2022, P2)

En un contexto electoral, considere un local de votación que posee 5 mesas de votación: “Mesa A”, “Mesa E”, “Mesa I”, “Mesa O”, y “Mesa U”. En cada una de ellas, hay una persona designada como vocal de mesa, quien puede ser “Nueva” (si está cumpliendo la función por primera vez), o “Antigua” (si ha desempeñado la labor anteriormente). Un vocal *Nuevo* tarda 8 minutos en atender a un votante, mientras que un vocal *Antiguo* tarda 6 minutos. La probabilidad de que haya un vocal *Nuevo* o *Antiguo* en cada una de las 5 mesas está dada por la siguiente tabla:

	Mesa A	Mesa E	Mesa I	Mesa O	Mesa U
Nueva	1/2	2/3	3/5	1/4	1/5
Antigua	1/2	1/3	2/5	3/4	4/5

- (a) Considere que un votante que llega al local es asignado con probabilidad de $1/3$ a la mesa que coincide con la vocal que más aparece en el nombre de dicho votante, y con probabilidad de $1/6$ es asignado al resto de las mesas.

Por ejemplo: si el nombre del votante contiene mayoría de E's (respecto al resto de las vocales), entonces la probabilidad de asignación a la Mesa E es de $1/3$, y de $1/6$ para el resto de las mesas. En caso de que haya un empate respecto a las vocales que más aparecen en el nombre, considere que se le asigna la probabilidad de $1/3$ a la primera vocal de éstas según el orden alfabético. Por ejemplo: si el nombre del votante es Alicia, entonces la probabilidad de asignación a la Mesa A es de $1/3$, y para el resto de las mesas, será de $1/6$. Si una persona que posee el mismo nombre que usted acude al local de votación, ¿cuál es la probabilidad que espere 6 minutos?

- (b) Considerando las reglas de asignación de la parte (a), indique el espacio muestral para una persona con el nombre *Federico*.

- (c) Considere ahora que una persona es asignada a una mesa con probabilidad proporcional al número de vocales de dicha mesa en el nombre de la persona. *Ejemplo: Úrsula será asignada a la Mesa A con probabilidad 1/3, a la Mesa U con probabilidad 2/3, y al resto de las mesas con probabilidad 0.* Encuentre la probabilidad de que una persona llamada Estefanía haya sido asignada a la Mesa E dado que esperó 8 minutos.

Solución

Introduzcamos los eventos asociados a la asignación y al tipo de vocal:

- M_j : ser asignado a la mesa $j \in \{A, E, I, O, U\}$.
- N : ser atendido por un vocal nuevo.
- G : ser atendido por un vocal antiguo.

(a). Esperar 6 minutos equivale a ser atendido por un vocal antiguo. Si j^* es la vocal seleccionada a partir del nombre, entonces $\mathbb{P}(M_{j^*}) = 1/3$ y $\mathbb{P}(M_j) = 1/6$ para $j \neq j^*$. Por la ley de probabilidades totales,

$$\mathbb{P}(G) = \sum_{j \in \{A, E, I, O, U\}} \mathbb{P}(G | M_j) \mathbb{P}(M_j).$$

Al usar la fila “Antigua” de la tabla, los cinco valores posibles son

j^*	A	E	I	O	U
$\mathbb{P}(G)$	197/360	187/360	191/360	53/90	43/72

Por ejemplo, si la vocal seleccionada es A,

$$\mathbb{P}(G) = \frac{1}{2} \frac{1}{3} + \frac{1}{6} \left(\frac{1}{3} + \frac{2}{5} + \frac{3}{4} + \frac{4}{5} \right) = \frac{197}{360}.$$

(b). Cada resultado debe indicar tanto la mesa asignada como el tipo de vocal que atiende. Por ello, un espacio muestral adecuado es

$$\Omega = \{A, E, I, O, U\} \times \{\text{Nueva, Antigua}\}.$$

Para Federico, la vocal que más se repite es E, de modo que $\mathbb{P}(M_E) = 1/3$ y cada una de las otras cuatro mesas tiene probabilidad 1/6; estas probabilidades no alteran el conjunto de resultados posibles.

(c). El nombre “Estefanía” contiene dos letras A, dos E y una I. En consecuencia,

$$\begin{aligned} \mathbb{P}(M_A) &= \frac{2}{5}, & \mathbb{P}(M_E) &= \frac{2}{5}, & \mathbb{P}(M_I) &= \frac{1}{5}, \\ \mathbb{P}(M_O) &= 0, & \mathbb{P}(M_U) &= 0. \end{aligned}$$

Esperar 8 minutos equivale al evento N . Primero calculamos su probabilidad total:

$$\begin{aligned} \mathbb{P}(N) &= \mathbb{P}(N | M_A) \mathbb{P}(M_A) + \mathbb{P}(N | M_E) \mathbb{P}(M_E) + \mathbb{P}(N | M_I) \mathbb{P}(M_I) \\ &= \frac{1}{2} \frac{2}{5} + \frac{2}{3} \frac{2}{5} + \frac{3}{5} \frac{1}{5} = \frac{44}{75}. \end{aligned}$$

El Teorema de Bayes entrega entonces

$$\mathbb{P}(M_E | N) = \frac{\mathbb{P}(N | M_E) \mathbb{P}(M_E)}{\mathbb{P}(N)} = \frac{(2/3)(2/5)}{44/75} = \frac{5}{11}.$$

Pregunta 21. (Control 1 - Otoño 2023, P1)

Suponga que usted tiene un amigo que siempre llega tarde a la clase de Probabilidades. Usted, que sabe la importancia de asistir a clases para aprender de buena manera los contenidos del curso, se preocupa por su amigo y decide preguntarle a qué se deben sus repetitivos atrasos. Su amigo le cuenta que en cada uno de sus viajes a la universidad, utiliza solo un medio de transporte de los tres que tiene disponibles: *bicicleta*, *automóvil* o *micro*. Cada mañana, él

debe escoger cuál de los tres medios de transporte usará, y le comenta lo siguiente:

- Debido al elevado precio de la bencina, utiliza el *automóvil* un 20 % de las veces.
- Como no es muy aficionado del deporte, solo hace un 20 % de los viajes a la universidad en *bicicleta*.
- La probabilidad de llegar tarde cuando viaja en *bicicleta* es mayor a 0.
- Por culpa de la congestión vehicular, la probabilidad de llegar tarde dado que utiliza el *automóvil* es el doble de la probabilidad de llegar tarde dado que utiliza la *bicicleta*.
- Si viaja en *micro*, la probabilidad de llegar atrasado es 4 veces la probabilidad de llegar atrasado dado que viaja en *bicicleta*.

- (a) A la siguiente clase, usted nota que su amigo vuelve a llegar tarde. Calcule numéricamente la probabilidad de que haya viajado en *micro*.

Usted se molesta con su amigo por su nueva irresponsabilidad, pero él le promete que llegará puntualmente a la próxima clase.

- (b) Considere que la probabilidad de llegar tarde es de 0,6. Calcule numéricamente la probabilidad de que su amigo **no llegue tarde**, dado que viajará en *automóvil*.

Solución

Introduzcamos los eventos de transporte y puntualidad:

- B : viajar en bicicleta.
- M : viajar en micro.
- A : viajar en automóvil.
- T : llegar tarde.

Escribamos $r = \mathbb{P}(T \mid B) > 0$. Los datos entregan

$$\mathbb{P}(B) = \mathbb{P}(A) = 0,2, \quad \mathbb{P}(M) = 0,6, \quad \mathbb{P}(T \mid A) = 2r, \quad \mathbb{P}(T \mid M) = 4r.$$

- (a). Por probabilidades totales,

$$\begin{aligned} \mathbb{P}(T) &= \mathbb{P}(T \mid B)\mathbb{P}(B) + \mathbb{P}(T \mid A)\mathbb{P}(A) + \mathbb{P}(T \mid M)\mathbb{P}(M) \\ &= 0,2r + 0,2(2r) + 0,6(4r) = 3r. \end{aligned}$$

El Teorema de Bayes entrega

$$\begin{aligned} \mathbb{P}(M \mid T) &= \frac{\mathbb{P}(T \mid M)\mathbb{P}(M)}{\mathbb{P}(T)} \\ &= \frac{4r(0,6)}{3r} = 0,8. \end{aligned}$$

- (b). Si $\mathbb{P}(T) = 0,6$, entonces $r = 0,2$. Por tanto,

$$\mathbb{P}(T^c \mid A) = 1 - \mathbb{P}(T \mid A) = 1 - 2r = 0,6.$$

Pregunta 22. (Control 1 - Primavera 2023, P2)

Don Eustacio vive en su departamento, en el cual posee tres plantas desde hace un tiempo. Don Eustacio se va a ir de vacaciones fuera de la ciudad por una semana, por lo que deja encargado del riego diario de las plantas a su vecina, Antonia. Antonia es algo olvidadiza, por lo que existe la posibilidad de que se le olvide ir todos los días. En particular, considere que hay un 50 % de probabilidad de que Antonia vaya a regar todos los días, un 30 % de que vaya a regar día por medio, y un 20 % de que no vaya a regar ningún día. Además, se sabe que si es que el riego de las plantas se realiza durante todos los días durante la semana indicada,

entonces la probabilidad de que cada planta florezca es de un 80 %. Sin embargo, en caso de que las plantas se rieguen día por medio durante la semana, dicha probabilidad es 60 %. Por último, si es que las plantas no se riegan durante la semana, entonces la probabilidad de florecer de cada planta es de 0 %. Note que dado un escenario específico de riego, las probabilidades de florecer de las plantas son independientes entre sí. Asuma que ninguna planta ha florecido previo a que don Eustacio se vaya de vacaciones.

- a) Al volver de vacaciones don Eustacio mira una de las plantas al azar, y observa que esta no ha florecido, ¿cuál es la probabilidad de que Antonia haya olvidado por completo ir a regar durante la semana?
- b) Al volver de vacaciones, don Eustacio ha mirado dos de sus plantas, de las cuales ninguna ha florecido. Don Eustacio va a la cocina a ver una tercera planta. ¿Cuál es la probabilidad de que la planta que se encuentra en la cocina haya florecido?

Solución

Introduzcamos los tres esquemas de riego:

- R_T : regar todos los días.
- R_M : regar día por medio.
- R_N : no regar ningún día.

(a). Si F indica que la planta florece, por probabilidades totales,

$$\begin{aligned}\mathbb{P}(F^c) &= \sum_{R \in \{R_T, R_M, R_N\}} \mathbb{P}(F^c | R) \mathbb{P}(R) \\ &= 0,5(0,2) + 0,3(0,4) + 0,2(1) = 0,42.\end{aligned}$$

Luego, el Teorema de Bayes entrega

$$\begin{aligned}\mathbb{P}(R_N | F^c) &= \frac{\mathbb{P}(F^c | R_N) \mathbb{P}(R_N)}{\mathbb{P}(F^c)} \\ &= \frac{1 \cdot 0,2}{0,42} = \frac{10}{21} \approx 0,4762.\end{aligned}$$

(b). Sean F_i los eventos de florecimiento de las tres plantas. Por independencia condicional dado el esquema de riego,

$$\begin{aligned}\mathbb{P}(F_3 \cap F_1^c \cap F_2^c) &= \sum_{R \in \{R_T, R_M, R_N\}} \mathbb{P}(F_3 \cap F_1^c \cap F_2^c | R) \mathbb{P}(R) \\ &= 0,5(0,8)(0,2)^2 + 0,3(0,6)(0,4)^2 = 0,0448, \\ \mathbb{P}(F_1^c \cap F_2^c) &= \sum_{R \in \{R_T, R_M, R_N\}} \mathbb{P}(F_1^c \cap F_2^c | R) \mathbb{P}(R) \\ &= 0,5(0,2)^2 + 0,3(0,4)^2 + 0,2 = 0,268.\end{aligned}$$

En consecuencia,

$$\mathbb{P}(F_3 | F_1^c \cap F_2^c) = \frac{0,0448}{0,268} = \frac{56}{335} \approx 0,1672.$$

Pregunta 23. (Control 1 - Otoño 2024, P2)

- (a) Pruebe que para cualquier par de eventos $A, B \subseteq \Omega$, y para cualquier partición de Ω $C_1, C_2, \dots, C_n$, se cumple la siguiente igualdad:

$$\mathbb{P}(A | B) = \sum_{i=1}^n \mathbb{P}(A | B \cap C_i) \cdot \mathbb{P}(C_i | B)$$

- (b) Se está realizando un concurso de patear penales, el concurso consiste en patear 3 penales, de modo que si anotas al menos 2 penales, ganas un premio. Previo a los tres lanzamientos, se debe escoger un balón de las dos disponibles que hay: uno de marca Nike, y otro de marca Adidas. Se sabe que el 70 % de las veces que se patea con el balón Nike, el lanzamiento termina en gol; mientras que con el balón Adidas el lanzamiento termina en gol solo un 50 % de las veces. Los balones han sido pintados por fuera, de forma que son indistinguibles para el pateador, quien no sabe qué pelota ha escogido (ni siquiera una vez que ha lanzado los tres penales). Por lo anterior, se puede asumir que la elección inicial de balón tiene igual probabilidad para cada uno de los balones disponibles. Note que los tres lanzamientos se hacen con el mismo balón escogido inicialmente.

Si es que un jugador ya ha pateado sus dos primeros penales (le falta solo el tercero), en los cuales anotó gol en solo uno de ellos, calcule la probabilidad de que termine ganando el premio.

Solución

(a). Supongamos $\mathbb{P}(B) > 0$. Como $C_1, \dots, C_n$ es una partición, los eventos $A \cap B \cap C_i$ son disjuntos y su unión es $A \cap B$. Definamos

$$I_+ = \{i : \mathbb{P}(B \cap C_i) > 0\}.$$

Para estos índices, todas las probabilidades condicionales que aparecen a continuación están bien definidas. Por tanto,

$$\begin{aligned} \mathbb{P}(A | B) &= \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \sum_{i=1}^n \frac{\mathbb{P}(A \cap B \cap C_i)}{\mathbb{P}(B)} \\ &= \sum_{i \in I_+} \frac{\mathbb{P}(A \cap B \cap C_i)}{\mathbb{P}(B \cap C_i)} \frac{\mathbb{P}(B \cap C_i)}{\mathbb{P}(B)} \\ &= \sum_{i \in I_+} \mathbb{P}(A | B \cap C_i) \mathbb{P}(C_i | B). \end{aligned}$$

Esta es la versión rigurosa de la identidad. La suma escrita en el enunciado es válida literalmente si además $\mathbb{P}(B \cap C_i) > 0$ para todo i . Cuando alguna celda tiene probabilidad cero, su peso $\mathbb{P}(C_i | B)$ también es cero, pero $\mathbb{P}(A | B \cap C_i)$ no está definida; por ello, ese índice debe omitirse en vez de asignarle un valor arbitrario.

(b). El enunciado no especifica la dependencia entre los tres lanzamientos, por lo que las probabilidades marginales de gol no bastan para determinar la respuesta. Bajo el supuesto adicional de que, condicionado al balón escogido, los lanzamientos son independientes, introduzcamos los eventos

- D : escoger el balón Adidas.
- N : escoger el balón Nike.
- H : anotar exactamente uno de los dos primeros penales.
- G_3 : anotar el tercer penal.

Entonces

$$\mathbb{P}(H | D) = 0,5, \quad \mathbb{P}(H | N) = 2(0,7)(0,3) = 0,42.$$

Como ambos balones se escogen con probabilidad $1/2$, la ley de probabilidades totales da

$$\mathbb{P}(H) = \mathbb{P}(H | D)\mathbb{P}(D) + \mathbb{P}(H | N)\mathbb{P}(N) = \frac{1}{2}(0,5) + \frac{1}{2}(0,42) = 0,46.$$

El Teorema de Bayes permite actualizar la probabilidad del balón escogido:

$$\begin{aligned} \mathbb{P}(D | H) &= \frac{\mathbb{P}(H | D)\mathbb{P}(D)}{\mathbb{P}(H)} = \frac{0,5(1/2)}{0,46} = \frac{25}{46}, \\ \mathbb{P}(N | H) &= \frac{\mathbb{P}(H | N)\mathbb{P}(N)}{\mathbb{P}(H)} = \frac{0,42(1/2)}{0,46} = \frac{21}{46}. \end{aligned}$$

Para ganar debe anotar el tercer penal. Condicionando respecto del balón,

$$\begin{aligned}\mathbb{P}(G_3 | H) &= \mathbb{P}(G_3 | D, H)\mathbb{P}(D | H) + \mathbb{P}(G_3 | N, H)\mathbb{P}(N | H) \\ &= 0,5 \frac{25}{46} + 0,7 \frac{21}{46} = \frac{68}{115} \approx 0,5913.\end{aligned}$$

Pregunta 24. (Control 1 - Primavera 2024, P2)

En el sector de los seguros de automóviles, se observa que las personas se distribuyen de manera intrínseca en tres niveles de riesgo: Bajo, Medio y Alto. Las probabilidades de que individuos en cada nivel estén involucrados en un **accidente durante un año** son del 0,05 para Bajo riesgo, 0,15 para riesgo Medio y 0,30 Alto riesgo, respectivamente. Adicionalmente, las personas tienen una probabilidad inherente de pertenecer a estas categorías, siendo 0,2 para Bajo Riesgo, 0,5 para Riesgo Medio, y 0,3 para Alto Riesgo.

- Calcule la probabilidad de que una persona elegida al azar tenga un accidente durante el año.
- Calcule la probabilidad de que una persona que no tuvo accidentes durante el año sea de riesgo Bajo o Medio.

Una compañía de seguros categoriza a las personas en tres niveles de riesgo basándose en su historial personal. La precisión de esta clasificación varía según la categoría real de riesgo de la persona. Específicamente, la probabilidad de que una persona clasificada como de Bajo riesgo sea realmente de Bajo riesgo es de 0,7. Además, existe una probabilidad de 0,28 de que una persona clasificada como de Bajo riesgo sea en realidad de riesgo Medio. Finalmente, la probabilidad de que alguien clasificado como de Bajo riesgo sea realmente de Alto riesgo es de 0,02. Además, la compañía clasifica a una persona como Bajo riesgo con probabilidad 0,1.

- Calcule la probabilidad de que una persona que fue clasificada como Bajo riesgo tenga al menos un accidente en **dos años**. Considere que, para un determinado nivel de riesgo¹ de una persona, los accidentes ocurridos en un año y otro son independientes. *Hint: Le podría ser de utilidad usar propiedades de una función de probabilidad cuando esta está condicionada sobre un evento.*
- Finalmente, calcule la probabilidad de que una persona que tuvo **al menos un accidente en dos años**, sea clasificada como Bajo riesgo. Si es que necesita algunas de las expresiones preguntadas de las partes anteriores ((a), (b), o (c)), las puede definir con una variable y usarlas.

Solución

Introduzcamos los eventos asociados al accidente y al riesgo real:

- E : sufrir un accidente durante un año.
- M : pertenecer al nivel de riesgo Medio.
- B : pertenecer al nivel de riesgo Bajo.
- A : pertenecer al nivel de riesgo Alto.

- Los niveles de riesgo forman una partición. Por la ley de probabilidades totales,

$$\begin{aligned}\mathbb{P}(E) &= \mathbb{P}(E | B)\mathbb{P}(B) + \mathbb{P}(E | M)\mathbb{P}(M) + \mathbb{P}(E | A)\mathbb{P}(A) \\ &= 0,05(0,2) + 0,15(0,5) + 0,30(0,3) = 0,175.\end{aligned}$$

- Como $B \cup M = A^c$ dentro de la partición de niveles,

$$\mathbb{P}(B \cup M | E^c) = 1 - \mathbb{P}(A | E^c).$$

¹Nos referimos al riesgo inherente de la persona, no la clasificación.

El Teorema de Bayes y el resultado anterior dan

$$\mathbb{P}(A \mid E^c) = \frac{\mathbb{P}(E^c \mid A)\mathbb{P}(A)}{\mathbb{P}(E^c)} = \frac{(1 - 0,30)0,30}{1 - 0,175} = \frac{14}{55}.$$

Por tanto,

$$\mathbb{P}(B \cup M \mid E^c) = 1 - \frac{14}{55} = \frac{41}{55} \approx 0,7455.$$

(c). Para los dos años, definamos

- C_B : ser clasificado como de Bajo riesgo.
- E_2 : sufrir un accidente durante el segundo año.
- E_1 : sufrir un accidente durante el primer año.

El cálculo requiere suponer, además, que, condicionado al riesgo real R , la clasificación C_B no aporta información sobre los accidentes futuros; en particular,

$$\mathbb{P}(E_1^c \cap E_2^c \mid R \cap C_B) = (1 - q_R)^2.$$

Sin esta hipótesis, los datos no determinan los incisos (c) y (d). Bajo este supuesto, para un nivel $R \in \{B, M, A\}$ con probabilidad anual q_R ,

$$\mathbb{P}(E_1 \cup E_2 \mid R \cap C_B) = 1 - (1 - q_R)^2.$$

Aplicando probabilidades totales respecto del riesgo verdadero y usando las proporciones entregadas para las personas clasificadas como de Bajo riesgo,

$$\begin{aligned} \mathbb{P}(E_1 \cup E_2 \mid C_B) &= \sum_{R \in \{B, M, A\}} \mathbb{P}(E_1 \cup E_2 \mid R \cap C_B) \mathbb{P}(R \mid C_B) \\ &= [1 - (1 - 0,05)^2](0,70) \\ &\quad + [1 - (1 - 0,15)^2](0,28) + [1 - (1 - 0,30)^2](0,02) \\ &= 0,15615. \end{aligned}$$

No corresponde multiplicar las probabilidades marginales de no accidentarse dado C_B : la mezcla de niveles de riesgo hace que los dos años no sean independientes una vez que solo se condiciona en la clasificación.

(d). Primero calculamos la probabilidad no condicionada de sufrir al menos un accidente en dos años. Nuevamente se condiciona en el riesgo verdadero antes de usar independencia:

$$\begin{aligned} \mathbb{P}(E_1 \cup E_2) &= \sum_{R \in \{B, M, A\}} \mathbb{P}(E_1 \cup E_2 \mid R) \mathbb{P}(R) \\ &= [1 - (1 - 0,05)^2](0,20) \\ &\quad + [1 - (1 - 0,15)^2](0,50) + [1 - (1 - 0,30)^2](0,30) \\ &= 0,31125. \end{aligned}$$

El Teorema de Bayes entrega entonces

$$\begin{aligned} \mathbb{P}(C_B \mid E_1 \cup E_2) &= \frac{\mathbb{P}(E_1 \cup E_2 \mid C_B) \mathbb{P}(C_B)}{\mathbb{P}(E_1 \cup E_2)} \\ &= \frac{0,15615(0,1)}{0,31125} \approx 0,0502. \end{aligned}$$

Pregunta 25. (Control 1 - Otoño 2025, P2)

- (a) Un niño va al colegio y su padre y su madre, le preparan la mochila con colaciones cada día. El padre siempre que se despidе de su hijo, coloca un chocolate en la mochila, sin que él se de cuenta. La madre, por su parte, cada mañana decide qué colación le mandará, tal que con probabilidad de $\frac{1}{2}$ le coloca un chocolate en su mochila y con probabilidad complementaria le manda una fruta. Así, la mochila siempre contiene exactamente dos

colaciones. Cuando el niño llega a la escuela, saca una colación al azar, de las dos que hay en su mochila, y se da cuenta que es un chocolate. ¿Cuál es la probabilidad de que la colación que queda dentro de la mochila también sea un chocolate?

- (b) Los padres del niño, orgullosos de su excelente desempeño escolar, deciden premiarlo con un auto de juguete. Como se trata de un modelo nuevo que aún no está disponible en el país, lo compran a través de AliExpress, pero el vendedor es desconocido y no ofrece garantías. El auto puede ser **original** o **falsificado**. Si es **original**, viene con 11 cables completamente funcionales (los 11 cables funcionan correctamente), en donde 5 de los cables corresponden a las luces, 4 para el audio y 2 para el movimiento. Si es **falsificado**, solo 7 de los 11 cables funcionan correctamente, por lo que 4 cables (de los 11) al azar NO funcionan. Por ejemplo, el juguete falsificado podría tener los siguientes cables buenos: 3 cables de luces, 3 audio, y 1 para movimiento (en cuyo caso los 4 cables que no funcionan serían: 2 de luces, 1 de audio y 1 de movimiento). El juguete solo enciende las luces si al menos 3 cables de luces funcionan correctamente. Basados en reseñas generales de la plataforma, estiman que la probabilidad de que el auto sea **original** es $\frac{1}{3}$, y con probabilidad complementaria será **falsificado**. ¿Cuál es la probabilidad de que el auto sea falsificado si, al probarlo, el niño observa que las luces se encienden?

Solución

(a). Introduzcamos los eventos asociados a la mochila y a la extracción:

- D : la mochila contiene dos chocolates.
- S : la primera colación extraída es un chocolate.
- H : la mochila contiene un chocolate y una fruta.

Los eventos D y H forman una partición, con $\mathbb{P}(D) = \mathbb{P}(H) = 1/2$. Además, $\mathbb{P}(S | D) = 1$ y $\mathbb{P}(S | H) = 1/2$. Por la ley de probabilidades totales y el Teorema de Bayes,

$$\begin{aligned}\mathbb{P}(D | S) &= \frac{\mathbb{P}(S | D)\mathbb{P}(D)}{\mathbb{P}(S | D)\mathbb{P}(D) + \mathbb{P}(S | H)\mathbb{P}(H)} \\ &= \frac{1(1/2)}{1(1/2) + (1/2)(1/2)} = \frac{2}{3}.\end{aligned}$$

(b). Introduzcamos los eventos asociados al origen del auto y a su funcionamiento:

- F : el auto es falsificado.
- L : las luces se encienden.
- $O = F^c$: el auto es original.

Se tiene $\mathbb{P}(F) = 2/3$, $\mathbb{P}(O) = 1/3$ y $\mathbb{P}(L | O) = 1$.

Si el auto es falsificado, los cuatro cables defectuosos se eligen uniformemente entre los once. Para que las luces se enciendan, a lo más dos de los cinco cables de luces pueden estar entre los cuatro defectuosos. Por ello,

$$\begin{aligned}\mathbb{P}(L | F) &= \frac{\sum_{k=0}^2 \binom{5}{k} \binom{6}{4-k}}{\binom{11}{4}} \\ &= \frac{15 + 100 + 150}{330} = \frac{53}{66}.\end{aligned}$$

La ley de probabilidades totales da

$$\mathbb{P}(L) = \mathbb{P}(L | O)\mathbb{P}(O) + \mathbb{P}(L | F)\mathbb{P}(F) = \frac{1}{3} + \frac{53}{66} \cdot \frac{2}{3} = \frac{86}{99}.$$

Finalmente, el Teorema de Bayes entrega

$$\mathbb{P}(F | L) = \frac{\mathbb{P}(L | F)\mathbb{P}(F)}{\mathbb{P}(L)} = \frac{(53/66)(2/3)}{86/99} = \frac{53}{86}.$$

Pregunta 26. (Control 1 - Primavera 2025, P1)

- (a) Con el propósito de conocer las preferencias presidenciales de la población de la Región Metropolitana previo a la primera vuelta de 2021, se aplicó una encuesta a personas de distintos rangos etarios. El gráfico de barras muestra, para cada rango, el porcentaje de encuestados que votaría por cada candidato; además, el gráfico de torta indica el porcentaje de encuestados por rango etario. Note que el número de rangos etarios de ambos gráficos no son iguales.

Candidato/a	Rango Etario			
	18-30	31-40	41-50	51-80
José Antonio Kast	20.5	35.1	33.4	33.3
Gabriel Boric	39.3	18.3	19.2	20.3
Yasna Provoste	8.6	11.4	17.3	21.4
Sebastián Sichel	11.9	11.9	11.3	11.1
Franco Parisi	11.8	16.4	9.2	6.0
Marco Enríquez-O.	4.6	5.6	7.0	5.1
Eduardo Artés	3.3	1.3	2.6	2.8

Cuadro 3.1: Preferencia por rango etario (%).

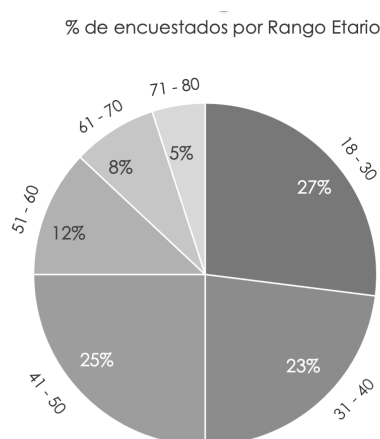


Figura 3.1: Distribución de preferencias.

Dado que una persona votó por Gabriel Boric, determine cuál es la probabilidad de que esta persona pertenezca al rango etario 18-30. Sea claro en la definición de los eventos a utilizar.

- (b) Se eligen aleatoriamente dos bolas de una urna que contiene 8 **blancas**, 4 **negras** y 2 **naranjas**. Suponga que ganamos \$2 por cada bola **negra** seleccionada y perdemos \$1 por cada bola **blanca** seleccionada. Sea X la ganancia obtenida. ¿Cuáles son los posibles valores de X ? Encuentre la probabilidad de perder plata, utilizando la función de distribución acumulada y función de probabilidad de masa de la variable aleatoria X .

Solución

(a). Introduzcamos los eventos asociados al voto y al rango etario:

- C : la persona votó por Gabriel Boric.
- E_1 : la persona tiene entre 18 y 30 años.
- E_2 : la persona tiene entre 31 y 40 años.
- E_3 : la persona tiene entre 41 y 50 años.
- E_4 : la persona tiene entre 51 y 80 años.

Los porcentajes de la tabla corresponden a

$$(\mathbb{P}(C | E_1), \dots, \mathbb{P}(C | E_4)) = (0,393, 0,183, 0,192, 0,203).$$

El gráfico de torta usa rangos más finos. Para compatibilizarlos con los de la tabla, el intervalo 51-80 reúne tres porciones, cuya masa total es $0,12 + 0,08 + 0,05 = 0,25$. Por tanto, las probabilidades

previas son

$$(\mathbb{P}(E_1), \dots, \mathbb{P}(E_4)) = (0,27, 0,23, 0,25, 0,25).$$

Por la ley de probabilidades totales,

$$\begin{aligned}\mathbb{P}(C) &= \sum_{j=1}^4 \mathbb{P}(C | E_j) \mathbb{P}(E_j) \\ &= 0,393(0,27) + 0,183(0,23) + 0,192(0,25) + 0,203(0,25) \\ &= 0,24695.\end{aligned}$$

El Teorema de Bayes da entonces

$$\mathbb{P}(E_1 | C) = \frac{\mathbb{P}(C | E_1) \mathbb{P}(E_1)}{\mathbb{P}(C)} = \frac{0,393(0,27)}{0,24695} \approx 0,430.$$

(b). Si se extraen dos bolas, X es dos veces el número de bolas negras menos el número de bolas blancas. Al enumerar las combinaciones de colores posibles se obtiene

$$\text{rango}(X) = \{-2, -1, 0, 1, 2, 4\}.$$

En efecto, dos blancas dan -2 ; una blanca y una naranja, -1 ; dos naranjas, 0 ; una negra y una blanca, 1 ; una negra y una naranja, 2 ; y dos negras, 4 .

Perder dinero equivale a $X < 0$. Como no hay masas entre -1 y 0 , esta probabilidad puede escribirse mediante la CDF y la PMF como

$$\begin{aligned}\mathbb{P}(X < 0) &= F_X(-1) = p_X(-2) + p_X(-1) \\ &= \frac{\binom{8}{2}}{\binom{14}{2}} + \frac{\binom{8}{1}\binom{2}{1}}{\binom{14}{2}} \\ &= \frac{28 + 16}{91} = \frac{44}{91}.\end{aligned}$$

Pregunta 27. (Control 1 - Otoño 2026, P1)

- (a) Sean A , B y C tres eventos definidos en un espacio de probabilidad. Se conocen los siguientes valores:

$$\mathbb{P}(A \cap B) = 0,30, \quad \mathbb{P}(A \cap C) = 0,25, \quad \mathbb{P}(A \cap B \cap C) = 0,10.$$

¿Cuál es el valor de $\mathbb{P}(A \cap (B \cup C))$?

- (b) Suponga que está en una fiesta lejos de su hogar y tiene que volverse en Uber a su casa. Dependiendo de qué tan buena está la fiesta, usted decide si irse a las 3 am o quedarse media hora más. Usted decide quedarse media hora más con probabilidad 0,6. Una vez que se va, es posible que el Uber lo venga a buscar de inmediato o que tenga que esperarlo media hora debido a la alta demanda. La probabilidad de que el Uber tarde media hora en llegar a buscarlo es 0,5 a las 3 am y 0,7 a las 3:30 am. Una vez que el Uber lo pasa a buscar, tarda media hora en llegar a su casa, independientemente de la hora.
- ¿Cuál es la probabilidad de llegar a las 4 am a su casa?
 - Si sabemos que llegó a las 4 am, ¿cuál es la probabilidad de haber pedido el Uber a las 3:30?

Solución

(a). La ley distributiva permite escribir $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$. Como ambas

intersecciones se traslapan en $A \cap B \cap C$, inclusión-exclusión da

$$\begin{aligned}\mathbb{P}(A \cap (B \cup C)) &= \mathbb{P}(A \cap B) + \mathbb{P}(A \cap C) - \mathbb{P}(A \cap B \cap C) \\ &= 0,30 + 0,25 - 0,10 = 0,45.\end{aligned}$$

(b)(i). Introduzcamos los eventos asociados al viaje:

- V : pedir el Uber a las 3 am.
- L : llegar a la casa a las 4 am.
- E : esperar media hora al Uber.

Llegar a las 4 am ocurre en dos casos disjuntos: pedir el Uber a las 3 am y esperar, o pedirlo a las 3:30 y no esperar. Por tanto,

$$L = (V \cap E) \cup (V^c \cap E^c).$$

Usando la regla de la cadena en ambos casos,

$$\begin{aligned}\mathbb{P}(L) &= \mathbb{P}(E | V)\mathbb{P}(V) + \mathbb{P}(E^c | V^c)\mathbb{P}(V^c) \\ &= 0,5(1 - 0,6) + (1 - 0,7)0,6 = 0,38.\end{aligned}$$

(b)(ii). El evento de pedir el Uber a las 3:30 y llegar a las 4 am es $V^c \cap E^c$. Por definición de probabilidad condicional,

$$\begin{aligned}\mathbb{P}(V^c | L) &= \frac{\mathbb{P}(V^c \cap E^c)}{\mathbb{P}(L)} \\ &= \frac{\mathbb{P}(E^c | V^c)\mathbb{P}(V^c)}{\mathbb{P}(L)} \\ &= \frac{0,3(0,6)}{0,38} = \frac{9}{19} \approx 0,474.\end{aligned}$$

Esta es también la aplicación directa del Teorema de Bayes, pues $\mathbb{P}(L | V^c) = \mathbb{P}(E^c | V^c) = 0,3$.

Pregunta 28. (Control 1 - Otoño 2026, P2)

En la genética, para cada posible característica (color de pelo, altura, etc.) una persona tiene dos genes: uno heredado del padre y uno de la madre. Dado que el padre y la madre tienen dos genes cada uno, cada gen de cada progenitor se hereda al hijo con probabilidad 0,5, de manera independiente entre la madre y el padre. Es decir, si la madre tiene genes XY y el padre tiene genes ZW, el/la descendiente tendrá genes XZ, XW, YZ o YW, cada opción con la misma probabilidad. Para cada característica, hay un gen dominante (que denotaremos con letra mayúscula) y uno recesivo (que denotaremos con letra minúscula). Los genes dominantes siempre “ganan” a los recesivos en términos de expresión: si alguien tiene un gen dominante y uno recesivo, el dominante se expresará. Si alguien tiene los dos genes del mismo tipo, ese será el tipo que se exprese.

Para las siguientes preguntas, considere sabido que Adriana y Bruno tienen genes Aa , y que Camila es hija de Adriana y Bruno.

- (a) ¿Cuál es la probabilidad de que Camila exprese la característica a ?
- (b) Fernanda es hija de Camila y Ernesto (Ernesto no tiene ninguna relación sanguínea con Camila). Se sabe que tanto Fernanda como Ernesto tienen genes Aa . ¿Cuál es la probabilidad de que Camila tenga genes AA ?

Indicación: Si le resulta útil, puede usar la siguiente versión condicional del Teorema de Bayes: Si A, B, C son eventos con $\mathbb{P}(A \cap C) > 0$ y $\mathbb{P}(B \cap C) > 0$, entonces $\mathbb{P}(A | B \cap C) = \frac{\mathbb{P}(B|A \cap C)\mathbb{P}(A|C)}{\mathbb{P}(B|C)}$.

Solución

(a). Para expresar la característica recesiva, Camila debe heredar el gen a de ambos progenitores. Estas herencias son independientes y cada una tiene probabilidad $1/2$. Por tanto,

$$\mathbb{P}(\text{Camila expresa } a) = \frac{1}{2} \frac{1}{2} = \frac{1}{4}.$$

(b). Para cada genotipo g , introduzcamos los eventos

- $C : g$: Camila tiene genotipo g .
- $F : g$: Fernanda tiene genotipo g .
- $E : g$: Ernesto tiene genotipo g .

Como Adriana y Bruno son ambos Aa , la distribución inicial del genotipo de Camila es

$$\mathbb{P}(C : AA) = \frac{1}{4}, \quad \mathbb{P}(C : Aa) = \frac{1}{2}, \quad \mathbb{P}(C : aa) = \frac{1}{4}.$$

Bajo el supuesto implícito sugerido por el enunciado de que los genotipos de Camila y Ernesto son independientes, y dado que Ernesto es Aa , Fernanda tiene genotipo Aa con probabilidad $1/2$ para cualquiera de los tres genotipos posibles de Camila:

$$\begin{aligned} \mathbb{P}(F : Aa \mid C : AA \cap E : Aa) &= \frac{1}{2}, \\ \mathbb{P}(F : Aa \mid C : Aa \cap E : Aa) &= \frac{1}{2}, \\ \mathbb{P}(F : Aa \mid C : aa \cap E : Aa) &= \frac{1}{2}. \end{aligned}$$

En el primer caso, Ernesto debe transmitir a ; en el tercero, debe transmitir A ; y en el caso intermedio los dos cruces heterocigotos producen un descendiente Aa en dos de las cuatro combinaciones equiprobables.

La ley de probabilidades totales muestra que $\mathbb{P}(F : Aa \mid E : Aa) = 1/2$. Como el genotipo de Camila es independiente del de Ernesto, el Teorema de Bayes condicional entrega

$$\begin{aligned} &\mathbb{P}(C : AA \mid (F : Aa) \cap (E : Aa)) \\ &= \frac{\mathbb{P}(F : Aa \mid (C : AA) \cap (E : Aa)) \mathbb{P}(C : AA \mid E : Aa)}{\mathbb{P}(F : Aa \mid E : Aa)} \\ &= \frac{(1/2)(1/4)}{1/2} \\ &= \frac{1}{4}. \end{aligned}$$

La notación genética Aa describe un genotipo no ordenado: no identifica cuál alelo provino de la madre o del padre. Tratar Aa y aA como eventos diferentes respondería una pregunta distinta de la formulada.

Capítulo 4

Principio de Conteo

4.1. Técnicas de conteo

Resumen

Principios de conteo.

Antes de contar se debe precisar si importa el orden, si se permiten repeticiones y qué configuraciones se consideran equivalentes.

- **Principio aditivo.** Si los casos son disjuntos y tienen $n_1, \dots, n_r$ resultados, su unión tiene $n_1 + \dots + n_r$ resultados.
- **Principio multiplicativo.** Si una construcción sucesiva ofrece n_i opciones en la etapa i , existen $\prod_{i=1}^r n_i$ resultados.
- **Conteo por complemento.** El número de configuraciones que satisfacen una condición es el total menos las que la incumplen.

Cuatro esquemas básicos.

Al escoger k elementos a partir de n tipos:

- **Ordenado con reemplazo.** Hay n^k secuencias.
- **Ordenado sin reemplazo.** Hay $(n)_k = n!/(n-k)!$ secuencias.
- **No ordenado sin reemplazo.** Hay $\binom{n}{k} = n!/[k!(n-k)!]$ subconjuntos.
- **No ordenado con reemplazo.** Hay $\binom{n+k-1}{k}$ multiconjuntos.

Estrellas y barras.

Las soluciones enteras no negativas de $x_1 + \dots + x_n = k$ se cuentan mediante

$$\binom{n+k-1}{n-1}.$$

Si se exige $x_i \geq 1$ para todo i , se escribe $y_i = x_i - 1$ y el número de soluciones es $\binom{k-1}{n-1}$, para $k \geq n$.

Coefficientes binomiales.

Las identidades más utilizadas son:

- **Simetría.** $\binom{n}{k} = \binom{n}{n-k}$.

- **Relación de Pascal.** $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$.
- **Suma de subconjuntos.** $\sum_{k=0}^n \binom{n}{k} = 2^n$.
- **Identidad de Vandermonde.**

$$\sum_j \binom{m}{j} \binom{n}{k-j} = \binom{m+n}{k},$$

con la suma sobre los índices factibles.

Distribución en grupos.

Si $n_1 + \dots + n_r = n$, el número de formas de distribuir n objetos distintos en grupos rotulados de esos tamaños es

$$\binom{n}{n_1, \dots, n_r} = \frac{n!}{n_1! \dots n_r!}.$$

Si los grupos no están rotulados, se debe dividir además por las permutaciones de los grupos indistinguibles.

Órdenes circulares y simetrías.

- **Permutación circular.** n objetos distintos alrededor de un círculo admiten $(n-1)!$ órdenes cuando solo se identifican rotaciones.
- **Lema de Burnside.** Si un grupo G de simetrías actúa sobre las configuraciones, el número de clases distintas es

$$\frac{1}{|G|} \sum_{g \in G} |\text{Fix}(g)|.$$

Restricciones de posición.

Los objetos que deben permanecer juntos pueden tratarse como un bloque. Las restricciones de separación se organizan mediante espacios o posiciones disponibles. En particular, la cantidad de formas de elegir k posiciones no consecutivas entre N posiciones en línea es

$$\binom{N-k+1}{k}.$$

4.1.1. Preguntas

Pregunta 29.

Un grupo de Taller de Liderazgo conformado por 15 mujeres y 5 hombres se deben separar en dos sub-grupos: 10 personas para el proyecto de PACE (P) y 10 para el proyecto de Reinserción (R). Además, a cada proyecto se debe asignar un jefe o una jefa de proyecto entre sus integrantes.

- Determine la cantidad de formas que hay de asignar las personas a los proyectos y a un jefe o jefa de proyecto en cada uno de estos.
- Si todas las asignaciones son igualmente probables, calcule la probabilidad de que el proyecto de PACE no incluya hombres.
- Calcule la probabilidad de que las dos jefaturas sean ocupadas por mujeres.
- Calcule la probabilidad de que, simultáneamente, el proyecto de PACE no incluya hombres y que las dos jefaturas sean ocupadas por mujeres.
- Usando los apartados anteriores, calcule la probabilidad de que, simultáneamente, el proyecto PACE incluya al menos un hombre y que al menos una jefatura sea ocupada por un hombre.

Solución

(a). La cantidad total de formas de asignar a 20 personas en dos grupos de 10 y seleccionar un jefe para cada grupo se calcula como sigue: primero se escogen las 10 personas de PACE, lo que puede hacerse de $\binom{20}{10}$ maneras. El grupo de Reinserción queda determinado por las 10 personas restantes. Una vez formados los grupos, hay 10 opciones para cada jefatura. Por el principio multiplicativo, el número total es

$$\binom{20}{10} \cdot 10 \cdot 10.$$

(b). La composición de PACE es uniforme entre los subconjuntos de 10 personas. Para que no incluya hombres, sus 10 integrantes deben escogerse entre las 15 mujeres. Las elecciones de jefatura se cancelan en el cociente, por lo que

$$\mathbb{P}(\text{PACE no incluye hombres}) = \frac{\binom{15}{10}}{\binom{20}{10}}.$$

(c). Las dos jefaturas corresponden a personas distintas. Bajo las asignaciones equiprobables, la primera jefatura es uniforme entre las 20 personas y, fijada esta, la segunda es uniforme entre las 19 restantes. Por lo tanto,

$$\mathbb{P}(\text{ambas jefaturas son ocupadas por mujeres}) = \frac{15}{20} \frac{14}{19} = \frac{21}{38}.$$

(d). Si PACE no incluye hombres, se escogen sus 10 integrantes entre las 15 mujeres. Su jefatura tiene entonces 10 opciones femeninas. Reinserción queda con las cinco mujeres restantes y los cinco hombres, de modo que su jefatura debe escogerse entre cinco mujeres. Así,

$$\mathbb{P}(\text{PACE sin hombres y ambas jefaturas femeninas}) = \frac{\binom{15}{10} \cdot 10 \cdot 5}{\binom{20}{10} \cdot 10 \cdot 10} = \frac{1}{2} \frac{\binom{15}{10}}{\binom{20}{10}}.$$

(e). Introduzcamos los eventos

- H : PACE incluye al menos un hombre.
- J : ambas jefaturas son ocupadas por mujeres.

Por las leyes de De Morgan e inclusión-exclusión,

$$\begin{aligned}\mathbb{P}(H \cap J^c) &= 1 - \mathbb{P}(H^c) - \mathbb{P}(J) + \mathbb{P}(H^c \cap J) \\ &= \frac{17}{38} - \frac{1}{2} \frac{\binom{15}{10}}{\binom{20}{10}}.\end{aligned}$$

Aquí se usó $\mathbb{P}(H^c) = \binom{15}{10}/\binom{20}{10}$, $\mathbb{P}(J) = 21/38$ y $\mathbb{P}(H^c \cap J) = \frac{1}{2} \binom{15}{10}/\binom{20}{10}$.

Pregunta 30.

- (a) Calcule la probabilidad de que nadie en esta sección tenga el mismo cumpleaños que el auxiliar (4 de Julio, asumiendo 365 días equiprobables y que la cantidad de alumnos inscritos es 101).

- (b) ¿Cuál es la probabilidad de que en un grupo de n personas, ningún cumpleaños se repita?

Indicación: Calcule primero la probabilidad de que ninguna persona cumpla años el mismo día que otra y luego realice estimaciones para distintos n (recomendable para $n = 23$ y $n = 57$). Investigue la “Paradoja del cumpleaños” sus aplicaciones.

Solución

(a). Bajo la hipótesis adicional usual de que los cumpleaños de las distintas personas son independientes, para cada estudiante la probabilidad de no cumplir años el 4 de julio es $364/365$. Por independencia,

$$\mathbb{P}(\text{ningún estudiante cumple el 4 de julio}) = \left(\frac{364}{365}\right)^{101}.$$

(b). Bajo la misma hipótesis, para que los n cumpleaños sean distintos, la primera persona puede cumplir años cualquier día, la segunda debe evitar ese día y así sucesivamente. Para $1 \leq n \leq 365$,

$$\mathbb{P}(\text{sin cumpleaños repetidos}) = \prod_{i=0}^{n-1} \frac{365-i}{365} = \frac{365!}{(365-n)! 365^n}.$$

Si $n > 365$, esta probabilidad es 0. En particular,

n	$\mathbb{P}(\text{sin repetición})$	$\mathbb{P}(\text{alguna coincidencia})$
23	0,4927028	0,5072972
57	0,0098775	0,9901225

El rápido crecimiento de la segunda probabilidad es la llamada paradoja del cumpleaños y fundamenta, por ejemplo, los ataques de colisión en criptografía.

Pregunta 31.

Un grupo de 4 alumnos a_1 , a_2 , a_3 y a_4 deben realizar una tarea que consta de 12 problemas (enumerados de 1 a 12). Debido a la carga académica, encuentran que va a ser difícil resolver en conjunto los problemas, por lo que deciden repartirse los problemas para resolver cada uno 3 problemas de forma independiente:

- (a) ¿De cuántas maneras se pueden repartir los problemas?
- (b) ¿Cuál es la probabilidad de que el alumno p_1 le toque hacer problemas consecutivos?
- (c) ¿Cuál es la probabilidad de que a todos les toque hacer problemas consecutivos?

Solución

(a). Se escogen sucesivamente los tres problemas de cada estudiante; el último grupo queda determinado. La cantidad de reparticiones es

$$\binom{12}{3} \binom{9}{3} \binom{6}{3} \binom{3}{3} = \frac{12!}{(3!)^4}.$$

(b). Un estudiante designado puede recibir cualquiera de los diez bloques $\{j, j+1, j+2\}$, con $j = 1, \dots, 10$. Como su grupo de tres problemas es uniforme entre los $\binom{12}{3}$ subconjuntos,

$$\mathbb{P}(\text{el estudiante recibe problemas consecutivos}) = \frac{10}{\binom{12}{3}} = \frac{1}{22}.$$

(c). Si los cuatro grupos son consecutivos y forman una partición de $\{1, \dots, 12\}$, necesariamente son

$$\{1, 2, 3\}, \quad \{4, 5, 6\}, \quad \{7, 8, 9\}, \quad \{10, 11, 12\}.$$

Estos cuatro bloques pueden asignarse a los estudiantes de $4!$ maneras. Por tanto,

$$\mathbb{P}(\text{todos reciben problemas consecutivos}) = \frac{4!}{\binom{12}{3} \binom{9}{3} \binom{6}{3} \binom{3}{3}}.$$

Pregunta 32.

Un grupo $n < 365$ amigos se reúne a celebrar cada día que uno (o varios) de ellos están de cumpleaños. Calcule la probabilidad de que se celebren:

- (a) Solo una vez.
- (b) Exactamente 2 veces.
- (c) Exactamente n veces.
- (d) Menos de n veces.

Solución

Bajo el modelo usual en que los cumpleaños son independientes y uniformes entre los 365 días del año, cada configuración es una n -tupla de elementos de $\{1, \dots, 365\}$. Las 365^n tuplas son equiprobables. (a). Para reunirse una sola vez, todas las coordenadas deben ser iguales. Hay una tupla por cada día del año, de modo que

$$\mathbb{P}(\text{una celebración}) = \frac{365}{365^n} = \frac{1}{365^{n-1}}.$$

(b). Primero se eligen los dos días distintos, lo que puede hacerse de $\binom{365}{2}$ maneras. Fijados esos días, entre las 2^n asignaciones de cumpleaños exactamente dos no usan ambos días: las dos funciones constantes. Por tanto,

$$\mathbb{P}(\text{dos celebraciones}) = \frac{\binom{365}{2}(2^n - 2)}{365^n}.$$

Para $n = 2$, esta expresión se reduce a $364/365$, como corresponde a la probabilidad de que dos personas cumplan años en días distintos.

(c). Reunirse exactamente n veces exige que todos los cumpleaños sean distintos. El número de asignaciones inyectivas es $365 \cdot 364 \cdots (365 - n + 1) = 365!/(365 - n)!$, por lo que

$$\mathbb{P}(n \text{ celebraciones}) = \frac{365!}{365^n (365 - n)!}.$$

(d). Hay menos de n celebraciones si y sólo si al menos dos personas comparten cumpleaños.

Este evento es el complemento del inciso anterior; luego,

$$\mathbb{P}(\text{menos de } n \text{ celebraciones}) = 1 - \frac{365!}{365^n(365 - n)!}.$$

Pregunta 33.

Suponga que usted forma parte de un grupo de 4 personas DIM y 4 personas DII que deben ir a una fiesta de disfraces para lo cual disponen de 8 disfraces (4 disfraces DIM y 4 disfraces DII) distintos.

- ¿De cuántas maneras se pueden asignar los disfraces?
- ¿De cuántas maneras se pueden asignar los disfraces si se debe respetar la especialidad?
- Si los disfraces por especialidad son idénticos ¿Cómo cambia su respuesta de las partes anteriores?
- Suponga ahora disfraces distintos y que 2 DIM y 1 DII jamás se pondrían un disfraz contrario a su especialidad. Si los disfraces son asignados al azar, calcule la probabilidad de que las 8 personas queden conformes.

Solución

(a). Como los ocho disfraces son distintos, asignarlos a las ocho personas equivale a construir una biyección. Hay $8!$ asignaciones.

(b). Al respetar la especialidad, se asignan independientemente los cuatro disfraces DIM y los cuatro disfraces DII. Hay $(4!)^2$ asignaciones.

(c). Si los disfraces de una misma especialidad son idénticos, una asignación queda determinada por las cuatro personas que reciben un disfraz DIM. Sin la restricción de especialidad hay $\binom{8}{4}$ asignaciones; al respetarla, existe una sola.

(d). Los dos integrantes DIM con restricción pueden recibir disfraces DIM distintos de $4!/2!$ maneras, y el integrante DII con restricción tiene cuatro disfraces posibles. Los cinco disfraces restantes se asignan libremente a las otras cinco personas, lo que da $5!$ posibilidades. Entre las $8!$ asignaciones equiprobables,

$$\mathbb{P}(\text{todas las personas quedan conformes}) = \frac{(4!/2!)45!}{8!} = \frac{1}{7}.$$

4.1.2. Preguntas de Evaluación

Pregunta 34. (Control 1 - Otoño 2021, P2)

Se ha inaugurado una nueva galería de arte en la cual se exhibirán grandes trabajos de pintores y fotógrafos del país. Para cada situación presentada a continuación conteste las siguientes preguntas. Para las partes 2 y 3, basta con dejar expresada la fórmula de la probabilidad, no es necesario computar el valor exacto. Recuerde que sí debe incluir su razonamiento.

- Uno de los grados atractivos de esta galería de arte será “El gran salón”, el cual cuenta con m posiciones específicas para exponer cuadros grandes y n para cuadros pequeños. Se sabe además que se tiene un total de r cuadros grandes y s cuadros pequeños a ser potencialmente incluidos en “El gran salón”. Sin embargo, no es posible incluir todos los cuadros grandes en “El gran salón”, pues la cantidad de cuadros grandes es mayor a la

cantidad de espacios disponibles para cuadros grandes. Similarmente ocurre lo mismo para los cuadros pequeños, es decir, no es posible poner todos los cuadros pequeños en “El Gran Salón”. Además, no se pueden exponer cuadros pequeños en posiciones destinadas para cuadros grandes, ni vice versa. ¿Cuántas configuraciones diferentes existen para “El gran salón”? (Notar que dos configuraciones son diferentes si es que en al menos una posición poseen un cuadro distinto. Además, una configuración tiene que tener todas las posiciones disponibles de “El Gran Salón” con un cuadro de su respectivo tipo).

- (b) Otro de los grandes salones es el “Salón circular”. Este es un salón donde se exponen fotografías, las cuales están una al lado de la otra completando la circunferencia del salón. Suponga que hay un total de L fotografías, donde cada fotografía está dos veces: una vez en colores y una vez en blanco y negro. El “Salón circular” tiene $2L$ posiciones disponibles. Similarmente al “El gran salón”, una persona todas las mañana ubica las fotografías al azar en el “Salón circular”. ¿Cuál es la probabilidad de que las fotografías del “Salón circular” queden intercaladas? Es decir, que una fotografía a color tenga a sus dos lados fotografías en blanco y negro, y similarmente, que a ambos costados de cada fotografía en blanco y negro hayan fotografías a color. *Hint*: Podría ser de utilidad pensar en “fijar la posición de uno de los cuadros”.
- (c) Otro de los grandes atractivos, es el “Salón de Pigmentos”. En este se pueden crear diferentes colores según desee la persona. Para crear un color, se tiene una máquina que extrae diferentes cantidades de los tres pigmentos base: “Rojo”, “Azul”, y “Amarillo”. El color está listo una vez que se tiene un total de V *unidades de volumen* de pigmentos, los cuales son mezclados hasta obtener un color homogéneo. La máquina extrae de a una *unidad de volumen* a la vez de cada pigmento, y puede repetir el proceso varias veces. No se pueden extraer cantidades fraccionales de *unidades de volumen* de los pigmentos. Su amigo Esteban visita el “Salón de Pigmentos”, y decide crear colores. Sin embargo, dado su gusto por los pigmentos “Rojo” y “Azul”, Esteban incluirá **al menos** R *unidades de volumen* del pigmento “Rojo”, y A *unidades de volumen* del pigmento “Azul”. No hay restricción sobre el número de *unidades de volumen* a usar del pigmento “Amarillo”. ¿Cuántos son el total de colores que Esteban puede crear? Notar que la creación de un color es conmutativa, es decir, el color depende solo de la cantidad de *unidades de volumen* que posea de cada tipo de pigmento, y no de la secuencia en que éstos fueron añadidos por la máquina.

Solución

(a). Para los cuadros grandes se deben elegir y ordenar m de los r cuadros disponibles, lo que da $r!/(r-m)!$ asignaciones. De manera análoga, los cuadros pequeños admiten $s!/(s-n)!$ asignaciones. Como ambas elecciones se realizan de manera independiente, el principio multiplicativo entrega

$$\frac{r!}{(r-m)!} \frac{s!}{(s-n)!}.$$

(b). Fijemos una fotografía a color para eliminar la simetría producida por las rotaciones del círculo. Las restantes $L-1$ fotografías a color pueden ordenarse en sus lugares de $(L-1)!$ maneras y las L fotografías en blanco y negro, en los lugares intermedios, de $L!$ maneras. Por otra parte, una vez fijada una fotografía, las $2L-1$ restantes admiten $(2L-1)!$ órdenes circulares. Por lo tanto,

$$\mathbb{P}(\text{fotografías intercaladas}) = \frac{(L-1)!L!}{(2L-1)!}.$$

Equivalentemente, se pueden ordenar por separado los dos círculos de L fotografías y contar las L formas de intercalarlos, lo que conduce al mismo numerador $(L-1)!^2 L = (L-1)!L!$.

(c). Sean x_1, x_2, x_3 las cantidades de pigmento rojo, azul y amarillo. Estas cantidades satisfacen

$$x_1 + x_2 + x_3 = V, \quad x_1 \geq R, \quad x_2 \geq A, \quad x_3 \geq 0.$$

Al definir $y_1 = x_1 - R$, $y_2 = x_2 - A$ y $y_3 = x_3$, el problema se reduce a contar las soluciones enteras no negativas de

$$y_1 + y_2 + y_3 = V - R - A.$$

Si $V \geq R + A$, el método de estrellas y barras da

$$\binom{V - R - A + 2}{2} = \binom{V - R - A + 2}{V - R - A}.$$

Si $V < R + A$, no existe ninguna configuración factible.

Pregunta 35. (Control 1 - Primavera 2021, P2)

Don Ennio tiene una pizzería con $N = C + V$ ingredientes disponibles: C contienen carne y V son vegetarianos. Además, dispone de S salsas, de las cuales S_c contienen carne. Suponga $1 \leq I \leq V$, $0 \leq S_c < S$ y que todas las elecciones indicadas son independientes.

- Doña Alessia escoge uniformemente I ingredientes distintos y una salsa. Calcule la probabilidad de que la pizza no contenga carne.
- Doña Nicoletta divide una pizza en $d \geq 1$ sectores iguales. Para cada sector escoge I ingredientes distintos; las elecciones de sectores diferentes pueden coincidir. Dos pizzas se consideran iguales si una se obtiene de la otra mediante una rotación, pero no se identifican por reflexión. Determine el número de pizzas distintas.
- Doña Alessia prepara $P \geq 1$ pizzas. Cada pizza contiene exactamente I ingredientes, con repetición permitida, y su multiconjunto de ingredientes se escoge uniformemente entre todos los posibles. Un cliente escoge de manera independiente y con la misma distribución un multiconjunto de I ingredientes. Calcule la probabilidad de que su elección coincida con al menos una de las P pizzas.

Solución

(a). La pizza no contiene carne si todos sus ingredientes son vegetarianos y la salsa tampoco contiene carne. Hay $\binom{N}{I}$ selecciones equiprobables de ingredientes y $\binom{V}{I}$ de ellas usan solo ingredientes vegetarianos. Por independencia entre ambas elecciones,

$$\mathbb{P}(\text{pizza sin carne}) = \frac{\binom{V}{I}}{\binom{N}{I}} \frac{S - S_c}{S}.$$

(b). Cada sector admite $m = \binom{N}{I}$ selecciones de ingredientes. Para identificar pizzas que solo difieren por una rotación, aplicamos el lema de Burnside: el número de configuraciones distintas es el promedio del número de configuraciones que deja fijas cada rotación. Una rotación de r sectores descompone los d sectores en $\gcd(d, r)$ ciclos. Para que la pizza quede fija, todos los sectores de un mismo ciclo deben tener la misma selección, de modo que esa rotación fija $m^{\gcd(d, r)}$ configuraciones. Así,

$$\frac{1}{d} \sum_{r=0}^{d-1} m^{\gcd(d, r)} = \frac{1}{d} \sum_{e|d} \varphi(e) m^{d/e},$$

donde φ es la función totiente de Euler.

(c). Como se permite repetir ingredientes y no importa el orden dentro de una pizza, hay

$$M = \binom{N + I - 1}{I}$$

multiconjuntos posibles. Condicional en la elección del cliente, cada pizza falla en coincidir con ella con probabilidad $1 - 1/M$. Las P pizzas se escogen independientemente, por lo que la probabilidad de que ninguna coincida es $(1 - 1/M)^P$. Tomando el complemento se obtiene

$$1 - \left(1 - \frac{1}{\binom{N+I-1}{I}}\right)^P.$$

Pregunta 36. (Control 1 - Otoño 2023, P3)

- (a) Don Pedro tiene un emprendimiento que consiste en armar computadores. Todas las piezas de los computadores que arma Don Pedro son iguales, a excepción de: la memoria RAM, el procesador, disco duro, tarjeta gráfica, y batería. La Tabla 4.1 muestra para cada una de estas partes los posibles componentes que Don Pedro usa.

Componente	Tipos				
Memoria RAM [Gb]	8	16	32		
Procesador [núcleos]	4	8	12	16	
Disco duro [Gb]	256	512	1024	2048	
Tarjeta gráfica	Tipo 1	Tipo 2	Tipo 3	Tipo 4	Tipo 5
Batería [horas]	4	6	8	12	18

Cuadro 4.1

Suponga que Don Pedro arma los computadores al azar. Específicamente, en cada computador que se arma, para cada uno de los componentes de la primera columna de la Tabla 4.1, se selecciona al azar (con igual probabilidad en cada caso) uno de los tipos posibles para la parte respectiva.

Luego de algunas semanas, se conoció un artículo que menciona que: (i) los computadores que tienen simultáneamente la tarjeta gráfica “Tipo 4” y la memoria “RAM de 8 Gb” no funcionan, y (ii) los computadores que tienen una “batería de 6 o menos horas” y un “procesador de 16 núcleos” tampoco funcionan. Note que tanto si se cumple (i) o (ii), el computador no funciona. Encuentre la probabilidad de que un computador armado por Don Pedro no funcione. Puede entregar una fracción como respuesta, tal que el numerador y denominador sean un número (no una expresión).

- (b) Suponga que Don Pedro le vende a Doña Inez 25 computadores, quien desea usarlos para armar una *estación de computadores*. Específicamente, Doña Inez posee una oficina con 25 escritorios, tal que cada uno tendrá un computador. No hay internet de ningún tipo, de modo que para conectar los computadores en una red, Doña Inez decide conectar por cable los computadores, de tal manera que cada computador queda conectado con los dos computadores más cercanos de cada lado (en el caso de los computadores de los extremos, estos se conectarán con menos computadores). En la Figura 4.1 se ilustra un ejemplo con 8 computadores.

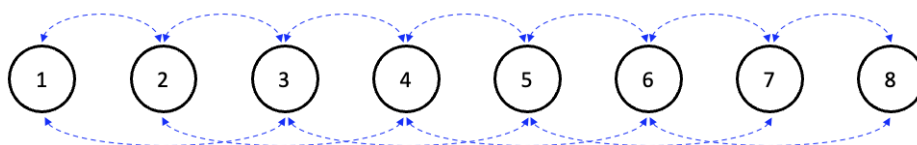


Figura 4.1

Asuma que los dos computadores de los extremos están buenos, mientras que del resto (i.e., de los 23 computadores restantes), hay 6 computadores que están malos. El sistema de computadores se dice que está en estado *funcional* si es que entre cualquier par de computadores buenos existe al menos un camino mediante cables que permite llegar de un computador al otro pasando solo por computadores buenos. Por ejemplo, usando los valores de los datos del problema, si los computadores en las posiciones de izquierda a derecha 2, 4, 6, 8, 10 y 12 están malos, entonces el sistema estaría *funcional*; sin embargo, si por ejemplo los computadores 2, 4, 6, 7, 10 y 12 están malos, entonces el sistema no estaría *funcional*. Suponga que los computadores están distribuidos al azar (excepto los dos computadores de los extremos que sabemos que están buenos). Determine la probabilidad de que el sistema esté *funcional*. No es necesario que compute el valor exacto.

Solución

(a). Cada computador queda determinado por una elección de cada fila de la tabla. Como todas esas elecciones son uniformes e independientes, los

$$3 \cdot 4 \cdot 4 \cdot 5 \cdot 5 = 1200$$

computadores posibles son equiprobables. Definamos los dos tipos de falla:

- A : ocurre la incompatibilidad descrita en (i). ■ B : ocurre la incompatibilidad descrita en (ii).

Para A , se fijan la memoria RAM y la tarjeta gráfica, mientras que el procesador, el disco duro y la batería pueden escogerse libremente; por tanto, $|A| = 4 \cdot 4 \cdot 5 = 80$. Análogamente,

$$|B| = 3 \cdot 4 \cdot 5 \cdot 2 = 120, \quad |A \cap B| = 4 \cdot 2 = 8.$$

El principio de inclusión-exclusión permite contar los computadores que no funcionan:

$$\begin{aligned} |A \cup B| &= |A| + |B| - |A \cap B| \\ &= 80 + 120 - 8 = 192. \end{aligned}$$

Como todos los computadores posibles tienen la misma probabilidad,

$$\mathbb{P}(A \cup B) = \frac{192}{1200} = \frac{4}{25}.$$

(b). Los seis computadores malos ocupan un subconjunto uniforme de seis posiciones entre las 23 posiciones interiores. Como cada computador se conecta con los dos más cercanos de cada lado, un único computador malo puede ser puentado mediante un cable, pero dos computadores malos consecutivos interrumpen todo camino entre los extremos. En consecuencia, el sistema es funcional si y solo si las seis posiciones malas no son consecutivas dos a dos.

El número de subconjuntos de seis posiciones no consecutivas dentro de 23 posiciones es

$$\binom{23 - 6 + 1}{6} = \binom{18}{6}.$$

Esto también puede verse distribuyendo los 19 computadores buenos en los siete espacios que quedan antes, entre y después de los seis computadores malos: cada espacio debe contener al menos un computador bueno, y las composiciones positivas de 19 en siete partes son $\binom{18}{6}$. Como existen $\binom{23}{6}$ ubicaciones equiprobables para los computadores malos,

$$\mathbb{P}(\text{el sistema es funcional}) = \frac{\binom{18}{6}}{\binom{23}{6}}.$$

Pregunta 37. (Control 1 - Primavera 2023, P3)

- (a) En una cumbre internacional participan 10 países, cada uno con un representante de Estado. Hay 10 sillas en fila en las cuales se sientan los representantes de Estado. ¿De cuántas formas se pueden sentar tal que los representantes de los países Estados Unidos y Reino Unido queden juntos?
- (b) De los 10 países que participan de la cumbre internacional, se hace una subcomisión de 3 representantes para votar sobre un tratado. El tratado se aprueba si de los 3 miembros de la subcomisión, hay al menos dos que aprueban. Considere que de los 10 países hay 6 que aprueban y 4 que rechazan. Si es que la subcomisión se conforma escogiendo al azar 3 de los 10 representantes, determine la probabilidad de que el tratado se apruebe.

Solución

(a). Consideremos a los representantes de Estados Unidos y Reino Unido como un solo bloque. El bloque y los otros ocho representantes pueden ordenarse de $9!$ maneras. Dentro del bloque hay dos órdenes posibles, por lo que el número total de disposiciones es

$$2 \cdot 9!.$$

(b). El tratado se aprueba si los tres integrantes lo apoyan o si exactamente dos lo apoyan y uno lo rechaza. Estos casos son disjuntos y aportan, respectivamente, $\binom{6}{3}$ y $\binom{6}{2}\binom{4}{1}$ subcomisiones favorables. Como existen $\binom{10}{3}$ subcomisiones posibles,

$$\mathbb{P}(\text{aprobar el tratado}) = \frac{\binom{6}{3} + \binom{6}{2}\binom{4}{1}}{\binom{10}{3}} = \frac{20 + 60}{120} = \frac{2}{3}.$$

Pregunta 38. (Control 1 - Otoño 2024, P3)

Durante la segunda guerra mundial, los alemanes usaban la máquina *Enigma* para enviar mensajes cifrados. Dicha máquina cuenta con un sistema de rotores. Un rotor es un disco compuesto por dos caras con las 26 letras del alfabeto en cada cara. Cada letra de una cara está conectada con otra letra de la contracara del mismo rotor¹. Un rotor puede ser visto como una función biyectiva $f : \{A, B, C, \dots, Z\} \rightarrow \{A, B, C, \dots, Z\}$. La máquina usa tres rotores que van conectados secuencialmente, mientras que el primer rotor va conectado con el teclado de la máquina. Existen cinco rotores diferentes que se pueden usar, de los cuales se seleccionan tres. Volviendo a la analogía anterior, cada rotor sería una función (posiblemente diferente), las cuales no son conmutativas.

Por otra parte, previo a los rotores, existen 10 cables que son conectados para intercambiar letras en el teclado. Específicamente, cada cable se conecta a dos letras diferentes, de tal manera que si un cable se conecta a las letras “L” y “F”, al presionar la tecla “L” esta es intercambiada por la letra “F”, y viceversa en caso de presionar la “F”. Dado que son 10 cables, no todas las letras quedarán conectadas con dichos cables. Además, una letra puede estar conectada con a lo más un cable. Se usan los 10 cables.

Determine el número total de configuraciones que se pueden realizar con la máquina *Enigma*. Debe justificar su respuesta. No es necesario que calcule el valor numérico de su respuesta.

Nota 1: respecto a la parte de los cables, le podría ser útil pensar el caso en donde se tiene un solo cable, y luego el caso en donde se tienen dos cables.

¹Considere que las conexiones de cada uno de los rotores son fijas, es decir, no se pueden cambiar.

Nota 2: considere para esta pregunta solo el enunciado, no su conocimiento previo acerca de la máquina *Enigma*.

Solución

Los tres rotores se eligen entre cinco y su orden importa porque las funciones no son conmutativas. Por tanto, el número de configuraciones de rotores es

$$5 \cdot 4 \cdot 3 = \frac{5!}{2!}.$$

Para los cables, primero se eligen las 20 letras que quedarán conectadas y luego se particionan en diez pares no ordenados. Una vez seleccionadas las letras, hay $20!$ secuencias posibles; dividir por 2^{10} elimina el orden dentro de cada par y dividir por $10!$ elimina el orden entre los diez cables. Así, el número de configuraciones de cables es

$$\binom{26}{20} \frac{20!}{2^{10}10!} = \frac{26!}{6!2^{10}10!}.$$

La misma cuenta puede escribirse de las formas equivalentes

$$\frac{1}{10!} \prod_{k=1}^{10} \binom{28-2k}{2} \quad \text{o} \quad \frac{1}{10!} \binom{26}{2, 2, \dots, 2, 6},$$

donde en el coeficiente multinomial aparecen diez grupos de tamaño 2.

Finalmente, las elecciones de rotores y cables son independientes. El número total de configuraciones de la máquina es

$$\frac{5!}{2!} \frac{26!}{6!2^{10}10!}.$$

Pregunta 39. (Control 1 - Primavera 2024, P3)

Finalmente, el semestre de primavera de 2024 ha concluido, y la emoción por el descanso que se aproxima es tan grande que decides organizar una salida a la playa con tus amigos. Una vez allí, se dan cuenta de que olvidaron llevar algo para comer. Sin embargo, la buena noticia no tarda en llegar, uno de tus amigos, eufórico por haber aprobado todos sus cursos, se ofrece a invitar *dulces de la Ligüa*. Justo en ese momento, aparece un vendedor que pasea por la playa con un canasto repleto de dulces, ofreciendo siete tipos distintos de dulces. El canasto del vendedor, que está completamente lleno, posee capacidad para 80 dulces.

- (a) ¿De cuántas formas distintas el vendedor podría llenar su canasto con dulces? Considere que los dulces se van moviendo dentro del canasto, de modo que no hay posiciones fijas dentro del canasto para los dulces. Así, dos formas de llenar el canasto, A y B , son distintas si y solo si hay al menos un tipo dulce cuya cantidad difiere entre A y B .

El vendedor le explica a tu amigo que existe una promoción especial, al comprar seis dulces al azar², se aplica un descuento sobre el total. Como gran aficionado a los *cachitos* (el *cachito* es un tipo de *dulce de la Ligüa*), sabes que si entre los dulces seleccionados hay alguno de este tipo, tus amigos te lo cederán. Para darles una idea de lo que podrían obtener, el vendedor especifica que su canasto contiene la siguiente variedad de dulces: 12 empolvados, 20 alfajores, 15 paletas, 6 cachitos, 9 cocadas, 10 almejas y 8 príncipes.

- (b) Compran una promoción que contiene seis dulces ¿Cuál es la probabilidad de que esta contenga al menos un cachito?

²El vendedor escoge los 6 dulces totalmente al azar entre los dulces dentro del canasto.

Solución

(a). Si y_i es la cantidad de dulces del tipo i , una forma de llenar el canasto corresponde a una solución entera no negativa de

$$y_1 + y_2 + y_3 + y_4 + y_5 + y_6 + y_7 = 80.$$

Por el método de estrellas y barras, el número de soluciones es

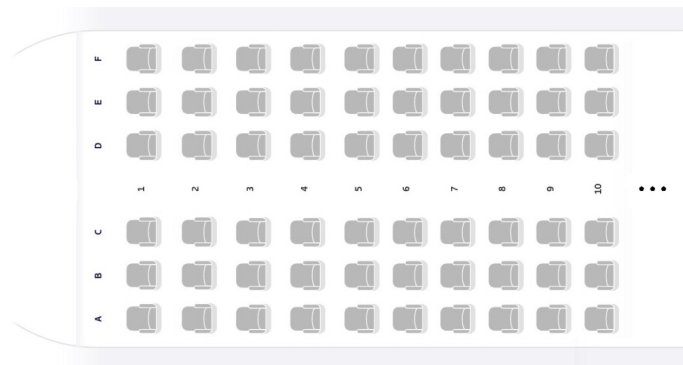
$$\binom{80 + 7 - 1}{7 - 1} = \binom{86}{6} = \binom{86}{80}.$$

(b). Entre los 80 dulces hay 74 que no son cachitos. Es más directo contar el evento complementario: ninguna de las seis elecciones es un cachito. Como la selección se realiza sin reposición,

$$\mathbb{P}(\text{al menos un cachito}) = 1 - \frac{\binom{74}{6}}{\binom{80}{6}}.$$

Pregunta 40. (Control 1 - Otoño 2025, P3)

Tres amigas deciden viajar juntas para descansar tras el semestre universitario. Para ahorrar, compran pasajes en una aerolínea de bajo costo que asigna los asientos al azar. Asuma que el vuelo va lleno. El avión tiene 20 filas de asientos, donde cada fila contiene dos secciones de tres asientos en línea, separados por un pasillo, que van indicados con letras de la A a la F. Los asientos A, B, C van hacia el costado izquierdo del pasillo, y los asientos D, E, F van al costado derecho. Además, solo los asientos A y F poseen ventana. Notar que en total son 120 asientos y pasajeros (incluyendo a las 3 amigas).



- Sabiendo que todas prefieren quedar en la ventana, ¿cuál es la probabilidad de que las 3 amigas queden en asientos de ventana?
- Aunque prefieren la ventana, valoran aún más estar juntas, es decir, las 3 en la misma fila hacia un mismo costado del pasillo. Es decir, que queden en una misma fila en los asientos A, B, C; o en la misma fila en los asientos D, E, F. ¿Cuál es la probabilidad de que queden juntas?
- A mitad del vuelo, la azafata ofrece 3 platos diferente para el almuerzo a cada pasajero. Tras tomar los 120 pedidos, la azafata comunica a la cocina la cantidad total de cada plato a preparar. ¿Cuántas formas diferentes de demanda puede recibir la cocina para estos 3 platos? Por ejemplo, si las opciones de plato son carne, pollo, y pasta; entonces una demanda posible es que se hayan pedido 50 platos de carne, 40 de pollo, y 30 de pasta.
- Calcule la probabilidad de que solo dos de las amigas queden juntas en el avión, es decir,

dos estén sentadas en asientos adyacentes en la misma fila y hacia un mismo costado del pasillo, y la tercera en cualquier otro asiento no adyacente a las dos amigas (es decir, en otra fila o en el otro costado del pasillo).

Solución

Como las amigas son personas distintas, el número total de asignaciones ordenadas de sus asientos es

$$P_3^{120} = 120 \cdot 119 \cdot 118.$$

(a). Existen 40 asientos de ventana: los asientos A y F de cada una de las 20 filas. Las asignaciones favorables son las permutaciones de tres de estos asientos, por lo que

$$\mathbb{P}(\text{tres en ventana}) = \frac{P_3^{40}}{P_3^{120}} = \frac{40 \cdot 39 \cdot 38}{120 \cdot 119 \cdot 118}.$$

(b). Hay 40 bloques de tres asientos ubicados en una misma fila y a un mismo costado del pasillo. Elegido el bloque, las tres amigas pueden distribuirse en sus asientos de $3!$ maneras. En consecuencia,

$$\mathbb{P}(\text{tres juntas}) = \frac{40 \cdot 3!}{120 \cdot 119 \cdot 118}.$$

(c). Si x_i es el número de pedidos del plato i , una demanda corresponde a una solución entera no negativa de $x_1 + x_2 + x_3 = 120$. Por estrellas y barras, el número de demandas es

$$\binom{120 + 3 - 1}{3 - 1} = \binom{122}{2} = 7381.$$

(d). En cada fila hay cuatro pares de asientos adyacentes dentro de un mismo costado: AB, BC, DE y EF. Por tanto, los asientos del par pueden escogerse de $20 \cdot 4$ maneras. Una vez elegido el par, hay $P_2^3 = 3 \cdot 2$ formas de escoger y ubicar a las dos amigas que se sentarán juntas. La tercera debe ocupar uno de los 117 asientos que quedan fuera del bloque de tres que contiene al par; así se evita que las tres queden juntas. Por ello,

$$\mathbb{P}(\text{exactamente dos juntas}) = \frac{20 \cdot 4 \cdot (3 \cdot 2) \cdot 117}{120 \cdot 119 \cdot 118}.$$

Pregunta 41. (Control 1 - Primavera 2025, P3)

El Spotify DJ dispone de un conjunto de 40 canciones diferentes, cada una clasificada en exactamente uno de tres géneros: Indie (7 canciones), Rock (13 canciones) o Pop (20 canciones). Una **playlist** es una lista ordenada de 12 canciones sin repetición. Spotify DJ ofrece dos servicios: el **gratis** y el **pagado**. La diferencia entre ambos es que en el **gratis**, la lista de 12 canciones es en orden alfabético según los nombres de las canciones; en el **pagado**, el cliente puede escoger el orden de las 12 canciones de la lista. Considere que dos listas son diferentes si es que una lista posee al menos una canción que la otra no posee, o si es que el orden de las canciones es diferente.

- (i) En el sistema **gratis**, ¿cuántas **playlist** diferentes se pueden crear tales que contengan exactamente i canciones Indie, r canciones Rock y p canciones Pop? (donde $i + r + p = 12$).
- (ii) En el sistema **pagado**, ¿cuántas **playlist** diferentes se pueden crear tales que contengan exactamente i canciones Indie, r canciones Rock y p canciones Pop? (donde $i + r + p = 12$).

De las 40 canciones, sean a y b dos canciones, tal que al ordenarlas alfabéticamente, existen k canciones cuyos nombres están entre las canciones a y b (k no incluye ni a a y ni a b). Considere que a es anterior a b alfabéticamente.

- (iii) Bajo el sistema **gratis** ¿Cuántos **playlist** existen que contengan a las canciones a y b juntas?

- (iv) Bajo el sistema **pagado** ¿Cuántos playlist existen que contengan a las canciones a y b juntas?

Solución

(i). En el servicio gratis, una vez escogidas las doce canciones su orden alfabético queda determinado. Se eligen i de las 7 canciones Indie, r de las 13 canciones Rock y p de las 20 canciones Pop. Por la regla del producto, el número de playlists es

$$\binom{7}{i} \binom{13}{r} \binom{20}{p}.$$

Esta cantidad se interpreta como cero si alguna selección excede el número de canciones disponibles de su género.

(ii). En el servicio pagado, cada selección del inciso anterior admite $12!$ órdenes, porque todas las canciones son distintas. Por tanto, el número de playlists es

$$\binom{7}{i} \binom{13}{r} \binom{20}{p} 12!.$$

(iii). En el orden alfabético, a y b quedan juntas si se incluyen ambas y no se incluye ninguna de las k canciones ubicadas entre ellas. Fijadas a y b , las otras diez canciones deben elegirse entre las $40 - 2 - k$ restantes que no están entre ambas. El número pedido es

$$\binom{40 - 2 - k}{10}.$$

(iv). En el servicio pagado, tratamos a y b como un bloque. El bloque admite $2!$ órdenes internos y puede comenzar en cualquiera de las primeras once posiciones de la playlist. Las otras diez posiciones se llenan eligiendo y ordenando diez canciones distintas entre las 38 restantes. Por la regla del producto, el número de playlists es

$$2! \cdot 11 \cdot \frac{38!}{28!}.$$

Pregunta 42. (Control 1 - Otoño 2026, P3)

En el curso *Taller de Liderazgo*, los grupos tienen 16 estudiantes.

- (a) En cada grupo, cinco estudiantes tendrán un rol especial (jefatura, subjefatura, coordinación, secretaría y tesorería). ¿De cuántas maneras se pueden asignar estos roles en un grupo?

Para las partes que siguen, considere que en el curso hay 160 estudiantes, por lo que se forman diez grupos. Los grupos no poseen identificador o etiqueta.

- (b) De los 160 estudiantes, 90 son hombres y 70 son mujeres. Se desea que en cada grupo haya 9 hombres y 7 mujeres. ¿De cuántas formas se pueden hacer los grupos? (No considere los roles dentro de cada grupo como en la parte (a), sino solo la asignación de estudiantes a grupos.)
- (c) Cada grupo debe entregar la Tarea 1 antes del día 14/04 a las 23:59, pero se permiten atrasos. El i -ésimo grupo en entregar recibe una bonificación de $17 - i$ décimas, pero cada grupo que entrega atrasado recibe un descuento de 1 punto. ¿Cuántas asignaciones posibles hay de décimas (positivas o negativas) a grupos?

Indicación: Note que la asignación está dada por dos elementos: (1) el orden en que entregan los grupos y (2) el número de grupos que entregan atrasados. Puede asumir que todos los grupos entregan a horas distintas.

Solución

(a). Los cinco roles son distintos. Para la jefatura hay 16 opciones; para la subjefatura quedan 15; y así sucesivamente hasta la tesorería, con 12 opciones. Por el principio multiplicativo,

$$P_5^{16} = 16 \cdot 15 \cdot 14 \cdot 13 \cdot 12 = \frac{16!}{11!}.$$

(b). Supongamos primero que los diez grupos están etiquetados. Las 70 mujeres pueden distribuirse en grupos de siete de

$$\binom{70}{7, 7, \dots, 7} = \frac{70!}{(7!)^{10}}$$

maneras, y los 90 hombres pueden distribuirse en grupos de nueve de

$$\binom{90}{9, 9, \dots, 9} = \frac{90!}{(9!)^{10}}$$

maneras. Al combinar ambas distribuciones se obtiene cada partición en grupos no etiquetados $10!$ veces, una por cada permutación de las etiquetas. Por lo tanto, el número solicitado es

$$\frac{1}{10!} \binom{70}{7, 7, \dots, 7} \binom{90}{9, 9, \dots, 9}.$$

Equivalentemente, se pueden particionar por separado hombres y mujeres en diez grupos no etiquetados y luego emparejar ambos tipos de subgrupos de $10!$ maneras; los factores se simplifican hasta la misma expresión.

(c). Primero se ordenan los diez grupos según su hora de entrega, lo que puede hacerse de $10!$ maneras. Fijado el orden, el número de grupos atrasados puede ser cualquier entero entre 0 y 10, es decir, existen 11 posiciones posibles para el corte entre entregas a tiempo y atrasadas. La indicación asegura que estos dos elementos determinan la asignación, por lo que el principio multiplicativo da

$$10! \cdot 11 = 11!.$$

Pregunta 43. (Control 2 - Primavera 2022, P1)

Estefanía posee una automotora en la cual se venden vehículos de la marca *ABC*. Dicha marca posee un total de m modelos de autos, los cuales pueden ser mostrados físicamente en la automotora de Estefanía. Sin embargo, por limitaciones de espacio, dicha automotora puede mostrar un máximo de l vehículos simultáneamente. Asuma que la automotora siempre muestra exactamente l vehículos.

- Sea $l < m$. Suponga que de cada modelo se puede mostrar a lo más una unidad en la automotora. ¿Cuántas combinaciones diferentes de modelos de vehículos se pueden mostrar simultáneamente en la automotora? Note que solo importa cuáles son los vehículos a mostrar, y no en qué “posiciones” serían mostrados.
- Suponga que los vehículos mostrados en la automotora se exponen en un círculo que va rotando continuamente, de modo que los vehículos se ubican en posiciones en el contorno de dicho círculo con el capó apuntando hacia el centro del círculo. Específicamente, se tienen $m + r$ vehículos a mostrar en la automotora, uno de cada modelo, excepto el modelo *AAA*, del cual la automotora posee $r + 1$ unidades. Asuma que el círculo donde la automotora ubicará los vehículos contiene exactamente $m + r$ posiciones, es decir, $l = m + r$. Determine de cuántas maneras diferentes se pueden exponer los vehículos.
- Suponga que ahora, a diferencia de la parte (b), no importan las posiciones en las cuales se muestran los vehículos. Considere que se pueden mostrar simultáneamente múltiples unidades de un mismo modelo en la automotora, y se posee una capacidad de l vehículos para mostrar simultáneamente, con $l > 3m$. Además, considere que de cada modelo deben

mostrarse al menos tres vehículos. ¿De cuántas maneras distintas puede la automotora mostrar los vehículos?

Solución

(a). Como se exhiben l modelos distintos y el orden de las posiciones no importa, cada configuración corresponde a escoger un subconjunto de l elementos entre los m modelos disponibles. Por tanto, el número de configuraciones es

$$\binom{m}{l}.$$

(b). Si $m \geq 2$, se puede fijar uno de los modelos distintos de AAA para eliminar la simetría por rotaciones. Quedan $m + r - 1$ lugares, entre los cuales las $r + 1$ unidades del modelo AAA son indistinguibles. En consecuencia, hay

$$\frac{(m + r - 1)!}{(r + 1)!}$$

maneras diferentes de exponerlos. Si $m = 1$, todos los vehículos son unidades indistinguibles del modelo AAA, por lo que existe una sola disposición circular.

(c). Sea x_i el número de vehículos del modelo i que se exhibe. Las configuraciones buscadas son las soluciones enteras de

$$x_1 + \cdots + x_m = l, \quad x_i \geq 3.$$

Al definir $y_i = x_i - 3$, la restricción mínima desaparece y queda

$$y_1 + \cdots + y_m = l - 3m, \quad y_i \geq 0.$$

Como $l > 3m$, el método de estrellas y barras entrega

$$\binom{(l - 3m) + m - 1}{m - 1} = \binom{l - 2m - 1}{m - 1} = \binom{l - 2m - 1}{l - 3m}.$$

Pregunta 44. (Examen v2 - Primavera 2025, P1)

- En días lluviosos, Joe llega tarde al trabajo con probabilidad 0.3; en días no lluviosos, llega tarde con probabilidad 0.1. Con probabilidad 0.7, lloverá mañana.
 - Encuentre la probabilidad de que Joe llegue temprano mañana.
 - Dado que Joe llegó temprano, ¿cuál es la probabilidad condicional de que haya llovido?
- Siete regalos distintos deben distribuirse entre 10 niños. ¿Cuántos resultados distintos son posibles si ningún niño recibe más de un regalo? Asuma que todos los regalos son diferentes. Debe justificar su respuesta.
- Un niño tiene 12 bloques, de los cuales 6 son negros, 4 son rojos, 1 es blanco y 1 es azul. Si el niño coloca los bloques en una fila, ¿cuántos arreglos distintos son posibles? Debe justificar su respuesta.

Solución

(1)(a). Definamos los eventos

- L : mañana es un día lluvioso;
- T : Joe llega tarde al trabajo.

Del enunciado, $\mathbb{P}(L) = 0,7$, $\mathbb{P}(T | L) = 0,3$ y $\mathbb{P}(T | L^c) = 0,1$. Como $\{L, L^c\}$ es una partición, la

ley de probabilidades totales da

$$\begin{aligned}\mathbb{P}(T^c) &= \mathbb{P}(T^c \mid L)\mathbb{P}(L) + \mathbb{P}(T^c \mid L^c)\mathbb{P}(L^c) \\ &= (1 - 0,3)(0,7) + (1 - 0,1)(0,3) \\ &= 0,76.\end{aligned}$$

(1)(b). Usando el resultado anterior, el Teorema de Bayes entrega

$$\begin{aligned}\mathbb{P}(L \mid T^c) &= \frac{\mathbb{P}(T^c \mid L)\mathbb{P}(L)}{\mathbb{P}(T^c)} \\ &= \frac{(0,7)(0,7)}{0,76} = \frac{49}{76} \approx 0,6447.\end{aligned}$$

(2). Cada resultado asigna los siete regalos distintos a siete niños diferentes. Para el primer regalo hay 10 destinatarios posibles, para el segundo quedan 9, y así sucesivamente. Por el principio multiplicativo,

$$10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 = \frac{10!}{(10-7)!} = 604\,800.$$

(3). Las permutaciones de los 12 bloques distinguen únicamente sus colores. Como los 6 bloques negros son indistinguibles entre sí y lo mismo ocurre con los 4 rojos, el número de arreglos es

$$\frac{12!}{6! 4! 1! 1!} = 27\,720.$$

Capítulo 5

Variables Aleatorias Discretas

5.1. Función de masa, distribución y momentos

Resumen

Variable aleatoria discreta.

Una variable aleatoria es una función $X: \Omega \rightarrow \mathbb{R}$. Es discreta cuando su rango

$$R_X = \{x \in \mathbb{R} : \mathbb{P}(X = x) > 0\}$$

es finito o numerable.

Función de masa de probabilidad.

La PMF de X es $p_X(x) = \mathbb{P}(X = x)$. Satisface:

- **No negatividad.** $p_X(x) \geq 0$ para todo $x \in R_X$.
- **Normalización.** $\sum_{x \in R_X} p_X(x) = 1$.
- **Probabilidad de un conjunto.** Para $A \subseteq \mathbb{R}$,

$$\mathbb{P}(X \in A) = \sum_{x \in A \cap R_X} p_X(x).$$

Función de distribución.

La CDF se define, para todo $x \in \mathbb{R}$, por

$$F_X(x) = \mathbb{P}(X \leq x) = \sum_{t \in R_X: t \leq x} p_X(t).$$

- **Monotonía.** Si $x \leq y$, entonces $F_X(x) \leq F_X(y)$.
- **Límites.** $\lim_{x \rightarrow -\infty} F_X(x) = 0$ y $\lim_{x \rightarrow \infty} F_X(x) = 1$.
- **Continuidad por la derecha.** $\lim_{t \downarrow x} F_X(t) = F_X(x)$.
- **Salto.** $p_X(x) = F_X(x) - F_X(x^-)$.
- **Intervalos.** $\mathbb{P}(a < X \leq b) = F_X(b) - F_X(a)$.

Esperanza.

Siempre que la suma esté bien definida,

$$\mathbb{E}[g(X)] = \sum_{x \in R_X} g(x)p_X(x).$$

En particular, $\mathbb{E}[X]$ es el promedio ponderado de los valores de X . Sus propiedades básicas son:

- **Linealidad.** $\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$; no requiere independencia.
- **Monotonía.** Si $X \leq Y$ casi seguramente, entonces $\mathbb{E}[X] \leq \mathbb{E}[Y]$.
- **Indicatriz.** Para un evento A , $\mathbb{E}[\mathbb{1}_{\{A\}}] = \mathbb{P}(A)$.
- **Producto de variables independientes.** Si X e Y son independientes, entonces $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$.

Varianza.

La varianza mide la dispersión en torno a la media:

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

- **No negatividad.** $\text{Var}(X) \geq 0$, con igualdad si y solo si X es constante casi seguramente.
- **Transformación afín.** $\text{Var}(aX + b) = a^2 \text{Var}(X)$.
- **Suma independiente.** Si X e Y son independientes, $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$.

La desviación estándar es $\text{DE}(X) = \sqrt{\text{Var}(X)}$; si $\mathbb{E}[X] \neq 0$, el coeficiente de variación es $\text{CV}(X) = \text{DE}(X)/\mathbb{E}[X]$.

5.1.1. Preguntas

Pregunta 45.

La adivina Solanda Yultana quiere demostrar su poder predictivo proponiendo el siguiente juego a las personas que tienen exactamente dos hijos (de cualquier sexo). Ella les hace la pregunta “Dime el nombre de una hija tuya”. Si la persona responde “No tengo hijas” entonces no hay nada que adivinar y se acaba el juego sin un ganador. En cambio, si responde un nombre, Solanda intenta adivinar el sexo del otro hijo o hija. Si acierta, la persona le da \$1000 a Solanda y si no acierta, Solanda entrega \$1000 a la persona. Si Solanda en realidad siempre dice “tu otro hijo es un hombre” ¿Cuál es la ganancia esperada de la adivina en cada adivinanza? ¿Y su varianza? (Suponga que para todas las personas con dos hijos la probabilidad de que cada hijo sea de sexo masculino es $1/2$ independiente del sexo del otro hijo).

Solución

Como se pregunta por la ganancia en cada adivinanza, condicionamos a que la persona tenga al menos una hija y entregue un nombre. Los casos (F, F) , (F, M) y (M, F) son equiprobables. Solanda pierde en el primero y gana en los otros dos. Si G es su ganancia,

$$\mathbb{P}(G = 1000) = \frac{2}{3}, \quad \mathbb{P}(G = -1000) = \frac{1}{3}.$$

Así,

$$\mathbb{E}[G] = \frac{1000}{3}, \quad \text{Var}(G) = \frac{8\,000\,000}{9}.$$

Pregunta 46.

Se tienen dos dados: uno de 4 caras y otro de 6 caras. Llamemos X al resultado obtenido al lanzar el dado de 4 caras e Y al resultado obtenido al lanzar el dado de 6 caras. Ahora definimos Z como el promedio de X e Y , es decir, $Z = \frac{X+Y}{2}$.

- Encuentra la PMF de X y de Y .
- Encuentra la desviación estándar de X , Y y Z .
- Se lleva a cabo un juego de apuestas. Si $X > Y$, se gana $\$2X$, pero si $X \leq Y$, se pierde $\$1$. Ahora, supongamos que se realiza este juego exactamente 60 veces. ¿Cuál sería la ganancia total esperada (positiva) o pérdida (negativa) después de estas 60 rondas?

Solución

(a). Como ambos dados son balanceados, cada cara de un mismo dado tiene la misma probabilidad. Por tanto,

$$p_X(x) = \begin{cases} \frac{1}{4}, & x \in \{1, 2, 3, 4\}, \\ 0, & \text{en otro caso,} \end{cases} \quad p_Y(y) = \begin{cases} \frac{1}{6}, & y \in \{1, 2, 3, 4, 5, 6\}, \\ 0, & \text{en otro caso.} \end{cases}$$

(b). Para calcular las varianzas, obtenemos primero los dos primeros momentos de cada dado:

$$\begin{aligned} \mathbb{E}[X] &= \frac{1}{4} \sum_{x=1}^4 x = \frac{5}{2}, & \mathbb{E}[X^2] &= \frac{1}{4} \sum_{x=1}^4 x^2 = \frac{15}{2}, \\ \mathbb{E}[Y] &= \frac{1}{6} \sum_{y=1}^6 y = \frac{7}{2}, & \mathbb{E}[Y^2] &= \frac{1}{6} \sum_{y=1}^6 y^2 = \frac{91}{6}. \end{aligned}$$

Al usar $\text{Var}(U) = \mathbb{E}[U^2] - \mathbb{E}[U]^2$, se obtiene

$$\text{Var}(X) = \frac{5}{4}, \quad \text{Var}(Y) = \frac{35}{12}.$$

En consecuencia,

$$\sigma_X = \frac{\sqrt{5}}{2}, \quad \sigma_Y = \sqrt{\frac{35}{12}}.$$

Como $Z = (X + Y)/2$ y los dados son independientes,

$$\text{Var}(Z) = \frac{1}{4} \left(\frac{5}{4} + \frac{35}{12} \right) = \frac{25}{24}, \quad \sigma_Z = \sqrt{\frac{25}{24}}.$$

(c). Sea W la ganancia de una ronda. De los 24 pares equiprobables, en 18 se cumple $X \leq Y$. Los seis pares ganadores son $(2, 1)$, $(3, 1)$, $(3, 2)$, $(4, 1)$, $(4, 2)$ y $(4, 3)$. Por ello,

$$\mathbb{E}[W] = -\frac{18}{24} + \frac{4 + 2 \cdot 6 + 3 \cdot 8}{24} = \frac{11}{12}.$$

Después de 60 rondas, la ganancia esperada es $60(11/12) = 55$ dólares.

Pregunta 47.

- (i) Suponga que una persona da un paso a la derecha con probabilidad $p \in [0, 1]$ y un paso a la izquierda con probabilidad $1 - p$. Para modelar la trayectoria que sigue esta persona, supondremos que se posiciona sobre un número entero, comenzando en el origen, y usaremos la variable aleatoria X que denota la posición final del personaje. Encuentre la probabilidad de que caiga en la posición $k \in \mathbb{Z}$ luego de $n \in \mathbb{N}$ saltos. Deduzca la función de probabilidad de la posición del borrachin y su rango. ¿Cómo calcularía la esperanza y varianza?
- (ii) En un club desconocido de Santiago, n personas dejan sus abrigo en un guardarropas. Al cerrar el club, estas personas sacan un abrigo de manera aleatoria y se lo llevan. ¿Cuál es el valor esperado de abrigo que fueron devueltos de manera correcta? Considere que la probabilidad de que una persona obtenga su abrigo es uniforme.

Solución

(i). Sea $D \sim \text{Binomial}(n, p)$ el número de pasos a la derecha. Los $n - D$ pasos restantes son hacia la izquierda, por lo que la posición final satisface

$$X = D - (n - D) = 2D - n.$$

En consecuencia, X solo puede tomar los valores $-n, -n + 2, \dots, n$. Para que $X = x$, el número de pasos a la derecha debe ser $(x + n)/2$. La función de masa queda entonces

$$p_X(x) = \begin{cases} \binom{n}{(x+n)/2} p^{(x+n)/2} (1-p)^{(n-x)/2}, & x \in \{-n, -n+2, \dots, n\}, \\ 0, & \text{en otro caso.} \end{cases}$$

La representación afín $X = 2D - n$ permite calcular directamente sus momentos:

$$\begin{aligned} \mathbb{E}[X] &= 2\mathbb{E}[D] - n = n(2p - 1), \\ \text{Var}(X) &= 4\text{Var}(D) = 4np(1 - p). \end{aligned}$$

(ii). Para cada persona i , definamos la indicadora

$$I_i = \mathbb{1}_{\{\text{la persona } i \text{ recibe su propio abrigo}\}}.$$

Como cualquiera de los n abrigos tiene la misma probabilidad de ser entregado a la persona i ,

$$\mathbb{E}[I_i] = \mathbb{P}(I_i = 1) = \frac{1}{n}.$$

Si $X = \sum_{i=1}^n I_i$ es el número de devoluciones correctas, la linealidad de la esperanza entrega

$$\mathbb{E}[X] = \sum_{i=1}^n \mathbb{E}[I_i] = n \left(\frac{1}{n} \right) = 1.$$

Las indicadoras no son independientes, pero esa independencia no es necesaria para aplicar la linealidad de la esperanza.

5.1.2. Preguntas de Evaluación

Pregunta 48. (Control 1 - Otoño 2021, P3)

Usted administra una granja de gallinas felices y está considerando comprar por adelantado alimento para la siguiente temporada, ya que el precio actual es relativamente bajo y se quiere proteger contra futuras alzas. Ha calculado que para la siguiente temporada necesita 10.000 kilos de alimento. Si compra ahora, el costo por kilo es de \$3.000 (suponga que no hay costo extra por almacenar el alimento hasta la siguiente temporada). De acuerdo a su conocimiento del mercado, pronostica que hay tres escenarios posibles para el valor del alimento en la siguiente temporada: (i) el bueno, en que el precio es \$1.000 por kilo, (ii) el intermedio, con \$4.000 por kilo, y (iii) el malo, con \$8.000 por kilo. Estos escenarios ocurren con probabilidad $3/4$, $1/8$ y $1/8$, respectivamente.

- Llame X al costo por kilo en la siguiente temporada. Calcule el rango, la PMF, la CDF, el valor esperado y la varianza de X .
- Suponga que si termina pagando más de \$50.000.000 la granja se va a la quiebra. Calcule (i) la probabilidad de que la granja quiebre si no compra nada ahora, y (ii) la probabilidad de que la granja quiebre si compra los 10.000 kilos ahora.
- Llame Y_z al costo total que terminaría pagando si compra $0 \leq z \leq 10.000$ kilos ahora, es decir, el costo de z kilos al precio de hoy, más el costo de $(10.000 - z)$ kilos al precio de la siguiente temporada. Calcule el rango y el valor esperado de Y_z (exprese estas respuestas en función de z). ¿Qué cantidad z le conviene comprar ahora si su objetivo es minimizar $\mathbb{E}[Y_z]$?
- Calcule la probabilidad de quebrar en términos de z . ¿Qué cantidad z le conviene comprar ahora si su objetivo es minimizar $\mathbb{E}[Y_z]$, pero evitando a toda costa la quiebra (es decir, sujeto a que la probabilidad de quiebra es 0)?

Solución

(a). El rango son los valores que puede tomar X , es decir, $R_X = \{1.000, 4.000, 8.000\}$. La PMF es

$$p_X(x) = \begin{cases} 3/4 & \text{si } x = 1.000 \\ 1/8 & \text{si } x \in \{4.000, 8.000\} \\ 0 & \text{en cualquier otro caso.} \end{cases}$$

La CDF es

$$F_X(x) = \begin{cases} 0 & \text{si } x < 1.000 \\ 3/4 & \text{si } 1.000 \leq x < 4.000 \\ 7/8 & \text{si } 4.000 \leq x < 8.000 \\ 1 & \text{si } 8.000 \leq x. \end{cases}$$

Los dos primeros momentos son

$$\begin{aligned} \mathbb{E}[X] &= 1000 \left(\frac{3}{4}\right) + 4000 \left(\frac{1}{8}\right) + 8000 \left(\frac{1}{8}\right) = 2250, \\ \mathbb{E}[X^2] &= 1000^2 \left(\frac{3}{4}\right) + 4000^2 \left(\frac{1}{8}\right) + 8000^2 \left(\frac{1}{8}\right) = 10\,750\,000. \end{aligned}$$

Por tanto,

$$\text{Var}(X) = 10\,750\,000 - 2250^2 = 5\,687\,500.$$

(b). (i) Si no se compra nada ahora, se paga $10.000 \cdot X$ pesos. De los tres valores que puede tomar X se exceden los \$50.000.000 solo cuando $X = 8.000$, por lo que la probabilidad de quebrar es exactamente $\mathbb{P}(X = 8.000) = 1/8$.

(ii) Si los 10.000 kilos se compran ahora, se termina pagando \$30.000.000, por lo que la probabilidad de quiebra es 0.

(c). Podemos escribir $Y_z = z \cdot 3.000 + (10.000 - z) \cdot X$. Por lo tanto, el rango de Y_z es

$$R_{Y_z} = \left\{ (2.000z + 10.000.000), (40.000.000 - 1.000z), (80.000.000 - 5.000z) \right\}.$$

Para calcular el valor esperado podemos usar la linealidad:

$$\begin{aligned} \mathbb{E}[Y_z] &= \mathbb{E}[z \cdot 3.000 + (10.000 - z) \cdot X] \\ &= z \cdot 3.000 + (10.000 - z) \cdot \mathbb{E}[X] \\ &= z \cdot 3.000 + (10.000 - z) \cdot 2.250 \\ &= 750z + 22.500.000. \end{aligned}$$

Esta cantidad es creciente en z , por lo que lo más conveniente es comprar $z = 0$ kilos ahora.

(d). La probabilidad de quebrar es exactamente $\mathbb{P}(Y_z > 50.000.000)$. Del rango de Y_z observamos que el único valor que podría potencialmente llevarnos a la quiebra es $(80.000.000 - 5.000z)$, que ocurre con probabilidad $1/8$. Este valor es mayor que 50.000.000 si $z < 6.000$, por lo que la probabilidad de quiebra es

$$\mathbb{P}(Y_z > 50.000.000) = \begin{cases} 1/8 & \text{si } z < 6.000 \\ 0 & \text{si no.} \end{cases}$$

Por lo tanto, queremos imponer que $z \geq 6.000$, y entonces, recordando que $\mathbb{E}[Y_z]$ era creciente en z , concluimos que el mínimo se alcanza en $z = 6.000$.

Pregunta 49. (Control 1 - Primavera 2021, P3)

Se ha inaugurado un salón de juegos en la facultad. Responda cada situación por separado.

- (a) Usted juega dos partidas independientes de pool. En la primera, la probabilidad de perder es $1 - p$; en la segunda, es $1 - q$, con $p, q \in (0, 1)$. En cada partida, ganar y empatar tienen la misma probabilidad. Se obtienen tres puntos por ganar, uno por empatar y cero por perder. Sea X el puntaje total. Calcule el soporte y la función de masa p_X .
- (b) Un tablero de dardos tiene radio R_1 y contiene dos círculos concéntricos de radios R_2 y R_3 , donde $0 < R_3 < R_2 < R_1$. Se obtienen 50 puntos en el círculo menor, 25 en el anillo intermedio, 10 en el anillo exterior y 0 fuera del tablero. El dardo cae en el tablero con

probabilidad $\tau \in (0, 1]$ y, condicional en caer dentro, su posición es uniforme respecto del área. Sea Y el puntaje de un lanzamiento. Calcule su soporte, su función de masa p_Y y su esperanza.

- (c) Se realizan tres lanzamientos independientes bajo el modelo del inciso anterior. Calcule la probabilidad de obtener al menos 100 puntos.
- (d) Suponga ahora que el dardo siempre cae dentro del tablero. Manteniendo fijo R_1 y permitiendo que los dos círculos interiores coincidan, indique cómo elegir los otros dos radios para maximizar la varianza del puntaje y compare esta configuración con cualquier configuración original que satisfaga $0 < R_3 < R_2 < R_1$.

Solución

(a). En la primera partida, ganar y empatar tienen probabilidad $p/2$ cada uno y perder tiene probabilidad $1 - p$. En la segunda, las probabilidades correspondientes son $q/2$, $q/2$ y $1 - q$. Al sumar los puntajes de ambas partidas, el soporte es

$$R_X = \{0, 1, 2, 3, 4, 6\}.$$

Usando independencia entre las partidas,

$$p_X(x) = \begin{cases} (1-p)(1-q), & x = 0, \\ \frac{p}{2}(1-q) + (1-p)\frac{q}{2}, & x = 1, \\ \frac{pq}{4}, & x = 2, \\ \frac{p}{2}(1-q) + (1-p)\frac{q}{2}, & x = 3, \\ \frac{pq}{4}, & x = 4, \\ \frac{pq}{4}, & x = 6, \\ 0, & \text{en otro caso.} \end{cases}$$

Los valores 1 y 3 tienen la misma probabilidad porque ambos combinan una derrota con uno de los dos resultados que tienen probabilidad mitad de p o de q .

(b). El soporte depende de si es posible caer fuera del tablero:

$$R_Y = \begin{cases} \{0, 10, 25, 50\}, & 0 < \tau < 1, \\ \{10, 25, 50\}, & \tau = 1. \end{cases}$$

Condicionado en que el dardo cae dentro del tablero, la probabilidad de cada zona es su razón de áreas. Por tanto,

$$\begin{aligned} p_Y(0) &= 1 - \tau, \\ p_Y(10) &= \tau \frac{R_1^2 - R_2^2}{R_1^2}, \\ p_Y(25) &= \tau \frac{R_2^2 - R_3^2}{R_1^2}, \\ p_Y(50) &= \tau \frac{R_3^2}{R_1^2}. \end{aligned}$$

La esperanza es

$$\mathbb{E}[Y] = 10p_Y(10) + 25p_Y(25) + 50p_Y(50).$$

(c). Escribamos $p_s = p_Y(s)$. Para alcanzar al menos 100 puntos en tres lanzamientos, puede haber exactamente dos puntajes de 50, tres puntajes de 50, o exactamente un puntaje de 50

acompañado por dos puntajes de 25. Estos casos son disjuntos, de modo que

$$\mathbb{P}(Y_1 + Y_2 + Y_3 \geq 100) = 3p_{50}^2(p_0 + p_{10} + p_{25}) + p_{50}^3 + 3p_{50}p_{25}^2.$$

(d). Cuando $\tau = 1$, el puntaje está siempre en el intervalo $[10, 50]$. Como $10 \leq Y \leq 50$, se cumple $(Y - 10)(50 - Y) \geq 0$. Si $\mu = \mathbb{E}[Y]$, al expandir y tomar esperanza obtenemos

$$\mathbb{E}[Y^2] \leq 60\mu - 500.$$

Esta cota permite acotar la varianza sin invocar resultados externos:

$$\begin{aligned}\text{Var}(Y) &= \mathbb{E}[Y^2] - \mu^2 \\ &\leq 60\mu - 500 - \mu^2 \\ &= 400 - (\mu - 30)^2 \leq 400.\end{aligned}$$

Para alcanzar la igualdad, Y debe tomar solamente los valores extremos 10 y 50, y además debe satisfacer $\mu = 30$. Esto exige concentrar probabilidad $1/2$ en cada extremo. La configuración se obtiene permitiendo que los círculos interiores coincidan y tomando

$$R_2 = R_3 = \frac{R_1}{\sqrt{2}}.$$

Entonces el anillo de 25 puntos tiene área cero, mientras que las regiones de 10 y 50 puntos ocupan cada una la mitad del tablero. Toda configuración original con $0 < R_3 < R_2 < R_1$ asigna probabilidad positiva al puntaje intermedio de 25 y, por ello, no satisface el caso de igualdad; su varianza es estrictamente menor que 400.

Pregunta 50. (Control 2 - Primavera 2022, P2)

La empresa *Júpiter* produce termómetros infrarrojos. Cada termómetro producido queda bien calibrado con una probabilidad de p , en cuyo caso el termómetro, al ser usado, siempre indicará la temperatura correcta (o real) del objeto al cual se le mide su temperatura. En caso de que el termómetro quede mal calibrado, este entregará un valor desviado por δ grados Celsius respecto a la temperatura real del objeto. Asuma que ambos casos, donde la lectura de la temperatura se desvía por **sobre** o **bajo** δ grados Celsius, son equiprobables cuando el termómetro queda mal calibrado. Llamaremos al parámetro δ el *error*.

La empresa *Bedebueno* testea los termómetros producidos por la empresa *Júpiter*. Específicamente, mide la temperatura de un objeto que está a una temperatura real de a grados Celsius con un termómetro producido por la empresa *Júpiter*. Llame X a la temperatura que entrega la lectura con un termómetro producido por la empresa *Júpiter*.

Asuma que $a > 0$, $\delta > 0$ y $p \in (0, 1)$.

- Entregue la CDF de la variable aleatoria X . Además, gráfiquela.
- Bedebueno* está implementando una certificación de calidad, la cual se entrega a la empresa que produce los termómetros, en el caso de que esta última cumpla con el criterio establecido. En particular, dicho criterio consiste en que los termómetros producidos por la empresa, al medir la temperatura de un objeto a temperatura real de a grados Celsius, deben tener una desviación estándar menor o igual a $\delta/2$. Indique las condiciones de los parámetros del problema (p , a y/o δ) de tal modo que *Júpiter* reciba la certificación.
- Suponga que la empresa *Júpiter* está considerando arrendar una tecnología que le permita controlar el *error* de la lectura del termómetro. El precio de arriendo de dicha tecnología es de $\$ \lambda/\delta$ por cada termómetro producido con un *error* de δ , donde $\lambda > 0$ es un parámetro fijo y conocido. Además, considere que ahora *Júpiter* debe pagar en reparación un monto de $\$ c \cdot \delta$ por cada termómetro malo que tenga una lectura de δ grados Celsius por **sobre** o **bajo** la temperatura real del objeto al cual se le mide su temperatura.

- (I) Dado un parámetro fijo de *error* igual a $\delta > 0$, obtenga el costo esperado de producir un termómetro. Incluya todos los costos asociados, es decir, de arriendo de tecnología y de reparación.

Hint: Le podría ser de utilidad escribir el costo como una expresión que dependa de la variable aleatoria X .

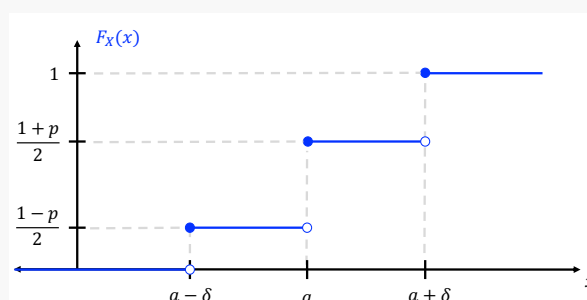
- (II) Considerando que *Júpiter* puede escoger cualquier valor no negativo para el parámetro de *error* a contratar con la nueva tecnología, indique el valor óptimo que debería utilizar.

Solución

(a). La lectura toma los valores $a - \delta$, a y $a + \delta$, con probabilidades $(1 - p)/2$, p y $(1 - p)/2$, respectivamente. Al acumular estas masas en orden creciente se obtiene

$$F_X(x) = \begin{cases} 0, & x < a - \delta, \\ \frac{1-p}{2}, & a - \delta \leq x < a, \\ \frac{1+p}{2}, & a \leq x < a + \delta, \\ 1, & x \geq a + \delta. \end{cases}$$

La gráfica es una función escalonada, continua por la derecha, con saltos de tamaños $(1 - p)/2$, p y $(1 - p)/2$ en esos tres valores:



- (b). Por la simetría de los dos posibles errores alrededor de a , la esperanza de la lectura es

$$\mathbb{E}[X] = \frac{1-p}{2}(a - \delta) + pa + \frac{1-p}{2}(a + \delta) = a.$$

Usando este valor, la varianza se calcula directamente como

$$\begin{aligned} \text{Var}(X) &= \mathbb{E}[(X - a)^2] \\ &= \frac{1-p}{2}\delta^2 + p \cdot 0^2 + \frac{1-p}{2}\delta^2 \\ &= (1-p)\delta^2. \end{aligned}$$

Por tanto, $\sigma_X = \delta\sqrt{1-p}$. Como $\delta > 0$, el criterio de certificación equivale a

$$\begin{aligned} \delta\sqrt{1-p} \leq \frac{\delta}{2} &\iff \sqrt{1-p} \leq \frac{1}{2} \\ &\iff p \geq \frac{3}{4}. \end{aligned}$$

La condición no depende de a ni del valor positivo de δ .

- (c)(i). Para un error fijo $\delta > 0$, el costo de un termómetro puede escribirse como

$$C = \frac{\lambda}{\delta} + c|X - a|.$$

El término de reparación es cero cuando el termómetro está bien calibrado y vale $c\delta$ cuando está mal calibrado. En consecuencia,

$$\begin{aligned}\mathbb{E}[C] &= \frac{\lambda}{\delta} + c\mathbb{E}[|X - a|] \\ &= \frac{\lambda}{\delta} + c\left(\frac{1-p}{2}\delta + p \cdot 0 + \frac{1-p}{2}\delta\right) \\ &= \frac{\lambda}{\delta} + c(1-p)\delta.\end{aligned}$$

(c)(ii). Si $c > 0$, debemos minimizar la función anterior sobre $\delta > 0$; $\delta = 0$ no es factible porque el costo de arriendo λ/δ no está definido. Su derivada es

$$-\frac{\lambda}{\delta^2} + c(1-p).$$

Al igualarla a cero se obtiene

$$\delta^* = \sqrt{\frac{\lambda}{c(1-p)}}.$$

Además, la segunda derivada es $2\lambda/\delta^3 > 0$, por lo que este punto crítico es el mínimo único. Si $c = 0$, no existe un óptimo finito: el ínfimo del costo es cero y se alcanza solo en el límite $\delta \rightarrow \infty$. Si se admitiera $c < 0$, el costo no estaría acotado inferiormente.

Pregunta 51. (Examen - Primavera 2022, P1)

Responda las siguientes preguntas.

- (a) Un grupo de 10 personas lanza un dado. Cada persona es de *Tipo A* con probabilidad 25 %, independiente del resto de las personas, y de *Tipo B* con probabilidad complementaria. Una persona de *Tipo A* lanza el dado una sola vez. Una persona de *Tipo B* lanzará nuevamente el dado si es que el resultado de su primer lanzamiento es menor o igual a 3. Note que una persona de *Tipo B* podría lanzar el dado una o dos veces dependiendo de lo que obtenga en el primer lanzamiento.
- Entregue la función de masa de probabilidad (PMF) del número que sale en el último lanzamiento del dado para una persona que se sabe que es de *Tipo B*.
 - Determine el valor esperado del promedio obtenido entre las 10 personas. Dicho promedio se computa sumando el valor obtenido del último lanzamiento de las 10 personas y luego dividiendo por 10.
- (b) Una empresa que se dedica a la fabricación de muebles hace bolsas de pellet con los residuos de madera que resultan de los diferentes procesos. El peso de las bolsas de pellet producidas por la empresa sigue una distribución Normal de media 18 [kg] y desviación estándar 2 [kg]. Cada bolsa dice en su exterior que el contenido es de 18 kilogramos. Hay una normativa que indica que un producto *está en norma* si es que el contenido es mayor o igual al 95 % del valor indicado en su envoltorio. Suponga que usted compra una bolsa de pellet, la cual *está en norma*. Determine la probabilidad de que dicha bolsa contenga más de 18 kilogramos de pellet. Para esta pregunta se pide que entregue el desarrollo y el valor numérico de la respuesta.
- (c) En una fila de 5 asientos, uno al lado del otro, se sentarán al azar 5 personas que denominaremos *A*, *B*, *C*, *D* y *E*. La persona *A* no se lleva bien con la persona *B*. ¿Cuál es la probabilidad de que *A* no quede sentado al lado de *B*? Para esta pregunta no es

necesario entregar el valor numérico de la respuesta; puede entregar la expresión con todos los valores en esta.

Solución

(a)(i). Sea X el último resultado de una persona de Tipo B. Los valores 1, 2, 3 solo pueden aparecer en el segundo lanzamiento, mientras que 4, 5, 6 pueden aparecer en cualquiera de los dos. Por ello,

$$p_X(x) = \begin{cases} \frac{1}{12}, & x \in \{1, 2, 3\}, \\ \frac{1}{4}, & x \in \{4, 5, 6\}, \\ 0, & \text{en otro caso.} \end{cases}$$

(a)(ii). El último resultado de una persona Tipo A tiene media 3,5, y el de una persona Tipo B tiene media

$$\frac{1+2+3}{12} + \frac{4+5+6}{4} = \frac{17}{4} = 4,25.$$

Por esperanza total, la media del resultado de cada persona es

$$0,25(3,5) + 0,75(4,25) = 4,0625.$$

La linealidad de la esperanza implica que el promedio de las 10 personas tiene el mismo valor esperado.

(b). Estar en norma equivale a $X \geq 17,1$. Si $Z = (X - 18)/2 \sim \mathcal{N}(0, 1)$,

$$\mathbb{P}(X > 18 \mid X \geq 17,1) = \frac{\mathbb{P}(X > 18)}{\mathbb{P}(X \geq 17,1)} = \frac{1 - \Phi(0)}{\Phi(0,45)} \approx \frac{0,5}{0,6736} \approx 0,7422.$$

(c). Hay $5!$ ordenamientos equiprobables. Para que A y B queden juntos se elige uno de los cuatro pares de asientos consecutivos, se ordena a A, B de $2!$ formas y al resto de $3!$ formas. Entonces

$$\mathbb{P}(A \text{ no queda junto a } B) = 1 - \frac{4 \cdot 2! \cdot 3!}{5!} = \frac{3}{5}.$$

Pregunta 52. (Examen - Primavera 2024, P1)

Responda las siguientes preguntas.

- (a) Martina tiene un clóset cápsula que contiene 5 poleras diferentes, 4 pantalones diferentes y 3 polerones diferentes. Con estas prendas puede crear diferentes combinaciones de looks.
- (i) Martina tiene un viaje y solo puede llevar 4 prendas en su maleta. ¿Cuántas combinaciones posibles puede llevar? Considere que es válido llevar las 4 prendas del mismo tipo o no llevar prendas de algún tipo.
 - (ii) Hay exactamente una prenda negra de cada tipo; por ende, hay 4 poleras, 3 pantalones y 2 polerones que no son negros. ¿De cuántas formas se puede vestir Martina si usa una polera, un pantalón y un polerón, de tal forma que a lo más tenga puesta una prenda negra?
- (b) En la sala silenciosa de la biblioteca hay 3 mesas: A , B y C . Cada mesa puede estar vacía, o ser ocupada por 1 o 2 estudiantes. Hay dos estudiantes que eligen al azar, es decir, con probabilidades equiprobables, y de manera independiente una mesa para estudiar. El bibliotecario quiere saber cuántas mesas quedarán ocupadas; para eso, le pide a usted que lo ayude a resolver las siguientes preguntas:
- (i) Determine el rango y la PMF del número de mesas ocupadas.
 - (ii) Calcule la esperanza y la varianza del número de mesas ocupadas.

- (c) La empresa Batpower fabrica baterías, las cuales pueden ser de **calidad alta**, de **calidad media** o de **calidad baja**. Un cliente importante compra un lote de 100 baterías. Cada batería, de manera independiente, tiene probabilidad 0,2 de ser de calidad alta, 0,3 de ser de calidad media y 0,5 de ser de calidad baja. La empresa debe pagar una compensación al cliente si el número de baterías de calidad baja es mayor o igual que 10 unidades. Si se sabe que el lote contiene exactamente 70 baterías de calidad alta, ¿cuál es la probabilidad de que Batpower deba pagar la compensación al cliente?

Solución

(a)(i). Hay 12 prendas distintas y se eligen 4, de modo que existen

$$\binom{12}{4} = 495$$

combinaciones.

(a)(ii). Sin prendas negras hay $4 \cdot 3 \cdot 2 = 24$ atuendos. Con exactamente una prenda negra hay

$$1 \cdot 3 \cdot 2 + 4 \cdot 1 \cdot 2 + 4 \cdot 3 \cdot 1 = 26.$$

En total hay $24 + 26 = 50$ posibilidades.

(b)(i). Si X es el número de mesas ocupadas, entonces $\mathcal{R}_X = \{1, 2\}$. Una vez que el primer estudiante elige, el segundo coincide con probabilidad $1/3$; por ello,

$$p_X(x) = \begin{cases} \frac{1}{3}, & x = 1, \\ \frac{2}{3}, & x = 2, \\ 0, & \text{en otro caso.} \end{cases}$$

(b)(ii). A partir de la PMF,

$$\mathbb{E}[X] = \frac{1}{3} + \frac{4}{3} = \frac{5}{3}, \quad \mathbb{E}[X^2] = \frac{1}{3} + \frac{8}{3} = 3.$$

Luego, $\text{Var}(X) = 3 - (5/3)^2 = 2/9$.

(c). Condicionalmente a que 70 baterías son de calidad alta, las otras 30 no son altas. Para cada una de ellas,

$$\mathbb{P}(\text{baja} \mid \text{no alta}) = \frac{0,5}{0,3 + 0,5} = 0,625.$$

Así, el número B de baterías de calidad baja, dado el antecedente, distribuye Binomial(30, 0,625), y

$$\mathbb{P}(B \geq 10 \mid 70 \text{ altas}) = 1 - \sum_{k=0}^9 \binom{30}{k} (0,625)^k (0,375)^{30-k}.$$

5.2. Distribuciones discretas

Resumen

Bernoulli.

Una variable $X \sim \text{Bernoulli}(p)$ indica el resultado de un ensayo con probabilidad de éxito p :

$$p_X(x) = p^x(1-p)^{1-x}, \quad x \in \{0, 1\}.$$

Su esperanza es $\mathbb{E}[X] = p$ y su varianza es $\text{Var}(X) = p(1-p)$.

Binomial.

Si X cuenta los éxitos en n ensayos Bernoulli independientes con el mismo parámetro p , entonces $X \sim \text{Binomial}(n, p)$ y

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, \dots, n.$$

Además, $\mathbb{E}[X] = np$ y $\text{Var}(X) = np(1-p)$.

Geométrica.

Si X cuenta los ensayos hasta el primer éxito, usando soporte $\{1, 2, \dots\}$,

$$p_X(k) = (1-p)^{k-1}p, \quad k \geq 1.$$

Su esperanza es $1/p$ y su varianza es $(1-p)/p^2$. Para enteros $s, t \geq 0$ satisface la propiedad de falta de memoria:

$$\mathbb{P}(X > s+t \mid X > s) = \mathbb{P}(X > t).$$

Binomial negativa.

Si X cuenta los ensayos hasta el m -ésimo éxito,

$$p_X(k) = \binom{k-1}{m-1} p^m (1-p)^{k-m}, \quad k \geq m.$$

Su esperanza es m/p y su varianza es $m(1-p)/p^2$. Para $m = 1$ se recupera la distribución geométrica.

Hipergeométrica.

Al extraer n objetos sin reemplazo de una población de tamaño N con K éxitos, el número X de éxitos tiene PMF

$$p_X(k) = \frac{\binom{K}{k} \binom{N-K}{n-k}}{\binom{N}{n}}$$

para los valores enteros factibles de k . Su esperanza y varianza son

$$\mathbb{E}[X] = n \frac{K}{N},$$

$$\text{Var}(X) = n \frac{K}{N} \left(1 - \frac{K}{N}\right) \frac{N-n}{N-1}.$$

El factor $(N-n)/(N-1)$ corrige la dependencia introducida por el muestreo sin reemplazo.

Poisson.

Si $X \sim \text{Poisson}(\lambda)$, con $\lambda > 0$,

$$p_X(k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k \in \mathbb{N}_0.$$

Se cumple $\mathbb{E}[X] = \text{Var}(X) = \lambda$.

Propiedades de cierre.

- **Suma binomial.** Si $X_i \sim \text{Binomial}(n_i, p)$ son independientes y comparten p , entonces $\sum_i X_i \sim \text{Binomial}(\sum_i n_i, p)$.
- **Suma Poisson.** Si $X_i \sim \text{Poisson}(\lambda_i)$ son independientes, entonces $\sum_i X_i \sim \text{Poisson}(\sum_i \lambda_i)$.
- **Adelgazamiento Poisson.** Si cada llegada de $N \sim \text{Poisson}(\lambda)$ se clasifica independientemente como éxito con probabilidad p , el número de éxitos es $\text{Poisson}(\lambda p)$ y es independiente del número de fracasos.
- **Aproximación Poisson.** Si n es grande, p es pequeño y $np \approx \lambda$, una $\text{Binomial}(n, p)$ se aproxima por una $\text{Poisson}(\lambda)$.

5.2.1. Preguntas

Pregunta 53.

En un pueblo aislado viven m habitantes con una gran afición por su equipo local, el cual se juega la posibilidad de clasificar al campeonato nacional en los próximos días. Las posibilidades de clasificar no son muchas, por lo cual la decisión del precio q a cobrar por cada entrada al espectáculo no es evidente: si se cobra un precio muy alto, poca gente podrá acudir al evento. Según estimaciones realizadas con antecedentes históricos, se sabe que si se cobra un precio q por la entrada, entonces cada habitante del pueblo asiste, independiente de los demás, con probabilidad e^{-q} al estadio. Suponga además que el estadio tiene capacidad m , es decir, todas las personas del pueblo podrían asistir.

a) ¿Qué pasa si la entrada es gratuita? ¿Y si el precio es muy alto?

- b) Si se cobra q , ¿cuál es el valor esperado y la varianza del ingreso monetario?
- c) ¿Qué precio se debe cobrar si se desea maximizar los ingresos esperados?

Solución

(a). Si $q = 0$, entonces $e^{-q} = 1$: todas las personas asisten con probabilidad uno, pero el ingreso es cero porque la entrada es gratuita. Si el precio aumenta sin cota, la asistencia individual satisface $e^{-q} \rightarrow 0$. Más precisamente, el ingreso esperado que se obtiene en el inciso siguiente cumple

$$\mathbb{E}[I] = mqe^{-q} \rightarrow 0 \quad \text{cuando } q \rightarrow \infty.$$

(b). Como cada habitante asiste independientemente con probabilidad e^{-q} , $X \sim \text{Binomial}(m, e^{-q})$. Si $I = qX$ es el ingreso,

$$\mathbb{E}[I] = qme^{-q}, \quad \text{Var}(I) = q^2me^{-q}(1 - e^{-q}).$$

(c). El ingreso esperado es $g(q) = mqe^{-q}$ y $g'(q) = me^{-q}(1 - q)$. Por tanto, g crece para $0 \leq q < 1$ y decrece para $q > 1$; el único máximo global se alcanza en $q^* = 1$.

Pregunta 54.

Los auxiliares de un ramo instalan un nuevo sistema para decidir la nota final de los alumnos: guardan en una caja siete pelotas numeradas desde el 1 hasta el 7, y para cada alumno sacan una pelota al azar, le ponen como nota el número que indica la pelota, y luego la devuelven a la caja. Suponga que la sección tiene 50 personas (ordenadas según su número de lista) y que la nota de aprobación es 4.

- a) Calcule la probabilidad de que el alumno 4 pase el ramo.
- b) Calcule la probabilidad de que al menos una persona pase el ramo.
- c) Calcule la probabilidad de que más de 45 personas pasen el ramo.
- d) Calcule el número esperado de estudiantes que pasan el ramo.
- e) Calcule la probabilidad de que el alumno 7 sea el tercero en pasar el ramo.
- f) Sea M el número de lista de la primera persona en obtener nota 7. Calcule el valor esperado y la varianza de M .
- Suponga que ahora tienen 10 pelotas con cada una de las notas, es decir, tienen 70 pelotas (10 pelotas con un 1, 10 pelotas con un 2, etc), y que ya no se devuelven a la caja las pelotas extraídas.
- g) Repita los cálculos desde a) hasta e) en este nuevo contexto.
- h) Sea M el número de lista de la primera persona en obtener nota 7. Calcule la probabilidad de que M sea menor o igual a 3.

Solución

(a). Con reposición, las notas de estudiantes distintos son independientes. Si

$$I_i = \mathbb{1}_{\{\text{la persona } i \text{ aprueba}\}},$$

entonces $I_i \sim \text{Bernoulli}(4/7)$. En particular,

$$\mathbb{P}(\text{la persona 4 aprueba}) = \frac{4}{7}.$$

(b). Sea $X = \sum_{i=1}^{50} I_i$. Como $X \sim \text{Bin}(50, 4/7)$,

$$\mathbb{P}(X \geq 1) = 1 - \mathbb{P}(X = 0) = 1 - \left(\frac{3}{7}\right)^{50} \approx 1 - 3,99 \times 10^{-19}.$$

(c). La probabilidad de que aprueben más de 45 personas es

$$\mathbb{P}(X > 45) = \sum_{k=46}^{50} \binom{50}{k} \left(\frac{4}{7}\right)^k \left(\frac{3}{7}\right)^{50-k} \approx 5,8 \times 10^{-8}.$$

(d). La media de una variable binomial es el número de ensayos por la probabilidad de éxito. Por tanto,

$$\mathbb{E}[X] = 50 \left(\frac{4}{7}\right) = \frac{200}{7} \approx 28,57.$$

(e). Sea Y el número de lista del tercer estudiante que aprueba. Entonces $Y \sim \text{NegBin}(3, 4/7)$, con la convención de contar ensayos hasta el tercer éxito, y

$$\mathbb{P}(Y = 7) = \binom{6}{2} \left(\frac{4}{7}\right)^3 \left(\frac{3}{7}\right)^4 \approx 0,0944.$$

(f). Si el procedimiento se prolonga indefinidamente y M es el índice de la primera nota 7, entonces

$$M \sim \text{Geom}\left(\frac{1}{7}\right), \quad \mathbb{E}[M] = 7, \quad \text{Var}(M) = 42.$$

En la sección real de 50 personas, nadie obtiene un 7 con probabilidad $(6/7)^{50} \approx 0,0004494$; sin una convención adicional, M no está definida en ese evento. Condicionando a que aparezca al menos un 7,

$$\mathbb{P}(M = m \mid M \leq 50) = \frac{(1/7)(6/7)^{m-1}}{1 - (6/7)^{50}}, \quad m = 1, \dots, 50,$$

cuya media y varianza son aproximadamente 6,9775 y 40,8754, respectivamente.

(g). Sin reposición, las respuestas dejan de ser independientes.

(g)(a). Aun así, cada posición recibe marginalmente una de las 70 pelotas de manera uniforme, por lo que $\mathbb{P}(\text{la persona 4 aprueba}) = 40/70 = 4/7$. Equivalentemente, si $X_3 \sim \text{HGeom}(70, 40, 3)$ cuenta los aprobados entre las tres primeras personas, probabilidades totales entrega

$$\sum_{j=0}^3 \frac{40-j}{67} \frac{\binom{40}{j} \binom{30}{3-j}}{\binom{70}{3}} = \frac{4}{7}.$$

(g)(b). Solo hay 30 pelotas con nota inferior a 4. Entre 50 estudiantes habrá al menos 20 aprobados, de modo que la probabilidad de que apruebe al menos uno es 1.

(g)(c). Como existen solo 40 pelotas con nota de aprobación, no pueden aprobar más de 40 personas. La probabilidad de que aprueben más de 45 es 0.

(g)(d). El número X de aprobados entre 50 estudiantes distribuye $\text{HGeom}(70, 40, 50)$. Por tanto,

$$\mathbb{E}[X] = 50 \frac{40}{70} = \frac{200}{7} \approx 28,57.$$

(g)(e). Para que la persona 7 sea la tercera en aprobar, debe haber exactamente dos aprobados entre las seis primeras y luego debe extraerse una de las 38 pelotas aprobatorias restantes. Así,

$$\mathbb{P}(\text{la persona 7 es la tercera en aprobar}) = \frac{\binom{40}{2} \binom{30}{4}}{\binom{70}{6}} \frac{38}{64} \approx 0,0968.$$

(h). El evento $M \leq 3$ es el complemento de que las tres primeras personas obtengan una nota distinta de 7. En consecuencia,

$$\mathbb{P}(M \leq 3) = 1 - \frac{60}{70} \frac{59}{69} \frac{58}{68} \approx 0,3749.$$

Pregunta 55.

Sean X, Y V.A.'s independientes de distribución $\text{Geom}(p)$, es decir $\mathbb{P}(X = n) = \mathbb{P}(Y = n) = p(1 - p)^{n-1}$.

- Calcule la función de probabilidad de $Z = X + Y$
- Calcule $\mathbb{P}(X = Y)$ y $\mathbb{P}(X < Y)$.

Solución

Para que la distribución geométrica esté bien definida, suponemos $0 < p \leq 1$.

(a). Como X e Y son independientes y toman valores en $\{1, 2, \dots\}$, para todo entero $n \geq 2$,

$$\begin{aligned} p_Z(n) &= \sum_{y=1}^{n-1} \mathbb{P}(X = n - y) \mathbb{P}(Y = y) \\ &= (n - 1)p^2(1 - p)^{n-2}. \end{aligned}$$

Esta función vale 0 fuera de $\{2, 3, \dots\}$.

(b). Por independencia,

$$\mathbb{P}(X = Y) = \sum_{y=1}^{\infty} p^2(1 - p)^{2(y-1)} = \frac{p}{2 - p}.$$

Como X e Y son idénticamente distribuidas, $\mathbb{P}(X < Y) = \mathbb{P}(Y < X)$. Por ello,

$$\mathbb{P}(X < Y) = \frac{1 - \mathbb{P}(X = Y)}{2} = \frac{1 - p}{2 - p}.$$

Pregunta 56.

Un cultivo de bacterias de un cierto tipo puede proliferar o bien extinguirse, lo que ocurre con probabilidad p y probabilidad $1 - p$ respectivamente. Para realizar un estudio, usted genera n de estos cultivos de manera independiente.

- Indique la función p_X de la variable X correspondiente al número de cultivos que proliferan.
- Debido a un corte de luz, el sistema de refrigeración de los cultivos deja de funcionar por unas horas, lo cual significa que cada cultivo que proliferó inicialmente se mantendrá vivo o se extinguirá con probabilidad q y $1 - q$ respectivamente, independiente del resto. ¿Cuál es la distribución de la variable aleatoria Y correspondiente a los cultivos que quedan vivos?

Solución

(a). Sea

$$X_i = \mathbb{1}_{\{\text{el cultivo } i \text{ prolifera}\}}.$$

Como los X_i son independientes y distribuyen $\text{Bernoulli}(p)$,

$$X = \sum_{i=1}^n X_i \sim \text{Binomial}(n, p).$$

Por lo tanto,

$$p_X(k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k}, & k = 0, 1, \dots, n, \\ 0, & \text{en otro caso.} \end{cases}$$

(b). Cada cultivo queda vivo si primero prolifera y luego sobrevive al corte. Por independencia, esto ocurre con probabilidad pq . Como los cultivos también son independientes entre sí,

$$Y \sim \text{Binomial}(n, pq),$$

es decir,

$$p_Y(j) = \begin{cases} \binom{n}{j} (pq)^j (1-pq)^{n-j}, & j = 0, 1, \dots, n, \\ 0, & \text{en otro caso.} \end{cases}$$

Pregunta 57.

Las profesoras auxiliares Josefa e Isidora están preocupados debido a que las notas del primer control se distribuyeron de manera normal con una media de 3.7 y una desviación estándar de 1. Después de analizar los resultados, se ha decidido que cualquier alumno con una nota inferior a 4 necesitará apoyo urgente.

Justo después de tomar esta decisión, las profesoras auxiliares reciben un correo electrónico de un estudiante preocupado por un ejercicio.

- (a) ¿Cuál es la probabilidad de que este alumno en particular necesite ser ayudado de manera urgente?

Los estudiantes se preparan intensamente para el próximo control, pero suelen recibir, en promedio, dos tareas por semana. Conscientes de este desafío, los auxiliares han elaborado una guía de estudios diseñada para ser completada en tres días, esto es, considerando, que los estudiantes no dedican todo su tiempo exclusivamente al estudio, ya que también deben atender otras obligaciones académicas y personales.

Dado lo anterior, se evalúa la probabilidad de que los estudiantes no reciban ninguna tarea durante los tres días previos al control. Esto les permitiría dedicar tiempo completo a resolver la guía de estudios y así mejorar su preparación para el examen.

- (b) ¿Cuál es la probabilidad de que no se asignen tareas en esos tres días críticos?

Lamentablemente justo durante los tres días previos al control, deben entregar una tarea importante que aún no han comenzado. Este contratiempo significa que no podrán dedicar el tiempo necesario a estudiar la guía de estudios diseñada específicamente para esos días. A pesar de esto, los estudiantes son conocidos por su excelencia académica, lo cual les permite tener una alta probabilidad de resolver correctamente las preguntas de la guía incluso sin la totalidad de estos días.

Dado que la probabilidad de que un estudiante resuelva correctamente una pregunta bajo estas circunstancias es de 0.7, y considerando que la guía consta de 14 preguntas.

- (c) ¿Cuál la probabilidad de que resuelvan al menos la mitad de las preguntas de la guía?

La nota final del control 2 es una V.A Aleatoria Y tal que si resuelven menos de 3 preguntas de la guía de estudio, su nota se mantiene en 1. Mientras que por cada dos preguntas adicionales resueltas correctamente, su nota sube un punto, hasta un máximo de 7.

- (d) Indique la PMF de la nota final del control.

Solución

(a). Si $X \sim \mathcal{N}(3, 7, 1)$ es la nota, entonces

$$\mathbb{P}(X < 4) = \Phi\left(\frac{4 - 3,7}{1}\right) = \Phi(0,3) \approx 0,6179.$$

(b). Conocer únicamente el promedio de dos tareas por semana no basta para calcular la probabilidad solicitada. Bajo el supuesto adicional usual de que las asignaciones siguen un proceso de Poisson con tasa constante de dos tareas por semana, el número N de tareas durante tres días tiene parámetro $2(3/7) = 6/7$. En ese modelo,

$$\mathbb{P}(N = 0) = e^{-6/7} \approx 0,4244.$$

(c). Sea X el número de preguntas resueltas correctamente. Si, además, los resultados de las 14 preguntas son independientes, entonces $X \sim \text{Binomial}(14, 0,7)$. Bajo esta hipótesis, resolver al menos la mitad tiene probabilidad

$$\mathbb{P}(X \geq 7) = 1 - \sum_{j=0}^6 \binom{14}{j} (0,7)^j (0,3)^{14-j}.$$

(d). El enunciado no especifica qué nota corresponde a exactamente tres respuestas correctas. Para completar la regla, adoptaremos la convención de que tres aciertos todavía mantienen la nota 1; otra convención produciría una función de masa diferente. Con esta elección,

$$Y = \begin{cases} 1, & 0 \leq X \leq 3, \\ 1 + \left\lfloor \frac{X-2}{2} \right\rfloor, & 4 \leq X \leq 13, \\ 7, & X = 14. \end{cases}$$

Para escribir la función de masa de Y , denotemos

$$b_j = \binom{14}{j} (0,7)^j (0,3)^{14-j}, \quad j = 0, \dots, 14.$$

Así, la nota es k cuando $X \in \{2k, 2k+1\}$, para $k = 2, \dots, 6$, y vale 7 cuando $X = 14$. En consecuencia,

$$p_Y(y) = \begin{cases} \sum_{j=0}^3 b_j, & y = 1, \\ b_{2y} + b_{2y+1}, & y = 2, \dots, 6, \\ b_{14}, & y = 7, \\ 0, & \text{en otro caso.} \end{cases}$$

5.2.2. Preguntas de Evaluación**Pregunta 58.** (Control 1 - Otoño 2022, P3)

Se tiene un curso con n estudiantes. Considere que cada una de las personas puede estar de cumpleaños en cualquiera de los 365 días del año con igual probabilidad, de manera independiente.

- Sea X el número de estudiantes que está de cumpleaños el mismo día que la profesora del curso. Identifique qué distribución sigue X , indicando el(los) valor(es) de sus parámetros, justifique brevemente su respuesta.
- Suponga ahora que en el curso se sabe que hay exactamente k personas que tienen su cumpleaños en el mismo mes que la profesora del curso. Un día de clases, en donde todos los estudiantes asisten, la profesora decide consultar a los s alumnos de la primera fila el

mes en el que están de cumpleaños. Sea Y la cantidad de estudiantes de la primera fila que están de cumpleaños en el mismo mes que la profesora. Asuma que los alumnos se sientan al azar en la sala, y que la capacidad de la sala es exactamente igual al número de alumnos que está tomando dicho curso. Calcule la probabilidad de que exactamente dos personas de la primera fila posean mes de cumpleaños igual al de la profesora.

- (c) La profesora quiere determinar si es que en el curso hay al menos una fecha de cumpleaños repetida entre los alumnos. Para esto, la profesora pregunta a un alumno al azar su fecha de cumpleaños, el alumno responde su fecha, y si es que hay otra persona del curso que tiene su fecha de cumpleaños en el mismo día, esta persona levanta la mano y la profesora termina de preguntar. En caso contrario (es decir si nadie tiene la misma fecha de cumpleaños que la persona consultada inicialmente), entonces la profesora escoge un segundo alumno al azar (diferente del primero), le pregunta su fecha de cumpleaños, y si es que hay otro alumno que tiene fecha de cumpleaños en el mismo día, entonces la profesora termina de preguntar. En caso contrario, la profesora escoge a otro alumno al azar (diferente de los dos anteriores), y repite el procedimiento. Así hasta que alguien diga que tiene una fecha de cumpleaños que coincide con la persona consultada, o bien hasta interrogar a todo el curso. Se define la VA Z como el número de veces que la profesora interroga a un alumno al azar consultando su fecha de cumpleaños (notar que dicha variable aleatoria es a lo más igual al número de alumnos). Determine la función de distribución acumulada de Z . Nota: No necesariamente se trata de una distribución con nombre vista en clases.

Solución

(a). Para cada estudiante, coincidir con el cumpleaños de la profesora es un evento de probabilidad $1/365$, independiente de los demás estudiantes. Por tanto,

$$X \sim \text{Binomial}\left(n, \frac{1}{365}\right).$$

(b). Los s estudiantes de la primera fila forman una muestra sin reemplazo de los n estudiantes, entre los cuales hay exactamente k cuyo cumpleaños ocurre en el mismo mes que el de la profesora. En consecuencia,

$$Y \sim \text{HGeom}(N = n, K = k, m = s),$$

y

$$\mathbb{P}(Y = 2) = \frac{\binom{k}{2} \binom{n-k}{s-2}}{\binom{n}{s}}.$$

(c). Para un entero $k < n$, el evento $\{Z > k\}$ significa que ninguna de las primeras k personas consultadas comparte su fecha de cumpleaños con otra persona del curso. Esto requiere dos condiciones: que esas k fechas sean distintas y que las $n - k$ personas restantes eviten todas esas fechas. La probabilidad de que las primeras k fechas sean distintas es

$$\frac{365 \cdot 364 \cdots (365 - k + 1)}{365^k}.$$

Condicionado a esas fechas, cada una de las $n - k$ personas restantes evita las k fechas con probabilidad $(365 - k)/365$. Por independencia,

$$\mathbb{P}(Z > k) = \frac{365 \cdot 364 \cdots (365 - k + 1)}{365^k} \left(\frac{365 - k}{365}\right)^{n-k}.$$

El producto se interpreta como cero cuando $k > 365$. Para $z \in \mathbb{R}$, sea $k = \lfloor z \rfloor$. La CDF queda

$$F_Z(z) = \begin{cases} 0, & z < 1, \\ 1 - \frac{365 \cdot 364 \cdots (365 - k + 1)}{365^k} \left(\frac{365 - k}{365} \right)^{n-k}, & 1 \leq z < n, \\ 1, & z \geq n. \end{cases}$$

Pregunta 59. (Control 2 - Primavera 2022, P3)

Dos amigos, Antonia y Benito, tienen una importante reunión en el otro extremo de la ciudad. La ciudad es convenientemente modelada por una grilla de 7×7 . Específicamente, Antonia y Benito se encuentran en el punto con coordenadas $(0, 0)$, mientras que la reunión se realizará en el punto con coordenadas $(7, 7)$. Para trasladarse al lugar de la reunión, en cada esquina, par de coordenadas cuya abscisa y ordenada son valores enteros, lanzan una moneda, la cual tiene probabilidad p de salir cara y $1 - p$ de salir sello. Cada vez que la moneda sale cara, avanzan por un arco hacia la derecha, es decir, avanzan hacia la derecha en una unidad. Si, por el contrario, la moneda sale sello, entonces avanzan por el arco hacia arriba. Asuma que, en caso de llegar a los puntos con abscisa u ordenada, pero no ambas, igual a 7, los siguientes tramos son recorridos en línea recta en la dirección que apunta hacia el punto de coordenadas $(7, 7)$.

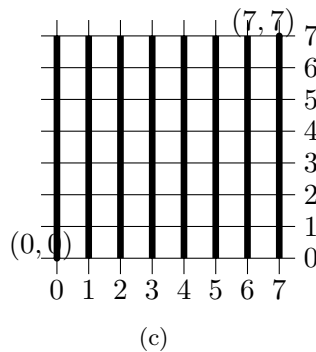
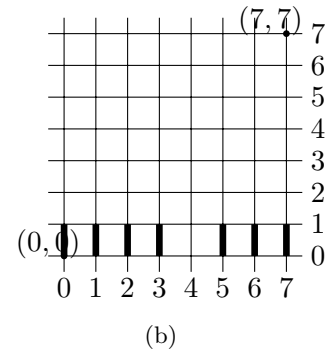
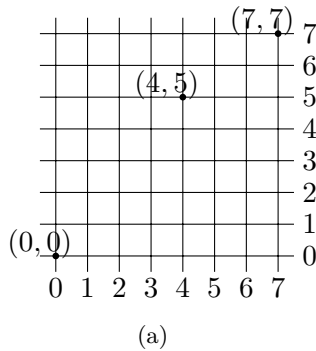


Figura 5.1: Mapas de la ciudad para las partes (a), (b) y (c), respectivamente.

- (a) Calcule la probabilidad de pasar por el punto con coordenadas $(x, y) = (4, 5)$ en el trayecto hacia la reunión.

Hint: Note que solo importa mirar los primeros $x + y$ movimientos en la grilla. Véase la figura (a).

- (b) Suponga que en las calles marcadas con línea gruesa hay alto tráfico. Calcule la probabilidad de que Antonia y Benito **no** hayan pasado por zonas de alto tráfico en su trayecto hacia la reunión. Véase la figura (b).
- (c) Ahora suponga que existe alto tráfico en todos los arcos en sentido vertical. Se sabe que Antonia manejó 6 arcos y Benito el resto, pero no se sabe cuáles manejaron cada uno. Se sabe además que Benito se enoja cuando maneja 3 o más arcos con tráfico. ¿Cuál es la probabilidad de que Benito haya llegado enojado a la reunión?

Hint: Note que, para responder la pregunta, no importa el trayecto específico que hicieron para llegar desde el punto de origen al punto de destino. Véase la figura (c).

Solución

(a). Pasar por $(4, 5)$ exige que, entre los primeros nueve movimientos, haya exactamente cuatro pasos hacia la derecha y cinco hacia arriba. Si R denota el número de pasos hacia la derecha en esos nueve movimientos, entonces $R \sim \text{Bin}(9, p)$. Por tanto,

$$\mathbb{P}(\text{pasar por } (4, 5)) = \mathbb{P}(R = 4) = \binom{9}{4} p^4 (1-p)^5.$$

(b). Los únicos arcos verticales sin tráfico en la fila inferior están a la altura de $x = 4$. Para evitar todos los arcos marcados, los primeros cuatro movimientos deben ser hacia la derecha, de modo que se llegue a $(4, 0)$, y el quinto debe ser hacia arriba, hasta $(4, 1)$. Desde allí ya no es necesario usar un arco con tráfico. Por independencia de los lanzamientos,

$$\mathbb{P}(\text{no pasar por zonas de alto tráfico}) = p^4 (1-p).$$

(c). Todo trayecto desde $(0, 0)$ hasta $(7, 7)$ contiene catorce arcos: siete verticales, con tráfico, y siete horizontales, sin tráfico. Como Antonia manejó seis, Benito manejó ocho. Si, además, las ocho posiciones del trayecto que maneja Benito se eligen uniformemente entre los catorce arcos, y V es el número de arcos verticales que maneja, entonces

$$V \sim \text{Hipergeom}(14, 7, 8),$$

donde 14 es el número total de arcos, 7 el número de arcos con tráfico y 8 el número manejado por Benito. Benito llega enojado cuando $V \geq 3$, por lo que

$$\begin{aligned} \mathbb{P}(V \geq 3) &= \sum_{k=3}^7 \mathbb{P}(V = k) \\ &= \sum_{k=3}^7 \frac{\binom{7}{k} \binom{7}{8-k}}{\binom{14}{8}}. \end{aligned}$$

Sin la hipótesis de selección uniforme, saber solamente cuántos arcos manejó cada persona no determina una probabilidad única.

Pregunta 60. (Control 2 - Otoño 2023, P2)

Usted tiene un amigo llamado *LeBron Jokic* el cual estará de cumpleaños el próximo mes. Su amigo es un fan del basketball y sobre todo de la NBA, motivo por el cual usted decide regalarle una polera de un equipo de la NBA. Al cumpleaños asistirán 10 amigos y amigas de *LeBron Jokic* (es decir, usted y 9 amigos/as más sin contar a *LeBron Jokic*). Conversando con los demás, se da cuenta que todos van a regalarle una polera de la NBA a *LeBron Jokic*. Dado que nadie sabe el equipo favorito del cumpleañosero, los amigos se ponen de acuerdo para comprarle una polera de un equipo distinto cada uno. Así, los 10 invitados van juntos a la misma tienda de poleras de basketball, de modo que cada uno compra una polera de un equipo al azar de la NBA, **de tal forma que ninguna polera se repite**.

Considere que hay 30 equipos en la NBA. Además, *LeBron Jokic* es fan de varios equipos pues le gustan distintos jugadores, en particular, simpatiza con cuatro equipos distintos.

Responda las siguientes preguntas. Además, **para cada parte indique, en caso de existir, la distribución de probabilidad (con su nombre y valor(es) de parámetro(s)) que modele el fenómeno preguntado. En caso de que ninguna distribución de las vistas en clases coincida con lo que se pregunta, indicar explícitamente que ninguna de las distribuciones vistas en clases sirve para modelar el fenómeno descrito.**

- (a) ¿Cuál es la probabilidad de que el cumpleaños reciba 2 o más poleras que le gusten?
- (b) ¿Cuál es la probabilidad de que el quinto regalo sea el primero que le guste? (Note que se está asumiendo que el cumpleaños está abriendo los regalos en un cierto orden).

Suponga el mismo contexto de antes, con la única diferencia que los 10 amigos no se coordinan en la compra de los regalos. Como resultado, cada uno de los 10 invitados va a comprar una polera de un equipo al azar de la NBA, sin embargo, **es posible que se repitan los equipos de las poleras compradas por los invitados.**

- (c) ¿Cual es la probabilidad de que el cumpleaños reciba 2 o más poleras que le gusten? (Note que al cumpleaños le gustan las poleras si es que calzan con uno de los cuatro equipos, independiente si las poleras que recibe son repetidas)
- (d) ¿Cuál es la probabilidad de que el quinto regalo sea el primero que le guste?

Solución

(a). Sea X el número de las diez poleras que pertenecen a alguno de los cuatro equipos favoritos. Como las poleras se escogen sin repetición,

$$X \sim \text{HGeom}(N = 30, K = 4, n = 10).$$

Por tanto,

$$\mathbb{P}(X \geq 2) = \sum_{k=2}^4 \frac{\binom{4}{k} \binom{26}{10-k}}{\binom{30}{10}}.$$

El límite superior es 4 porque solo existen cuatro equipos favoritos y ningún equipo se repite.

(b). Extienda el orden de compra a una permutación aleatoria de los 30 equipos y sea T la posición del primer equipo favorito. Entonces T tiene distribución hipergeométrica negativa, si esta se incluye entre las distribuciones estudiadas, y el evento $T = 5$ depende solo de los primeros cinco lugares. Por ello,

$$\mathbb{P}(T = 5) = \frac{26}{30} \frac{25}{29} \frac{24}{28} \frac{23}{27} \frac{4}{26}.$$

(c). Ahora las compras son independientes y cada polera pertenece a un equipo favorito con probabilidad $4/30$. En consecuencia,

$$X \sim \text{Binomial}\left(10, \frac{4}{30}\right),$$

y

$$\mathbb{P}(X \geq 2) = \sum_{k=2}^{10} \binom{10}{k} \left(\frac{4}{30}\right)^k \left(\frac{26}{30}\right)^{10-k}.$$

La suma puede llegar hasta 10 porque ahora se permiten equipos repetidos.

(d). Bajo compras independientes, el número T de regalos abiertos hasta obtener el primero de un equipo favorito sigue una distribución geométrica con parámetro $4/30$. Por tanto,

$$\mathbb{P}(T = 5) = \left(\frac{26}{30}\right)^4 \frac{4}{30}.$$

Pregunta 61. (Control 2 - Primavera 2023, P1)

Andrea posee el álbum de los Juegos Panamericanos 2023, el cual requiere de N láminas diferentes para ser completado. Cada sobre tiene un costo de s [\$] y trae exactamente una lámina del álbum. Todas las láminas tienen la misma probabilidad de salir en los sobres. En caso de que Andrea logre llenar el álbum, ella lo canjea por un premio que asciende a un valor de r [\$]. La utilidad de Andrea se define como la suma de ingresos, en caso de completar el álbum, menos el costo asociado a la compra de sobres. Los enunciados y supuestos de las partes a continuación aplican solo para la parte respectiva.

- (a) Suponga que $N = 2$, es decir, el álbum solo requiere de dos láminas diferentes para ser completado. Andrea posee el álbum vacío. Con el objetivo de llenar el álbum, Andrea comprará máximo tres sobres. Note que podría ser el caso que compre menos de tres sobres, lo cual ocurriría si llena el álbum antes. Grafique la función de distribución acumulada de la utilidad considerando que $s = \$2$ y $r = \$5$.
- (b) Asuma ahora que $N = 10$. Suponga ahora que Andrea comprará sobres hasta completar el álbum. Determine la utilidad esperada.

Hint: Le podría ser útil considerar la variable aleatoria asociada al número de sobres necesarios para obtener la i -ésima lámina, desde el momento en que se acaba de obtener la $(i - 1)$ -ésima lámina. El caso $i = 1$ sería el número de sobres necesarios para obtener la primera lámina del álbum.

Solución

Los cálculos siguientes requieren suponer que las láminas de sobres distintos son independientes. La equiprobabilidad marginal de los tipos de lámina, por sí sola, no determina las probabilidades pedidas.

(a). La primera lámina siempre es nueva. La segunda completa el álbum con probabilidad $1/2$; en ese caso, Andrea compra dos sobres y su utilidad es $5 - 2(2) = 1$. Si la segunda lámina repite la primera, lo que ocurre con probabilidad $1/2$, Andrea compra un tercer sobre. Condicionado a ello, el tercer sobre completa el álbum con probabilidad $1/2$. Así, la utilidad U tiene la función de masa

$$p_U(u) = \begin{cases} \frac{1}{4}, & u = -6, \\ \frac{1}{4}, & u = -1, \\ \frac{1}{2}, & u = 1, \\ 0, & \text{en otro caso.} \end{cases}$$

En efecto, $U = -6$ si compra tres sobres y no completa el álbum, $U = -1$ si lo completa con el tercer sobre y $U = 1$ si lo completa con el segundo. Al acumular estas probabilidades,

$$F_U(u) = \begin{cases} 0, & u < -6, \\ \frac{1}{4}, & -6 \leq u < -1, \\ \frac{1}{2}, & -1 \leq u < 1, \\ 1, & u \geq 1. \end{cases}$$

La función es escalonada y continua por la derecha. Su gráfica es



(b). Sea L_i el número de sobres que Andrea compra desde que tiene $i - 1$ tipos distintos hasta obtener un tipo nuevo, para $i = 1, \dots, 10$. Cuando ya posee $i - 1$ tipos, quedan $11 - i$ tipos nuevos entre los diez posibles. Por tanto,

$$L_i \sim \text{Geom}\left(\frac{11 - i}{10}\right), \quad \mathbb{E}[L_i] = \frac{10}{11 - i}.$$

Si L es el número total de sobres necesarios para completar el álbum, entonces $L = \sum_{i=1}^{10} L_i$. La linealidad de la esperanza entrega

$$\begin{aligned} \mathbb{E}[L] &= \sum_{i=1}^{10} \frac{10}{11 - i} \\ &= 10 \sum_{j=1}^{10} \frac{1}{j} \approx 29,29. \end{aligned}$$

Andrea recibe el premio r con certeza cuando finalmente completa el álbum y paga s por cada sobre. En consecuencia,

$$\mathbb{E}[U] = r - s\mathbb{E}[L] = r - 10s \sum_{j=1}^{10} \frac{1}{j}.$$

Pregunta 62. (Control 2 - Otoño 2024, P1)

Todos los domingos Don Casimiro prepara su desayuno favorito: huevos con tocino. En la preparación del tocino, este es cortado en 20 trozos pequeños de igual tamaño, donde cada trozo posee dos lados. Luego, Don Casimiro hecha un poco de aceite a la sartén, y luego de bajar el fuego al mínimo, hecha todos los trozos de tocino a la sartén. Cada trozo queda con uno de sus lados tocando hacia abajo (en contacto directo con la sartén) y otro lado hacia arriba. Al cabo de un rato, Don Casimiro revuelve la sartén, de tal forma que cada trozo de tocino, independiente de los otros trozos, se *da vuelta*¹ con probabilidad 0,3, y con probabilidad complementaria no se da vuelta. Considere que Don Casimiro revuelve la sartén un total de 4 veces.

Para que un trozo de tocino quede *bien cocinado*, se requiere que haya sido cocinado por ambos lados hacia abajo en algún instante.

Nota: Si prefiere, puede considerar que los ingredientes son “not-huevos” y “not-tocino”.

- Determine la probabilidad de que un trozo de tocino quede *bien cocinado*.
- Determine la distribución de probabilidad del número de trozos de tocino que quedan *bien*

¹Que un trozo se *dé vuelta* significa que el lado del trozo que estaba hacia abajo queda hacia arriba, y el lado que estaba hacia arriba queda hacia abajo.

cocinados. En caso de ser una de las distribuciones con nombre vista en clases, entregue su nombre y el valor de sus parámetros; en caso contrario, entregue su PMF. **Hint:** Si estima que el valor de la respuesta de la parte (a) le sirve para contestar esta pregunta, puede denotar dicho valor por una letra y dejar así su respuesta en función de esta.

- (c) Considere que para esta parte Don Casimiro no trozará el tocino, sino que lo cocinara en un solo “trozo”. Específicamente, Don Casimiro hecha el tocino (entero) a la sartén, de tal forma que el trozo queda con uno de sus dos lados hacia abajo, y el otro hacia arriba. Con probabilidad un tercio, Don Casimiro revolverá la sartén una vez, y con probabilidad complementaria lo hará dos veces. Cada vez que se revuelve la sartén, hay una probabilidad de 0,3 que el tocino se de vuelta, y con probabilidad complementaria no se da vuelta. El tocino queda bien cocinado si es que ambos lados estuvieron hacia abajo (es decir, en contacto directo con la sartén). Considere una variable aleatoria para el número de veces que se da vuelta el tocino, y otra variable aleatoria para el número de trozos cocinados (dado que es un solo trozo, dicha variable aleatoria puede solo tomar el valor 0 o 1). Si es que se sabe que el tocino **no quedó** bien cocinado, compute la probabilidad de que Don Casimiro haya revuelto la sartén una sola vez. Defina claramente las variables aleatorias que use. **Hint:** le podría ser de utilidad computar la probabilidad de que el tocino **no quede** bien cocido dado el número de veces que se revuelve la sartén, para los diferentes valores que puede tomar revolver la sartén.

Solución

(a). Suponiendo, además, que para cada trozo los resultados de las cuatro revoluciones son independientes, el trozo queda bien cocinado si se da vuelta al menos una vez. Es más sencillo calcular el evento complementario: que no se dé vuelta en ninguna de las cuatro ocasiones. Como cada intento tiene probabilidad 0,7 de no producir un giro,

$$p = \mathbb{P}(\text{trozo bien cocinado}) = 1 - (0,7)^4.$$

(b). Sea B el número de trozos bien cocinados. Bajo el mismo supuesto de independencia entre revoluciones, cada uno de los 20 trozos queda bien cocinado con la probabilidad $p = 1 - (0,7)^4$ obtenida en el inciso anterior. Además, los giros de trozos diferentes son independientes. Por tanto,

$$B \sim \text{Binomial}(20, 1 - (0,7)^4).$$

En particular, su función de masa es

$$p_B(k) = \binom{20}{k} (1 - (0,7)^4)^k ((0,7)^4)^{20-k}, \quad k = 0, 1, \dots, 20.$$

(c). Sea R el número de veces que Don Casimiro revuelve la sartén, de modo que $\mathbb{P}(R = 1) = 1/3$ y $\mathbb{P}(R = 2) = 2/3$. Sea V el número de veces que el tocino se da vuelta y sea

$$Y = \mathbb{1}_{\{\text{el tocino queda bien cocinado}\}}.$$

Suponiendo nuevamente que los resultados de revoluciones distintas son independientes, condicionado a $R = r$ cada una de las r ocasiones produce un giro con probabilidad 0,3, de modo que

$$V \mid (R = r) \sim \text{Binomial}(r, 0,3), \quad Y = \mathbb{1}_{\{V \geq 1\}}.$$

Como uno de los lados está inicialmente en contacto con la sartén, el tocino no queda bien cocinado exactamente cuando nunca se da vuelta. Por ello,

$$\mathbb{P}(Y = 0 \mid R = 1) = 0,7, \quad \mathbb{P}(Y = 0 \mid R = 2) = (0,7)^2.$$

La probabilidad solicitada se obtiene mediante el Teorema de Bayes. Usando la ley de probabili-

dades totales en el denominador,

$$\begin{aligned}\mathbb{P}(R = 1 \mid Y = 0) &= \frac{\mathbb{P}(Y = 0 \mid R = 1)\mathbb{P}(R = 1)}{\mathbb{P}(Y = 0 \mid R = 1)\mathbb{P}(R = 1) + \mathbb{P}(Y = 0 \mid R = 2)\mathbb{P}(R = 2)} \\ &= \frac{0,7 \cdot \frac{1}{3}}{0,7 \cdot \frac{1}{3} + (0,7)^2 \cdot \frac{2}{3}} = \frac{5}{12}.\end{aligned}$$

Pregunta 63. (Control 2 - Primavera 2024, P1)

Una empresa proveedora de frutas envía un contenedor con n de manzanas a un supermercado. El supermercado se encarga de empacar (al azar) todas las frutas del contenedor en cajas de m manzanas cada una (asuma por simplicidad que n es divisible por m). La probabilidad de que una manzana esté mala es q , independiente del resto de las manzanas. Diremos que una caja está *buen*a si es que todas sus frutas están buenas; en caso contrario diremos que la caja está *mal*a.

- (a) Calcule la probabilidad de que una caja esté *buen*a.
- (b) Determine la distribución de probabilidad del número de cajas buenas. Si la distribución es conocida (entre las vistas en clase), indique su nombre y los valores de sus parámetros. De lo contrario, determine su función de masa de probabilidad y su rango.

Considere para las partes (c) y (d) que en el contenedor hay exactamente c manzanas malas, tales que $1 < c < n$.

- (c) Encuentre la probabilidad de que la primera caja que se empaca esté *mal*a.
- (d) Calcule el número esperado de cajas *buen*as. Hint: Le podría ser útil usar para cada caja i una VA que valga 1 si esta está buena, 0 si no.

Solución

(a). Cada manzana está buena con probabilidad $1 - q$. Como los estados de las m manzanas de una caja son independientes, la caja está buena con probabilidad

$$\mathbb{P}(\text{caja buena}) = (1 - q)^m.$$

(b). Hay n/m cajas. Sus indicadores de caja buena dependen de conjuntos disjuntos de manzanas y, por la independencia entre los estados de todas las frutas, son independientes. Por tanto, si Y denota el número de cajas buenas, entonces

$$Y \sim \text{Binomial}\left(\frac{n}{m}, (1 - q)^m\right).$$

(c). Sea X el número de manzanas malas en la primera caja. Como se seleccionan m manzanas sin reemplazo desde un total de n , de las cuales exactamente c están malas,

$$X \sim \text{Hipergeom}(n, c, m).$$

La caja es mala si $X > 0$. Conviene calcular el complemento: de las $\binom{n}{m}$ selecciones posibles para la primera caja, $\binom{n-c}{m}$ contienen solo manzanas buenas. En consecuencia,

$$\mathbb{P}(X > 0) = 1 - \mathbb{P}(X = 0) = 1 - \frac{\binom{n-c}{m}}{\binom{n}{m}}.$$

(d). Para cada caja i , defina $Z_i = \mathbf{1}_{\{\text{la caja } i \text{ está buena}\}}$. Por simetría, cada caja contiene una

muestra uniforme de m manzanas del contenedor, de modo que

$$\mathbb{E}[Z_i] = \mathbb{P}(Z_i = 1) = \frac{\binom{n-c}{m}}{\binom{n}{m}}.$$

Si $Y = \sum_{i=1}^{n/m} Z_i$ es el número de cajas buenas, la linealidad de la esperanza da

$$\mathbb{E}[Y] = \sum_{i=1}^{n/m} \mathbb{E}[Z_i] = \frac{n}{m} \frac{\binom{n-c}{m}}{\binom{n}{m}}.$$

Esta igualdad no requiere que los indicadores Z_i sean independientes.

Pregunta 64. (Control 2 - Otoño 2025, P1)

Se sabe que cada disquete producido por cierta compañía tiene una probabilidad de ser defectuoso de 0,01, independientemente del resto. La compañía vende los disquetes en paquetes de 10 unidades. Un paquete se denomina como defectuoso si posee **dos o más disquetes** defectuosos. La compañía ofrece la devolución del dinero en caso de que el paquete esté defectuoso. **Si en su desarrollo aparece una variable aleatoria que sigue alguna de las distribuciones vistas en clases, especifique su nombre y parámetros.**

- ¿Cuál es la probabilidad de que un paquete esté defectuoso?
- Un cliente compra tres paquetes. ¿Cuál es la probabilidad de que devuelva exactamente un paquete?
- Si un cliente compra hasta que un paquete resulte defectuoso, ¿cuál es el valor esperado del número de paquetes que compra?

Solución

(a). Sea X el número de disquetes defectuosos en un paquete. La independencia entre los diez disquetes y la probabilidad común de defecto implican que

$$X \sim \text{Bin}(10, 0,01).$$

Un paquete es defectuoso cuando $X \geq 2$. Es más directo calcular la probabilidad del complemento:

$$\begin{aligned} q &:= \mathbb{P}(X \geq 2) \\ &= 1 - \mathbb{P}(X = 0) - \mathbb{P}(X = 1) \\ &= 1 - (0,99)^{10} - 10(0,01)(0,99)^9 \\ &\approx 0,0042662. \end{aligned}$$

Equivalentemente, $q = \sum_{k=2}^{10} \binom{10}{k} (0,01)^k (0,99)^{10-k}$.

(b). Los tres paquetes contienen conjuntos disjuntos de disquetes, por lo que son independientes. Si Y es el número de paquetes defectuosos entre los tres, entonces

$$Y \sim \text{Bin}(3, q).$$

Devolver exactamente un paquete corresponde a $Y = 1$, de modo que

$$\mathbb{P}(Y = 1) = \binom{3}{1} q(1-q)^2 = 3q(1-q)^2 \approx 0,01269.$$

(c). Sea Z el número de paquetes comprados hasta encontrar el primero defectuoso, incluyendo ese último paquete. Cada paquete constituye un nuevo ensayo independiente con probabilidad de éxito q , por lo que

$$Z \sim \text{Geom}(q).$$

Con la convención de contar el ensayo en que aparece el primer éxito,

$$\mathbb{E}[Z] = \frac{1}{q} \approx 234,40.$$

Pregunta 65. (Control 2 - Primavera 2025, P1)

En la noche de Halloween, un niño sale solo a pedir dulces, que posteriormente compartirá con sus dos hermanos menores. En el vecindario hay un total de n casas que el niño recorrerá. Cada casa entrega un dulce (BomBomBún) con probabilidad p , o no entrega dulce con probabilidad $1 - p$. Para todas las partes, considere que la entrega del dulce entre casas es independiente.

- (a) En esta parte se asume que $n = 10$, es decir, hay 10 casas donde el niño pide dulces.
- ¿Cuál es la probabilidad de que el niño consiga **al menos un dulce para él y uno para cada uno de sus dos hermanos**?
 - Calcule el valor esperado de la cantidad de dulces que obtendrá el niño al visitar las casas. Justifique su respuesta.
- (b) Asuma que $n = 10$. Sin embargo, sabemos que a cada vecino le puede gustar más o le puede gustar menos celebrar Halloween, por lo que no todos van a entregar BonBonBum con la misma probabilidad. Suponga para esta parte que la probabilidad de que el niño reciba un dulce no es la misma en cada casa. Sea p_j la probabilidad de que la j -ésima casa entregue un dulce al niño ($j = 1, 2, \dots, 10$).

Calcule el valor esperado de la cantidad total de dulces que obtiene el niño al visitar las 10 casas. Hint: Le podría ser de utilidad definir otra(s) variable(s) aleatoria(s) como se hizo en clases para computar, de manera alternativa, la esperanza de una distribución Binomial.

- (c) Suponga ahora que el vecindario es muy grande, con una cantidad *infinita de casas*, y que cada casa entrega un BonBonBum con probabilidad p , y con probabilidad complementaria no entrega nada.

El niño continuará recorriendo casas hasta que logre obtener un total de tres dulces (uno para él y uno para cada uno de sus dos hermanos). Una vez que obtenga su tercer dulce, el niño no recorrerá más casas pidiendo dulces.

Calcule el número esperado de casas que el niño visitará.

Solución

(a). Sea X el número de dulces que obtiene el niño al visitar las diez casas. Como cada casa entrega un dulce con probabilidad p y las entregas son independientes,

$$X \sim \text{Binomial}(10, p).$$

(a)(i). Para disponer de al menos un dulce para cada uno de los tres hermanos, incluido el niño, se necesitan al menos tres dulces. Conviene usar el complemento:

$$\begin{aligned} \mathbb{P}(X \geq 3) &= 1 - \mathbb{P}(X \leq 2) \\ &= 1 - \sum_{k=0}^2 \binom{10}{k} p^k (1-p)^{10-k}. \end{aligned}$$

(a)(ii). La esperanza de una variable binomial con parámetros 10 y p es

$$\mathbb{E}[X] = 10p.$$

También puede obtenerse sumando las esperanzas de los diez indicadores de entrega.

(b). Para cada casa j , defina

$$I_j = \mathbb{1}_{\{\text{la casa } j \text{ entrega un dulce}\}}.$$

La cantidad total de dulces es

$$X = \sum_{j=1}^{10} I_j.$$

Como $\mathbb{E}[I_j] = \mathbb{P}(I_j = 1) = p_j$, la linealidad de la esperanza entrega

$$\begin{aligned} \mathbb{E}[X] &= \mathbb{E} \left[\sum_{j=1}^{10} I_j \right] \\ &= \sum_{j=1}^{10} \mathbb{E}[I_j] = \sum_{j=1}^{10} p_j. \end{aligned}$$

La linealidad no requiere que los indicadores tengan la misma probabilidad de éxito.

(c). Sea T el número de casas visitadas hasta obtener el tercer dulce. Cada casa corresponde a un ensayo Bernoulli independiente con probabilidad de éxito p . Para $p > 0$,

$$T \sim \text{BinomialNegativa}(3, p),$$

bajo la convención que cuenta el número total de ensayos necesarios para conseguir tres éxitos. Su esperanza es

$$\mathbb{E}[T] = \frac{3}{p}.$$

Si $p = 0$, el niño nunca obtiene los tres dulces y el tiempo de espera es infinito.

Pregunta 66. (Control 2 - Otoño 2026, P1)

- Suponga que usted quiere comenzar a ir al gimnasio y debe tomar una decisión: inscribirse en el gimnasio A o en el gimnasio B . Sabe que en ambos gimnasios se imparten clases de 17 categorías diferentes (zumba, crossfit, spinning, etc.), y dentro de estas 17 categorías, hay 9 en particular que son de su gusto.

Cada gimnasio elige las categorías de manera diferente. El gimnasio A escoge cada día una categoría al azar que se realizará ese día, independientemente de los demás días (cada categoría con la misma probabilidad), permitiendo que se haga más de una clase de la misma categoría en una misma semana. El gimnasio B elige, al principio de cada semana, las categorías para las 7 clases de esa semana al azar, de manera tal que no se hagan dos clases de la misma categoría (todas las secuencias de 7 categorías distintas tienen la misma probabilidad).

Con esto en mente, usted decide aplicar sus conocimientos de probabilidades para tomar una decisión. Durante el desarrollo de esta pregunta, si utiliza alguna fórmula de una distribución conocida, especifique el **nombre de la distribución** y sus **parámetros**.

Calcule **para cada gimnasio**:

- La probabilidad de que **no le guste** ninguna de las clases impartidas en una semana.
- El número esperado de clases impartidas en una semana que **no le gustan**.

2. Considere la variable aleatoria continua X , cuya función de densidad de probabilidad es

$$f_X(x) = \begin{cases} 4x^3, & x \in (0, 1], \\ 0, & \text{en otro caso.} \end{cases}$$

Sea Y la variable aleatoria definida como $Y = 1/X$. ¿Cuál es la función de densidad de probabilidad de Y ?

Solución

(1)(a). En el gimnasio A , las siete elecciones son independientes y cada clase pertenece a una categoría del gusto del estudiante con probabilidad $9/17$. Si X_A es el número de clases que le gustan,

$$X_A \sim \text{Binomial}\left(7, \frac{9}{17}\right).$$

Por tanto,

$$\mathbb{P}(X_A = 0) = \left(\frac{8}{17}\right)^7.$$

En el gimnasio B , las siete categorías se escogen sin reemplazo entre las 17 disponibles. Si X_B es el número de categorías escogidas que le gustan,

$$X_B \sim \text{Hipergeom}(17, 9, 7).$$

Para que no le guste ninguna clase, las siete categorías deben escogerse entre las ocho que no le gustan. Así,

$$\mathbb{P}(X_B = 0) = \frac{\binom{8}{7}}{\binom{17}{7}}.$$

(1)(b). Sea $Y_A = 7 - X_A$ el número de clases que no le gustan en el gimnasio A . Usando la esperanza de la distribución binomial,

$$\begin{aligned} \mathbb{E}[Y_A] &= 7 - \mathbb{E}[X_A] \\ &= 7 - 7 \frac{9}{17} = \frac{56}{17}. \end{aligned}$$

Análogamente, sea $Y_B = 7 - X_B$ el número de clases que no le gustan en el gimnasio B . Como la esperanza de una variable hipergeométrica con parámetros (N, K, n) es nK/N ,

$$\begin{aligned} \mathbb{E}[Y_B] &= 7 - \mathbb{E}[X_B] \\ &= 7 - \frac{7 \cdot 9}{17} = \frac{56}{17}. \end{aligned}$$

Aunque las selecciones se realizan de manera distinta, ambos gimnasios tienen el mismo número esperado de clases que no le gustan; la diferencia entre ellos aparece en la dependencia y en la variabilidad semanal.

(2). Como $X \in (0, 1]$, la transformación $Y = 1/X$ toma valores en

$$R_Y = [1, \infty).$$

La función $g(x) = 1/x$ es estrictamente decreciente y derivable en $(0, 1]$. Su inversa y la magnitud de su derivada son

$$g^{-1}(y) = \frac{1}{y}, \quad \left| \frac{d}{dy} g^{-1}(y) \right| = \frac{1}{y^2}.$$

Aplicando la fórmula de cambio de variable, para $y \geq 1$ se obtiene

$$\begin{aligned} f_Y(y) &= f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| \\ &= 4 \left(\frac{1}{y} \right)^3 \frac{1}{y^2} = \frac{4}{y^5}. \end{aligned}$$

Por tanto,

$$f_Y(y) = \begin{cases} \frac{4}{y^5}, & y \in [1, \infty), \\ 0, & \text{en otro caso.} \end{cases}$$

Pregunta 67. (Control 2 - Otoño 2026, P2)

Ha salido la nueva expansión de cartas coleccionables de Pokémon, y usted quiere coleccionarlas todas. Las cartas se venden en sobres que contienen una sola carta cada uno. Suponga que la expansión contiene un total de n cartas diferentes y el contenido de cada sobre es uniformemente aleatorio e independiente del resto; es decir, en cada sobre cada carta viene con probabilidad $1/n$, independiente de los demás sobres. Suponga que usted ya posee i cartas de la colección (con $i < n$).

- Se define una variable aleatoria como el número de sobres que compra hasta obtener una carta nueva (diferente a las que ya posee). Describa la función de masa de probabilidad de esta variable aleatoria. Si corresponde a una distribución vista en clases, especifique su nombre y parámetros.
- Un amigo suyo que también colecciona cartas Pokémon le dice: “si definimos una variable aleatoria como el número de sobres que compro hasta obtener 3 cartas nuevas (diferentes a las que ya tengo y diferentes entre sí), entonces esta variable corresponde a una binomial negativa”. ¿Es verdadera o falsa la observación de su amigo? Argumente su respuesta.
- Suponga para esta parte que no posee ninguna lámina del álbum, es decir, $i = 0$. Calcule el número esperado de sobres que se deben abrir hasta completar la colección. *Indicación: Le puede ser útil expresar la variable aleatoria de interés en función de variables aleatorias como las de la parte (a).*

Solución

(a). Denote por X_i el número de sobres que se abren hasta obtener una carta nueva cuando ya se poseen i cartas distintas. Mientras no aparece una carta nueva, la colección no cambia y cada sobre tiene probabilidad constante

$$p_i = \frac{n-i}{n}$$

de contener una carta nueva. Como los sobres son independientes, X_i tiene distribución geométrica con parámetro p_i , contando el sobre en que aparece el éxito:

$$X_i \sim \text{Geom}\left(\frac{n-i}{n}\right), \quad p_{X_i}(k) = \left(\frac{i}{n}\right)^{k-1} \frac{n-i}{n}, \quad k = 1, 2, \dots$$

(b). La afirmación es falsa. Una distribución binomial negativa supone ensayos Bernoulli con una probabilidad de éxito constante. Aquí, la probabilidad de obtener una carta que aún no se posee comienza en

$$\frac{n-i}{n},$$

pero después de obtener la primera y la segunda cartas nuevas cambia, respectivamente, a

$$\frac{n-i-1}{n} \quad \text{y} \quad \frac{n-i-2}{n}.$$

Por ello, si $i+3 \leq n$, el número de sobres necesario es una suma de tres variables geométricas con parámetros distintos, no una binomial negativa. Si quedan menos de tres cartas por obtener, el evento descrito ni siquiera es posible.

(c). La variable X_i del inciso (a) tiene esperanza

$$\mathbb{E}[X_i] = \frac{1}{p_i} = \frac{n}{n-i}.$$

Partiendo sin cartas, el tiempo Y para completar la colección es la suma de los tiempos empleados en pasar de i a $i+1$ cartas distintas:

$$Y = \sum_{i=0}^{n-1} X_i.$$

La linealidad de la esperanza, que no requiere independencia entre estos tiempos, entrega

$$\mathbb{E}[Y] = \sum_{i=0}^{n-1} \frac{n}{n-i} = n \sum_{j=1}^n \frac{1}{j} = nH_n,$$

donde H_n es el n -ésimo número armónico.

Pregunta 68. (Examen alternativo - Primavera 2021, P2)

Una empresa esta evaluando la tecnología a contratar para un sistema de producción de ampolletas. La tecnología a escoger es un valor real no negativo tal que la calidad de la probabilidad de falla de una ampolleta dependerá de dicha calidad. En particular, se tiene que para un nivel de tecnología t , la probabilidad de producir una ampolleta defectuosa es de $\frac{1}{t+10}$. Las ampolletas se producen en cajas de a 25 unidades. Si es que una caja posee 4 o más ampolletas defectuosas, la empresa debe pagar una multa de \$25000; en caso contrario (i.e. en donde las unidades defectuosas por caja es menor a 4), no se paga multa alguna. Por otra parte, se tiene que el nivel de tecnología tiene un costo de producción por una caja de ampolletas de $\$ \frac{t^2}{4}$.

- Determine el valor de tecnología óptima (si no hay fórmula cerrada para obtener este valor, podría dejar planteado el problema de optimización). No use aproximaciones de las expresiones que involucren probabilidades.
- Resuelva la parte (a) pero considerando ahora la desigualdad de Chebyshev para reemplazar términos que sean una probabilidad.
- Asuma que el nivel de tecnología es $t = 0$. Suponga que la empresa, luego de producir una caja de ampolletas, revisa esta para detectar si es que tiene 4 o más ampolletas con fallas. En caso de que la empresa detecte que hay una caja en tales condiciones (con 4 o más ampolletas defectuosas), se separa la caja del resto, la cual se venderá a un precio menor). Suponga que deseamos modelar el número de cajas a producir para obtener 10 cajas de ampolletas buenas (una caja de ampolletas es considerada *buenas* si tiene menos de 4 ampolletas defectuosas). ¿Qué distribución sigue la variable aleatoria que describe el fenómeno descrito? Justifique su respuesta.

Solución

(a). Para cada $t \geq 0$, sea X_t el número de ampolletas defectuosas en una caja. Suponiendo independencia entre las unidades,

$$X_t \sim \text{Binomial}\left(25, \frac{1}{t+10}\right).$$

El costo esperado por caja es el costo de la tecnología más la multa ponderada por la probabilidad

de falla:

$$\begin{aligned} C(t) &= \frac{t^2}{4} + 25\,000 \mathbb{P}(X_t \geq 4) \\ &= \frac{t^2}{4} + 25\,000 [1 - \mathbb{P}(X_t \leq 3)] \\ &= \frac{t^2}{4} + 25\,000 \left[1 - \sum_{k=0}^3 \binom{25}{k} \left(\frac{1}{t+10} \right)^k \left(1 - \frac{1}{t+10} \right)^{25-k} \right]. \end{aligned}$$

Por tanto, el nivel óptimo queda determinado exactamente por

$$t^* \in \operatorname{argmin}_{t \geq 0} C(t).$$

(b). La media y la varianza de X_t son

$$\mu_t = \frac{25}{t+10}, \quad \operatorname{Var}(X_t) = 25 \frac{1}{t+10} \left(1 - \frac{1}{t+10} \right) = \frac{25(t+9)}{(t+10)^2}.$$

Como $t \geq 0$, se tiene $4 - \mu_t > 0$. Además,

$$\{X_t \geq 4\} \subseteq \{|X_t - \mu_t| \geq 4 - \mu_t\}.$$

La desigualdad de Chebyshev permite entonces acotar

$$\begin{aligned} \mathbb{P}(X_t \geq 4) &\leq \mathbb{P}(|X_t - \mu_t| \geq 4 - \mu_t) \\ &\leq \frac{\operatorname{Var}(X_t)}{(4 - \mu_t)^2} = \frac{25(t+9)}{(4t+15)^2}. \end{aligned}$$

Al reemplazar la probabilidad por esta cota superior, se obtiene el problema conservador

$$\operatorname{minimize}_{t \geq 0} \left\{ \frac{t^2}{4} + 25\,000 \frac{25(t+9)}{(4t+15)^2} \right\}.$$

La derivada del objetivo es

$$\frac{t}{2} - 625\,000 \frac{4t+57}{(4t+15)^3}.$$

El primer término es creciente y el segundo término, en valor absoluto, es decreciente para $t \geq 0$, por lo que existe un único punto crítico. Este satisface

$$t(4t+15)^3 = 1\,250\,000(4t+57),$$

y su minimizador es, numéricamente, $t_{\text{Ch}}^* \approx 43,256$. Este valor minimiza la cota de Chebyshev del costo esperado, no necesariamente el costo exacto de la parte (a).

(c). Cuando $t = 0$, el número X de ampolletas defectuosas en una caja distribuye Binomial(25, 0,1). La probabilidad de que una caja sea buena es

$$q = \mathbb{P}(X \leq 3) = \sum_{k=0}^3 \binom{25}{k} (0,1)^k (0,9)^{25-k}.$$

Si las cajas se producen de manera independiente y N es el número total de cajas que deben producirse para obtener la décima caja buena, entonces N sigue una distribución binomial negativa con 10 éxitos y probabilidad de éxito q :

$$N \sim \operatorname{NegBin}(10, q), \quad \mathbb{P}(N = n) = \binom{n-1}{9} q^{10} (1-q)^{n-10}, \quad n = 10, 11, \dots$$

Pregunta 69. (Examen - Primavera 2025, P2)

1. Aníbal y Belén se encuentran en un parque y modelan el número de asistentes al parque en un día como una distribución de Poisson con parámetro igual a la temperatura en Celsius. Además, la temperatura de cada día toma un valor **entero** equiprobable entre 11 y 25 grados. Actualmente hay 12 personas en el parque ¿Cuál es la probabilidad de que la temperatura sea mayor a 20 grados?
2. Se entregarán cinco premios distintos (mejor desempeño académico, mejor liderazgo, etc.) a estudiantes seleccionados de una clase de 30. ¿Cuántos resultados diferentes son posibles si:
 - (a) un estudiante puede recibir cualquier cantidad de premios?
 - (b) cada estudiante puede recibir a lo más un premio?

Solución

(1). La temperatura X es uniforme sobre los enteros de 11 a 25:

$$p_X(x) = \begin{cases} 1/15, & x \in \{11, 12, \dots, 25\}, \\ 0, & \text{en otro caso.} \end{cases}$$

Sea N el número de personas en el parque. Condicionado a $X = x$,

$$N \mid X = x \sim \text{Poisson}(x).$$

Por el Teorema de Bayes, para $x \in \{11, \dots, 25\}$,

$$\begin{aligned} p_{X|N}(x \mid 12) &= \frac{p_{N|X}(12 \mid x)p_X(x)}{\sum_{y=11}^{25} p_{N|X}(12 \mid y)p_X(y)} \\ &= \frac{x^{12}e^{-x}}{\sum_{y=11}^{25} y^{12}e^{-y}}, \end{aligned}$$

porque los factores $12!$ y 15 se cancelan. En consecuencia,

$$\mathbb{P}(X > 20 \mid N = 12) = \frac{\sum_{x=21}^{25} x^{12}e^{-x}}{\sum_{y=11}^{25} y^{12}e^{-y}}.$$

(2)(a). Cada uno de los cinco premios distintos puede entregarse a cualquiera de los 30 estudiantes. Como un estudiante puede recibir varios premios, las cinco elecciones son independientes en el conteo. El número de resultados es

$$30^5.$$

(2)(b). Ahora los ganadores deben ser distintos. Hay 30 opciones para el primer premio, 29 para el segundo y así sucesivamente hasta 26 para el quinto. El número de resultados es

$$30 \cdot 29 \cdot 28 \cdot 27 \cdot 26 = \frac{30!}{25!}.$$

Pregunta 70. (Examen v2 - Primavera 2025, P2)

1. Se lanza un dado dos veces y se define la variable aleatoria como el valor del primer dado menos el del segundo. Indique el rango y la función de probabilidad de masa.
2. Los puntajes de bowling de Ximena se distribuyen aproximadamente de manera normal con media de 170 y desviación estándar de 20, mientras que los puntajes de Jaime se distribuyen aproximadamente de manera normal con media de 160 y desviación estándar

de 15. Si Jaime y Ximena juegan una partida cada uno, y suponiendo que sus puntajes son variables aleatorias independientes, aproxime la probabilidad de que el puntaje de Jaime sea mayor.

3. Una apuesta consiste en que un jugador lanza una moneda balanceada hasta que aparezca cara. El jugador recibe como ganancia 1,5 elevado al número total de lanzamientos realizados. Determine el valor esperado de la ganancia. Si no es posible obtener un valor cerrado, entregue el desarrollo más avanzado posible. Debe ser claro en su desarrollo y definir toda VA que utilice.

Solución

(1). Sean D_1 y D_2 los resultados del primer y segundo lanzamiento, respectivamente, y defina $X = D_1 - D_2$. Su rango es

$$R_X = \{-5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5\}.$$

Para cada $x \in R_X$ existen $6 - |x|$ pares ordenados (D_1, D_2) cuya diferencia es x , entre los 36 resultados equiprobables. Por tanto,

$$p_X(x) = \mathbb{P}(X = x) = \begin{cases} \frac{6 - |x|}{36}, & x \in \{-5, -4, \dots, 5\}, \\ 0, & \text{en otro caso.} \end{cases}$$

(2). Modelamos los puntajes de Ximena y Jaime mediante

$$X \sim \mathcal{N}(170, 20^2), \quad J \sim \mathcal{N}(160, 15^2),$$

que corresponde a la aproximación indicada en el enunciado. Bajo este modelo y por independencia, la diferencia $D = J - X$ satisface

$$D \sim \mathcal{N}(160 - 170, 15^2 + 20^2) = \mathcal{N}(-10, 25^2).$$

En consecuencia,

$$\begin{aligned} \mathbb{P}(J > X) &= \mathbb{P}(D > 0) \\ &= \mathbb{P}\left(Z > \frac{0 - (-10)}{25}\right) \\ &= 1 - \Phi(0,4) \approx 0,3446, \end{aligned}$$

donde $Z \sim \mathcal{N}(0, 1)$.

(3). Sea N el número de lanzamientos hasta obtener la primera cara. Como la moneda es balanceada,

$$N \sim \text{Geom}\left(\frac{1}{2}\right), \quad p_N(n) = \mathbb{P}(N = n) = \left(\frac{1}{2}\right)^n, \quad n = 1, 2, \dots$$

La ganancia es $G = (3/2)^N$. Por definición de esperanza,

$$\begin{aligned} \mathbb{E}[G] &= \sum_{n=1}^{\infty} \left(\frac{3}{2}\right)^n p_N(n) \\ &= \sum_{n=1}^{\infty} \left(\frac{3}{4}\right)^n = \frac{3/4}{1 - 3/4} = 3. \end{aligned}$$

Pregunta 71. (Examen - Otoño 2026, P2)

Un grupo de estudiantes de ingeniería organiza una venta de paquetes de galletas artesanales durante tres semanas consecutivas para financiar sus vacaciones de verano.

- (a) En cada semana, las ventas pueden ser de dos tipos: en una semana regular venden 5 paquetes, mientras que en una semana de alta demanda venden 12 paquetes. La

probabilidad de una semana regular es 70 % y la de una semana de alta demanda es 30 %. Suponga que las ventas de distintas semanas son independientes. El total de paquetes vendidos al final de la campaña se puede escribir como

$$5 \cdot 3 + (12 - 5)Y,$$

donde Y es una variable aleatoria cuya distribución fue vista en clases.

- (i) Indique el nombre y los parámetros de la distribución de Y . Justifique.
- (ii) Calcule el valor esperado de las ventas totales al finalizar las tres semanas.

En las partes siguientes, considere la posibilidad de robos. Si ocurre un robo, los ladrones se llevan todo el dinero acumulado hasta esa semana inclusive. Inicialmente, los estudiantes no tienen protección. Si sufren un primer robo, contratan de inmediato un seguro para las semanas posteriores. El seguro cuesta 1 unidad monetaria por cada semana restante y repone el dinero robado si ocurren nuevos robos.

- (b) Suponga que la utilidad por ventas es constante e igual a 8 unidades monetarias por semana. Además, cada semana existe una probabilidad de 4 % de sufrir un robo, independientemente entre semanas. Se desea determinar la PMF de la utilidad total al finalizar las tres semanas.
 - (i) Determine la utilidad total si no ocurre ningún robo y la probabilidad de este escenario.
 - (ii) Determine la utilidad total si el primer robo ocurre exactamente en la semana t , donde $t \in \{1, 2, 3\}$, y la probabilidad de este escenario.
 - (iii) Determine la PMF de la utilidad total al finalizar las tres semanas.
- (c) Suponga ahora que los ingresos por ventas de cada semana distribuyen normal con media 10 y varianza 3^2 , independientemente entre semanas.

El riesgo de robo es distinto al de la parte anterior: el primer robo puede ocurrir únicamente al finalizar la semana 1 o al finalizar la semana 2, con probabilidades 50 % y 40 %, respectivamente. Además, existe un 10 % de probabilidad de no sufrir robos durante todo el período. Las condiciones del seguro son las mismas.

Al finalizar la campaña, los estudiantes observan que la utilidad total, luego de considerar ventas, posibles robos y costos del seguro, es exactamente 25 unidades monetarias.

- (i) Determine la distribución de la utilidad total en cada uno de los tres escenarios: robo al finalizar la semana 1, robo al finalizar la semana 2 o ningún robo. Si es una distribución conocida, indique su nombre y parámetros.
- (ii) Determine la probabilidad de que no haya ocurrido ningún robo durante todo el período. Puede entregar una fórmula analítica.

Solución

(a)(i). La variable Y cuenta las semanas de alta demanda entre tres ensayos independientes, cada uno con probabilidad 0,3. Por tanto,

$$Y \sim \text{Binomial}(3, 0,3).$$

(a)(ii). Si $T = 15 + 7Y$ es la venta total, entonces

$$\mathbb{E}[T] = 15 + 7\mathbb{E}[Y] = 15 + 7(3 \cdot 0,3) = 21,3.$$

(b)(i). Sin robos, la utilidad es 24 y, por independencia, este escenario tiene probabilidad $0,96^3$.

(b)(ii). La probabilidad de que el primer robo ocurra en $t \in \{1, 2, 3\}$ es

$$0,96^{t-1}(0,04).$$

El robo elimina todo lo acumulado hasta t . En las $3 - t$ semanas restantes, el ingreso neto asegurado es $8 - 1 = 7$ por semana, de modo que la utilidad final es $7(3 - t)$.

(b)(iii). Si U es la utilidad total,

$$p_U(u) = \begin{cases} 0,96^2(0,04), & u = 0, \\ 0,96(0,04), & u = 7, \\ 0,04, & u = 14, \\ 0,96^3, & u = 24, \\ 0, & \text{en otro caso.} \end{cases}$$

(c)(i). Sean $X_i \sim \mathcal{N}(10, 9)$ los ingresos semanales independientes y sea R el escenario del primer robo. Entonces

$$\begin{aligned} R = 1 : & \quad U = X_2 + X_3 - 2, \quad U \mid R = 1 \sim \mathcal{N}(18, 18), \\ R = 2 : & \quad U = X_3 - 1, \quad U \mid R = 2 \sim \mathcal{N}(9, 9), \\ R = \infty : & \quad U = X_1 + X_2 + X_3, \quad U \mid R = \infty \sim \mathcal{N}(30, 27). \end{aligned}$$

(c)(ii). Como U es continua, el antecedente $U = 25$ se trata mediante densidades. El Teorema de Bayes en su forma mixta da

$$\mathbb{P}(R = \infty \mid U = 25) = \frac{0,1 f_{U \mid R}(25 \mid \infty)}{0,5 f_{U \mid R}(25 \mid 1) + 0,4 f_{U \mid R}(25 \mid 2) + 0,1 f_{U \mid R}(25 \mid \infty)}.$$

Al evaluar las tres densidades normales,

$$\mathbb{P}(R = \infty \mid U = 25) = \frac{\frac{0,1}{\sqrt{54\pi}} e^{-25/54}}{\frac{0,5}{\sqrt{36\pi}} e^{-49/36} + \frac{0,4}{\sqrt{18\pi}} e^{-128/9} + \frac{0,1}{\sqrt{54\pi}} e^{-25/54}} \approx 0,286.$$

Capítulo 6

Variables Aleatorias Continuas

6.1. Función de distribución y densidad

Resumen

Función de densidad.

Si X admite una densidad f_X , las probabilidades se obtienen integrando sobre conjuntos. La densidad satisface:

- **No negatividad.** $f_X(x) \geq 0$ para todo $x \in \mathbb{R}$.
- **Normalización.** $\int_{\mathbb{R}} f_X(x) dx = 1$.
- **Probabilidad de un conjunto.** $\mathbb{P}(X \in A) = \int_A f_X(x) dx$.
- **Probabilidad puntual.** $\mathbb{P}(X = x) = 0$; por ello, incluir o excluir los extremos de un intervalo no cambia su probabilidad.

Función de distribución.

La CDF $F_X(x) = \mathbb{P}(X \leq x)$ satisface

- **Monotonía.** Si $x \leq y$, entonces $F_X(x) \leq F_X(y)$.
- **Límites.** $F_X(x) \rightarrow 0$ cuando $x \rightarrow -\infty$ y $F_X(x) \rightarrow 1$ cuando $x \rightarrow \infty$.
- **Continuidad por la derecha.** $\lim_{t \downarrow x} F_X(t) = F_X(x)$.
- **Continuidad.** Si X admite densidad, F_X es continua en $\mathbb{R}$.
- **Acumulación de densidad.** $F_X(x) = \int_{-\infty}^x f_X(t) dt$ y, donde F_X es derivable, $f_X(x) = F_X'(x)$.
- **Probabilidad de intervalos.** $\mathbb{P}(a < X \leq b) = F_X(b) - F_X(a) = \int_a^b f_X(x) dx$.

Esperanza y momentos.

Para una función g integrable,

$$\mathbb{E}[g(X)] = \int_{\mathbb{R}} g(x) f_X(x) dx,$$

$$\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2.$$

Si $X \geq 0$ y $\mathbb{E}[X] < \infty$, también se cumple la identidad de colas

$$\mathbb{E}[X] = \int_0^{\infty} \mathbb{P}(X > t) dt.$$

Cuantiles, mediana y moda.

- **Cuantil.** Para $\alpha \in (0, 1)$, $q_\alpha = \inf\{x : F_X(x) \geq \alpha\}$.
- **Mediana.** Es un cuantil de nivel $1/2$.
- **Moda.** Es un punto en que f_X alcanza su máximo; una transformación puede cambiarla y debe maximizarse la densidad transformada.

6.1.1. Preguntas

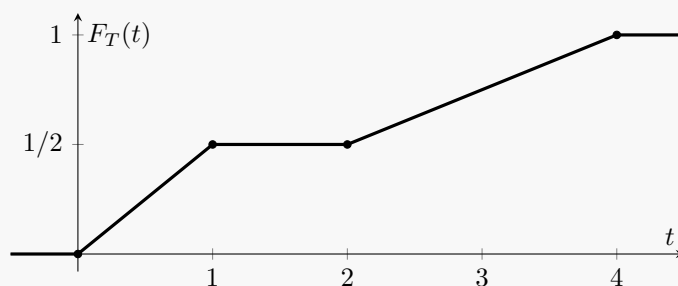
Pregunta 72.

Investigaciones del Metro han revelado que el tiempo de espera en minutos, hasta que pase un tren por una estación, es una v.a. T con función de distribución (acumulada) F_T definida por: $F_T(t) = 0$ para $t < 0$; $F_T(t) = t/2$ para $0 \leq t \leq 1$; $F_T(t) = 1/2$ para $1 < t < 2$; $F_T(t) = t/4$ para $2 \leq t \leq 4$; $F_T(t) = 1$ para $t > 4$.

- Grafique la función de distribución F_T , determine si es continua y calcule su densidad de ser posible.
- Calcule la probabilidad de que una persona tenga que esperar más de 3 minutos; menos de 3 minutos; entre 1 y 3 minutos.
- Calcule la probabilidad de que la persona espere: más de 3 minutos, sabiendo que ya lleva 1 minuto esperando; menos de 3 minutos, dado que lleva 1 minuto esperando.
- Calcule la esperanza y varianza de T .

Solución

(a). La función es lineal por tramos y no presenta saltos: los valores laterales coinciden en $t = 0, 1, 2, 4$. Su gráfico es



Como F_T es continua y derivable salvo en los puntos donde cambia su pendiente, una densidad de T es

$$f_T(t) = \begin{cases} \frac{1}{2}, & 0 < t < 1, \\ \frac{1}{4}, & 2 < t < 4, \\ 0, & \text{en otro caso.} \end{cases}$$

Los valores asignados a la densidad en los extremos de los intervalos no afectan las probabilidades.

(b). Como T es continua, no es necesario distinguir entre desigualdades estrictas y no estrictas. Evaluando la función de distribución,

$$\mathbb{P}(T > 3) = 1 - F_T(3) = 1 - \frac{3}{4} = \frac{1}{4},$$

$$\mathbb{P}(T < 3) = F_T(3) = \frac{3}{4},$$

$$\mathbb{P}(1 < T < 3) = F_T(3) - F_T(1) = \frac{3}{4} - \frac{1}{2} = \frac{1}{4}.$$

(c). Llevar un minuto esperando significa que ocurrió $T > 1$, cuya probabilidad es

$$\mathbb{P}(T > 1) = 1 - F_T(1) = \frac{1}{2}.$$

Como $\{T > 3\} \subseteq \{T > 1\}$, la primera probabilidad condicional es

$$\mathbb{P}(T > 3 \mid T > 1) = \frac{\mathbb{P}(T > 3)}{\mathbb{P}(T > 1)} = \frac{1/4}{1/2} = \frac{1}{2}.$$

Para esperar menos de tres minutos después de haber esperado ya uno, el evento relevante es $1 < T < 3$. Por tanto,

$$\mathbb{P}(T < 3 \mid T > 1) = \frac{\mathbb{P}(1 < T < 3)}{\mathbb{P}(T > 1)} = \frac{1/4}{1/2} = \frac{1}{2}.$$

(d). La densidad es positiva únicamente en $(0, 1)$ y $(2, 4)$. Por ello, el primer momento es

$$\begin{aligned}\mathbb{E}[T] &= \int_{\mathbb{R}} t f_T(t) dt \\ &= \frac{1}{2} \int_0^1 t dt + \frac{1}{4} \int_2^4 t dt \\ &= \frac{1}{4} + \frac{3}{2} = \frac{7}{4}.\end{aligned}$$

Para calcular la varianza necesitamos también el segundo momento:

$$\begin{aligned}\mathbb{E}[T^2] &= \int_{\mathbb{R}} t^2 f_T(t) dt \\ &= \frac{1}{2} \int_0^1 t^2 dt + \frac{1}{4} \int_2^4 t^2 dt \\ &= \frac{1}{6} + \frac{1}{4} \left(\frac{4^3 - 2^3}{3} \right) = \frac{29}{6}.\end{aligned}$$

Finalmente,

$$\begin{aligned}\text{Var}(T) &= \mathbb{E}[T^2] - \mathbb{E}[T]^2 \\ &= \frac{29}{6} - \left(\frac{7}{4} \right)^2 = \frac{85}{48}.\end{aligned}$$

6.1.2. Preguntas de Evaluación

Pregunta 73. (Control 2 - Otoño 2023, P1)

Un zapatero repara zapatos. El tiempo que demora en reparar un par de zapatos no es siempre el mismo. El ayudante del zapatero intenta comprender mejor cuánto tiempo le toma reparar un par de zapatos. Luego de hacer varias mediciones, el ayudante le informa que el tiempo que le toma al zapatero reparar un par de zapatos es una VA cuya PDF es una función lineal afín que toma un valor $b > 0$ en el tiempo 0, y luego decrece linealmente hasta valer cero para algún valor (desconocido por el ayudante).

- Indique el rango de la VA (en función de b). Escriba la función de densidad de probabilidad del tiempo en reparar un par de zapatos en función de b .
- Se sabe que la probabilidad de que el zapatero repare un par de zapatos en menos de 1 minuto es de $5/9$. Determine la PDF de la VA del tiempo que le toma al zapatero reparar un par de zapatos. Hint: En algún momento, le podría ser útil notar que el máximo valor que toma la VA debe ser mayor a 1.
- Suponga ahora que luego de reparar cada par de zapatos, el zapatero se toma un descanso tal que el *tiempo de reparación* más el posterior *tiempo de descanso* suman 10 minutos. Luego de dicho descanso, el zapatero inicia la reparación de otro par de zapatos, y

posteriormente se toma un descanso tal que el tiempo de reparación (del último par de zapatos) más el tiempo de descanso suman 10 minutos. Así sucesivamente. Compare la varianza del tiempo que le toma al zapatero reparar un par de zapatos, con la varianza del tiempo que se toma de descanso luego de reparar un par de zapatos. Es decir, determine qué varianza es mayor que la otra, o bien si son iguales. Su respuesta debe tener un fundamento matemático. Hint: Le podría ser útil dejar una VA en función de otra.

Solución

(a). Sea $a > 0$ el instante en que la recta corta el eje horizontal. El área bajo la densidad es la de un triángulo de base a y altura b ; por lo tanto, $ab/2 = 1$ y $a = 2/b$. La recta tiene intercepto b y pendiente $-b/a = -b^2/2$, de modo que

$$f_X(x) = \begin{cases} b - \frac{b^2}{2}x, & 0 \leq x \leq \frac{2}{b}, \\ 0, & \text{en otro caso.} \end{cases}$$

Así, el rango puede escribirse como $\mathcal{R}_X = [0, 2/b]$; incluir o excluir el extremo derecho no modifica la distribución.

(b). Como la distribución es continua, $\mathbb{P}(X < 1) = \mathbb{P}(X \leq 1)$. Además, el dato exige que 1 pertenezca al interior del soporte. Integrando la densidad obtenida en el inciso anterior,

$$\begin{aligned} \mathbb{P}(X \leq 1) &= \int_0^1 \left(b - \frac{b^2}{2}x \right) dx \\ &= b - \frac{b^2}{4} = \frac{5}{9}. \end{aligned}$$

Por consiguiente, b satisface

$$\frac{b^2}{4} - b + \frac{5}{9} = 0,$$

cuyas raíces son

$$b = \frac{1 \pm \sqrt{1 - 5/9}}{1/2} = \frac{1 \pm 2/3}{1/2} \in \left\{ \frac{2}{3}, \frac{10}{3} \right\}.$$

El extremo superior del soporte es $2/b$ y debe ser mayor que 1, por lo que $b < 2$. Esto descarta $b = 10/3$ y deja $b = 2/3$. Al sustituir,

$$f_X(x) = \begin{cases} \frac{2}{3} - \frac{2}{9}x, & 0 \leq x \leq 3, \\ 0, & \text{en otro caso.} \end{cases}$$

(c). Sea Y el tiempo de descanso y mantengamos X para el tiempo de reparación. Como ambos tiempos suman diez minutos, $Y = 10 - X$. La varianza no cambia al sumar una constante y se multiplica por el cuadrado del factor de escala; entonces

$$\text{Var}(Y) = \text{Var}(10 - X) = (-1)^2 \text{Var}(X) = \text{Var}(X).$$

Por lo tanto, los dos tiempos tienen la misma varianza.

Pregunta 74. (Control 2 - Otoño 2024, P2)

Todos los días Itadori viaja desde su casa a la universidad. Primero camina hacia la estación de metro más cercana a su casa, luego toma el metro para trasladarse hacia la estación Parque O'higgins, para posteriormente caminar hacia el paradero de Parque O'higgins, y finalmente, decide si tomar una micro o caminar por el parque para llegar a la universidad. El tiempo que demora Itadori en trasladarse desde su casa a la estación de metro más cercana a su casa es de 1 minuto. Una vez que llega a dicha estación de metro, el tiempo que demora en pasar el metro distribuye Uniforme(0, 1). Considere que el recorrido en el metro toma

un tiempo de 1 minuto. Una vez que Itadori está en estación Parque O'Higgins se demora un tiempo despreciable en llegar al paradero. La micro que va hacia la universidad con probabilidad $\frac{1}{2}$ está cuando Itadori llega al paradero, y el recorrido en dicha micro tarda 1 minuto en llegar a la universidad. Si la micro no está, Itadori decide caminar por el parque, lo cual le toma 5 minutos. Sea X la variable aleatoria que corresponde al tiempo total en minutos que le toma a Itadori trasladarse desde su casa a la universidad. Se tiene que

$$f_X(x) = \frac{1}{2} \mathbb{1}_{\{x \in [3,4] \cup [7,8]\}}. \quad (6.1)$$

Nota: note que para responder las preguntas no es necesario poner atención al texto inicial, pues la Ecuación (6.1) indica explícitamente la función de densidad de probabilidad del tiempo de traslado.

- (a) Entregue el rango de la variable aleatoria X .
- (b) Entregue la CDF de la variable aleatoria X de forma analítica y gráfica.
- (c) Calcule la esperanza de la variable aleatoria X .
- (d) Considere ahora que la variable aleatoria X tiene una nueva distribución. Esta es

$$f_X(x) = \frac{1}{2} \mathbb{1}_{\{x \in [4,5] \cup [6,7]\}}.$$

¿Cómo cambia la varianza con respecto a la distribución inicial? Puede responder esta pregunta dando argumentos matemáticos con los cálculos requeridos, o bien fundamentar su respuesta con palabras basado en los conceptos.

Solución

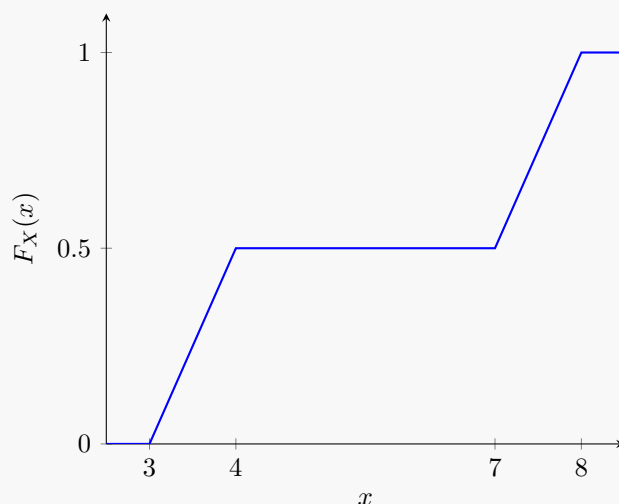
(a). La densidad es positiva únicamente en los intervalos $[3, 4]$ y $[7, 8]$. Por tanto, el rango o soporte de X es

$$R_X = [3, 4] \cup [7, 8].$$

(b). La función de distribución se obtiene acumulando la densidad desde $-\infty$ hasta x . En los intervalos donde la densidad es $1/2$, la CDF crece linealmente; entre 4 y 7 permanece constante. Así,

$$F_X(x) = \begin{cases} 0, & x < 3, \\ \frac{x-3}{2}, & 3 \leq x < 4, \\ \frac{1}{2}, & 4 \leq x < 7, \\ \frac{1}{2} + \frac{x-7}{2} = \frac{x-6}{2}, & 7 \leq x < 8, \\ 1, & x \geq 8. \end{cases}$$

La representación gráfica correspondiente es



(c). Usando la definición de esperanza para una variable continua,

$$\begin{aligned}\mathbb{E}[X] &= \int_3^4 \frac{x}{2} dx + \int_7^8 \frac{x}{2} dx \\ &= \frac{1}{4} (4^2 - 3^2) + \frac{1}{4} (8^2 - 7^2) = \frac{7}{4} + \frac{15}{4} = \frac{11}{2}.\end{aligned}$$

(d). Ambas distribuciones tienen esperanza $11/2$, pues los dos pares de intervalos son simétricos alrededor de ese valor. Para comparar sus varianzas, calculamos primero el segundo momento de la distribución inicial:

$$\begin{aligned}\mathbb{E}[X^2] &= \int_3^4 \frac{x^2}{2} dx + \int_7^8 \frac{x^2}{2} dx \\ &= \frac{4^3 - 3^3}{6} + \frac{8^3 - 7^3}{6} = \frac{103}{3}.\end{aligned}$$

Por tanto,

$$\text{Var}_{\text{inicial}}(X) = \frac{103}{3} - \left(\frac{11}{2}\right)^2 = \frac{49}{12}.$$

Para la nueva distribución, el segundo momento es

$$\begin{aligned}\mathbb{E}[X^2] &= \int_4^5 \frac{x^2}{2} dx + \int_6^7 \frac{x^2}{2} dx \\ &= \frac{5^3 - 4^3}{6} + \frac{7^3 - 6^3}{6} = \frac{94}{3}.\end{aligned}$$

En consecuencia,

$$\text{Var}_{\text{nueva}}(X) = \frac{94}{3} - \left(\frac{11}{2}\right)^2 = \frac{13}{12}.$$

La nueva varianza es menor porque la misma masa se concentra en intervalos más cercanos a la media.

Pregunta 75. (Control 2 (recuperativo) - Otoño 2024, P2)

Se tiene la siguiente función de densidad de probabilidad, asociado a los milímetros de lluvia caídos dentro del terreno A :

$$f_X(x) = \frac{2}{x} \mathbf{1}_{\{x \in [a, b]\}},$$

donde $a, b > 0$ son parámetros desconocidos. Considere que la varianza de esta distribución es igual a 1.

- (a) Calcule el promedio de milímetros de lluvia que caen en dicho terreno. Se debe entregar el valor numérico o bien una expresión que solo posea valores conocidos. Puede dejar su respuesta en función de a y b , mientras encuentre el valor de estos parámetros o deje correctamente planteadas las expresiones a partir de las cuales éstos se puedan encontrar (note que en caso que deba integrar, debe resolver estas integrales).

Considere la variable aleatoria que modela los milímetros de lluvia caídos en el terreno B . Esta VA distribuye Normal, con valor esperado igual a 4 y la misma varianza que la del terreno A . (Note que pueden haber valores negativos de lluvia con densidad positiva, sin embargo, dichos valores deberían tener una probabilidad muy baja.)

- (b) Para los milímetros de lluvia caídos en el terreno B , se desea identificar un intervalo centrado en la media (i.e., en la esperanza), de tal forma que este intervalo contenga 90 % de los escenarios de lluvia. Encuentre este intervalo.

Solución

(a). Sea $\mu = \mathbb{E}[X]$. Primero usamos que toda densidad debe integrar uno:

$$1 = \int_a^b \frac{2}{x} dx = 2 \ln \left(\frac{b}{a} \right).$$

De esta igualdad se obtiene

$$\frac{b}{a} = e^{1/2}, \quad \text{es decir,} \quad b = a\sqrt{e}.$$

La esperanza se calcula directamente a partir de la densidad:

$$\mu = \int_a^b x \frac{2}{x} dx = 2(b - a).$$

Como la varianza es uno, usamos $\text{Var}(X) = \mathbb{E}[X^2] - \mu^2$. El segundo momento es

$$\mathbb{E}[X^2] = \int_a^b x^2 \frac{2}{x} dx = b^2 - a^2.$$

Por tanto, los parámetros satisfacen

$$1 + 4(b - a)^2 = b^2 - a^2.$$

Al sustituir $b = a\sqrt{e}$,

$$1 = a^2 [(e - 1) - 4(\sqrt{e} - 1)^2].$$

Como $a > 0$,

$$a = \frac{1}{\sqrt{(e - 1) - 4(\sqrt{e} - 1)^2}}, \quad b = \frac{\sqrt{e}}{\sqrt{(e - 1) - 4(\sqrt{e} - 1)^2}}.$$

En consecuencia,

$$\begin{aligned} \mu &= 2(b - a) \\ &= \frac{2(\sqrt{e} - 1)}{\sqrt{(e - 1) - 4(\sqrt{e} - 1)^2}} \approx 6,943. \end{aligned}$$

Numéricamente, $a \approx 5,351$ y $b \approx 8,822$.

(b). Sea Y la cantidad de lluvia en el terreno B . Como tiene media 4 y la misma varianza que X ,

$$Y \sim \mathcal{N}(4, 1).$$

Buscamos un intervalo de la forma $[4 - \delta, 4 + \delta]$ que contenga probabilidad 0,9. Al estandarizar,

$$\begin{aligned} 0,9 &= \mathbb{P}(4 - \delta \leq Y \leq 4 + \delta) \\ &= \mathbb{P}(-\delta \leq Y - 4 \leq \delta) \\ &= \Phi(\delta) - \Phi(-\delta) \\ &= 2\Phi(\delta) - 1. \end{aligned}$$

Por ello, $\Phi(\delta) = 0,95$ y

$$\delta = \Phi^{-1}(0,95) \approx 1,645.$$

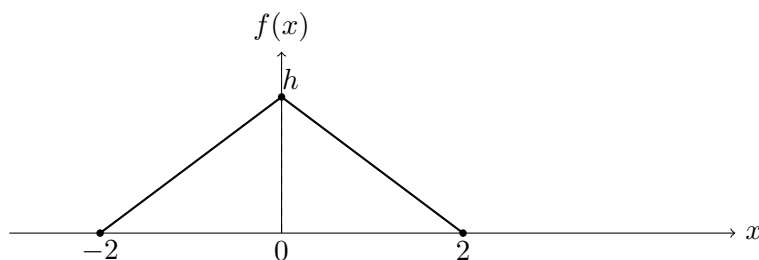
El intervalo solicitado es

$$[4 - \delta, 4 + \delta] \approx [2,355, 5,645].$$

Pregunta 76. (Control 2 - Primavera 2024, P2)

Alice desea comunicar a Bob la respuesta a una pregunta de “sí” o “no” a través de un canal de comunicación ruidoso. Este canal codifica el “sí” como 1 y el “no” como 0. Debido al ruido presente en el canal, Bob no recibe exactamente lo que Alice envía, sino una versión afectada por dicho ruido¹. Bob interpreta el mensaje recibido según la siguiente regla: si el valor recibido es mayor a $1/2$, lo considera un “sí”; de lo contrario, lo interpreta como un “no”.

Asuma para las partes (a) y (b) que la variable aleatoria del ruido sigue una distribución triangular con parámetro desconocido h (ver figura).



- (a) Determine el valor de h . Indique el rango y la PDF de la variable aleatoria del ruido.
- (b) Suponga que Alice emite la respuesta “No”. Calcule la probabilidad de que Bob entienda **incorrectamente** el mensaje de Alice.

Asuma para las partes (c) y (d) que el ruido sigue una distribución $\mathcal{N}(0, 2^2)$.

- (c) Calcule la probabilidad de que Bob entienda **incorrectamente** el mensaje de Alice. Al igual que en la parte (b), considere que Alice envía un “no”.
- (d) Siguiendo el caso en que el ruido distribuye Normal. ¿Qué pasa con el resultado de (c) si la desviación estándar σ es muy pequeña? ¿Qué pasa si σ es muy grande? Puede responder esta pregunta sin realizar cálculos, sino que solo justificando claramente su respuesta.

Solución

(a). La densidad tiene la forma de un triángulo con base 4 y altura h . Como el área bajo una densidad debe ser uno,

$$\frac{4h}{2} = 1,$$

de donde $h = 1/2$. El rango del ruido es $[-2, 2]$. Las ecuaciones de las rectas que forman el

¹El ruido es independiente del mensaje transmitido por Alice.

triángulo entregan

$$f_X(x) = \begin{cases} \frac{x+2}{4}, & -2 \leq x < 0, \\ \frac{2-x}{4}, & 0 \leq x \leq 2, \\ 0, & \text{en otro caso.} \end{cases}$$

(b). Sea X el ruido. Como Alice envía un “no”, el valor transmitido es cero y la señal recibida por Bob coincide con X . Bob comete un error si interpreta la señal como un “sí”, lo cual ocurre cuando $X > 1/2$. En ese intervalo se usa la rama derecha de la densidad:

$$\begin{aligned} \mathbb{P}\left(X > \frac{1}{2}\right) &= \int_{1/2}^2 \frac{2-x}{4} dx \\ &= \left[\frac{2x - x^2/2}{4} \right]_{1/2}^2 \\ &= \frac{9}{32}. \end{aligned}$$

(c). Ahora $X \sim \mathcal{N}(0, 2^2)$. El error sigue correspondiendo al evento $\{X > 1/2\}$. Al estandarizar mediante $Z = X/2 \sim \mathcal{N}(0, 1)$,

$$\begin{aligned} \mathbb{P}\left(X > \frac{1}{2}\right) &= 1 - \mathbb{P}\left(X \leq \frac{1}{2}\right) \\ &= 1 - \mathbb{P}\left(Z \leq \frac{1}{4}\right) \\ &= 1 - \Phi\left(\frac{1}{4}\right) \approx 0,4013. \end{aligned}$$

(d). Para una desviación estándar general $\sigma > 0$, la probabilidad de error cuando Alice envía un “no” es

$$\mathbb{P}\left(X > \frac{1}{2}\right) = 1 - \Phi\left(\frac{1}{2\sigma}\right).$$

Cuando σ es muy pequeña, el ruido se concentra cerca de cero y $1/(2\sigma)$ es muy grande. Por ello, la probabilidad de superar $1/2$ tiende a cero. Cuando σ es muy grande, $1/(2\sigma)$ se aproxima a cero y la probabilidad de error tiende a

$$1 - \Phi(0) = \frac{1}{2}.$$

Pregunta 77. (Control 2 - Primavera 2025, P2)

- (a) La administración de la estación de servicio (gasolinera) desea dimensionar el estanque de almacenamiento de manera tal que, aun en semanas con ventas muy altas, la probabilidad de quedarse sin combustible antes del próximo abastecimiento sea extremadamente baja.

La estación se reabastece una vez por semana y en cada reabastecimiento se llena el estanque a su capacidad total.

El volumen de demanda semanal (en miles de litros) posee la siguiente función de densidad de probabilidad:

$$f(x) = \begin{cases} 5(1-x)^4, & 0 < x < 1, \\ 0, & \text{en otro caso.} \end{cases}$$

Determine cuál debe ser la **capacidad total del estanque** (en miles de litros) para que la probabilidad de que el suministro se agote en una semana sea exactamente uno por

ciento.

- (b) Una pareja, conformada por Andrea y Carlos, va al casino y juega a la ruleta. A ambos les gusta jugar a la ruleta, pero tienen diferentes estrategias. Andrea juega una ficha por turno apostando siempre al número 24; mientras que Carlos apuesta en cada turno cuatro fichas de modo que pondrá una en cada número del 1 al 4. Tanto Andrea como Carlos jugarán hasta que alguno de los números a los que cada uno apuesta salga (por ende asuma que tienen mucho dinero). La ruleta posee 38 números en total, y todos los números tienen la misma probabilidad de salir.

Asuma que solo un miembro de la pareja jugará, tal que la probabilidad de que sea Andrea es 30 %, y Carlos con la probabilidad complementaria. Notar que el o la jugadora que juega, jugará todas las rondas hasta ganar.

Si es que se sabe que el miembro de la pareja que jugó lo hizo por 10 turnos, ¿cuál es la probabilidad de que haya sido Andrea? Para responder esto, **debe** considerar una VA que vale 0 si es que juega Andrea, y 1 si es que juega Carlos.

Solución

(a). Sea X la demanda semanal, medida en miles de litros. Para $0 \leq x \leq 1$, su CDF es

$$\begin{aligned} F_X(x) &= \int_0^x 5(1-t)^4 dt \\ &= 1 - (1-x)^5. \end{aligned}$$

Si C es la capacidad del estanque, este se agota cuando $X > C$. Como se requiere una probabilidad de agotamiento igual a 0,01,

$$\begin{aligned} 0,01 &= \mathbb{P}(X > C) \\ &= 1 - F_X(C) \\ &= (1 - C)^5. \end{aligned}$$

Por tanto, la capacidad necesaria es

$$C = 1 - 0,01^{1/5}$$

miles de litros. Este valor pertenece a $(0, 1)$, como exige el soporte de la demanda.

(b). Definamos la variable exigida en el enunciado:

$$Z = \begin{cases} 0, & \text{si juega Andrea,} \\ 1, & \text{si juega Carlos.} \end{cases}$$

Así, $\mathbb{P}(Z = 0) = 0,3$ y $\mathbb{P}(Z = 1) = 0,7$. Sea T el número de turnos hasta ganar. Andrea acierta en cada turno con probabilidad $p_A = 1/38$, mientras que Carlos acierta con probabilidad $p_C = 4/38$. Condicionado en la persona que juega,

$$T \mid Z = 0 \sim \text{Geom}(p_A), \quad T \mid Z = 1 \sim \text{Geom}(p_C),$$

donde la geométrica cuenta el turno del primer acierto. En particular,

$$\mathbb{P}(T = 10 \mid Z = 0) = (1 - p_A)^9 p_A, \quad \mathbb{P}(T = 10 \mid Z = 1) = (1 - p_C)^9 p_C.$$

El Teorema de Bayes, con la ley de probabilidades totales en el denominador, entrega

$$\begin{aligned} \mathbb{P}(Z = 0 \mid T = 10) &= \frac{\mathbb{P}(T = 10 \mid Z = 0)\mathbb{P}(Z = 0)}{\sum_{z=0}^1 \mathbb{P}(T = 10 \mid Z = z)\mathbb{P}(Z = z)} \\ &= \frac{0,3(37/38)^9(1/38)}{0,3(37/38)^9(1/38) + 0,7(34/38)^9(4/38)} \\ &\approx 0,1866. \end{aligned}$$

6.2. Transformaciones y distribuciones continuas

Resumen

Método de la CDF.

Para cualquier transformación $Y = g(X)$, primero se puede determinar su rango y calcular

$$F_Y(y) = \mathbb{P}(g(X) \leq y).$$

Después, si F_Y es derivable, $f_Y(y) = F'_Y(y)$. Este método también funciona cuando g no es inyectiva o cuando la desigualdad describe varias regiones del rango de X .

Cambio de variable.

Si $Y = g(X)$ y g es monótona, diferenciable e invertible en el soporte de X ,

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right|,$$

para y en el rango de Y . Si g no es uno a uno y sus preimágenes regulares son x_j , entonces

$$f_Y(y) = \sum_{j: g(x_j)=y} \frac{f_X(x_j)}{|g'(x_j)|}.$$

Uniforme.

Si $X \sim \text{Uniforme}(a, b)$, con $a < b$,

$$f_X(x) = \frac{1}{b-a} \mathbb{1}_{\{a \leq x \leq b\}}.$$

Su esperanza es $(a+b)/2$ y su varianza es $(b-a)^2/12$.

Exponencial.

Con $\lambda > 0$ interpretado como tasa, si $X \sim \text{Exponencial}(\lambda)$,

$$f_X(x) = \lambda e^{-\lambda x} \mathbb{1}_{\{x \geq 0\}},$$

$$F_X(x) = 1 - e^{-\lambda x}, \quad x \geq 0.$$

Su esperanza es $1/\lambda$ y su varianza es $1/\lambda^2$. Para $s, t \geq 0$ satisface la falta de memoria:

$$\mathbb{P}(X > s+t \mid X > s) = \mathbb{P}(X > t).$$

Normal.

Si $X \sim \mathcal{N}(\mu, \sigma^2)$, entonces

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right],$$

con $\mathbb{E}[X] = \mu$ y $\text{Var}(X) = \sigma^2$. Si Φ denota la CDF de una normal estándar,

$$Z = \frac{X-\mu}{\sigma} \sim \mathcal{N}(0, 1),$$

$$\mathbb{P}(X \leq x) = \Phi\left(\frac{x-\mu}{\sigma}\right),$$

$$\Phi(-z) = 1 - \Phi(z).$$

Gamma.

Para forma $\alpha > 0$ y tasa $\lambda > 0$,

$$f_X(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} \mathbb{1}_{\{x \geq 0\}}.$$

Si $X \sim \text{Gamma}(\alpha, \lambda)$, entonces $\mathbb{E}[X] = \alpha/\lambda$ y $\text{Var}(X) = \alpha/\lambda^2$. La distribución exponencial corresponde a $\alpha = 1$.

Transformaciones y sumas conocidas.

- **Transformación normal afín.** Si $X \sim \mathcal{N}(\mu, \sigma^2)$, entonces $aX + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$.
- **Suma de normales.** Si $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$ son independientes, entonces $\sum_i X_i \sim \mathcal{N}(\sum_i \mu_i, \sum_i \sigma_i^2)$.
- **Mínimo de exponenciales.** Si $X_i \sim \text{Exponencial}(\lambda_i)$ son independientes, entonces $\min_i X_i \sim \text{Exponencial}(\sum_i \lambda_i)$.
- **Escalamiento exponencial.** Si $X \sim \text{Exponencial}(\lambda)$ y $c > 0$, entonces $cX \sim \text{Exponencial}(\lambda/c)$.
- **Suma Gamma.** Si $X_i \sim \text{Gamma}(\alpha_i, \lambda)$ son independientes y comparten la tasa, entonces $\sum_i X_i \sim \text{Gamma}(\sum_i \alpha_i, \lambda)$.
- **Transformación lognormal.** Si $X \sim \mathcal{N}(\mu, \sigma^2)$ y $Y = e^X$, entonces, para $y > 0$,

$$F_Y(y) = \Phi\left(\frac{\log y - \mu}{\sigma}\right).$$

6.2.1. Preguntas

Pregunta 78.

Su artista favorita sacará una nueva canción. El gusto de usted por la nueva canción se puede representar por una variable aleatoria continua X , donde $X = 1$ representa que le gusta mucho y $X = -1$ representa que usted la detesta. La PDF de X está dada por

$$f_X(x) = \begin{cases} x^2 & \text{si } x \in [-1, 0] \\ \frac{4}{3}x & \text{si } x \in [0, 1] \\ 0 & \text{en otro caso} \end{cases}$$

- Verifique que f_X es una función de densidad bien definida.
- Diremos que usted *tolera* la canción (pero no le encanta), si es que $0 \leq X \leq 0.3$. Calcule la probabilidad de que usted *tolere* la canción.
- El efecto que provoca la canción en usted (ya sea gusto o disgusto) se representa por la variable $Y = |X|$. Ocupe el TCV para calcular la PDF de Y .
- Calcule el valor esperado de Y .
- Diremos que la canción no tiene un efecto relevante en su vida si es que Y es menor o igual a 0.5. Calcule la probabilidad de que la canción sí tenga un efecto relevante en su vida.

Solución

(a). La función es no negativa. Además,

$$\int_{-\infty}^{\infty} f_X(x) dx = \int_{-1}^0 x^2 dx + \int_0^1 \frac{4}{3}x dx = \frac{1}{3} + \frac{2}{3} = 1.$$

Por tanto, f_X es una densidad.

(b). Como X es continua, la probabilidad se obtiene integrando la densidad en el intervalo:

$$\mathbb{P}(0 \leq X \leq 0.3) = \int_0^{0.3} \frac{4}{3}x dx = 0.06.$$

(c). Para cada $0 < y < 1$, la transformación $Y = |X|$ tiene las preimágenes $x = y$ y $x = -y$, ambas con jacobiano absoluto 1. En consecuencia,

$$f_Y(y) = \begin{cases} \frac{4}{3}y + y^2, & 0 \leq y \leq 1, \\ 0, & \text{en otro caso.} \end{cases}$$

(d). Usando la densidad anterior,

$$\mathbb{E}[Y] = \int_0^1 y \left(\frac{4}{3}y + y^2 \right) dy = \frac{4}{9} + \frac{1}{4} = \frac{25}{36}.$$

(e). El efecto es relevante cuando $Y > 0.5$. Por tanto,

$$\begin{aligned} \mathbb{P}(Y > 0.5) &= 1 - \int_0^{1/2} \left(\frac{4}{3}y + y^2 \right) dy \\ &= 1 - \frac{1}{6} - \frac{1}{24} = \frac{19}{24} \approx 0.7917. \end{aligned}$$

Pregunta 79.

- (i) El tiempo que se mantiene funcionando el internet desde cierto inicio inicial se modela según una variable $T \sim \text{Exp}(\lambda)$ cuya densidad es $f_T(z) = \lambda e^{-\lambda z}$ para $z \in [0, \infty)$.
- (a) Calcule la probabilidad de que el internet dure menos que un tiempo $a > 0$.
- (b) La duración del cable, desde el mismo tiempo inicial, se puede modelar con otra variable $S \sim \text{Exp}(\lambda t)$ que asumimos independiente a T . Calcule la probabilidad de que al menos uno de los dos servicios deje de funcionar en un tiempo $a > 0$.
- (c) La producción de su trabajo online distribuye según S^2 , calcule la densidad y esperanza de su producción.
- (ii) Suponga que X tiene una distribución exponencial con parámetro $\lambda = 1$, para $c > 0$ demuestre que $Y = \frac{X}{c}$ sigue una distribución exponencial con parámetro $\lambda = c$.
- (iii) Hay un teorema (visto en clases) que indica que si $X \sim \mathcal{N}(\mu, \sigma^2)$, entonces la variable aleatoria $Y = aX + b$ también distribuye Normal. Demuestre dicho teorema, y además determine la esperanza y varianza de Y .

Solución

(i)(a). Como $T \sim \text{Exp}(\lambda)$ y $a > 0$,

$$\mathbb{P}(T < a) = 1 - e^{-\lambda a}.$$

(i)(b). Para que $\text{Exp}(\lambda t)$ esté bien definida como una distribución exponencial, suponemos $t > 0$. Como $S \sim \text{Exp}(\lambda t)$ es independiente de $T \sim \text{Exp}(\lambda)$,

$$\mathbb{P}(S \leq a \text{ o } T \leq a) = 1 - e^{-\lambda a(t+1)}.$$

(i)(c). Sea $P = S^2$. Como $S \geq 0$, para $p \geq 0$ se tiene

$$F_P(p) = \mathbb{P}(S^2 \leq p) = \mathbb{P}(S \leq \sqrt{p}) = 1 - e^{-\lambda t \sqrt{p}}.$$

Al derivar esta función de distribución y usar $\mathbb{E}[S^2] = \text{Var}(S) + \mathbb{E}[S]^2$, se obtiene

$$f_P(p) = \begin{cases} \frac{\lambda t}{2\sqrt{p}} e^{-\lambda t \sqrt{p}}, & p > 0, \\ 0, & p \leq 0. \end{cases} \quad \mathbb{E}[P] = \frac{2}{(\lambda t)^2}.$$

(ii). Para $Y = X/c$, con $X \sim \text{Exp}(1)$ y $c > 0$,

$$F_Y(y) = \begin{cases} 0, & y < 0, \\ 1 - e^{-cy}, & y \geq 0, \end{cases} \quad f_Y(y) = \begin{cases} 0, & y < 0, \\ ce^{-cy}, & y \geq 0. \end{cases}$$

Por lo tanto, $Y \sim \text{Exp}(c)$.

(iii). Si $a \neq 0$, la transformación tiene inversa $x = (y - b)/a$ y jacobiano absoluto $1/|a|$. Por ello,

$$\begin{aligned} f_Y(y) &= \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right) \\ &= \frac{1}{\sqrt{2\pi a^2 \sigma^2}} \exp\left\{-\frac{(y - a\mu - b)^2}{2a^2 \sigma^2}\right\}. \end{aligned}$$

Esta es la densidad de una variable normal con media $a\mu + b$ y varianza $a^2 \sigma^2$. Además, por linealidad de la esperanza y las propiedades de la varianza,

$$Y \sim \mathcal{N}(a\mu + b, a^2 \sigma^2), \quad \mathbb{E}[Y] = a\mu + b, \quad \text{Var}(Y) = a^2 \sigma^2.$$

Si $a = 0$, entonces $Y = b$ casi seguramente, con varianza nula.

Pregunta 80.

Considere el siguiente método para dibujar una cuerda al azar en una circunferencia de radio R . Sobre un radio se escoge un punto al azar, de acuerdo con la distribución Uniforme en $[0, R]$, y luego, en dicho punto, se traza una perpendicular al radio, obteniendo así una cuerda. Sea L la variable aleatoria que designa el largo de la cuerda obtenida.

- Obtener la densidad de probabilidad de L .
- Calcular la esperanza de L y del área del triángulo formado uniendo los extremos de la cuerda con el centro de la circunferencia.

Solución

(a). Sea $V \sim \text{Unif}(0, R)$ la distancia del centro al punto escogido. La longitud de la cuerda es $L = 2\sqrt{R^2 - V^2}$. Para $0 \leq \ell \leq 2R$,

$$\begin{aligned} F_L(\ell) &= \mathbb{P}\left(2\sqrt{R^2 - V^2} \leq \ell\right) \\ &= \mathbb{P}\left(V \geq \sqrt{R^2 - \frac{\ell^2}{4}}\right) \\ &= 1 - \frac{1}{R}\sqrt{R^2 - \frac{\ell^2}{4}}, \end{aligned}$$

y, por tanto,

$$f_L(\ell) = \begin{cases} \frac{\ell}{4R\sqrt{R^2 - \ell^2/4}}, & 0 < \ell < 2R, \\ 0, & \text{en otro caso.} \end{cases}$$

(b). Usando $L = 2\sqrt{R^2 - V^2}$,

$$\mathbb{E}[L] = \frac{1}{R} \int_0^R 2\sqrt{R^2 - v^2} dv = \frac{\pi R}{2}.$$

El área del triángulo es $A = V\sqrt{R^2 - V^2}$, de modo que

$$\mathbb{E}[A] = \frac{1}{R} \int_0^R v\sqrt{R^2 - v^2} dv = \frac{R^2}{3}.$$

6.2.2. Preguntas de Evaluación**Pregunta 81.** (Control 2 - Otoño 2023, P3)

- Josefa es una investigadora muy aplicada y necesita hacerse un examen de colesterol. Suponga que una persona tiene su colesterol *saludable* cuando este es menor o igual a 180. Antes de acudir al laboratorio Josefa decide investigar un poco, y descubrió que los resultados de los exámenes de colesterol se ven afectados por factores externos del día en que se toma el examen (cómo durmió la noche anterior, qué comió, etc.). Como consecuencia de esto, el **resultado del examen** es un valor que distribuye $\text{Normal}(\mu, 25^2)$, donde μ es el **valor real** desconocido del colesterol de Josefa. Suponga que Josefa posee un **valor real** de colesterol de 210. ¿Cuál es la probabilidad de que la lectura del examen entregue un valor *saludable*? Puede entregar la fórmula de su resultado en función de números y operaciones matemáticas, no funciones de variables aleatorias, integrales ni sumatorias.
- Hay un teorema visto en clases que indica que, si $X \sim \text{Normal}(\mu, \sigma^2)$, entonces la variable aleatoria $Y = aX + b$ también distribuye Normal. Demuestre dicho teorema y, además,

determine la esperanza y varianza de Y .

Solución

(a). Sea C la lectura del examen. Bajo el valor real indicado en el enunciado,

$$C \sim \text{Normal}(210, 25^2).$$

Una lectura saludable corresponde al evento $C \leq 180$. Al estandarizar,

$$\begin{aligned}\mathbb{P}(C \leq 180) &= \mathbb{P}\left(\frac{C - 210}{25} \leq \frac{180 - 210}{25}\right) \\ &= \Phi(-1,2) \approx 0,1151.\end{aligned}$$

(b). Escribamos $X = \mu + \sigma Z$, donde $Z \sim \text{Normal}(0, 1)$. Entonces

$$Y = aX + b = (a\mu + b) + a\sigma Z.$$

Si $a \neq 0$, la variable $\text{sgn}(a)Z$ sigue siendo normal estándar, debido a la simetría de su distribución. Por ello,

$$Y \sim \text{Normal}(a\mu + b, a^2\sigma^2).$$

La misma expresión se obtiene directamente para sus momentos:

$$\begin{aligned}\mathbb{E}[Y] &= a\mathbb{E}[X] + b = a\mu + b, \\ \text{Var}(Y) &= \text{Var}(aX + b) = a^2 \text{Var}(X) = a^2\sigma^2.\end{aligned}$$

Si $a = 0$, entonces $Y = b$ casi seguramente. Este es el caso degenerado, con esperanza b y varianza cero; puede escribirse como $\text{Normal}(b, 0)$ cuando se admite la distribución normal degenerada.

Pregunta 82. (Control 2 - Primavera 2023, P2)

Es lunes y usted no escuchó su alarma en la mañana, por lo que se encuentra notoriamente atrasado/a para llegar a sus clases. Para llegar a la universidad tiene dos opciones: (i) caminar al paradero y tomar una micro (es decir, un bus), o (ii) caminar directamente a la universidad. En caso de (i), debe caminar hasta el paradero (que está de camino a la universidad, por lo que no lo aleja de su destino) y esperar a que esta pase. El tiempo de caminata entre la casa y el paradero es de a minutos, mientras que el tiempo de caminata del paradero a la universidad es de c minutos. Además, el tiempo que se demora en pasar una micro (por el paradero) contando desde el tiempo en que pasó la última micro (por el mismo paradero), se puede modelar como una VA con distribución Exponencial de parámetro λ [minutos]⁻¹. Asuma que el tiempo de traslado de la micro es prácticamente nulo, incluso si la micro para a recoger pasajeros. Suponga que al salir de su casa, usted ve pasar la micro justo frente a sus narices (esto aplica para ambas preguntas que se muestran a continuación).

- (a) Usted decide la opción (ii), es decir, caminar desde su casa a la universidad. ¿Cuál es la probabilidad de que se demore menos que con la opción (i)?

Solución

(a). La caminata completa desde la casa hasta la universidad tarda $a + c$ minutos. Si se elige la micro, se caminan primero a minutos hasta el paradero y luego se espera a la siguiente micro; su tiempo de traslado se considera nulo. Sea W el tiempo de espera desde la llegada al paradero. Los tiempos entre micros son exponenciales con parámetro λ . Por la propiedad de falta de memoria, el tiempo restante hasta la próxima micro al llegar al paradero satisface

$$W \sim \text{Exponencial}(\lambda),$$

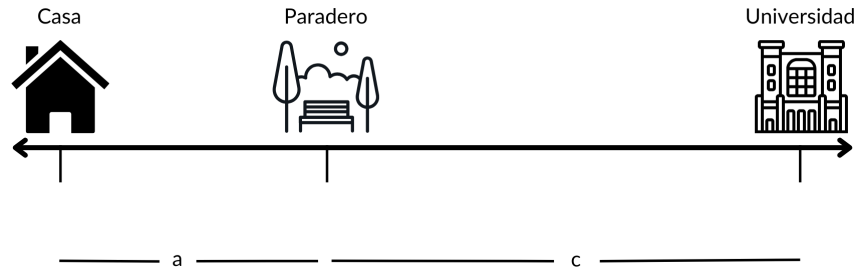


Figura 6.1: Esquema del problema. a y c son las unidades de tiempo que toma el traslado **caminando** para cada segmento, respectivamente.

independientemente de las micros que hayan pasado durante los a minutos de caminata. Caminar es más rápido que tomar la micro exactamente cuando

$$a + c < a + W,$$

es decir, cuando $W > c$. Por tanto,

$$\mathbb{P}(\text{caminar tarda menos}) = \mathbb{P}(W > c) = e^{-\lambda c}.$$

Pregunta 83. (Control 2 - Primavera 2023, P3)

El cambio climático ha afectado de manera importante el desarrollo normal de las lluvias, provocando sequías o inundaciones, dependiendo de la cantidad de precipitaciones. La dirección meteorológica ha recopilado datos históricos sobre la cantidad anual de lluvia en milímetros para los últimos 30 años. Estos datos siguen una distribución normal con una media de 90 mm y una desviación estándar de 20 mm. Don Osvaldo ha decidido plantar un huerto a principios de este año, el cual solo se regará por las precipitaciones. Específicamente, el huerto necesita una cantidad de agua igual a 100 ± 30 mm, pues si el huerto recibe menos de 70 mm, las plantas se secarán por falta de agua, pero en caso de que las lluvias del año sean mayores a 130 mm, las plantas se ahogarán por exceso de agua.

- Encuentre la probabilidad de que el huerto de Don Osvaldo no muera debido a las lluvias del año que viene. Debe entregar finalmente el valor numérico de la respuesta, más todo el desarrollo para llegar a este.
- Suponga ahora que Don Osvaldo tiene 15 huertos diferentes en la misma ciudad. Al igual como se menciona más arriba, cada uno de estos huertos no sobrevive si es que la cantidad de agua caída está fuera del intervalo $[70, 130]$ mm. Note que, si la lluvia de un año es de x mm, entonces todos los huertos recibirán x mm de lluvia. Sea W el número de huertos que sobrevive después del año. ¿Qué distribución sigue W ? Si sigue una distribución conocida, entregue su nombre y sus parámetros; en caso contrario, indique la distribución de dicha variable aleatoria, su rango y PMF. Si le es útil usar la probabilidad de la parte (a), puede llamar q a esta probabilidad y dejar su respuesta en función de esto.

Solución

(a). Sea X la cantidad anual de lluvia, en milímetros. Según el modelo,

$$X \sim \text{Normal}(90, 20^2).$$

El huerto sobrevive cuando la precipitación queda entre 70 y 130 mm. Al estandarizar ambos

extremos,

$$\begin{aligned}\mathbb{P}(70 \leq X \leq 130) &= \mathbb{P}\left(\frac{70-90}{20} \leq \frac{X-90}{20} \leq \frac{130-90}{20}\right) \\ &= \mathbb{P}(-1 \leq Z \leq 2) \\ &= \Phi(2) - \Phi(-1) \\ &\approx 0,9772 - 0,1587 = 0,8185,\end{aligned}$$

donde $Z \sim \text{Normal}(0, 1)$.

(b). Sea $q = \mathbb{P}(70 \leq X \leq 130)$, calculada en la parte anterior. Como todos los huertos están en la misma ciudad y reciben exactamente la misma cantidad de lluvia, sobreviven todos o no sobrevive ninguno. Por tanto,

$$W = 15\mathbb{1}_{\{70 \leq X \leq 130\}},$$

y su rango es $\{0, 15\}$. Su PMF es

$$p_W(w) = \begin{cases} 1 - q, & w = 0, \\ q, & w = 15, \\ 0, & \text{en otro caso.} \end{cases}$$

Equivalentemente, W tiene la misma distribución que $15B$, con $B \sim \text{Bernoulli}(q)$. No es una binomial con quince ensayos independientes, pues las supervivencias de los huertos están completamente determinadas por una misma precipitación anual.

Pregunta 84. (Control 2 - Otoño 2024, P3)

- (a) Considere dos variables aleatorias, X y Z . La variable aleatoria Z sigue una distribución Normal Estándar. La relación entre estas variables aleatorias es tal que $X = g(Z) = e^{\sigma Z + \mu}$.

Considere que $\mu = 1$ y $\sigma = \frac{1}{2}$. Evalúe la función de distribución acumulada de X en $\frac{5}{2}$. (**Hint:** le puede ser útil recordar alguno de los ejercicios propuestos).

- (b) Cada día, usted arrienda un carrito para vender jugos de diversos sabores a las afuera de un parque de diversiones. Usted se instala a las afueras del parque O'Higgins con su carrito desde las 8:00 hasta las 18:00, o bien hasta que la máquina que hace jugos del carrito deje de funcionar. El tiempo en horas que dura dicha máquina, desde las 8:00 de cada día, sigue una distribución Exponencial de parámetro 4 [horas⁻¹]. Considere que si es que la máquina se hecha a perder un día, esta se repara durante la noche para estar disponible la jornada del día siguiente. La ganancia por concepto de ventas se estima en u [\$/hora]. De esta manera, la ganancia de un día en el cual la máquina dura 2,5 horas, es de $2,5 \cdot u$. Por otra parte, el arriendo diario del carrito es de c [\$] (independiente si el carrito falla o no durante el día de arriendo). Determine la utilidad (i.e., ganancias menos costos) promedio que se obtiene en un día.

Solución

- (a). Como la función exponencial es estrictamente creciente, la desigualdad puede despejarse sin

cambiar su sentido. Sustituyendo $\mu = 1$ y $\sigma = 1/2$,

$$\begin{aligned} F_X\left(\frac{5}{2}\right) &= \mathbb{P}\left(X \leq \frac{5}{2}\right) \\ &= \mathbb{P}\left(e^{Z/2+1} \leq \frac{5}{2}\right) \\ &= \mathbb{P}\left(Z \leq \frac{\ln(5/2) - 1}{1/2}\right) \\ &= \Phi\left(2 \ln\left(\frac{5}{2}\right) - 2\right) \approx \Phi(-0,1674) \approx 0,4335. \end{aligned}$$

(b). Sea T la duración de la máquina desde las 8:00, medida en horas. Entonces

$$T \sim \text{Exponencial}(4),$$

donde 4 es la tasa. Como la jornada dura como máximo 10 horas, las horas efectivamente trabajadas son $\min\{T, 10\}$ y la utilidad diaria es

$$U = u \min\{T, 10\} - c.$$

Aplicando linealidad de la esperanza,

$$\mathbb{E}[U] = u \mathbb{E}[\min\{T, 10\}] - c.$$

La esperanza truncada puede calcularse separando los casos en que la máquina falla antes de 10 horas y en que sigue funcionando al final de la jornada:

$$\begin{aligned} \mathbb{E}[\min\{T, 10\}] &= \int_0^{10} 4te^{-4t} dt + 10\mathbb{P}(T > 10) \\ &= \int_0^{10} 4te^{-4t} dt + 10e^{-40} \\ &= \frac{1 - e^{-40}}{4}. \end{aligned}$$

Por tanto,

$$\mathbb{E}[U] = \frac{u}{4} (1 - e^{-40}) - c \approx \frac{u}{4} - c.$$

Pregunta 85. (Control 2 (recuperativo) - Otoño 2024, P1)

El Telescopio Espacial Hubble está investigando un nuevo sector espacial y ha recibido nuevas imágenes de objetos celestes. Las imágenes astronómicas, especialmente las obtenidas por observatorios, suelen tener un rango dinámico muy amplio. Esto significa que hay una gran diferencia entre las regiones más brillantes (como estrellas o galaxias cercanas) y las más oscuras (como nebulosas o galaxias distantes). Capturar estos detalles puede ser un desafío, ya que las diferencias en la intensidad de la luz pueden ser extremas. Para mejorar la calidad de las imágenes y lograr una vista más detallada del cosmos, se toma la raíz cuadrada de la *intensidad de luz* de los píxeles. Esta transformación comprime el rango dinámico de la imagen, lo cual permite que los detalles en las áreas más oscuras sean más visibles sin saturar las áreas más brillantes. Considere la variable aleatoria X que modela la *intensidad de luz* de los píxeles de las imágenes que distribuye Rayleigh. Específicamente, su PDF es:

$$f_X(x) = xe^{-\frac{x^2}{2}} \mathbb{1}_{\{x \geq 0\}}.$$

Calcule la moda de la raíz cuadrada de la *intensidad de luz* de los píxeles, es decir, calcule el valor que maximiza la función de densidad de probabilidad de la variable aleatoria transformada (raíz cuadrada de la *intensidad de luz* de los píxeles). Justifique por qué el punto encontrado es un máximo.

Solución

Sea $Y = \sqrt{X}$. Como la transformación es creciente sobre $[0, \infty)$, su inversa es $x = y^2$ para $y \geq 0$. El cambio de variable entrega

$$\begin{aligned} f_Y(y) &= f_X(y^2) \left| \frac{d}{dy} y^2 \right| \\ &= 2y^3 e^{-y^4/2} \mathbf{1}_{\{y \geq 0\}}. \end{aligned}$$

La moda maximiza esta densidad, no la variable aleatoria. Para $y > 0$,

$$f'_Y(y) = 2y^2 e^{-y^4/2} (3 - 2y^4).$$

Los factores $2y^2$ y $e^{-y^4/2}$ son positivos en el interior del soporte. Por ello, $f'_Y(y) > 0$ cuando $0 < y < (3/2)^{1/4}$ y $f'_Y(y) < 0$ cuando $y > (3/2)^{1/4}$. Además, $f_Y(0) = 0$ y $f_Y(y) \rightarrow 0$ cuando $y \rightarrow \infty$. En consecuencia, el único punto donde la densidad pasa de creciente a decreciente es su máximo global, y la moda es

$$\left(\frac{3}{2}\right)^{1/4}.$$

Pregunta 86. (Control 2 - Otoño 2025, P2)

- (a) Sea X una variable aleatoria que sigue una distribución Exponencial de parámetro $\lambda > 0$. Definimos $Y = 1 - e^{-X}$.
1. Determine el rango de la variable aleatoria Y .
 2. Determine la función de distribución acumulada de la variable aleatoria Y .
- (b) Primero, Andrea tira un dado común. El número que aparece le indica cuántas monedas lanzará después. Se sabe que Andrea obtuvo 4 caras, pero no se sabe el número de monedas lanzadas. ¿Cuál es la probabilidad de que haya obtenido un 5 al tirar el dado?

Solución

(a)(1). Como $X \geq 0$, se tiene $0 < e^{-X} \leq 1$. Por consiguiente,

$$0 \leq Y = 1 - e^{-X} < 1.$$

Así, el rango de la transformación es $[0, 1)$. Si se usa la convención de describir solo los valores interiores donde la variable continua tiene densidad positiva, se escribe $(0, 1)$; el extremo 0 tiene probabilidad cero y esta diferencia no altera la CDF.

(a)(2). Para $0 < y < 1$, la función $x \mapsto 1 - e^{-x}$ es creciente, por lo que la desigualdad puede despejarse sin cambiar su sentido:

$$\begin{aligned} F_Y(y) &= \mathbb{P}(1 - e^{-X} \leq y) \\ &= \mathbb{P}(X \leq -\log(1 - y)) \\ &= F_X(-\log(1 - y)) \\ &= 1 - e^{-\lambda[-\log(1 - y)]} \\ &= 1 - (1 - y)^\lambda. \end{aligned}$$

Al incorporar los valores fuera del rango,

$$F_Y(y) = \begin{cases} 0, & y \leq 0, \\ 1 - (1 - y)^\lambda, & 0 < y < 1, \\ 1, & y \geq 1. \end{cases}$$

(b). Bajo la convención de que los lanzamientos de monedas son independientes y cada moneda

es balanceada, sea N el resultado del dado y sea H el número de caras obtenidas al lanzar las N monedas. El dado es común, de modo que $\mathbb{P}(N = k) = 1/6$ para $k = 1, \dots, 6$, y

$$H \mid (N = k) \sim \text{Bin}\left(k, \frac{1}{2}\right).$$

Por el Teorema de Bayes,

$$\mathbb{P}(N = 5 \mid H = 4) = \frac{\mathbb{P}(H = 4 \mid N = 5)\mathbb{P}(N = 5)}{\mathbb{P}(H = 4)}.$$

La ley de probabilidades totales muestra que solo contribuyen los resultados 4, 5 y 6 del dado:

$$\begin{aligned}\mathbb{P}(H = 4) &= \sum_{k=4}^6 \mathbb{P}(H = 4 \mid N = k)\mathbb{P}(N = k) \\ &= \frac{1}{6} \left[\binom{4}{4} \frac{1}{2^4} + \binom{5}{4} \frac{1}{2^5} + \binom{6}{4} \frac{1}{2^6} \right] \\ &= \frac{29}{6 \cdot 2^6}.\end{aligned}$$

Como $\mathbb{P}(H = 4 \mid N = 5) = \binom{5}{4}/2^5 = 5/2^5$, al sustituir se obtiene

$$\mathbb{P}(N = 5 \mid H = 4) = \frac{(5/2^5)(1/6)}{29/(6 \cdot 2^6)} = \frac{10}{29}.$$

Pregunta 87. (Control 2 - Otoño 2025, P3)

Una heladería produce helados a temperatura T , donde T es una variable aleatoria que sigue una distribución Normal de media -10 [C] y desviación estándar 2 [C]. Se dice que los helados cumplen la normativa para ser vendidos si su temperatura es menor o igual que -7 [C]. El precio de un helado es de 500 [\$], mientras que el costo de producción es de 100 [\$]. Considere que un helado no puede ser vendido si no cumple con la normativa, mientras que todo helado que la cumple es vendido.

- (a) ¿Cuál es la probabilidad de que un helado producido cumpla la normativa?
- (b) Sea la variable aleatoria

$$W = \mathbb{1}_{\{\text{el helado cumple la normativa}\}}.$$

Encuentre la varianza de la utilidad que obtiene la heladería por un helado producido. Podría serle útil escribir la utilidad en función de la variable aleatoria W .

Solución

(a). Como $T \sim \text{Normal}(-10, 2^2)$, la variable estandarizada

$$Z = \frac{T + 10}{2}$$

sigue una distribución Normal(0, 1). Por tanto,

$$\begin{aligned}p &:= \mathbb{P}(T \leq -7) \\ &= \mathbb{P}\left(Z \leq \frac{-7 - (-10)}{2}\right) \\ &= \Phi(1.5) \approx 0.9332.\end{aligned}$$

(b). La indicatriz W vale uno si el helado puede venderse y cero en caso contrario. Por ello,

$$W \sim \text{Bernoulli}(p), \quad \text{Var}(W) = p(1 - p).$$

La utilidad es el ingreso por venta menos el costo de producción:

$$U = 500W - 100.$$

El costo constante no aporta variabilidad y multiplicar por 500 multiplica la varianza por 500^2 . En consecuencia,

$$\begin{aligned}\text{Var}(U) &= 500^2 \text{Var}(W) \\ &= 250000 p(1-p) \\ &= 250000(0,9332)(1-0,9332) \\ &\approx 15584,44.\end{aligned}$$

Pregunta 88. (Control 2 - Primavera 2025, P3)

El conocido inversionista chileno Matías Herrera, gerente de la firma de inversiones *Andes Capital*, estudia el comportamiento diario de los retornos de una acción tecnológica que cotiza en la bolsa de Nueva York. Específicamente, el retorno diario de una acción X puede modelarse como una variable aleatoria Normal estándar.

Un analista financiero está interesado sólo en los **retornos moderados**, los cuales corresponden a días en los que la ganancia es menor a a , o bien la pérdida es menor a $-a$ (donde $a > 0$ es conocido). Es decir, en los días donde $|X| < a$.

Se define la variable aleatoria Y como la variable X condicionada al evento $\{-a < X < a\}$, es decir:

$$Y = X \mid (-a < X < a).$$

De esta manera, se tiene que la función de densidad de probabilidad de Y está dada por

$$f_Y(y) = \begin{cases} \frac{\phi(y)}{\Phi(a) - \Phi(-a)}, & -a < y < a, \\ 0, & \text{en otro caso,} \end{cases}$$

donde ϕ es la PDF de la normal estándar y Φ su CDF.

- ¿Cómo se compara la varianza de Y con respecto a la varianza de X ? Justifique su respuesta, lo cual puede hacer con palabras o matemáticamente.
- Sea $a = 3/2$, encuentre la probabilidad de que el **retorno moderado** sea menor o igual a $1/2$. Debe entregar una expresión algebraica con valores numéricos (ver instrucciones del control).

Solución

(a). Como el intervalo de truncamiento $(-a, a)$ y la densidad normal estándar son simétricos respecto de cero, $\mathbb{E}[Y] = 0$. Por consiguiente, $\text{Var}(Y) = \mathbb{E}[Y^2]$. Si

$$c = \mathbb{P}(-a < X < a) = \Phi(a) - \Phi(-a) = 2\Phi(a) - 1,$$

entonces

$$\text{Var}(Y) = \frac{1}{c} \int_{-a}^a y^2 \phi(y) dy.$$

Como $\phi'(y) = -y\phi(y)$, la integración por partes entrega

$$\begin{aligned}\int_{-a}^a y^2 \phi(y) dy &= [-y\phi(y)]_{-a}^a + \int_{-a}^a \phi(y) dy \\ &= c - 2a\phi(a).\end{aligned}$$

Por tanto,

$$\text{Var}(Y) = 1 - \frac{2a\phi(a)}{2\Phi(a) - 1} < 1 = \text{Var}(X),$$

pues $a > 0$, $\phi(a) > 0$ y $2\Phi(a) - 1 > 0$. La fórmula formaliza la intuición de que el truncamiento elimina ambas colas y concentra la distribución alrededor de cero.

(b). Para $a = 3/2$, el punto $1/2$ pertenece al soporte truncado. Integrando la densidad condicional,

$$\begin{aligned}\mathbb{P}\left(Y \leq \frac{1}{2}\right) &= \int_{-3/2}^{1/2} \frac{\phi(y)}{\Phi(3/2) - \Phi(-3/2)} dy \\ &= \frac{\Phi(1/2) - \Phi(-3/2)}{\Phi(3/2) - \Phi(-3/2)}.\end{aligned}$$

Usando $\Phi(1/2) = 0,6915$, $\Phi(3/2) = 0,9332$ y, por simetría, $\Phi(-3/2) = 0,0668$, resulta

$$\mathbb{P}\left(Y \leq \frac{1}{2}\right) = \frac{0,6915 - 0,0668}{0,9332 - 0,0668} = \frac{0,6247}{0,8664} \approx 0,7209.$$

Pregunta 89. (Control 2 - Otoño 2026, P3)

Suponga que usted está en camino a la universidad y acaba de llegar al paradero de la micro que le sirve. Sea T la variable aleatoria que describe el tiempo de espera hasta que llega la micro (en minutos).

- (a) Si se tiene que $\mathbb{P}(T > t) = \begin{cases} \left(\frac{4}{t}\right)^2 & \text{si } t \geq 4, \\ 1 & \text{si } t < 4, \end{cases}$ encuentre la esperanza de T .
- (b) Suponga ahora que T distribuye Normal con desviación estándar igual a 7 minutos, y se sabe que la probabilidad de esperar más de 20 minutos es 0,075. ¿Cuál es la esperanza de T ?
- (c) Suponga finalmente que T distribuye Exponencial con parámetro $1/2$. Para ver cuánto tiempo falta para que la micro llegue, usted revisa la aplicación Red en su teléfono, donde dice que la micro llegará entre 4 y 10 minutos. Asumiendo que esta información es acertada, ¿cuál es la probabilidad de que la micro llegue en menos de 7 minutos?

Solución

(a). Como T es no negativa, puede usarse la identidad de colas

$$\mathbb{E}[T] = \int_0^{\infty} \mathbb{P}(T > t) dt.$$

Separando en el punto $t = 4$,

$$\begin{aligned}\mathbb{E}[T] &= \int_0^4 1 dt + \int_4^{\infty} \left(\frac{4}{t}\right)^2 dt \\ &= 4 + 16 \left[-\frac{1}{t}\right]_4^{\infty} \\ &= 8.\end{aligned}$$

Equivalentemente, la CDF es $F_T(t) = 0$ para $t < 4$ y $F_T(t) = 1 - 16/t^2$ para $t \geq 4$; al derivarla se obtiene $f_T(t) = 32/t^3$ en $[4, \infty)$, que conduce a la misma esperanza.

(b). La información de cola equivale a

$$\mathbb{P}(T \leq 20) = 1 - 0,075 = 0,925.$$

Si $T \sim \mathcal{N}(\mu, 7^2)$, entonces

$$\Phi\left(\frac{20 - \mu}{7}\right) = 0,925.$$

Por tanto,

$$\frac{20 - \mu}{7} = \Phi^{-1}(0,925) \approx 1,44,$$

y

$$\mathbb{E}[T] = \mu = 20 - 7\Phi^{-1}(0,925) \approx 9,92 \text{ minutos.}$$

(c). El dato de la aplicación restringe el tiempo a $4 < T < 10$. Como T es continua,

$$\begin{aligned} \mathbb{P}(T < 7 \mid 4 < T < 10) &= \frac{\mathbb{P}(4 < T < 7)}{\mathbb{P}(4 < T < 10)} \\ &= \frac{F_T(7) - F_T(4)}{F_T(10) - F_T(4)}, \end{aligned}$$

donde $F_T(t) = 1 - e^{-\lambda t}$ para $t \geq 0$. En consecuencia,

$$\mathbb{P}(T < 7 \mid 4 < T < 10) = \frac{e^{-4\lambda} - e^{-7\lambda}}{e^{-4\lambda} - e^{-10\lambda}} = \frac{1 - e^{-3\lambda}}{1 - e^{-6\lambda}}.$$

Con $\lambda = 1/2$, la probabilidad solicitada es

$$\frac{1 - e^{-3/2}}{1 - e^{-3}}.$$

Pregunta 90. (Control 3 - Primavera 2022, P3)

Sea X una variable aleatoria que sigue una distribución Normal(μ, σ^2) e $Y = e^X$. En este caso se dice que $Y \sim \text{LogNormal}(\mu, \sigma^2)$. Encuentre la CDF de la VA Y en términos de Φ , donde Φ es la CDF de una distribución Normal Estándar (i.e., Normal(0, 1)).

Hint: Le puede ser de utilidad que $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp(-t^2/2) dt$ y el siguiente cambio de variable $z = \frac{\ln(y) - \mu}{\sigma}$.

Solución

Como $Y = e^X$ es siempre positiva, $F_Y(y) = 0$ para $y \leq 0$. Para $y > 0$, la función exponencial es estrictamente creciente, de modo que

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \leq y) = \mathbb{P}(e^X \leq y) \\ &= \mathbb{P}(X \leq \ln y) \\ &= \mathbb{P}\left(\frac{X - \mu}{\sigma} \leq \frac{\ln(y) - \mu}{\sigma}\right) = \Phi\left(\frac{\ln(y) - \mu}{\sigma}\right). \end{aligned}$$

Por lo tanto,

$$F_Y(y) = \begin{cases} 0, & y \leq 0, \\ \Phi\left(\frac{\ln(y) - \mu}{\sigma}\right), & y > 0. \end{cases}$$

Pregunta 91. (Examen - Otoño 2024, P1)

Parte A. En la orilla de una playa llegan olas cada 5 minutos. La altura de cada ola sigue una distribución exponencial de parámetro $1/30$ [cm⁻¹]. Las alturas de las olas son independientes entre sí. Justo después de que llega una ola, Daniela, una niña de 5 años, empieza a hacer una torre de arena en la orilla. Daniela avanza la altura de la torre a una velocidad de 20 [cm/5 minutos]. Cuando una ola impacta la torre, esta se derrumba si la ola posee una altura mayor que la de la torre, en cuyo caso Daniela empieza inmediatamente la construcción de una nueva torre desde cero. En caso contrario, Daniela continúa construyendo la torre.

Daniela juega 10 minutos. Nos referiremos a la *altura final* como la altura que alcanza la torre al cabo de los 10 minutos, justo después de que llega la ola del minuto 10. Defina A_i como el evento «la torre se ha derrumbado con la i -ésima ola», para $i \in \{1, 2\}$. La primera ola llega en el minuto 5 y la segunda en el minuto 10.

- En términos de los eventos definidos, escriba el evento correspondiente a que la altura final sea de 40 centímetros.
- Sea X la variable aleatoria asociada a la altura final de la torre. Determine el rango de X y justifique.
- Determine la PMF de X evaluada en 40. Puede ser útil usar la regla de la cadena y definir Y_i como la altura de la i -ésima ola.

Parte B. Responda lo siguiente:

- Sea X una variable aleatoria con distribución uniforme. ¿Qué parámetros se deben modificar para aumentar la varianza de X y cómo?
- Explique por qué $\Phi(-a) = 1 - \Phi(a)$, donde Φ es la CDF de una normal estándar.

Solución

(A)(a). La torre alcanza 40 cm exactamente cuando ninguna ola la derriba:

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

(A)(b). Después de la segunda ola, la torre puede quedar en 0 cm si esa ola la derriba, en 20 cm si solo la primera la derribó, o en 40 cm si ninguna la derribó. Por tanto,

$$\mathcal{R}_X = \{0, 20, 40\}.$$

(A)(c). Si Y_1, Y_2 son las alturas independientes de las olas, ambas $\text{Exp}(1/30)$, entonces

$$\begin{aligned} \mathbb{P}(X = 40) &= \mathbb{P}(A_1^c \cap A_2^c) \\ &= \mathbb{P}(A_1^c) \mathbb{P}(A_2^c \mid A_1^c) \\ &= \mathbb{P}(Y_1 \leq 20) \mathbb{P}(Y_2 \leq 40) \\ &= (1 - e^{-2/3})(1 - e^{-4/3}). \end{aligned}$$

(B)(a). Si $X \sim \text{Uniforme}(a, b)$, entonces

$$\text{Var}(X) = \frac{(b - a)^2}{12}.$$

La varianza aumenta al ampliar la longitud $b - a$ del intervalo, ya sea disminuyendo a , aumentando b , o mediante ambas modificaciones.

(B)(b). La densidad normal estándar es par. Con el cambio de variable $u = -x$,

$$\Phi(-a) = \int_{-\infty}^{-a} \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx = \int_a^{\infty} \frac{e^{-u^2/2}}{\sqrt{2\pi}} du = 1 - \Phi(a).$$

Pregunta 92. (Examen - Primavera 2024, P2)

Responda las siguientes preguntas.

- (a) Rafa Nadal lo ha contratado para analizar su saque, en particular, para calcular la probabilidad de que el saque sea bueno.

Consideraremos solo dos dimensiones, altura y profundidad, en el problema, y no tres como sería en la realidad. Además, por simplicidad, asumiremos que la altura inicial desde la cual la pelota es lanzada en el saque es cero. La velocidad inicial V de la pelota se modela como una variable aleatoria Uniforme $[0, w]$, donde w es un parámetro conocido, y con un ángulo Θ que distribuye Uniforme $[0, \pi/2]$.

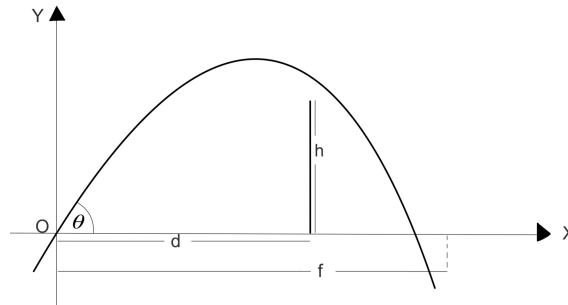
Un saque correcto se logra cuando se cumplen las siguientes condiciones: (i) la pelota pasa la malla por encima, y (ii) la pelota cae dentro del área delimitada pasada la malla. Considere que la malla posee una altura h y se encuentra a una distancia d del origen, ubicado en el punto inicial de la pelota, y que el área en la cual la pelota debe caer es el segmento $[d, f]$, como muestra la figura.

Para una realización $(V, \Theta) = (v, \theta)$, la parábola que describe la pelota es

$$y = \tan(\theta)x - \frac{g^2}{2v^2 \cos^2(\theta)}x^2,$$

donde y es la altura, x es la distancia horizontal de la pelota respecto del origen y g es la aceleración de gravedad, conocida. Note que una de las raíces de la parábola es $x = 0$. Calcule la probabilidad de que el saque de Rafa Nadal sea bueno.

Nota: No es necesario que compute el valor exacto; con dejar la expresión en función de parámetros conocidos es suficiente.



- (b) La fábrica de Willy Wonka produce barras de chocolate. El tiempo que tarda una máquina en producir una barra sigue una distribución uniforme con parámetros desconocidos. Cada día debe producir 50 barras, una a la vez. El encargado trabaja 480 minutos diarios y realiza horas extra si el tiempo total excede ese límite. Por el Teorema Central del Límite, se sabe que existe un 12 % de probabilidad de que esto ocurra. Además, la desviación estándar del tiempo de producción de una barra es de 5 minutos. Determine el tiempo promedio que tarda la máquina en producir una barra y compute su valor.

Solución

(a). Bajo la independencia implícita entre V y Θ , su densidad conjunta es $2/(\pi w)$ sobre $[0, w] \times [0, \pi/2]$. La pelota supera la malla cuando

$$h \leq d \tan \theta - \frac{g^2 d^2}{2v^2 \cos^2 \theta}.$$

La raíz positiva de la trayectoria es

$$x_* = \frac{2v^2 \cos \theta \sin \theta}{g^2},$$

por lo que además se requiere $d \leq x_* \leq f$. En consecuencia,

$$\mathbb{P}(\text{saque bueno}) = \int_0^{\pi/2} \int_0^w \mathbf{1}_{\left\{h \leq d \tan \theta - \frac{g^2 d^2}{2v^2 \cos^2 \theta}, \quad d \leq \frac{2v^2 \cos \theta \sin \theta}{g^2} \leq f\right\}} \frac{2}{\pi w} dv d\theta.$$

El punto $v = 0$, donde la expresión de la trayectoria no está definida, tiene medida cero.

(b). Sea X_i el tiempo por barra, con media μ y desviación estándar 5, y defina

$$Z = \frac{\sum_{i=1}^{50} X_i - 50\mu}{5\sqrt{50}}.$$

El TCL permite aproximar la distribución de Z por una $\mathcal{N}(0, 1)$. Como $\mathbb{P}(\sum_i X_i > 480) = 0,12$,

$$\Phi\left(\frac{480 - 50\mu}{5\sqrt{50}}\right) = 0,88.$$

Si $z_{0,88} = \Phi^{-1}(0,88) \approx 1,175$, entonces

$$\mu = \frac{480 - 5\sqrt{50} z_{0,88}}{50} \approx 8,77 \text{ minutos.}$$

Capítulo 7

Distribución Conjunta

7.1. Distribuciones conjuntas

Resumen

Distribución conjunta.

La CDF conjunta de X e Y es

$$F_{X,Y}(x, y) = \mathbb{P}(X \leq x, Y \leq y).$$

- **Monotonía.** Es no decreciente en cada coordenada.
- **Límites.** Converge a cero si alguna coordenada tiende a $-\infty$, y a uno si ambas tienden a $+\infty$.
- **Continuidad por la derecha.** Se cumple en cada coordenada.
- **Rectángulos.**

$$\begin{aligned} \mathbb{P}(a < X \leq b, c < Y \leq d) \\ &= F_{X,Y}(b, d) - F_{X,Y}(a, d) \\ &\quad - F_{X,Y}(b, c) + F_{X,Y}(a, c). \end{aligned}$$

PMF y PDF conjuntas.

- **Caso discreto.** $p_{X,Y}(x, y) = \mathbb{P}(X = x, Y = y)$, $p_{X,Y} \geq 0$ y $\sum_x \sum_y p_{X,Y}(x, y) = 1$. Para una región D ,

$$\mathbb{P}((X, Y) \in D) = \sum_{(x,y) \in D} p_{X,Y}(x, y).$$

- **Caso continuo.** $f_{X,Y}(x, y) \geq 0$ y $\iint_{\mathbb{R}^2} f_{X,Y}(x, y) dx dy = 1$. Además,

$$\mathbb{P}((X, Y) \in D) = \iint_D f_{X,Y}(x, y) dx dy.$$

Distribuciones marginales.

La ley de una componente se obtiene eliminando la otra:

$$p_X(x) = \sum_y p_{X,Y}(x, y),$$

$$f_X(x) = \int_{\mathbb{R}} f_{X,Y}(x, y) dy,$$

y análogamente para Y .

Distribuciones condicionales.

Cuando el denominador es positivo,

$$p_{X|Y}(x | y) = \frac{p_{X,Y}(x, y)}{p_Y(y)},$$

$$f_{X|Y}(x | y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}.$$

Para cada valor fijo de y , la ley condicional debe normalizarse sobre el soporte compatible con ese valor.

Independencia.

Las variables X e Y son independientes si y solo si su ley conjunta factoriza:

$$p_{X,Y}(x, y) = p_X(x)p_Y(y),$$

$$f_{X,Y}(x, y) = f_X(x)f_Y(y),$$

según corresponda, para todos los valores del soporte.

Esperanza conjunta.

Para una función g integrable,

$$\mathbb{E}[g(X, Y)] = \sum_x \sum_y g(x, y)p_{X,Y}(x, y)$$

en el caso discreto; en el continuo se reemplaza la suma doble por $\iint_{\mathbb{R}^2} g(x, y)f_{X,Y}(x, y) dx dy$.

Covarianza y correlación.

Cuando existen los segundos momentos y las varianzas son positivas,

$$\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y],$$

$$\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}.$$

Las propiedades de covarianza siguientes siguen siendo válidas cuando los momentos correspondientes existen:

- **Simetría.** $\text{Cov}(X, Y) = \text{Cov}(Y, X)$.
- **Transformación afín.** $\text{Cov}(aX + b, cY + d) = ac \text{Cov}(X, Y)$.

■ **Varianza de una suma.**

$$\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y).$$

- **Independencia.** Independencia implica covarianza nula; la recíproca no vale en general.

- **Cota de la correlación.** $-1 \leq \text{Corr}(X, Y) \leq 1$.

Cambio de variables.

Si $\mathbf{U} = g(\mathbf{X})$ es invertible y diferenciable,

$$f_{\mathbf{U}}(\mathbf{u}) = f_{\mathbf{X}}(g^{-1}(\mathbf{u})) |\det Dg^{-1}(\mathbf{u})|,$$

para $\mathbf{u}$ en el soporte transformado. El soporte debe transformarse junto con la densidad.

Multinomial.

Si $\mathbf{X} = (X_1, \dots, X_r)$ cuenta las categorías de n ensayos independientes con probabilidades $p_i \geq 0$ tales que $\sum_{i=1}^r p_i = 1$, entonces, para $\sum_{i=1}^r x_i = n$,

$$\mathbb{P}(\mathbf{X} = \mathbf{x}) = \frac{n!}{\prod_{i=1}^r x_i!} \prod_{i=1}^r p_i^{x_i},$$

$$\mathbb{E}[X_i] = np_i,$$

$$\text{Var}(X_i) = np_i(1 - p_i),$$

$$\text{Cov}(X_i, X_j) = -np_i p_j, \quad i \neq j.$$

Normal multivariada.

Si $\mathbf{X} \sim \mathcal{N}_d(\boldsymbol{\mu}, \Sigma)$, entonces $\mathbb{E}[\mathbf{X}] = \boldsymbol{\mu}$ y $\text{Cov}(\mathbf{X}) = \Sigma$, donde Σ es simétrica y semidefinida positiva. Para una matriz A y un vector $\mathbf{b}$,

$$A\mathbf{X} + \mathbf{b} \sim \mathcal{N}(A\boldsymbol{\mu} + \mathbf{b}, A\Sigma A^T).$$

En particular,

$$\mathbf{a}^T \mathbf{X} \sim \mathcal{N}(\mathbf{a}^T \boldsymbol{\mu}, \mathbf{a}^T \Sigma \mathbf{a}).$$

En un vector conjuntamente normal, covarianza nula equivale a independencia entre las componentes correspondientes; esta equivalencia no es válida para distribuciones arbitrarias.

Estadísticos de orden.

Para una muestra continua i.i.d. $X_1, \dots, X_n$ con CDF F_X y PDF f_X , se escribe $X_{(1)} \leq \dots \leq X_{(n)}$. Entonces

$$F_{X_{(1)}}(x) = 1 - [1 - F_X(x)]^n,$$

$$F_{X_{(n)}}(x) = F_X(x)^n,$$

y para $k = 1, \dots, n$,

$$f_{X_{(k)}}(x) = \frac{n!}{(k-1)!(n-k)!} \times F_X(x)^{k-1} [1 - F_X(x)]^{n-k} f_X(x).$$

7.1.1. Preguntas

Pregunta 93.

Sean X e Y dos variables aleatorias con la siguiente PMF conjunta:

$$p_{X,Y}(x, y) = \begin{cases} c(x^2 + y^2) & \text{si } x \in \{1, 2, 4\} \text{ e } y \in \{1, 3\} \\ 0 & \text{si no} \end{cases}$$

- ¿Cuál es el valor de c ?
- ¿Cuánto es $\mathbb{P}(Y < X)$?
- ¿Cuánto es $\mathbb{P}(Y > X)$?
- ¿Cuánto es $\mathbb{P}(Y = X)$?
- ¿Cuánto es $\mathbb{P}(Y = 3)$?
- Encuentre la PMF marginal de X e Y .
- ¿Cuánto es $\mathbb{P}(X = -Y)$?
- ¿Son X e Y variables independientes?

Solución

Los seis puntos del soporte tienen pesos proporcionales a 2, 10, 5, 13, 17 y 25, cuya suma es 72.

(a). La normalización exige $72c = 1$, de modo que $c = 1/72$.

(b). Los pares con $Y < X$ son $(2, 1)$, $(4, 1)$ y $(4, 3)$; por ello,

$$\mathbb{P}(Y < X) = \frac{5 + 17 + 25}{72} = \frac{47}{72}.$$

(c). Los pares con $Y > X$ son $(1, 3)$ y $(2, 3)$; por ello,

$$\mathbb{P}(Y > X) = \frac{10 + 13}{72} = \frac{23}{72}.$$

(d). El único par con $Y = X$ es $(1, 1)$, de modo que $\mathbb{P}(Y = X) = 2/72 = 1/36$.

(e). Sumando sobre los puntos con segunda coordenada 3,

$$\mathbb{P}(Y = 3) = \frac{10 + 13 + 25}{72} = \frac{2}{3}.$$

(f). Las funciones de masa marginales son

$$p_X(x) = \begin{cases} 1/6, & x = 1, \\ 1/4, & x = 2, \\ 7/12, & x = 4, \\ 0, & \text{en otro caso,} \end{cases} \quad p_Y(y) = \begin{cases} 1/3, & y = 1, \\ 2/3, & y = 3, \\ 0, & \text{en otro caso.} \end{cases}$$

(g). Como X e Y son positivas en todo el soporte, $\mathbb{P}(X = -Y) = 0$.

(h). No son independientes. Por ejemplo,

$$p_{X,Y}(1, 1) = \frac{2}{72} \neq \frac{1}{6} \frac{1}{3} = p_X(1)p_Y(1).$$

Pregunta 94.

Un grupo de inversionistas realizó un estudio en el cual se dieron cuenta que la mayoría de la comida que se vende alrededor de la facultad corresponde a productos salados. Con esta información encontraron una oportunidad de negocio, por lo que decidieron instalar un local que vendiera productos dulces llamado *Pastelería la Mona*. En esta pastelería se venden diversos postres, dentro de los cuales su producto estrella son los alfajores.

La pastelería fue un éxito en sus primeros meses de operación. En el reporte trimestral se presentaron los siguientes datos: a la pastelería llegan X personas por día, esta cantidad sigue una distribución Poisson de tasa 10 [clientes/día]. De los clientes que llegan a la tienda, cada uno compra un único postre independiente de lo que vayan a comprar los demás, en particular, cada cliente que llega a la tienda va a comprar un alfajor con probabilidad 0,3, y con probabilidad complementaria compra uno de los otros postres.

1. Identifique y calcule la distribución conjunta de la cantidad de personas que llega a la tienda en un día y la cantidad de personas que compra alfajores (en el mismo día).
2. Calcule la distribución de la cantidad de personas que va a comprar alfajores a la pastelería.
3. Si en un día llegan 4 personas a comprar alfajores a la tienda, ¿Cuál es la probabilidad de que ese mismo día hayan ido entre 5 y 10 personas a la tienda?

Solución

(a). Condicionalmente a $X = x$, $Y \mid X = x \sim \text{Bin}(x, 0,3)$. Para enteros $x \geq 0$ y $0 \leq y \leq x$,

$$\begin{aligned} p_{X,Y}(x, y) &= p_{Y \mid X}(y \mid x) p_X(x) \\ &= \binom{x}{y} (0,3)^y (0,7)^{x-y} \frac{e^{-10} 10^x}{x!} \\ &= \frac{e^{-10} 3^y 7^{x-y}}{y!(x-y)!}, \end{aligned}$$

y la probabilidad es cero para los demás pares.

(b). Para todo entero $y \geq 0$,

$$\begin{aligned} p_Y(y) &= \sum_{x=y}^{\infty} p_{X,Y}(x, y) = \frac{e^{-10} 3^y}{y!} \sum_{x=y}^{\infty} \frac{7^{x-y}}{(x-y)!} \\ &= \frac{e^{-3} 3^y}{y!}. \end{aligned}$$

Por lo tanto, $Y \sim \text{Poi}(3)$.

(c). Para enteros $x \geq y$,

$$p_{X \mid Y}(x \mid y) = \frac{e^{-7} 7^{x-y}}{(x-y)!}.$$

Dado $Y = 4$, la variable $X - 4$ distribuye Poisson de parámetro 7. Luego,

$$\mathbb{P}(5 \leq X \leq 10 \mid Y = 4) = \sum_{j=1}^6 \frac{e^{-7} 7^j}{j!} \approx 0,4488.$$

Pregunta 95.

Suponga que las variables aleatorias X e Y tienen distribución conjunta

$$f_{X,Y}(x, y) = \begin{cases} \frac{3}{2}(x^2 + y^2) & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & \text{en otro caso} \end{cases}$$

- Encuentre las densidades marginales de X e Y .
- Encuentre las densidades condicionales $f_{X \mid Y}(x \mid y)$ y $f_{Y \mid X}(y \mid x)$.
- Encuentre $\mathbb{E}[X \mid Y = y]$

Solución

(a). Las densidades marginales son

$$\begin{aligned} f_Y(y) &= \int_0^1 \frac{3}{2}(x^2 + y^2) dx = \frac{1 + 3y^2}{2}, & 0 \leq y \leq 1, \\ f_X(x) &= \int_0^1 \frac{3}{2}(x^2 + y^2) dy = \frac{1 + 3x^2}{2}, & 0 \leq x \leq 1. \end{aligned}$$

Fuera de esos intervalos, las densidades son cero.

(b). Al dividir la densidad conjunta por la marginal correspondiente,

$$\begin{aligned} f_{X \mid Y}(x \mid y) &= \frac{3(x^2 + y^2)}{1 + 3y^2}, & 0 \leq x \leq 1, \\ f_{Y \mid X}(y \mid x) &= \frac{3(x^2 + y^2)}{1 + 3x^2}, & 0 \leq y \leq 1. \end{aligned}$$

(c). Para $0 \leq y \leq 1$,

$$\begin{aligned}\mathbb{E}[X | Y = y] &= \int_0^1 x f_{X|Y}(x | y) dx \\ &= \frac{3}{1 + 3y^2} \left(\frac{1}{4} + \frac{y^2}{2} \right) = \frac{3(1 + 2y^2)}{4(1 + 3y^2)}.\end{aligned}$$

Pregunta 96.

Sean X y Y dos variables aleatorias **conjuntamente** normales tales que $X \sim \mathcal{N}(-1, 4)$, $Y \sim \mathcal{N}(1, 1)$ y $\text{Corr}(X, Y) = -0.5$.

- Calcule $\mathbb{P}(X + 2Y \leq 3)$.
- Calcule $\text{Cov}(X - Y, X + 2Y)$.

Solución

Como $\text{Corr}(X, Y) = -0.5$,

$$\text{Cov}(X, Y) = -0.5 \sqrt{\text{Var}(X) \text{Var}(Y)} = -1.$$

(a). Toda combinación lineal de un vector normal conjunto es normal. En particular,

$$\mathbb{E}[X + 2Y] = -1 + 2 = 1,$$

$$\text{Var}(X + 2Y) = 4 + 4 \text{Var}(Y) + 4 \text{Cov}(X, Y) = 4.$$

Por lo tanto, $X + 2Y \sim \mathcal{N}(1, 4)$ y

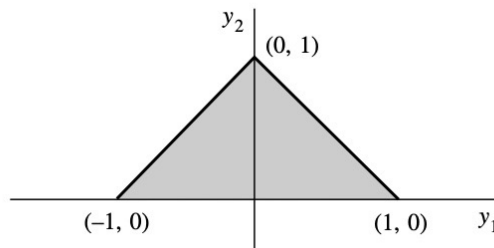
$$\mathbb{P}(X + 2Y \leq 3) = \Phi\left(\frac{3 - 1}{2}\right) = \Phi(1) \approx 0.8413.$$

(b). Por bilinealidad de la covarianza,

$$\begin{aligned}\text{Cov}(X - Y, X + 2Y) &= \text{Var}(X) + \text{Cov}(X, Y) - 2 \text{Var}(Y) \\ &= 4 - 1 - 2 = 1.\end{aligned}$$

Pregunta 97.

Suponga que el vector aleatorio $Y = (Y_1, Y_2)$ es un punto escogido al azar sobre el triángulo sombreado del diagrama siguiente.



- Obtenga la función de densidad de Y .
- Calcule las densidades marginales de Y_1, Y_2 y averigüe si estas v.a. son independientes.
- Calcule $\mathbb{P}(2Y_1 > Y_2), \mathbb{P}(Y_1 + Y_2 < \frac{1}{2})$.

Solución

Sea

$$T = \{(y_1, y_2) : -1 \leq y_1 \leq 1, 0 \leq y_2 \leq 1 - |y_1|\}.$$

El área de T es 1. **(a).** Como el punto se escoge uniformemente,

$$f_{Y_1, Y_2}(y_1, y_2) = \mathbb{1}_T(y_1, y_2).$$

(b). Las densidades marginales son

$$f_{Y_1}(y_1) = (1 - |y_1|)\mathbb{1}_{[-1,1]}(y_1), \quad f_{Y_2}(y_2) = 2(1 - y_2)\mathbb{1}_{[0,1]}(y_2).$$

El soporte T no es un producto cartesiano y la densidad conjunta no es el producto de las marginales; por tanto, Y_1 e Y_2 no son independientes.

(c). Como la densidad conjunta vale 1 en T , basta integrar las alturas de las regiones. Las rectas $y_2 = 2y_1$ y $y_2 = 1 - y_1$ se intersectan en $y_1 = 1/3$, por lo que

$$\mathbb{P}(2Y_1 > Y_2) = \int_0^{1/3} 2y_1 dy_1 + \int_{1/3}^1 (1 - y_1) dy_1 = \frac{1}{3}.$$

Para el segundo evento, la frontera $y_2 = 1/2 - y_1$ comienza a cortar la frontera izquierda del triángulo en $y_1 = -1/4$. Así,

$$\mathbb{P}(Y_1 + Y_2 < 1/2) = \int_{-1}^{-1/4} (1 + y_1) dy_1 + \int_{-1/4}^{1/2} \left(\frac{1}{2} - y_1\right) dy_1 = \frac{9}{16}.$$

Pregunta 98.

Los auxiliares y los profesores han quedado de encontrarse en la U para preparar una clase, entre las 7 y las 8 PM. Además, han convenido que ninguno esperará más de 15 minutos al otro. Suponga que ambos llegan independientemente y al azar entre las 7 y las 8 PM. **Indicación:** para modelar los tiempos de llegada defina el origen (0) como las 7 PM, además los profesores cuentan como un grupo y llegan juntos, mientras que los auxiliares cuentan como otro grupo que también llegan juntos.

- (a) Calcule la probabilidad de que se encuentren.
- (b) Los profesores proponen a los auxiliares de ampliar el intervalo de llegada con el propósito de aumentar la probabilidad de encontrarse (con la misma regla de no esperar más de 15 minutos). Entonces ahora pueden llegar al azar entre las 7 y las 9 PM. ¿Cree usted que esto es una buena idea? Fundamente su respuesta calculando la probabilidad de encontrarse en este caso.

Solución

Sean X e Y las horas de llegada, medidas desde las 19:00. Se encuentran si $|X - Y| \leq 1/4$. **(a).** En el intervalo de una hora, (X, Y) es uniforme en $[0, 1]^2$. La región favorable es la banda entre $y = x - 1/4$ e $y = x + 1/4$, cuya área es

$$1 - 2\frac{1}{2}\left(\frac{3}{4}\right)^2 = \frac{7}{16}.$$

(b). En el intervalo de dos horas, (X, Y) es uniforme en $[0, 2]^2$. El área favorable es $4 - (7/4)^2 = 15/16$ y el área total es 4. Por tanto,

$$\mathbb{P}(|X - Y| \leq 1/4) = \frac{15}{64}.$$

La ampliación disminuye la probabilidad de encuentro y no es conveniente.

Pregunta 99.

Se dispone de $N = 3$ cajas numeradas, dentro de las cuales hay una cantidad de bolitas dada por el número de la caja, es decir, la k -ésima caja tiene k bolitas numeradas del 1 al k . El experimento consiste en elegir una caja al azar y luego una bolita al azar dentro de esa caja. Sea X la v.a. correspondiente al número de la caja elegida y sea Y la v.a. correspondiente al número de la bolita elegida.

- (a) Escriba la función de probabilidad condicional $p_{Y|X}(y | x)$ y marginal $p_X(x)$.
- (b) Determine la función de probabilidad conjunta $p_{X,Y}(x, y)$, la condicional $p_{X|Y}(x | y)$ y la marginal $p_Y(y)$.
- (c) Indique si X y Y son v.a. independientes, justificando su respuesta.
- (d) Calcule $\mathbb{E}[Y | X = x]$, $\mathbb{E}[X]$, $\mathbb{E}[Y]$.

Solución

(a). La caja se escoge uniformemente y, una vez fijada, también se escoge uniformemente una de sus bolitas. Por ello,

$$p_X(x) = \begin{cases} \frac{1}{3}, & x \in \{1, 2, 3\}, \\ 0, & \text{en otro caso,} \end{cases} \quad p_{Y|X}(y | x) = \begin{cases} \frac{1}{x}, & y \in \{1, \dots, x\}, \\ 0, & \text{en otro caso.} \end{cases} \quad x \in \{1, 2, 3\}.$$

(b). La regla de la cadena entrega

$$p_{X,Y}(x, y) = \begin{cases} \frac{1}{3x}, & x \in \{1, 2, 3\}, y \in \{1, \dots, x\}, \\ 0, & \text{en otro caso.} \end{cases}$$

Además,

$$p_Y(1) = \frac{11}{18}, \quad p_Y(2) = \frac{5}{18}, \quad p_Y(3) = \frac{1}{9}, \quad p_Y(y) = 0 \text{ si } y \notin \{1, 2, 3\}.$$

Para $y \in \{1, 2, 3\}$, el Teorema de Bayes da

$$p_{X|Y}(x | y) = \frac{p_{X,Y}(x, y)}{p_Y(y)}.$$

Por ejemplo, $p_{X|Y}(1 | 1) = (1/3)/(11/18) = 6/11$. La matriz de $p_{X|Y}(x | y)$, con filas $x = 1, 2, 3$ y columnas $y = 1, 2, 3$, es

$$\begin{pmatrix} 6/11 & 0 & 0 \\ 3/11 & 3/5 & 0 \\ 2/11 & 2/5 & 1 \end{pmatrix}.$$

(c). No son independientes; por ejemplo,

$$p_{X,Y}(1, 1) = \frac{1}{3} \neq \frac{1}{3} \cdot \frac{11}{18} = p_X(1)p_Y(1).$$

(d). Como $Y | X = x$ es uniforme en $\{1, \dots, x\}$,

$$\mathbb{E}[Y | X = x] = \frac{x+1}{2}.$$

Finalmente, $\mathbb{E}[X] = 2$ y $\mathbb{E}[Y] = 3/2$.

Pregunta 100.

Sea (X, Y) un punto escogido al azar en la región (no acotada) del plano, dada por:

$$R = \{(x, y) \in \mathbb{R}^2 : 0 \leq x < +\infty, 0 \leq y \leq e^{-x}\}$$

Es decir, (X, Y) tiene densidad constante $f_{X,Y}(x, y) = C$ para $(x, y) \in R$ y $f_{X,Y}(x, y) = 0$ para $(x, y) \notin R$.

- Calcule C y obtenga las densidades marginales de X e Y y averigüe si son independientes.
- Determine las densidades y esperanzas condicionales $f_{Y|X}(y | x)$, $f_{X|Y}(x | y)$, $\mathbb{E}[Y | X = x]$.
- Calcule $\mathbb{P}(X + Y \leq 1)$ y $\mathbb{E}[XY]$.

Solución

(a). La condición de normalización exige

$$1 = C \int_0^\infty \int_0^{e^{-x}} dy dx = C \int_0^\infty e^{-x} dx = C,$$

de modo que $C = 1$. Las marginales son

$$f_X(x) = \begin{cases} e^{-x}, & x \geq 0, \\ 0, & \text{en otro caso,} \end{cases} \quad f_Y(y) = \begin{cases} -\ln y, & 0 < y \leq 1, \\ 0, & \text{en otro caso.} \end{cases}$$

Las variables no son independientes, pues el soporte conjunto $\{(x, y) : x \geq 0, 0 \leq y \leq e^{-x}\}$ no es un producto cartesiano.

(b). Las densidades condicionales son

$$f_{Y|X}(y | x) = \begin{cases} e^x, & 0 \leq y \leq e^{-x}, \\ 0, & \text{en otro caso,} \end{cases}$$

y, para $0 < y < 1$,

$$f_{X|Y}(x | y) = \begin{cases} \frac{1}{-\ln y}, & 0 \leq x \leq -\ln y, \\ 0, & \text{en otro caso.} \end{cases}$$

Por consiguiente, para $x \geq 0$ y $0 < y < 1$,

$$\mathbb{E}[Y | X = x] = \frac{e^{-x}}{2}, \quad \mathbb{E}[X | Y = y] = \frac{-\ln y}{2}.$$

(c). Integrando sobre $\{(x, y) : 0 \leq x \leq 1, 0 \leq y \leq \min(e^{-x}, 1 - x)\}$ y observando que $1 - x \leq e^{-x}$,

$$\mathbb{P}(X + Y \leq 1) = \int_0^1 (1 - x) dx = \frac{1}{2}.$$

Para el segundo cálculo,

$$\mathbb{E}[XY] = \int_0^\infty \int_0^{e^{-x}} xy dy dx = \frac{1}{2} \int_0^\infty xe^{-2x} dx = \frac{1}{8}.$$

Pregunta 101.

Para cierto producto hay distintos tipos de impurezas. Solo es posible medir la fracción de impurezas en todo el producto (dada por la v.a. X), y medir la fracción de impurezas dañinas dentro de todas las impurezas detectadas (dada por la v.a. Y). Suponga que la distribución conjunta de X y Y es:

$$f_{X,Y}(x, y) = \begin{cases} 2(1-x) & \text{para } 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & \text{en cualquier otro caso} \end{cases}$$

- (a) Sea U la v.a. que representa la fracción de impurezas dañinas en todo el producto. Encuentre el valor esperado de U .
- (b) Usando el método del Jacobiano, encuentre la densidad conjunta entre X y U .
- (c) Encuentre $\text{Cov}(X, Y)$ y $\text{Cov}(X, U)$. Comente respecto a la independencia de estas v.a.

Solución

El soporte conjunto es el cuadrado $[0, 1]^2$. Al integrar,

$$f_X(x) = 2(1-x), \quad 0 \leq x \leq 1, \quad f_Y(y) = 1, \quad 0 \leq y \leq 1.$$

Como $f_{X,Y}(x, y) = f_X(x)f_Y(y)$ en ese cuadrado, X e Y son independientes. Además, $\mathbb{E}[X] = 1/3$, $\mathbb{E}[X^2] = 1/6$ y $\mathbb{E}[Y] = 1/2$. **(a).** Como $U = XY$, la independencia permite escribir

$$\mathbb{E}[U] = \mathbb{E}[X]\mathbb{E}[Y] = \frac{1}{3} \frac{1}{2} = \frac{1}{6}.$$

(b). La transformación $(x, y) \mapsto (x, u) = (x, xy)$ tiene inversa $y = u/x$ y jacobiano absoluto $1/x$. El soporte transformado satisface $0 < x \leq 1$ y $0 \leq u \leq x$; por tanto,

$$f_{X,U}(x, u) = \begin{cases} \frac{2(1-x)}{x}, & 0 < x \leq 1, 0 \leq u \leq x, \\ 0, & \text{en otro caso.} \end{cases}$$

(c). La independencia de X e Y implica $\text{Cov}(X, Y) = 0$. Por otra parte,

$$\begin{aligned} \mathbb{E}[XU] &= \mathbb{E}[X^2Y] = \mathbb{E}[X^2]\mathbb{E}[Y] = \frac{1}{12}, \\ \text{Cov}(X, U) &= \mathbb{E}[XU] - \mathbb{E}[X]\mathbb{E}[U] = \frac{1}{12} - \frac{1}{18} = \frac{1}{36}. \end{aligned}$$

Así, X e Y son independientes, mientras que X y U no lo son, pues una covarianza no nula descarta independencia.

7.1.2. Preguntas de Evaluación**Pregunta 102.** (Control 2 (recuperativo) - Otoño 2024, P3)

- (a) Considere las variables aleatorias X e Y con PMF conjunta como muestra la tabla.

	$Y = 1$	$Y = 2$
$X = 1$	$\frac{1}{3}$	$\frac{1}{12}$
$X = 2$	$\frac{1}{6}$	0
$X = 4$	$\frac{1}{12}$	$\frac{1}{3}$

Sea $Z = X - 2Y$. Calcule $\mathbb{P}(Z \mid X = 2)$.

- (b) Pedrito disfruta hacer castillos de arena en la playa cerca de donde llegan las olas del mar. Pedrito construye castillos con una velocidad de 20 [cm/5 minutos]. Los castillos reciben olas cada 5 minutos, es decir, la primera ola la recibe en el minuto 5, la siguiente ola en el minuto 10, y así sucesivamente. Al recibir una ola el castillo, este se derrumba si es que dicha ola posea una altura mayor o igual a la del castillo en ese momento; en el caso que la ola posea una altura menor al castillo, el castillo queda intacto. Considere que la altura de una ola sigue una distribución Exponencial de parámetro $1/30$ [cm⁻¹]. El tamaño de las olas son independientes entre sí.

Pedrito construye un castillo nuevo cada 5 minutos. Cada vez que llega una ola al castillo, Pedrito se toma una foto con su castillo en caso de que dicha ola no haya derribado el castillo. Luego, inmediatamente después, Pedrito inicia la construcción de un nuevo castillo en los siguientes 5 minutos. Considere que el tiempo que le toma a Pedrito tomarse la foto es despreciable. Pedrito ha jugado 60 minutos en la playa. Determine la distribución de probabilidad del número de fotos que Pedrito se ha tomado durante este período (incluyendo la posible foto con el castillo que se termina en el minuto 60 luego de que llega la ola en este instante). Si la distribución es conocida, es decir, vista en clases, entregue el nombre y valor de sus parámetros, en caso contrario, entregue su PMF. Justifique su respuesta.

Solución

- (a). Interpretamos lo solicitado como la distribución condicional de Z dado $X = 2$. De la tabla,

$$\mathbb{P}(X = 2) = \mathbb{P}(X = 2, Y = 1) + \mathbb{P}(X = 2, Y = 2) = \frac{1}{6}.$$

Condicionado a $X = 2$, la única posibilidad con probabilidad positiva es $Y = 1$. En ese caso,

$$Z = X - 2Y = 2 - 2 = 0.$$

Por tanto,

$$p_{Z|X}(z | 2) = \begin{cases} 1, & z = 0, \\ 0, & z \neq 0. \end{cases}$$

En particular, $\mathbb{P}(Z = 0 | X = 2) = 1$.

- (b). Cada castillo alcanza una altura de 20 cm antes de recibir su ola. Si H denota la altura de una ola, entonces

$$H \sim \text{Exponencial} \left(\frac{1}{30} \right).$$

Pedrito se toma una foto cuando la ola tiene altura menor que 20 cm. La probabilidad de este evento es

$$\begin{aligned} p &= \mathbb{P}(H < 20) = F_H(20) \\ &= 1 - e^{-(1/30)20} = 1 - e^{-2/3}. \end{aligned}$$

Durante 60 minutos se construyen $60/5 = 12$ castillos. Como las alturas de las olas son independientes, los eventos de obtener una foto también lo son. Si N es el número de fotos,

$$N \sim \text{Binomial} \left(12, 1 - e^{-2/3} \right).$$

Su función de masa es

$$p_N(k) = \binom{12}{k} \left(1 - e^{-2/3} \right)^k \left(e^{-2/3} \right)^{12-k}, \quad k = 0, 1, \dots, 12.$$

Pregunta 103. (Control 2 - Primavera 2024, P3)

Se sabe que una caja de 5 transistores contiene 2 que están defectuosos. (Los transistores están separados en la caja, no hay conexiones). Los transistores deben ser probados, uno a la vez, hasta que se identifiquen las unidades defectuosas. Se define N_1 como el número de pruebas realizadas hasta que el primer transistor defectuoso se identifica, y por N_2 el número ensayos adicionales hasta que se identifique el segundo transistor defectuoso. Por ejemplo, si es que el primer transistor revisado está bueno, el segundo está malo, y luego el tercero está malo, entonces $N_1 = 2$ y $N_2 = 1$.

Para todas las partes, puede usar las respuestas de las partes anteriores como una variable (aunque no haya resuelto las partes anteriores). En dicho caso, explique claramente la variable que está usando para representar la respuesta de alguna parte anterior.

Para las partes b), c), y d), escriba como **parte** de su desarrollo la probabilidad solicitada en función de las variables aleatorias N_1 y N_2 .

- (a) Argumente la veracidad de la siguiente afirmación: “ N_1 sigue una distribución Geométrica”. Tanto si cree que es verdadera o falsa, justifique su respuesta.
- (b) Determine la probabilidad de que en la segunda prueba realizada, se encuentra el primer transistor defectuoso.
- (c) Se sabe que en la segunda prueba realizada se detectó el primer transistor defectuoso. Calcule la probabilidad de que la prueba siguiente (es decir, en la tercera prueba) también se detecte un transistor defectuoso.
- (d) Encontrar la probabilidad de que: “ N_1 valga 2 y N_2 valga 1”. (Notar que se pide una sola probabilidad).

Solución

Los cálculos requieren suponer que el orden de prueba es uniforme respecto de cuáles son los dos transistores defectuosos; equivalentemente, las $\binom{5}{2}$ ubicaciones posibles de estos transistores son equiprobables. Sin este mecanismo aleatorio, saber únicamente que hay dos unidades defectuosas no determina las probabilidades pedidas.

(a). La afirmación es falsa. Como hay solamente tres transistores buenos, el primer transistor defectuoso debe aparecer a más tardar en la cuarta prueba. Por tanto,

$$R_{N_1} = \{1, 2, 3, 4\}.$$

Una variable geométrica tiene rango infinito y surge de ensayos con reemplazo o con una probabilidad de éxito constante. Aquí las pruebas se realizan sin reemplazo y la probabilidad de encontrar un transistor defectuoso cambia después de cada resultado.

(b). El evento $\{N_1 = 2\}$ requiere que el primer transistor sea bueno y que el segundo sea defectuoso. Al muestrear sin reemplazo,

$$\mathbb{P}(N_1 = 2) = \frac{3}{5} \frac{2}{4} = \frac{3}{10}.$$

(c). Condicionado a $N_1 = 2$, entre los primeros dos transistores apareció uno bueno y uno defectuoso. Quedan tres transistores: dos buenos y uno defectuoso. Que el siguiente transistor sea defectuoso equivale a $\{N_2 = 1\}$. Por tanto,

$$\mathbb{P}(N_2 = 1 \mid N_1 = 2) = \frac{1}{3}.$$

(d). La probabilidad conjunta se descompone mediante la regla de la cadena:

$$\begin{aligned}\mathbb{P}(N_1 = 2, N_2 = 1) &= \mathbb{P}(N_2 = 1 \mid N_1 = 2)\mathbb{P}(N_1 = 2) \\ &= \frac{1}{3} \frac{3}{10} = \frac{1}{10}.\end{aligned}$$

Pregunta 104. (Control 3 - Primavera 2022, P1)

Usted ha sido designado para conducir un estudio sobre qué variables son las más relevantes para aprobar un ramo en la FCFM. Luego de un arduo estudio, ha descubierto que la variable más relevante para explicar las notas (representadas por la variable aleatoria Y) de los estudiantes son las horas de estudio semanales (representadas por X) que le dedican al curso. En su estudio ha descubierto que la PDF de X y de $Y \mid X$ son las siguientes:

$$f_X(x) = \begin{cases} \frac{3x(x+5)}{1750}, & 0 \leq x \leq 10, \\ 0, & \text{en cualquier otro caso,} \end{cases}$$

y

$$f_{Y|X}(y \mid x) = \begin{cases} \frac{5 \left(\frac{x}{10} + \frac{y-1}{6} \right)}{3(x+5)}, & 0 \leq x \leq 10, \quad 1 \leq y \leq 7, \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

- ¿Cuál es la probabilidad de que un estudiante apruebe el curso, es decir, que consiga una nota mayor o igual a 4?
- Es de interés para usted saber la distribución de las horas de estudio dadas las notas. Encuentre la densidad condicional de X dado Y .
- También quiere descubrir qué tan relacionadas están las dos variables aleatorias X e Y . Encuentre la covarianza entre las horas de estudio y las notas.

Solución

(a). La densidad conjunta se obtiene mediante la regla del producto:

$$f_{X,Y}(x, y) = f_{Y|X}(y \mid x)f_X(x) = \frac{x}{350} \left(\frac{x}{10} + \frac{y-1}{6} \right), \quad 0 \leq x \leq 10, \quad 1 \leq y \leq 7.$$

Al integrar respecto de x ,

$$f_Y(y) = \int_0^{10} f_{X,Y}(x, y) dx = \frac{y+3}{42}, \quad 1 \leq y \leq 7.$$

Por tanto,

$$\mathbb{P}(Y \geq 4) = 1 - \int_1^4 \frac{y+3}{42} dy = 1 - \frac{11}{28} = \frac{17}{28}.$$

(b). Para $0 \leq x \leq 10$ y $1 \leq y \leq 7$, el cociente entre la densidad conjunta y la marginal de Y entrega

$$f_{X|Y}(x \mid y) = \frac{f_{X,Y}(x, y)}{f_Y(y)} = \frac{3x}{25(y+3)} \left(\frac{x}{10} + \frac{y-1}{6} \right).$$

La densidad es cero fuera del soporte de X .

(c). A partir de las densidades anteriores,

$$\mathbb{E}[X] = \frac{50}{7}, \quad \mathbb{E}[Y] = \frac{31}{7}, \quad \mathbb{E}[XY] = \int_1^7 \int_0^{10} xy f_{X,Y}(x, y) dx dy = \frac{220}{7}.$$

En consecuencia,

$$\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \frac{220}{7} - \frac{50}{7} \frac{31}{7} = -\frac{10}{49}.$$

Pregunta 105. (Control 3 - Otoño 2023, P1)

Don Cristóbal tiene un campo donde cosecha trigo. La cantidad de trigo que se puede cosechar es variable cada año. Dicha cantidad, depende en cierta medida de la cantidad de lluvia que haya en la temporada. Un practicante de don Cristóbal determinó que los milímetros de lluvia que caen durante la temporada en el sector donde está sembrado el trigo tiene una distribución Uniforme entre 0 y 450. Por otra parte, la cantidad de toneladas de trigo que se cosecha en una temporada en la que hay X milímetros de lluvia, distribuye Uniforme entre 0 y $X \cdot (600 - X)/450$. En las siguientes preguntas, cuando se remita a “dos VA”, se hace referencia al par “milímetros de lluvia en una temporada” y “toneladas de trigo cosechados en la misma temporada”.

- Grafique el rango conjunto de las VA descritas. Indique los valores en los ejes (abscisa y ordenada) de los límites de dicha región.
- Encuentre la PDF conjunta de las dos VA.
- Comente sobre la veracidad de la siguiente afirmación “La distribución conjunta es uniforme en el Rango conjunto”. Justifique clara y brevemente su respuesta.
- Considere que el precio de mercado de un campo es igual a 700 millones de pesos por la cantidad de toneladas promedio de trigo que produce por temporada. Don Cristóbal no sabe cuánto es dicho precio, pero luego de breve un razonamiento afirma “Dado que la producción del campo está entre 0 y z , (donde z es la máxima producción posible de trigo del campo), he decidido poner un precio de venta al campo igual a $700 \cdot z/2$ millones de pesos”. (Note que usted ya debería haber obtenido el valor z en la parte (a).) Doña Agustina tiene intenciones de comprar el campo de don Cristóbal, sin embargo, ella no comprará el campo si es que el precio de este está por sobre el precio de mercado. Determine si Doña Agustina debería o no comprar el campo. Use la Ley de las esperanzas Iteradas para responder la pregunta.

Solución

(a). Sea X la lluvia, medida en milímetros, y sea Y la producción, medida en toneladas. Condicionado en $X = x$, la producción varía entre cero y

$$h(x) = \frac{x(600 - x)}{450}.$$

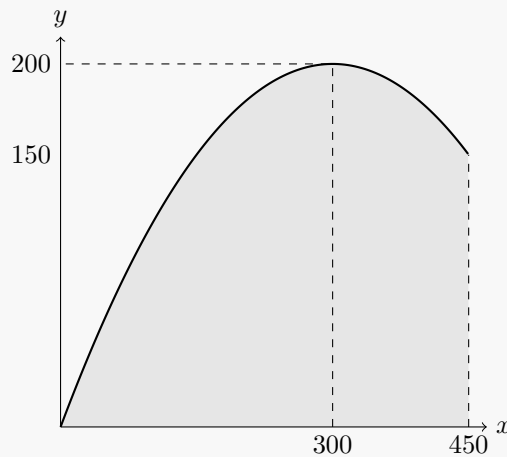
Por tanto, el soporte conjunto es

$$R_{X,Y} = \left\{ (x, y) : 0 \leq x \leq 450, 0 \leq y \leq \frac{x(600 - x)}{450} \right\}.$$

La frontera superior es una parábola cóncava. Su máximo se alcanza en $x = 300$ y vale

$$z = h(300) = 200.$$

El soporte es la región sombreada bajo la parábola:



En particular, $h(0) = 0$, $h(450) = 150$ y el máximo es $h(300) = 200$.

(b). La densidad de X es $f_X(x) = 1/450$ para $0 \leq x \leq 450$. Para $x > 0$, la distribución condicional de Y es uniforme en $[0, h(x)]$, de modo que

$$f_{Y|X}(y | x) = \frac{450}{x(600 - x)} \mathbf{1}_{\{0 \leq y \leq h(x)\}}.$$

La regla de la cadena entrega

$$f_{X,Y}(x, y) = \begin{cases} \frac{1}{x(600 - x)}, & 0 < x \leq 450, 0 \leq y \leq \frac{x(600 - x)}{450}, \\ 0, & \text{en otro caso.} \end{cases}$$

El valor asignado en la frontera $x = 0$ no afecta la distribución.

(c). La distribución conjunta no es uniforme sobre su soporte. En efecto, dentro de la región la densidad es $1/[x(600 - x)]$, que cambia con x ; una densidad uniforme tendría que ser constante en toda la región.

(d). Dado X , la producción Y es uniforme entre 0 y $h(X)$, por lo que su esperanza condicional es la mitad del extremo superior:

$$\mathbb{E}[Y | X] = \frac{X(600 - X)}{900}.$$

La ley de las esperanzas iteradas da

$$\begin{aligned} \mathbb{E}[Y] &= \mathbb{E}\left[\frac{X(600 - X)}{900}\right] \\ &= \frac{2}{3}\mathbb{E}[X] - \frac{1}{900}\mathbb{E}[X^2]. \end{aligned}$$

Como $X \sim \text{Uniforme}(0, 450)$,

$$\mathbb{E}[X] = 225, \quad \mathbb{E}[X^2] = \frac{450^2}{3} = 67\,500.$$

Por tanto,

$$\mathbb{E}[Y] = \frac{2}{3}(225) - \frac{67\,500}{900} = 75.$$

El precio de mercado es entonces $700 \cdot 75 = 52\,500$ millones de pesos. Don Cristóbal usa $z = 200$ y fija un precio de $700(z/2) = 700 \cdot 100 = 70\,000$ millones. Como el precio de venta supera al precio de mercado, Doña Agustina no debería comprar el campo bajo el criterio indicado.

Pregunta 106. (Control 3 - Otoño 2023, P2)

Considere las variables aleatorias independientes Z y W que distribuyen $\text{Normal}(0, 1)$. También considere la VA $X = aZ + b$ y la VA $Y = cZ + dW$, con $a, b, c, d \in \mathbb{R}$ constantes. El objetivo de esta pregunta es demostrar que la variable aleatoria $Y - X$ distribuye Normal. Para eso desarrolle los siguientes pasos:

- (a) Calcule la esperanza de $Y - X$.
- (b) Calcule la varianza de $Y - X$.
- (c) Demuestre que $Y - X$ distribuye Normal.

Hint: Recuerde que si $V \sim \text{Normal}(\mu, \sigma^2)$, entonces V se puede escribir como $V = \mu + \sigma U$, con $U \sim \text{Normal}(0, 1)$.

Solución

(a). Usando la linealidad de la esperanza y que Z y W están centradas,

$$\mathbb{E}[Y - X] = \mathbb{E}[(c - a)Z + dW - b] = -b.$$

(b). Como Z y W son independientes, su covarianza es cero. Por ello,

$$\begin{aligned}\text{Var}(Y - X) &= \text{Var}((c - a)Z + dW - b) \\ &= (c - a)^2 \text{Var}(Z) + d^2 \text{Var}(W) + 2(c - a)d \text{Cov}(Z, W) \\ &= (c - a)^2 + d^2.\end{aligned}$$

(c). La variable $Y - X = (c - a)Z + dW - b$ es una combinación lineal afín de dos normales independientes. En consecuencia,

$$Y - X \sim \text{Normal}(-b, (c - a)^2 + d^2).$$

Si $(c - a)^2 + d^2 = 0$, la variable es constante e igual a $-b$; corresponde al caso Normal degenerado.

Pregunta 107. (Control 3 - Primavera 2023, P1)

Una flecha cae en un blanco de forma circular de radio a , en un punto aleatorio tal que regiones de igual área tienen igual probabilidad de recibir la flecha. Considere ejes cartesianos en el blanco, con origen en el centro, tal que definimos (X, Y) como la posición donde cae la flecha.

- (a) ¿Cuál es la PDF conjunta?
- (b) Determine la PDF marginal de Y .
- (c) Determine la PDF condicional de $X | Y$.

Solución

(a). La densidad es constante sobre el disco $D = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq a^2\}$. Como el área de D es πa^2 , la condición de normalización entrega

$$f_{X,Y}(x, y) = \frac{1}{\pi a^2} \mathbb{1}_{\{x^2 + y^2 \leq a^2\}}.$$

(b). Para $|y| \leq a$, la coordenada x recorre el intervalo $[-\sqrt{a^2 - y^2}, \sqrt{a^2 - y^2}]$. Por tanto,

$$f_Y(y) = \int_{-\sqrt{a^2 - y^2}}^{\sqrt{a^2 - y^2}} \frac{1}{\pi a^2} dx = \frac{2\sqrt{a^2 - y^2}}{\pi a^2} \mathbb{1}_{\{|y| \leq a\}}.$$

(c). Para $|y| < a$, el cociente entre la densidad conjunta y la marginal de Y da

$$f_{X|Y}(x | y) = \frac{1}{2\sqrt{a^2 - y^2}} \mathbb{1}_{\{|x| \leq \sqrt{a^2 - y^2}\}}.$$

Así, condicional en $Y = y$, la variable X es uniforme sobre la cuerda horizontal determinada por ese valor de y .

Pregunta 108. (Control 3 - Primavera 2023, P2)

Doña Anastasia posee un taller de ropa donde se trabajan bordados. Debido a la creciente demanda, Anastasia decide contratar un practicante para realizar ciertas labores, dentro de las cuales, esta enhebrar una aguja. Dicha labor es la que requiere más tiempo, pues muchas veces se requiere de varios intentos para poder hacerlo. En particular, se sabe que el *número de intentos necesarios para enhebrar la aguja* se puede modelar como una VA Geométrica de parámetro $1/(1 + Y)$, donde Y es la *habilidad* del practicante. Considere que la *habilidad* de un practicante es fija (es decir, no cambia en el tiempo), y distribuye Uniforme entre cero y uno.

- a) Determine el número de intentos esperado que le tomará en enhebrar una aguja al futuro practicante que contratará doña Anastasia.
- b) Determine la covarianza entre el *número de intentos necesarios para enhebrar la aguja* y la *habilidad* del practicante. **Hint:** La esperanza de $g(X, Y)$ cuando X es una VA discreta e Y es una VA continua se define por $\mathbb{E}[g(X, Y)] = \int_{R_Y} \sum_{R_X} g(x, y) \cdot f_{XY}(x, y) dy$, y por otro lado, se puede aplicar Regla de la Cadena para obtener la ‘densidad’ conjunta en función de la PMF y PDF que se tienen, similar a la tarea 3, es decir, $f_{XY}(x, y) = p_{X|Y}(x, y)f_Y(y)$.

Solución

Sea X el número de intentos necesarios y sea Y la habilidad del practicante. Condicionado en $Y = y$,

$$X | Y = y \sim \text{Geom}\left(\frac{1}{1 + y}\right),$$

donde la geométrica cuenta el intento exitoso. Por ello,

$$\mathbb{E}[X | Y = y] = 1 + y.$$

(a). La ley de la esperanza iterada permite promediar primero sobre el número de intentos para una habilidad fija y luego sobre la habilidad:

$$\begin{aligned} \mathbb{E}[X] &= \mathbb{E}[\mathbb{E}[X | Y]] \\ &= \mathbb{E}[1 + Y] \\ &= 1 + \frac{1}{2} \\ &= \frac{3}{2}, \end{aligned}$$

donde se usó $Y \sim \text{Uniforme}(0, 1)$.

(b). Para calcular el momento cruzado, condicionamos nuevamente en Y . Como Y queda fijo dentro de la esperanza condicional,

$$\begin{aligned}\mathbb{E}[XY] &= \mathbb{E}[\mathbb{E}[XY | Y]] \\ &= \mathbb{E}[Y \mathbb{E}[X | Y]] \\ &= \mathbb{E}[Y(1 + Y)] \\ &= \mathbb{E}[Y] + \mathbb{E}[Y^2] \\ &= \frac{1}{2} + \frac{1}{3} = \frac{5}{6}.\end{aligned}$$

Además, $\mathbb{E}[X] = 3/2$ y $\mathbb{E}[Y] = 1/2$. Por tanto,

$$\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \frac{5}{6} - \frac{3}{2} \cdot \frac{1}{2} = \frac{1}{12}.$$

La covarianza positiva refleja que una mayor realización de Y reduce el parámetro geométrico $1/(1 + Y)$ y, por ello, aumenta el número esperado de intentos.

Pregunta 109. (Control 3 - Primavera 2023, P3)

- a) Es noviembre y el fin de semestre está cada vez mas cerca, y si bien a usted le quedan pocas semanas de trabajo por este año, para Ucampus es todo lo contrario. Una de las tareas que deben realizar una vez terminado el semestre es analizar la demanda de cursos para el próximo, en particular hay tres cursos altamente demandados, cuyas capacidades son c_1, c_2, c_3 , respectivamente. Suponga que existen N estudiantes, de los cuales cada estudiante inscribirá el curso i con probabilidad $p_i \geq 0$, con $i \in \{1, 2, 3\}$, donde $p_1 + p_2 + p_3 = 1$. Cada alumno inscribe exactamente uno de los 3 cursos. Además, la elección de cada alumno es independiente de la elección del resto. No se posee información de cómo se asignan los estudiantes, no obstante, se sabe que todos los alumnos quedan asignados al curso que inscriben si y solo si el número total de inscritos por curso es menor o igual a la capacidad de dicho curso. Escriba la probabilidad de que todos los alumnos puedan ser asignados al curso que inscribieron.
- b) Verdadero o Falso. Justifique en cada caso:
- (i) Sean X, Y dos distribuciones Normales con parámetros conocidos, entonces, es posible determinar la densidad conjunta de la Normal Multivariada.
 - (ii) Si (X, Y, Z) es una Normal Multivariada, entonces $(X + Y + Z)^2$ distribuye Normal.

Solución

(a). Para enteros no negativos x_1, x_2, x_3 cuya suma es N , la distribución conjunta es multinomial:

$$\mathbb{P}(X_1 = x_1, X_2 = x_2, X_3 = x_3) = \frac{N!}{x_1!x_2!x_3!} \prod_{i=1}^3 p_i^{x_i}.$$

Por lo tanto, la probabilidad de no exceder las capacidades es

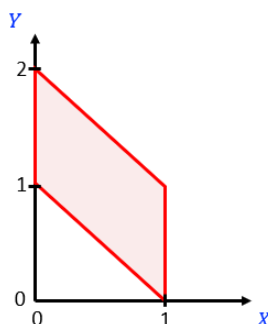
$$\sum_{\substack{x_1+x_2+x_3=N \\ 0 \leq x_i \leq c_i, i=1,2,3}} \frac{N!}{x_1!x_2!x_3!} \prod_{i=1}^3 p_i^{x_i}.$$

(b)(i). La afirmación es falsa. Conocer solamente las marginales normales no determina la covarianza ni garantiza que el vector sea conjuntamente normal.

(b)(ii). La afirmación es falsa. El cuadrado de una variable normal no degenerada no es normal; en el caso centrado y estandarizado tiene una distribución χ^2 con un grado de libertad.

Pregunta 110. (Control 3 - Otoño 2024, P1)

Sean las VA X e Y con densidad uniforme en el área descrita en la figura.



- (a) Calcular la PDF conjunta.
- (b) Calcule la PDF de X condicional en Y . Además, si dicha distribución calza con alguna de las vistas en clases, mencione su nombre y parámetros.

Solución

(a). La región de la figura puede escribirse como

$$A = \{(x, y) : 0 \leq x \leq 1, 1 - x \leq y \leq 2 - x\}.$$

Su área es 1, pues está formada por dos triángulos de área $1/2$. Como la densidad es uniforme y debe integrar 1,

$$f_{X,Y}(x, y) = \mathbb{1}_{\{0 \leq x \leq 1, 1-x \leq y \leq 2-x\}}.$$

(b). Para un valor fijo de y , las restricciones de la región equivalen a $\max\{1 - y, 0\} \leq x \leq \min\{2 - y, 1\}$. En consecuencia,

$$f_Y(y) = \begin{cases} y, & 0 < y \leq 1, \\ 2 - y, & 1 < y < 2, \\ 0, & \text{en otro caso.} \end{cases}$$

Dividiendo la densidad conjunta por esta marginal,

$$f_{X|Y}(x | y) = \begin{cases} \frac{1}{y} \mathbb{1}_{\{1-y \leq x \leq 1\}}, & 0 < y \leq 1, \\ \frac{1}{2-y} \mathbb{1}_{\{0 \leq x \leq 2-y\}}, & 1 < y < 2. \end{cases}$$

Por tanto, $X | Y = y$ es Uniforme $[1 - y, 1]$ si $0 < y \leq 1$ y Uniforme $[0, 2 - y]$ si $1 < y < 2$.

Pregunta 111. (Control 3 - Otoño 2024, P3)

- (a) Para la fiesta de pascua de resurrección, doña María esconderá huevitos de chocolate en el interior de los cajones de su cómoda. Dicha cómoda posee cuatro cajones, uno arriba de otro. Doña María esconderá 7 huevitos (indistinguibles) de chocolate de la siguiente manera: **para cada huevito**, ella lanzará una moneda hasta tener una cara, o hasta tener 3 sellos. En caso de tener una cara en el i -ésimo lanzamiento (donde $i \in \{1, 2, 3\}$), el huevito será escondido en el i -ésimo cajón¹. En caso de que los tres primeros lanzamientos sean sello, el huevito será guardado en el último cajón, es decir, en el cajón de más abajo.

¹ i -ésimo cajón de arriba hacia abajo, en consecuencia, el cajón de más arriba es el primer cajón, luego el segundo cajón es el cajón que está debajo del cajón de más arriba, y así sucesivamente.

Note que si, para un huevito, el primer lanzamiento es cara, entonces no se sigue lanzando la moneda para dicho huevito. Similarmente si el primer lanzamiento es sello y el segundo es cara. Determine la probabilidad de que el primer, segundo, tercer, y cuarto cajón posean exactamente, dos, tres, cero y dos huevitos, respectivamente. Justifique claramente cómo llega a su respuesta. Puede dejar su respuesta expresada, sin necesidad de calcular exactamente el valor.

- (b) Considere que a diferencia de la parte anterior, doña María posee una cómoda de n cajones, dentro de los cuales esconderá k huevitos de chocolate, donde $k < n$. Específicamente, cada cajón tendrá a lo más un huevito de chocolate, y además, todos los cajones tendrán la misma probabilidad de poseer un huevito. Defina una variable aleatoria que sea igual al número de huevitos de chocolate en el primer cajón, y otra variable aleatoria que sea igual al número de huevitos de chocolate en el segundo cajón. Determine la covarianza entre estas dos variables.
- (c) Dado el contexto del problema de la parte (b), explique sobre la intuición del signo de la covarianza. Específicamente, explique por qué tiene sentido que dicha covarianza posea cierto signo (que puede ser positivo, negativo, o bien ninguno en caso de que la covarianza sea cero). Note que puede responder esta pregunta sin haber resuelto la parte (b), pues se pide explicar la intuición sobre el signo que debería tener dicha covarianza.

Solución

(a). El enunciado no especifica la probabilidad de cara ni la independencia de los lanzamientos, por lo que la probabilidad solicitada no queda determinada de manera única. Sea $q = \mathbb{P}(\text{cara})$. Bajo el supuesto adicional de que todos los lanzamientos son independientes, las probabilidades de asignar un huevito a los cuatro cajones son

$$(p_1, p_2, p_3, p_4) = (q, (1-q)q, (1-q)^2q, (1-q)^3).$$

Si X_i denota la cantidad de huevitos en el cajón i , el vector de cantidades por cajón sigue entonces una distribución multinomial, y

$$\mathbb{P}((X_1, X_2, X_3, X_4) = (2, 3, 0, 2)) = \frac{7!}{2!3!0!2!} q^2 ((1-q)q)^3 ((1-q)^2q)^0 ((1-q)^3)^2.$$

Si además la moneda es balanceada, $q = 1/2$, y se recupera la expresión numérica de la pauta original.

(b). Sea X_i el indicador de que el cajón i contiene un huevito. Como hay exactamente k huevitos y todos los cajones tienen la misma probabilidad de estar ocupados,

$$\mathbb{E}[X_i] = \mathbb{P}(X_i = 1) = \frac{k}{n}.$$

Sin embargo,

$$\text{Cov}(X_1, X_2) = \mathbb{P}(X_1 = 1, X_2 = 1) - \left(\frac{k}{n}\right)^2,$$

y el enunciado no determina la probabilidad conjunta. Por ejemplo, si $n = 4$ y $k = 2$, escoger $\{1, 2\}$ o $\{3, 4\}$ con probabilidad $1/2$ mantiene marginales iguales a $1/2$, pero produce $\text{Cov}(X_1, X_2) = 1/4$. Si se añade el supuesto de que todos los subconjuntos de k cajones son equiprobables, entonces

$$\mathbb{P}(X_1 = 1, X_2 = 1) = \frac{k}{n} \frac{k-1}{n-1}, \quad \text{Cov}(X_1, X_2) = -\frac{k(n-k)}{n^2(n-1)}.$$

(c). Con la información original, el signo de la covarianza tampoco queda determinado. Bajo el supuesto adicional de selección uniforme, la covarianza es negativa: saber que el primer cajón está ocupado deja sólo $k-1$ huevitos para los $n-1$ cajones restantes, de modo que

$$\mathbb{P}(X_2 = 1 \mid X_1 = 1) = \frac{k-1}{n-1} < \frac{k}{n} = \mathbb{P}(X_2 = 1),$$

donde la desigualdad usa $k < n$.

Pregunta 112. (Control 3 - Primavera 2024, P1)

Una ambulancia viaja a velocidad constante a lo largo de una carretera de longitud L . En cierto momento, ocurre un accidente en un punto distribuido uniformemente a lo largo de la carretera, es decir, la distancia desde el lugar del accidente hasta uno de los extremos fijos de la carretera está distribuida uniformemente en el intervalo $[0, L]$. La ubicación de la ambulancia en el momento del accidente también sigue una distribución Uniforme en $[0, L]$. Además, se sabe que la ambulancia puede asistir eficazmente a los accidentados y salvar sus vidas solo si la distancia entre su posición y el lugar del accidente es menor o igual a $\frac{L}{4}$. Suponiendo que la posición de la ambulancia y la del accidente son independientes:

- Dibuje un gráfico de dos dimensiones, que muestre en sus ejes la posición de la ambulancia y la posición del accidente, indicando claramente la región en la que la ambulancia puede responder eficazmente al rescate de los accidentados.
- Expresa (sin calcular), mediante integrales resolubles, la probabilidad de que la ambulancia logre acudir al rescate de manera efectiva.

Hint: Podría serle útil recordar que para un vector aleatorio $\mathbf{X}$ con rango $R_{\mathbf{X}} \subseteq \mathbb{R}^n$, entonces $\mathbb{P}(\mathbf{X} \in A) = \int_{\mathbb{R}^n} \mathbb{1}_{\{x \in A\}} f(\mathbf{x}) d\mathbf{x}$, donde $A \subseteq \mathbb{R}^n$.

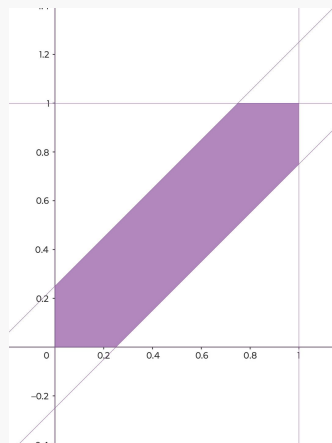
Indicación: Las respuestas deben escribirse en términos de integrales explícitas. No deje respuestas en forma de símbolos como $\mathbb{P}(\cdot)$ o $f_X(\cdot)$.

Solución

(a). Si X es la posición del accidente e Y la posición de la ambulancia, la región eficaz es

$$\{(x, y) \in [0, L]^2 : |x - y| \leq L/4\}.$$

Es la franja del cuadrado $[0, L]^2$ comprendida entre las rectas $x - y = L/4$ y $x - y = -L/4$.



(b). La independencia y las distribuciones uniformes implican

$$f_{X,Y}(x, y) = \frac{1}{L^2} \mathbb{1}_{\{0 \leq x \leq L, 0 \leq y \leq L\}}.$$

Integrando sobre la franja descrita en el inciso anterior, la probabilidad pedida queda expresada

por

$$\int_0^L \int_{\max\{0, y-L/4\}}^{\min\{L, y+L/4\}} \frac{1}{L^2} dx dy.$$

Pregunta 113. (Control 3 - Primavera 2024, P3)

- (a) Sean X e Y dos variables aleatorias continuas conjuntamente, con función de densidad conjunta

$$f_{X,Y}(x, y) = \begin{cases} x^2 + \frac{1}{2}y, & -1 \leq x \leq 1, 0 \leq y \leq 1, \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Encuentre y responda:

- (I) La función de densidad condicional de X dado Y .
 - (II) ¿Son X e Y independientes? Justifique su respuesta.
- (b) En una carrera, compiten dos corredores: uno del equipo A y otro del equipo B . El tiempo (en segundos) que cada competidor tarda en llegar a la meta es una variable aleatoria Exponencial. Para el corredor del equipo A , el tiempo sigue una distribución Exponencial con parámetro de tasa λ , mientras que para el corredor del equipo B , el tiempo sigue una distribución Exponencial con parámetro de tasa μ . Ambas variables aleatorias son independientes entre sí. Calcule la probabilidad de que el ganador de la carrera llegue a la meta en menos de t segundos. Además, indique si este tiempo sigue alguna de las distribuciones estudiadas en clase.

Solución

- (a). La función presentada en el enunciado no está normalizada, pues

$$\int_0^1 \int_{-1}^1 \left(x^2 + \frac{y}{2}\right) dx dy = \frac{7}{6}.$$

Por tanto, estrictamente no es una densidad conjunta. Si se interpreta la expresión dada como un núcleo proporcional a la densidad, la constante de normalización es $6/7$. Esta constante se cancela al condicionar y, para $0 \leq y \leq 1$, se obtiene

$$f_{X|Y}(x | y) = \frac{x^2 + y/2}{2/3 + y} \mathbb{1}_{\{-1 \leq x \leq 1\}}.$$

La expresión depende de y ; por ejemplo, su valor en $x = 0$ cambia con y . En consecuencia, bajo esa interpretación, X e Y no son independientes.

- (b). Sean $T_A \sim \text{Exponencial}(\lambda)$ y $T_B \sim \text{Exponencial}(\mu)$ los tiempos independientes de ambos corredores. El tiempo del ganador es $T = \min\{T_A, T_B\}$. Para $t \geq 0$,

$$\begin{aligned} \mathbb{P}(T < t) &= 1 - \mathbb{P}(T_A \geq t, T_B \geq t) \\ &= 1 - \mathbb{P}(T_A \geq t) \mathbb{P}(T_B \geq t) \\ &= 1 - e^{-(\lambda+\mu)t}. \end{aligned}$$

Así, $T \sim \text{Exponencial}(\lambda + \mu)$; para $t < 0$, la probabilidad solicitada es cero.

Pregunta 114. (Control 3 - Otoño 2025, P1)

Para esta pregunta, considere la función de densidad de probabilidad de una distribución Logística, de parámetros *localización* μ y *escala* s , es

$$f(t) = \frac{e^{-(t-\mu)/s}}{s(1 + e^{-(t-\mu)/s})^2} \quad \text{para todo } t \in \mathbb{R}.$$

Andrea se somete a un examen que mide la densidad de una cierta proteína en su sangre. Se considera que un resultado es **bueno** si la densidad es menor a 0.8 mg/l. Se sabe que la densidad **real** de esta proteína en sangre sigue una distribución **Uniforme** entre 0 y 1 mg/l. Sin embargo, el examen no permite observar directamente el valor real, sino que entrega una **estimación ruidosa**: el resultado observado sigue una distribución **logística**, con **parámetro de localización** igual a la densidad real de la proteína y **parámetro de escala** igual a 1. El resultado del examen de Andrea fue 0.85 mg/l. ¿Cuál es la probabilidad de que su densidad real de proteína en sangre sea **bueno** (es decir, menor a 0.8 mg/l)?

Le puede ser útil la siguiente integral

$$\int \frac{e^{-z}}{(1 + e^{-z})^2} dz = \frac{1}{1 + e^{-z}} + C.$$

Observación: La respuesta debe expresarse como un valor numérico o una **fórmula cerrada** que utilice únicamente operaciones algebraicas (suma, resta, multiplicación, división, potencias) y funciones elementales como logaritmos, exponenciales o funciones trigonométricas.

Solución

Sea $X \sim \text{Uniforme}(0, 1)$ la densidad real de proteína y sea Y el resultado observado. Condicionamente a $X = x$, la escala logística es uno y

$$f_{Y|X}(y | x) = \frac{e^{-(y-x)}}{(1 + e^{-(y-x)})^2}.$$

Como $f_X(x) = 1$ para $0 \leq x \leq 1$, el Teorema de Bayes para densidades permite escribir

$$\mathbb{P}(X < 0,8 | Y = 0,85) = \frac{\int_0^{0,8} f_{Y|X}(0,85 | x) dx}{\int_0^1 f_{Y|X}(0,85 | x) dx}.$$

Sea $G(u) = 1/(1 + e^{-u})$. Con el cambio de variable $u = 0,85 - x$ y usando que $G'(u) = e^{-u}/(1 + e^{-u})^2$, el numerador y el denominador son, respectivamente,

$$G(0,85) - G(0,05) \quad \text{y} \quad G(0,85) - G(-0,15).$$

Por tanto,

$$\mathbb{P}(X < 0,8 | Y = 0,85) = \frac{\frac{1}{1 + e^{-0,85}} - \frac{1}{1 + e^{-0,05}}}{\frac{1}{1 + e^{-0,85}} - \frac{1}{1 + e^{0,15}}}.$$

Pregunta 115. (Control 3 - Otoño 2025, P2)

Una viña está interesada en evaluar la calidad de su producción de vino, la cual depende principalmente de dos factores climáticos: la lluvia y la temperatura. Un estudio reciente afirma que la cantidad de milímetros [mm] de lluvia y la temperatura promedio en grados Celsius [C] durante la temporada de cosecha siguen la siguiente distribución Normal multivariada:

$$\mathcal{N}\left(\begin{pmatrix} 120 \text{ [mm]} \\ 26 \text{ [C]} \end{pmatrix}, \begin{pmatrix} 3600 \text{ [mm}^2\text{]} & -72 \text{ [mm}\cdot\text{C]} \\ -72 \text{ [mm}\cdot\text{C]} & 9 \text{ [C}^2\text{]} \end{pmatrix}\right).$$

Por otra parte, una consultora ha elaborado una fórmula para computar el **puntaje** de la calidad del vino después de una temporada como la siguiente función lineal afín:

$$\text{puntaje} = 20 - 0.4 \cdot \text{Milímetros de lluvia} + 3 \cdot \text{Temperatura en C}.$$

Además, un vino es considerado de excelente calidad si es que su **puntaje** es mayor a 90 puntos. Para cada parte, entregue el **valor numérico** de la respuesta final.

- Calcule el puntaje esperado de un vino.
- Calcule la varianza del puntaje de un vino.
- Calcule la probabilidad de que los vinos producidos después de una temporada sean de excelente calidad.

Solución

Sean X la lluvia, Y la temperatura y $S = 20 - 0.4X + 3Y$ el puntaje.

(a). Por linealidad de la esperanza,

$$\mathbb{E}[S] = 20 - 0.4\mathbb{E}[X] + 3\mathbb{E}[Y] = 20 - 0.4(120) + 3(26) = 50.$$

(b). La matriz de covarianza entrega $\text{Var}(X) = 3600$, $\text{Var}(Y) = 9$ y $\text{Cov}(X, Y) = -72$. Luego,

$$\begin{aligned} \text{Var}(S) &= (-0.4)^2 \text{Var}(X) + 3^2 \text{Var}(Y) + 2(-0.4)(3) \text{Cov}(X, Y) \\ &= 0.16(3600) + 9(9) + 2(-0.4)(3)(-72) \\ &= 829.8. \end{aligned}$$

(c). Como S es una combinación lineal afín de las componentes de un vector Normal multivariado,

$$S \sim \text{Normal}(50, 829.8).$$

Al estandarizar,

$$\begin{aligned} \mathbb{P}(S > 90) &= 1 - \Phi\left(\frac{90 - 50}{\sqrt{829.8}}\right) \\ &= 1 - \Phi(1.3886) \approx 0.0825. \end{aligned}$$

Pregunta 116. (Control 3 - Primavera 2025, P1)

- Se tiene que $f(x, y) = 1/x$, para $0 < y < x < 1$, y 0 en caso contrario.
 - Muestre que f es una función de probabilidad de densidad conjunta.
 - Asumiendo que f es la función de densidad conjunta de X e Y , encuentre la densidad marginal de Y .
- Don Enairo está considerando vender su auto por medio de un remate. Llegan 8 oferentes al remate, donde cada uno ofrece un valor al azar entre 0 y 10 millones de pesos, por lo que

se puede modelar cada oferta con una distribución Uniforme en el intervalo mencionado. Considere que las ofertas entre los oferentes son independientes entre sí. Enairo venderá su auto al oferente que ofrezca el monto más alto, quien pagará tal monto a Enairo. Estime el valor esperado del precio de venta del vehículo.

Solución

(a)(i). La función es no negativa porque $x > 0$ en su soporte. Para verificar que su integral total es uno, se usan los límites $0 < x < 1$ y $0 < y < x$:

$$\begin{aligned}\int_0^1 \int_0^x \frac{1}{x} dy dx &= \int_0^1 \frac{1}{x} [y]_0^x dx \\ &= \int_0^1 1 dx \\ &= 1.\end{aligned}$$

Por tanto, f es una densidad conjunta.

(a)(ii). Para un valor fijo $y \in (0, 1)$, la restricción $y < x < 1$ determina los límites de integración en x . Así,

$$\begin{aligned}f_Y(y) &= \int_y^1 \frac{1}{x} dx \\ &= \ln 1 - \ln y \\ &= -\ln y.\end{aligned}$$

En consecuencia,

$$f_Y(y) = \begin{cases} -\ln y, & 0 < y < 1, \\ 0, & \text{en otro caso.} \end{cases}$$

El valor de una densidad en los extremos 0 y 1 puede elegirse arbitrariamente sin cambiar la distribución.

(b). Sean $X_1, \dots, X_8$ las ofertas independientes, cada una con distribución Uniforme(0, 10), y sea

$$M = \max\{X_1, \dots, X_8\}$$

el precio de venta. Para $0 \leq m \leq 10$,

$$\begin{aligned}F_M(m) &= \mathbb{P}(X_1 \leq m, \dots, X_8 \leq m) \\ &= \prod_{i=1}^8 \mathbb{P}(X_i \leq m) \\ &= \left(\frac{m}{10}\right)^8.\end{aligned}$$

Al derivar,

$$f_M(m) = \frac{8m^7}{10^8}, \quad 0 < m < 10.$$

Por tanto,

$$\begin{aligned}\mathbb{E}[M] &= \int_0^{10} m \frac{8m^7}{10^8} dm \\ &= \frac{8}{10^8} \left[\frac{m^9}{9} \right]_0^{10} \\ &= \frac{80}{9}.\end{aligned}$$

El precio esperado es entonces $80/9$ millones de pesos.

Pregunta 117. (Control 3 - Primavera 2025, P2)

- (a) Un amigo suyo que está realizando su práctica en el Ministerio de Salud le pide ayuda con un problema, ya que usted conoce bien los contenidos del curso de probabilidades. Su amigo necesita modelar para cada hospital $i \in \{1, \dots, n\}$ la demanda diaria de pacientes (X_i) . El equipo de investigación ha decidido que el vector de demandas $\mathbf{X} = (X_1, X_2, \dots, X_n)^T$ se puede modelar como una Normal Multivariada. Para construir el modelo, su amigo recolectó la siguiente información:

- Para cada hospital i , la demanda diaria esperada es a_i .
- La desviación estándar de la demanda diaria en cada hospital i está dada por b_i .
- La demanda entre dos hospitales i y j está correlacionada. Esta correlación depende de la distancia d_{ij} entre ellos (con $d_{ii} = 0$) y un parámetro $\gamma > 0$, según la fórmula:

$$\text{Corr}(X_i, X_j) = e^{-\gamma d_{ij}}$$

Ayude a su amigo a encontrar los parámetros de la Normal Multivariada de $\mathbf{X}$.

- (b) Sea X una variable aleatoria que sigue una distribución $X \sim \text{Uniforme}(1, 2)$. Sea Y otra variable aleatoria, tal que la distribución condicional de Y dado $X = x$ es $Y | X = x \sim \text{Exp}(x)$. Calcule la covarianza entre X e Y .

Solución

(a). El vector de medias reúne las demandas esperadas:

$$\boldsymbol{\mu} = \mathbb{E}[\mathbf{X}] = (a_1, \dots, a_n)^T.$$

Para construir la matriz de covarianza se usa

$$\text{Cov}(X_i, X_j) = \text{Corr}(X_i, X_j) \sqrt{\text{Var}(X_i)} \sqrt{\text{Var}(X_j)}.$$

Como las desviaciones estándar son b_i y b_j ,

$$\Sigma_{ij} = b_i b_j e^{-\gamma d_{ij}}, \quad i, j \in \{1, \dots, n\}.$$

Esta fórmula también cubre la diagonal: $d_{ii} = 0$ implica $\Sigma_{ii} = b_i^2$. Por tanto, el modelo es

$$\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}),$$

con

$$\boldsymbol{\Sigma} = \begin{pmatrix} b_1^2 & b_1 b_2 e^{-\gamma d_{12}} & \dots & b_1 b_n e^{-\gamma d_{1n}} \\ b_2 b_1 e^{-\gamma d_{21}} & b_2^2 & \dots & b_2 b_n e^{-\gamma d_{2n}} \\ \vdots & \vdots & \ddots & \vdots \\ b_n b_1 e^{-\gamma d_{n1}} & b_n b_2 e^{-\gamma d_{n2}} & \dots & b_n^2 \end{pmatrix}.$$

Para que esta especificación defina una normal multivariada, la matriz resultante debe ser simétrica y semidefinida positiva. Esto se cumple para el núcleo exponencial aplicado a distancias euclidianas; si d_{ij} fuese una disimilitud arbitraria, esa propiedad tendría que verificarse.

(b). Usaremos

$$\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y].$$

Como $X \sim \text{Uniforme}(1, 2)$, se tiene $\mathbb{E}[X] = 3/2$. Además, la distribución condicional $Y | X =$

$x \sim \text{Exp}(x)$ tiene media $1/x$. La ley de la esperanza iterada entrega

$$\begin{aligned}\mathbb{E}[Y] &= \mathbb{E}[\mathbb{E}[Y \mid X]] \\ &= \mathbb{E}\left[\frac{1}{X}\right] \\ &= \int_1^2 \frac{1}{x} dx \\ &= \ln 2.\end{aligned}$$

Para el momento cruzado, X es constante al condicionar en sí mismo:

$$\begin{aligned}\mathbb{E}[XY] &= \mathbb{E}[\mathbb{E}[XY \mid X]] \\ &= \mathbb{E}[X \mathbb{E}[Y \mid X]] \\ &= \mathbb{E}\left[X \frac{1}{X}\right] \\ &= 1.\end{aligned}$$

Por tanto,

$$\text{Cov}(X, Y) = 1 - \frac{3}{2} \ln 2.$$

El valor es negativo, coherentemente con que una tasa exponencial mayor produce un tiempo condicional esperado menor.

Pregunta 118. (Control 3 - Otoño 2026, P1)

Tras el éxito del evento *Sunset* en el casino de la facultad, patrocinado por Red Bull, la marca ha decidido lanzar una tienda virtual exclusiva durante el Cyberday. Para optimizar su estrategia, el equipo de analistas modela el comportamiento de los usuarios mediante dos variables aleatorias continuas: X , que representa el volumen de visitas a la página web, e Y , que representa el volumen de compras concretadas. Ambas variables están expresadas en una escala normalizada entre 0 y 1. Además, como solo pueden concretar compras quienes visitan la página, se cumple que $0 \leq Y \leq X \leq 1$.

Se ha determinado que la densidad marginal de visitas es $f_X(x) = 2x\mathbb{1}_{\{x \in (0,1)\}}$. Además, el equipo notó que para un volumen dado de visitas x , el número de compradores y no presenta preferencias específicas, siendo igualmente probable encontrar cualquier cantidad de compras entre cero y el total de visitantes.

- (a) A partir del comportamiento descrito, determine la función de densidad de probabilidad conjunta del número de visitantes a la página web y del número de compras. Especifique claramente el dominio de la función.
- (b) Red Bull enfrenta una estructura de costos y beneficios: cada unidad de visita a la página web le cuesta a la empresa \$3, mientras que cada compra le otorga un ingreso de \$10.
 - (I) Calcule el número esperado de visitas y de compras.
 - (II) Defina la utilidad de la empresa como los ingresos (producto de las ventas) menos los costos (producto de las visitas a la página web). Calcule el valor esperado de la utilidad.
- (c) Calcule la covarianza entre las visitas al sitio y las compras realizadas.

Solución

(a). Condicional en $X = x$, el volumen de compras es uniforme en $[0, x]$, de modo que $f_{Y|X}(y | x) = 1/x$ para $0 < x \leq 1$ y $0 \leq y \leq x$. La regla del producto entrega

$$f_{X,Y}(x, y) = f_{Y|X}(y | x)f_X(x) = 2\mathbf{1}_{\{0 \leq y \leq x \leq 1\}}.$$

(b)(i). La esperanza de las visitas es

$$\mathbb{E}[X] = \int_0^1 x(2x) dx = \frac{2}{3}.$$

Como $Y | X \sim \text{Uniforme}[0, X]$, se tiene $\mathbb{E}[Y | X] = X/2$. La ley de la esperanza total da

$$\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y | X]] = \frac{1}{2}\mathbb{E}[X] = \frac{1}{3}.$$

(b)(ii). Si $U = 10Y - 3X$ es la utilidad, entonces

$$\mathbb{E}[U] = 10\mathbb{E}[Y] - 3\mathbb{E}[X] = \frac{4}{3}.$$

(c). Usando nuevamente la esperanza condicional,

$$\mathbb{E}[XY] = \mathbb{E}[X\mathbb{E}[Y | X]] = \frac{1}{2}\mathbb{E}[X^2] = \frac{1}{2} \int_0^1 x^2(2x) dx = \frac{1}{4}.$$

Por tanto,

$$\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \frac{1}{4} - \frac{2}{3} \frac{1}{3} = \frac{1}{36}.$$

Pregunta 119. (Control 3 - Otoño 2026, P3)

- (a) En el tradicional *Torneo de la Mona*, el Departamento de Ingeniería Industrial (DII) ha inscrito a 50 competidores. Históricamente, la probabilidad de que un competidor de Industrias obtenga medalla de oro es 0,1, de plata es 0,2 y de bronce es 0,3 (y la probabilidad restante es que no obtenga medalla), independiente de los demás competidores.

(I) Calcule la probabilidad de que el departamento termine la jornada con exactamente 7 medallas de oro, 5 de plata y 4 de bronce.

(II) ¿Cuál es la probabilidad de que el DII obtenga al menos una medalla entre todos los competidores?

- (b) Dos estudiantes de Industrias han decidido vender donas en el patio para financiar su próximo viaje. Sean X e Y los errores de peso (en gramos) detectados en la producción de dos lotes distintos. Se asume que el vector (X, Y) sigue una distribución Normal Multivariada con vector de medias $\mu = (0, 0)^T$ y una matriz de covarianza $\Sigma \in \mathbb{R}^{2 \times 2}$.

Determine la probabilidad de que el error del primer lote sea mayor que el del segundo, es decir, $\mathbb{P}(X > Y)$.

Sugerencia: Determine primero cómo distribuye la variable aleatoria $W = X - Y$.

- (c) Sean X e Y dos variables aleatorias independientes con distribución Bernoulli del mismo parámetro $p \in (0, 1)$. Un analista afirma que el promedio de ambas, definido como $Z = \frac{X+Y}{2}$, también sigue una distribución Bernoulli.

Evalúe la veracidad de esta afirmación.

Solución**(a)(i).** Sea

$$(X_1, X_2, X_3, X_4) \sim \text{Multinomial}(50; 0,1, 0,2, 0,3, 0,4),$$

donde las componentes cuentan medallas de oro, plata, bronce y competidores sin medalla, respectivamente. Como $50 - 7 - 5 - 4 = 34$,

$$\mathbb{P}(X_1 = 7, X_2 = 5, X_3 = 4, X_4 = 34) = \frac{50!}{7! 5! 4! 34!} (0,1)^7 (0,2)^5 (0,3)^4 (0,4)^{34}.$$

(a)(ii). El evento complementario es que los 50 competidores queden sin medalla. Por independencia,

$$\mathbb{P}(\text{al menos una medalla}) = 1 - (0,4)^{50}.$$

(b). Una combinación lineal de las componentes de un vector Normal multivariado es Normal. Así,

$$W = X - Y \sim \text{Normal}(0, \tau^2), \quad \tau^2 = (1 \quad -1) \Sigma \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Si $\tau^2 > 0$, la simetría respecto de cero implica

$$\mathbb{P}(X > Y) = \mathbb{P}(W > 0) = \frac{1}{2}.$$

Si $\tau^2 = 0$, entonces $W = 0$ casi seguramente y la probabilidad es 0. Por ello, el valor $1/2$ requiere que la diferencia no sea degenerada.

(c). La afirmación es falsa. Además de los valores 0 y 1, la variable Z toma el valor $1/2$ cuando exactamente una de las dos variables vale 1. Como

$$\mathbb{P}\left(Z = \frac{1}{2}\right) = 2p(1-p) > 0,$$

su soporte no es el de una variable Bernoulli.

Pregunta 120. (Examen - Otoño 2021, P1)

En una planta refinadora de níquel hay una cuota de $L > 0$ millones de litros de agua al mes. Sean X e Y los consumos, en millones de litros, de dos etapas del proceso. A partir de datos históricos se modela el par (X, Y) mediante una densidad constante sobre la región

$$T = \{(x, y) : x \geq 0, y \geq 0, x + y \leq L\}.$$

- Calcule la densidad conjunta $f_{X,Y}$.
- Calcule las densidades marginales f_X y f_Y .
- Calcule $\text{Cov}(X, Y)$.
- Defina $Z = Y/(L - X)$. Calcule la densidad conjunta de (X, Z) y muestre que X y Z son independientes.

Solución

(a). El triángulo T tiene base y altura iguales a L , por lo que su área es $L^2/2$. Como la densidad es constante sobre esa región, la constante debe ser el inverso del área:

$$f_{X,Y}(x, y) = \begin{cases} \frac{2}{L^2}, & x \geq 0, y \geq 0, x + y \leq L, \\ 0, & \text{en otro caso.} \end{cases}$$

(b). Para un valor fijo $x \in [0, L]$, la coordenada y varía entre 0 y $L - x$. Por tanto,

$$f_X(x) = \begin{cases} \frac{2(L-x)}{L^2}, & 0 \leq x \leq L, \\ 0, & \text{en otro caso.} \end{cases}$$

La región es simétrica al intercambiar x e y , de modo que

$$f_Y(y) = \begin{cases} \frac{2(L-y)}{L^2}, & 0 \leq y \leq L, \\ 0, & \text{en otro caso.} \end{cases}$$

(c). A partir de las marginales, $\mathbb{E}[X] = \mathbb{E}[Y] = L/3$. El momento cruzado se obtiene de la densidad conjunta:

$$\begin{aligned} \mathbb{E}[XY] &= \frac{2}{L^2} \int_0^L \int_0^{L-x} xy \, dy \, dx \\ &= \frac{L^2}{12}. \end{aligned}$$

En consecuencia,

$$\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \frac{L^2}{12} - \frac{L^2}{9} = -\frac{L^2}{36}.$$

(d). La transformación es $Z = Y/(L - X)$. Salvo en la frontera $X = L$, que tiene probabilidad cero, la transformación inversa es

$$x = x, \quad y = z(L - x),$$

y el valor absoluto de su jacobiano es $L - x$. La restricción $0 \leq y \leq L - x$ se transforma en $0 \leq z \leq 1$. Así,

$$f_{X,Z}(x, z) = \frac{2(L-x)}{L^2} \mathbb{1}_{\{0 \leq x \leq L, 0 \leq z \leq 1\}}.$$

Al integrar respecto de x se obtiene $f_Z(z) = 1$ para $0 \leq z \leq 1$. Por tanto,

$$f_{X,Z}(x, z) = f_X(x)f_Z(z),$$

en el soporte conjunto, lo que demuestra que X y Z son independientes.

Pregunta 121. (Examen - Primavera 2021, P1)

¡Jo Jo Jo! Ha llegado la Navidad al IN3141 y para ello necesitamos su ayuda. El equipo docente decidió realizar una cena navideña para el curso. Para lo anterior, se hace una colecta para recaudar dinero con el cual se compran los bebestibles. Dado que la fiesta navideña se realiza todos los años, se sabe que la cantidad total de litros de bebestible L que se compra distribuye uniforme continua entre cero y k . Además, **de los bebestibles**, hay uno que es de particular interés para la fecha, el *cola de mono*. Se sabe que los litros de bebestible que corresponden específicamente a *cola de mono* es una variable aleatoria X que sigue una distribución uniforme continua entre cero y la cantidad total de bebestible. Solo a modo de aclaración, note que la cantidad de litros de *cola de mono* es parte de la cantidad total de bebestible; sin embargo, no necesariamente todos los litros de bebestible son de *cola de mono*.

Hint: Le pueden ser de utilidad las siguientes integrales:

$$\int \ln(y) \, dy = y(\ln(y) - 1) + C, \quad \int y \ln(y) \, dy = \frac{y^2(2 \ln(y) - 1)}{4} + C.$$

donde C es una constante de integración.

(a) Identifique el rango conjunto de las variables aleatorias (X, L) .

- (b) Calcule la PDF conjunta del par (X, L) . Obvie los casos en que $l = 0$ o $x = 0$.
- (c) Calcule la densidad marginal de X . Obvie el caso en que $x = 0$.
- (d) Calcule la covarianza entre X y L .

Solución

(a). Como $0 \leq X \leq L \leq k$,

$$\mathcal{R}_{X,L} = \{(x, l) \in [0, k]^2 : x \leq l\}.$$

(b). Para $0 < l \leq k$, la regla del producto da

$$f_{X,L}(x, l) = f_{X|L}(x | l)f_L(l) = \frac{1}{kl} \mathbb{1}_{\{0 \leq x \leq l \leq k\}}.$$

(c). Si $0 < x \leq k$, se integra sobre $x \leq l \leq k$:

$$f_X(x) = \int_x^k \frac{1}{kl} dl = \frac{1}{k} \ln\left(\frac{k}{x}\right),$$

y $f_X(x) = 0$ fuera de $(0, k]$.

(d). Condicionalmente a L , $\mathbb{E}[X | L] = L/2$. Por ello,

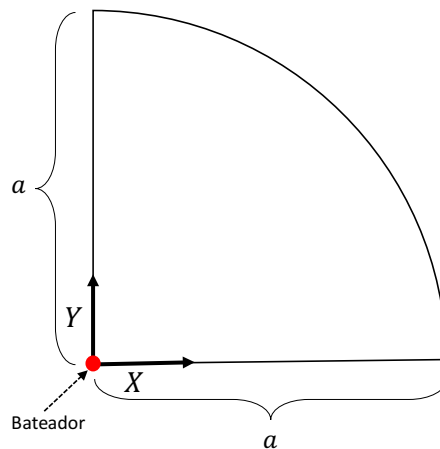
$$\mathbb{E}[L] = \frac{k}{2}, \quad \mathbb{E}[X] = \frac{\mathbb{E}[L]}{2} = \frac{k}{4}, \quad \mathbb{E}[XL] = \mathbb{E}\left[\frac{L^2}{2}\right] = \frac{k^2}{6}.$$

Así,

$$\text{Cov}(X, L) = \mathbb{E}[XL] - \mathbb{E}[X]\mathbb{E}[L] = \frac{k^2}{6} - \frac{k^2}{8} = \frac{k^2}{24}.$$

Pregunta 122. (Examen alternativo - Primavera 2021, P1)

Un bateador de béisbol batea la bola de tal manera que cae con igual probabilidad en el área comprendida en la sección circular con un ángulo de 90 grados (donde justo en la intersección de los dos segmentos rectos de la sección circular está el bateador), tal que el largo del radio del círculo asociado al segmento circular es de a metros. Sean X e Y las coordenadas cartesianas con origen en el lugar donde está el bateador, y con X e Y iguales a la posición en la que cae la bola luego se ser bateada. Ver Figura.



- (a) Identifique el rango conjunto de las variables aleatorias (X, Y) .
- (b) Calcule la PDF conjunta del par (X, Y) .

- (c) Calcule la densidad marginal de X
- (d) Suponga que hay un amigo del bateador se ubica en la posición $(a/3, a/3)$ (utilizando las mismas coordenadas asociadas a las VA X e Y). Suponga que el amigo del bateador tratará de agarrar la bola una vez que esta ha sido bateada por el bateador. Se sabe que el amigo de bateador puede tomar la bola en el aire (una vez que se ha bateado) solamente si es que la bola cae a un radio de $a/4$ de donde se ubica el amigo del bateador. Calcule la probabilidad de que el amigo del bateador tome la bola en el aire luego de que esta haya sido bateada por el bateador.

Solución

Sea

$$R = \{(x, y) \in \mathbb{R}^2 : x \geq 0, y \geq 0, x^2 + y^2 \leq a^2\}.$$

(a). El rango conjunto es el cuarto de disco R .

(b). Como el punto de caída es uniforme y $\text{area}(R) = \pi a^2/4$,

$$f_{X,Y}(x, y) = \begin{cases} \frac{4}{\pi a^2}, & (x, y) \in R, \\ 0, & \text{en otro caso.} \end{cases}$$

(c). Para $0 \leq x \leq a$,

$$f_X(x) = \int_0^{\sqrt{a^2 - x^2}} \frac{4}{\pi a^2} dy = \frac{4\sqrt{a^2 - x^2}}{\pi a^2},$$

y $f_X(x) = 0$ fuera de ese intervalo.

(d). El disco de alcance, centrado en $(a/3, a/3)$ y de radio $a/4$, queda dentro del primer cuadrante porque $a/3 - a/4 > 0$, y también dentro del círculo grande porque $\sqrt{2}a/3 + a/4 < a$. Por tanto, está completamente contenido en R , y la probabilidad es la razón de áreas

$$\frac{\pi(a/4)^2}{\pi a^2/4} = \frac{1}{4}.$$

Pregunta 123. (Examen - Otoño 2022, P1)

Ennio está realizando un estudio en el cual analiza para cada día: el tráfico vehicular, el cual se mide en una escala de 0 a 2, y el nivel de contaminación, el cual se mide en una escala de 0 a 1. Luego de recopilar información de variadas fuentes, Ennio ha determinado que la densidad de probabilidad conjunta entre ambos elementos descritos es

$$f_{T,C}(t, c) = \begin{cases} \frac{9}{10}tc^2 + \frac{1}{5}, & 0 \leq t \leq 2, 0 \leq c \leq 1, \\ 0, & \text{en caso contrario,} \end{cases}$$

donde la VA T representa el nivel de tráfico vehicular en un día y la VA C representa el nivel de contaminación en el mismo día. Suponga que el estudio es cierto, en el sentido de que los días futuros al estudio son tales que ambas variables aleatorias siguen la distribución de probabilidad descrita. Asuma que lo que ocurre en un día, respecto al tráfico y contaminación, es independiente de lo que ocurre respecto a las realizaciones de dichas VAs en el resto de los días.

- (a) Ennio está interesado en la distribución de probabilidad del tráfico en un día cuya contaminación es de $2/3$. Ayude a Ennio a resolver el problema. Específicamente, dado que la contaminación es de $2/3$, encuentre la distribución de probabilidad del tráfico del día.

- (b) Se sabe que el tiempo de traslado en minutos que le toma a don Ennio desde su casa a su trabajo, en auto, es de $20 + 4e^T$, donde T es el nivel de tráfico del día. Determine la función de densidad de probabilidad del tiempo que le tomará a don Ennio trasladarse a su trabajo en un día cualquiera.
- (c) Suponga que Ennio ha clasificado los días de la siguiente manera: si el nivel de contaminación es mayor o igual a $5/6$, entonces era un día para *no hacer deporte*. Si, por el contrario, el nivel de contaminación es menor que $5/6$ y además el nivel de tráfico es menor que $4/5$, entonces era un día para *andar en bicicleta*, dado que no hay tráfico. En el último caso, el restante, el día era para *trotar*. Determine la probabilidad de que en un mes de 30 días haya habido exactamente 12 días para *no hacer deporte*, 10 días para *andar en bicicleta* y 8 días para *trotar*. Para esta parte puede dejar expresiones matemáticas, sin necesidad de encontrar el valor exacto. No obstante, debe quedar claro qué significa cada elemento de las expresiones que incluya en su respuesta.

Solución

(a). Para $0 \leq c \leq 1$,

$$f_C(c) = \int_0^2 \left(\frac{9}{10}tc^2 + \frac{1}{5} \right) dt = \frac{9}{5}c^2 + \frac{2}{5}.$$

Por consiguiente,

$$f_{T|C}(t | 2/3) = \begin{cases} \frac{2t+1}{6}, & 0 \leq t \leq 2, \\ 0, & \text{en otro caso.} \end{cases}$$

(b). Primero,

$$f_T(t) = \int_0^1 \left(\frac{9}{10}tc^2 + \frac{1}{5} \right) dc = \frac{3}{10}t + \frac{1}{5}, \quad 0 \leq t \leq 2.$$

Sea $Y = 20 + 4e^T$. Esta transformación es creciente, $g^{-1}(y) = \ln((y-20)/4)$ y $(g^{-1})'(y) = 1/(y-20)$. Entonces

$$f_Y(y) = \begin{cases} \frac{\frac{3}{10} \ln\left(\frac{y-20}{4}\right) + \frac{1}{5}}{y-20}, & 24 \leq y \leq 20 + 4e^2, \\ 0, & \text{en otro caso.} \end{cases}$$

(c). Las probabilidades de las tres categorías son

$$\begin{aligned} p_n &= \int_{5/6}^1 \int_0^2 f_{T,C}(t, c) dt dc = \frac{23}{72}, \\ p_a &= \int_0^{5/6} \int_0^{4/5} f_{T,C}(t, c) dt dc = \frac{17}{90}, \\ p_t &= 1 - p_n - p_a = \frac{59}{120}. \end{aligned}$$

Por la independencia entre días, los conteos siguen una distribución multinomial. La probabilidad solicitada es

$$\binom{30}{12, 10, 8} p_n^{12} p_a^{10} p_t^8.$$

Pregunta 124. (Examen - Primavera 2022, P2)

- a) Ximena e Yvonne, amantes de los llaveros, deciden ir a visitar una feria de las pulgas con el propósito de vender parte de su colección y generar algo de dinero. Suponga que en dicha feria, se pueden vender o comprar llaveros. Debido al corto tiempo en el que están en la feria, cada una logra o bien comprar o vender a lo más 1 llavero, o bien ninguna de las anteriores (en cuyo caso no se compra ni vende). Sean X e Y las variables aleatorias que indican la **cantidad de llaveros vendidos** por Ximena e Yvonne, respectivamente (notar que si alguna de estas VA es igual a -1 , entonces corresponde a una compra de un llavero). La función de probabilidad de masa conjunta de que Ximena venda a llaveros e Yvonne venda b llaveros está dada por la siguiente tabla:

		$X = a$		
		-1	0	1
$Y = b$	-1	2/45	9/45	4/45
	0	7/45	5/45	3/45
	1	6/45	1/45	8/45

Cuadro 7.1: Valores de $p_{XY}(a, b)$.

Donde vender -1 llaveros significa haber comprado un llavero en la feria, 0 significa que no se ha comprado ni vendido, y 1 significa que se ha vendido.

- (I) Calcule la covarianza entre las VA asociadas a la venta de llaveros de Ximena (X) e Yvonne (Y). ¿Son X e Y independientes? Justifique su respuesta.
- (II) Sabiendo que Yvonne vendió un llavero, ¿Es más probable que Ximena compre o venda un llavero?
- (III) Ximena e Yvonne siguen yendo a la feria durante 7 días. La PMF conjunta de lo que ocurre cada día está dada por la tabla de la parte a) (i). La cantidad de días que ocurrió que Ximena vende a llaveros e Yvonne vende b llaveros en un día está dada por la siguiente tabla de frecuencias:

		$X = a$		
		-1	0	1
$Y = b$	-1	0	3	0
	0	1	0	0
	1	1	0	2

Por ejemplo, hay 3 días en que Ximena no vende ni compra nada e Yvonne compra un llavero. Calcule la probabilidad asociada al escenario descrito por la tabla.

- b) La densidad de una v.a $T \sim \text{Triangular}[0, 2]$ está dada por la función:

$$f_T(t) = \begin{cases} t & \text{si } t \in [0, 1] \\ 2 - t & \text{si } t \in (1, 2] \\ 0 & \text{si no.} \end{cases}$$

- (i) Calcule la varianza de T .

- (II) Para $Y = T + 1$, compare las desviaciones estándar de Y con la de T . ¿Qué puede decir al respecto?

Solución

(a)(i). Al sumar la tabla sobre los valores de la otra variable se obtiene

$$p_X(x) = p_Y(x) = \frac{1}{3}, \quad x \in \{-1, 0, 1\}.$$

Por simetría, $\mathbb{E}[X] = \mathbb{E}[Y] = 0$. Para el momento cruzado solo contribuyen las cuatro celdas en que ninguna variable vale cero:

$$\begin{aligned} \mathbb{E}[XY] &= \sum_{x=-1}^1 \sum_{y=-1}^1 xy p_{X,Y}(x, y) \\ &= \frac{2 - 6 - 4 + 8}{45} \\ &= 0. \end{aligned}$$

Por tanto,

$$\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = 0.$$

Una covarianza nula no basta para concluir independencia. En este caso,

$$p_{X,Y}(-1, -1) = \frac{2}{45} \neq p_X(-1)p_Y(-1) = \frac{1}{9} = \frac{5}{45},$$

por lo que X e Y no son independientes.

(a)(ii). Condicionado en que Yvonne vendió un llavero, es decir, en $Y = 1$, las probabilidades relevantes son

$$\begin{aligned} p_{X|Y}(1 | 1) &= \frac{p_{X,Y}(1, 1)}{p_Y(1)} = \frac{8/45}{1/3} = \frac{8}{15}, \\ p_{X|Y}(-1 | 1) &= \frac{p_{X,Y}(-1, 1)}{p_Y(1)} = \frac{6/45}{1/3} = \frac{2}{5}. \end{aligned}$$

Como $8/15 > 2/5$, es más probable que Ximena venda un llavero a que compre uno.

(a)(iii). Para cada $(x, y) \in \{-1, 0, 1\}^2$, sea $N_{x,y}$ el número de días, entre los siete, en que ocurre el resultado $(X, Y) = (x, y)$. Suponiendo independencia entre días, el vector de las nueve frecuencias sigue una distribución multinomial con siete ensayos y probabilidades dadas por la tabla. En el escenario observado, las únicas frecuencias positivas son

$$N_{0,-1} = 3, \quad N_{-1,0} = 1, \quad N_{-1,1} = 1, \quad N_{1,1} = 2.$$

Por consiguiente, su probabilidad es

$$\begin{aligned} &\frac{7!}{3!1!1!2!} \left(\frac{9}{45}\right)^3 \left(\frac{7}{45}\right) \left(\frac{6}{45}\right) \left(\frac{8}{45}\right)^2 \\ &= 420 \left(\frac{9}{45}\right)^3 \left(\frac{7}{45}\right) \left(\frac{6}{45}\right) \left(\frac{8}{45}\right)^2 \approx 0,00220. \end{aligned}$$

El coeficiente multinomial cuenta las formas de ubicar en los siete días los cuatro tipos de resultado con frecuencias 3, 1, 1 y 2.

(b)(i). La varianza se obtiene a partir de los dos primeros momentos. Usando la densidad por tramos,

$$\begin{aligned} \mathbb{E}[T] &= \int_0^1 t^2 dt + \int_1^2 t(2-t) dt = 1, \\ \mathbb{E}[T^2] &= \int_0^1 t^3 dt + \int_1^2 t^2(2-t) dt = \frac{7}{6}. \end{aligned}$$

Luego,

$$\text{Var}(T) = \mathbb{E}[T^2] - \mathbb{E}[T]^2 = \frac{1}{6}.$$

La misma respuesta se obtiene observando que T tiene la distribución de $U + V$, donde U y V son uniformes independientes en $[0, 1]$:

$$\text{Var}(T) = \text{Var}(U) + \text{Var}(V) = \frac{1}{12} + \frac{1}{12} = \frac{1}{6}.$$

(b)(ii). Sumar una constante desplaza una variable aleatoria, pero no modifica su dispersión. Como $Y = T + 1$,

$$\text{Var}(Y) = \text{Var}(T + 1) = \text{Var}(T).$$

Al tomar raíces cuadradas, las desviaciones estándar de Y y T coinciden; ambas son $1/\sqrt{6}$.

Pregunta 125. (Examen - Otoño 2023, P3)

Parte 1. Demuestre, matemáticamente o con argumentos válidos, la veracidad de las siguientes afirmaciones. En caso de ser falsa, puede presentar un contraejemplo.

- (a) Sean X e Y variables aleatorias independientes. Entonces $\text{Var}(Y - X) = \text{Var}(Y) + \text{Var}(X)$.
- (b) Sean X e Y variables aleatorias independientes. La varianza de $\mathbb{E}[X | Y]$ es siempre 0, pues $\mathbb{E}[X | Y]$ es un número real.
- (c) Sean $X_1, \dots, X_n$ variables aleatorias. La igualdad

$$\mathbb{E}[X_1 + \dots + X_n] = \mathbb{E}[X_1] + \dots + \mathbb{E}[X_n]$$

se cumple solo si las variables son independientes.

Parte 2. En el pequeño pueblo de Charlestown existe una gran avenida de largo L , famosa por la cantidad de accidentes que ocurren en ella. Por esto, el Hospital de Charlestown, cuya entrada está en dicha avenida, ha destinado una ambulancia que viaja exclusivamente por esta calle a velocidad constante. En cierto momento ocurre un accidente en este camino en un punto que distribuye de manera uniforme; es decir, la ubicación donde ocurre el accidente sigue una distribución Uniforme(0, L), donde el punto 0 corresponde al inicio de la avenida y el punto L corresponde al final de la avenida. Asuma que la ubicación de la ambulancia al momento del accidente también distribuye uniforme a lo largo de la avenida y de forma independiente a la ubicación del accidente.

- (a) Calcule la función de distribución acumulada de la distancia entre la ambulancia y el punto del accidente, es decir, calcule $\mathbb{P}(|X - Y| \leq a)$, con Y y X el punto del accidente y el de la ambulancia, respectivamente.

Hint: Puede servirle mostrar que

$$\mathbb{P}(|Y - X| \leq a) = \mathbb{P}(X \leq Y + a, Y < X) + \mathbb{P}(Y \leq X + a, X \leq Y).$$

- (b) Calcule la distancia esperada entre la ambulancia y el punto del accidente.

Solución

(1)(a). La afirmación es verdadera. Por independencia, $\text{Cov}(X, Y) = 0$, y

$$\text{Var}(Y - X) = \text{Var}(Y) + \text{Var}(X) - 2\text{Cov}(Y, X) = \text{Var}(Y) + \text{Var}(X).$$

(1)(b). La conclusión es verdadera bajo la independencia supuesta, aunque la justificación

enunciada no es correcta en general: una esperanza condicional suele ser una variable aleatoria. Aquí,

$$\mathbb{E}[X | Y] = \mathbb{E}[X] \quad \text{c.s.,}$$

que sí es constante, y por eso $\text{Var}(\mathbb{E}[X | Y]) = 0$.

(1)(c). La afirmación es falsa. La linealidad de la esperanza no requiere independencia, siempre que las esperanzas involucradas existan.

(2)(a). La densidad conjunta es $1/L^2$ sobre $[0, L]^2$. Para $0 < a < L$, la región $|x - y| \leq a$ es el cuadrado menos dos triángulos de cateto $L - a$, por lo que

$$\mathbb{P}(|X - Y| \leq a) = 1 - \frac{(L - a)^2}{L^2} = \frac{a}{L} \left(2 - \frac{a}{L}\right).$$

En consecuencia, la CDF de $D = |X - Y|$ es

$$F_D(a) = \begin{cases} 0, & a < 0, \\ \frac{a}{L} \left(2 - \frac{a}{L}\right), & 0 \leq a \leq L, \\ 1, & a \geq L. \end{cases}$$

(2)(b). Integrando sobre el cuadrado,

$$\mathbb{E}[|X - Y|] = \frac{2}{L^2} \int_0^L \int_0^x (x - y) dy dx = \frac{L}{3}.$$

Pregunta 126. (Examen - Primavera 2023, P2)

Azul es una fábrica que produce azulejos de forma rectangular. Una vez al año fabrican 1000 piezas iguales de un azulejo de *edición limitada*. Dicho azulejo es creado con dimensiones al azar. Específicamente, el **ancho** del azulejo distribuye uniforme entre 0 y 2 decímetros (dm), y el **largo** distribuye uniforme entre 0 y 2 dm.

- (1) (a) Suponga que, para esta parte, el **ancho** es independiente del **largo**. Encuentre la probabilidad de que el azulejo posea un área menor que 1 dm^2 .
- (b) Suponga que ahora se desea que el azulejo edición limitada sea más cuadrado, es decir, que sean más similares las dimensiones del **ancho** y el **largo**. Para lograr esto, la fábrica decide cambiar su forma de dar dimensión a sus azulejos. Específicamente, el **ancho** del azulejo distribuye uniforme entre 0 y 2 dm, y el **largo** toma un valor también uniforme entre la mitad del **ancho** y tres medios el **ancho**. Determine el área esperada de un azulejo.
- (2) Para cada una de las siguientes afirmaciones indique si es verdadera o falsa. Debe justificar su respuesta en cada caso, ya sea entregando una demostración de que lo afirmado es verdadero o dando un contraejemplo en caso de que sea falso.
 - (a) La esperanza del logaritmo de una variable aleatoria es siempre menor o igual que el logaritmo de su esperanza.
 - (b) Si la correlación entre dos variables aleatorias es estrictamente positiva, entonces también lo es su covarianza.

Solución

(1)(a). Por independencia, la densidad conjunta es $1/4$ sobre $[0, 2]^2$. Separando según si $1/y$

supera 2,

$$\begin{aligned}\mathbb{P}(XY < 1) &= \frac{1}{4} \int_0^2 \min\left\{2, \frac{1}{y}\right\} dy \\ &= \frac{1}{4} \int_0^{1/2} 2 dy + \frac{1}{4} \int_{1/2}^2 \frac{1}{y} dy = \frac{1}{4} + \frac{1}{4} \ln 4.\end{aligned}$$

(1)(b). Como $\mathbb{E}[Y | X] = X$, la esperanza iterada entrega

$$\mathbb{E}[XY] = \mathbb{E}[X \mathbb{E}[Y | X]] = \mathbb{E}[X^2] = \text{Var}(X) + \mathbb{E}[X]^2 = \frac{1}{3} + 1 = \frac{4}{3}.$$

(2)(a). Tal como está escrita, la afirmación requiere condiciones de dominio e integrabilidad: si X puede ser no positiva, $\log X$ ni siquiera está definida como variable real. Si $X > 0$ casi seguramente y ambas cantidades son finitas, entonces sí es verdadera, pues $\log$ es cóncava y Jensen da

$$\mathbb{E}[\log X] \leq \log(\mathbb{E}[X]).$$

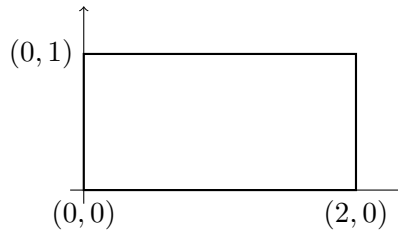
(2)(b). La afirmación es verdadera. Cuando la correlación está definida, ambas varianzas son positivas y

$$\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \text{Var}(Y)}}.$$

El denominador es positivo, de modo que correlación y covarianza tienen el mismo signo.

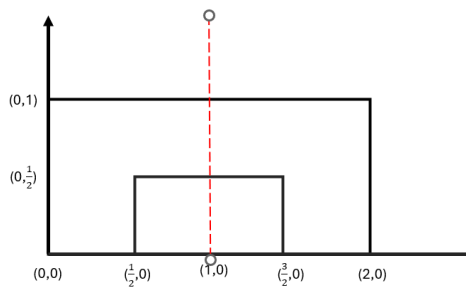
Pregunta 127. (Examen - Otoño 2024, P2)

Violeta es una niña pequeña que está entrenando para ser delantera de fútbol. Su entrenamiento consiste en practicar tiros a un arco de 2 metros de ancho y 1 metro de alto. Violeta es zurda, por lo que la probabilidad de que el balón entre en la mitad derecha del arco es 2 veces la probabilidad de que entre en la mitad izquierda. En cada mitad del arco, la probabilidad se distribuye uniformemente. Considere que la esquina inferior izquierda del arco está situada en el origen.



(a) Calcule la PDF conjunta del punto (X, Y) en el que la pelota entra al arco.

En el entrenamiento también se encuentra Camila, que es arquera y cubre un área rectangular en el centro del arco de 1 metro de ancho por 0,5 metros de alto. Con probabilidad 1, si la pelota entra en el área de la arquera, esta la atrapa.



(b) Calcule la probabilidad de que Violeta haga un gol.

Solución

(a). Sea c la densidad en la mitad izquierda; en la derecha es $2c$. Como ambas mitades tienen área 1,

$$c + 2c = 1, \quad c = \frac{1}{3}.$$

Por tanto,

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{3}, & 0 \leq x < 1, \quad 0 \leq y \leq 1, \\ \frac{2}{3}, & 1 \leq x \leq 2, \quad 0 \leq y \leq 1, \\ 0, & \text{en otro caso.} \end{cases}$$

(b). Según la figura, el área cubierta por la arquera es $\{(x,y) : 1/2 \leq x \leq 3/2, \quad 0 \leq y \leq 1/2\}$. Así,

$$\begin{aligned} \mathbb{P}(\text{no gol}) &= \int_0^{1/2} \int_{1/2}^1 \frac{1}{3} dx dy + \int_0^{1/2} \int_1^{3/2} \frac{2}{3} dx dy \\ &= \frac{1}{4}. \end{aligned}$$

En consecuencia, $\mathbb{P}(\text{gol}) = 1 - \frac{1}{4} = \frac{3}{4}$.

Pregunta 128. (Examen - Otoño 2025, P2)

- En el curso de Probabilidades, los profesores Charles y Felipe realizan **dos controles**. Cada control consiste en 70 preguntas de alternativas, y la nota de un control corresponde al número de respuestas correctas dividido por 10 (sin punto base). Sofía, alumna del curso, responde cada pregunta del primer control correctamente con probabilidad 0,7, de forma independiente entre preguntas. Sin embargo, ella sabe que su rendimiento en el primer control influirá en cuánto estudia para el segundo: si obtiene una baja nota, estudiará más; si le va bien, estudiará menos. En particular, si Sofía responde X preguntas correctamente en el primer control, entonces la probabilidad de que responda correctamente una pregunta del segundo control es $1 - \frac{X}{70}$. Las respuestas del segundo control también son independientes entre sí. La nota del segundo control se calcula del mismo modo: número de respuestas correctas dividido por 10. El curso se aprueba si el promedio de las dos notas es mayor o igual a 3,95.
 - Determine la distribución conjunta de las notas de ambos controles. Si la distribución es conocida, es decir, vista en clases, entregue el nombre y valor de sus parámetros; en caso contrario, entregue su función de masa de probabilidad. Justifique su respuesta.
 - Calcule la probabilidad de que Sofía apruebe el curso, considerando que ha obtenido una nota de 4,2. Puede dejar su respuesta

2. Un ascensor en un edificio puede ser abordado por una cantidad aleatoria de personas. La cantidad de personas que entran en el ascensor en cada viaje sigue una distribución de Poisson con media 3. El peso de cada persona, es también una variable aleatoria, independiente del número de personas que sube al ascensor y del peso de los demás pasajeros, que sigue una distribución Uniforme entre 50 y 100 kg. Encuentre la varianza de la suma de los pesos de las personas que van adentro del ascensor.

Solución

(1)(a). Sea X el número de respuestas correctas del primer control y sea Y el del segundo. Entonces

$$X \sim \text{Binomial}(70, 0,7), \quad Y | X = x \sim \text{Binomial}\left(70, 1 - \frac{x}{70}\right).$$

La dependencia entre ambos controles se incorpora mediante la distribución condicional de Y . Por la regla de la cadena, para $x, y \in \{0, \dots, 70\}$,

$$\begin{aligned} p_{X,Y}(x, y) &= p_X(x) p_{Y|X}(y | x) \\ &= \binom{70}{x} (0,7)^x (0,3)^{70-x} \binom{70}{y} \left(1 - \frac{x}{70}\right)^y \left(\frac{x}{70}\right)^{70-y}. \end{aligned}$$

Fuera de ese soporte, la PMF vale cero. Si $N_1 = X/10$ y $N_2 = Y/10$ son las notas, entonces

$$p_{N_1, N_2}\left(\frac{x}{10}, \frac{y}{10}\right) = p_{X,Y}(x, y),$$

para los mismos valores enteros de x e y .

(1)(b). Una nota de 4,2 en el primer control equivale a $X = 42$. En tal caso,

$$Y | X = 42 \sim \text{Binomial}(70, 0,4).$$

El curso se aprueba cuando

$$\frac{X/10 + Y/10}{2} \geq 3,95,$$

equivalentemente, cuando $X + Y \geq 79$. Condicionado en $X = 42$, esto exige $Y \geq 37$. Por tanto,

$$\mathbb{P}(\text{aprobar} | X = 42) = \sum_{y=37}^{70} \binom{70}{y} (0,4)^y (0,6)^{70-y}.$$

(2). Sea $N \sim \text{Poisson}(3)$ el número de pasajeros y sean $W_1, W_2, \dots$ sus pesos, independientes e idénticamente distribuidos como Uniforme(50, 100) e independientes de N . El peso total es

$$S = \sum_{i=1}^N W_i,$$

con la suma igual a cero cuando $N = 0$. Para el peso individual,

$$\mathbb{E}[W_i] = 75, \quad \text{Var}(W_i) = \frac{(100 - 50)^2}{12} = \frac{625}{3}.$$

Condicionado en N ,

$$\mathbb{E}[S | N] = 75N, \quad \text{Var}(S | N) = \frac{625}{3}N.$$

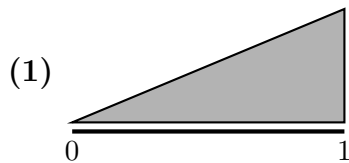
Como una variable Poisson de media tres tiene esperanza y varianza iguales a tres, la ley de la varianza total entrega

$$\begin{aligned} \text{Var}(S) &= \mathbb{E}[\text{Var}(S | N)] + \text{Var}(\mathbb{E}[S | N]) \\ &= \frac{625}{3} \mathbb{E}[N] + 75^2 \text{Var}(N) \\ &= 625 + 16\,875 \\ &= 17\,500. \end{aligned}$$

Pregunta 129. (Examen - Otoño 2026, P1)

Responda las siguientes preguntas.

- (1) Una ciudad se modela como un segmento entre las posiciones 0 y 1. La ubicación donde se inicia un incendio tiene densidad de probabilidad con la forma indicada en la figura (1), mientras que la ubicación del camión de bomberos al momento de una alarma tiene densidad con la forma indicada en la figura (2). Ambas ubicaciones son independientes. Los bomberos deben llegar al incendio en un máximo de 5 minutos. El camión se desplaza a velocidad constante de 0,1 unidades de distancia por minuto. Calcule la probabilidad de que lleguen a tiempo.



- (2) La Copa Mundial de la FIFA 2026 amplió los cupos a 48 selecciones. En la fase de grupos, los equipos se dividen en 12 grupos de 4 integrantes. De cada grupo avanzan a dieciseisavos de final los dos primeros y los ocho mejores terceros. En dieciseisavos se juegan 16 partidos entre los 32 clasificados, cada equipo juega una sola vez y no hay empates.
- (a) Considerando exclusivamente la fase de grupos, con los grupos ya conformados por cuatro equipos fijos, ¿cuántas configuraciones diferentes existen respecto de las selecciones que pasan a dieciseisavos de final?
- (b) Una vez definidos los 32 equipos y los cruces de dieciseisavos, ¿de cuántas formas pueden darse los resultados que determinan a los 16 clasificados a octavos de final?

Solución

(1). Sea X la ubicación del incendio y Y la del camión. La normalización del triángulo de la figura (1) da $f_X(x) = 2x$ para $0 \leq x \leq 1$, mientras que la figura (2) da $f_Y(y) = 1$ en el mismo intervalo. Por independencia,

$$f_{X,Y}(x,y) = 2x \mathbb{1}_{\{0 \leq x,y \leq 1\}}.$$

Llegar en a lo más cinco minutos equivale a

$$\frac{|X - Y|}{0,1} \leq 5 \iff |X - Y| \leq \frac{1}{2}.$$

Para $0 \leq x \leq 1/2$, los valores admisibles de y forman el intervalo $[0, x + 1/2]$; para $1/2 \leq x \leq 1$, forman $[x - 1/2, 1]$. Por tanto,

$$\begin{aligned} \mathbb{P}(|X - Y| \leq 1/2) &= \int_0^{1/2} 2x \left(x + \frac{1}{2}\right) dx + \int_{1/2}^1 2x \left(\frac{3}{2} - x\right) dx \\ &= \frac{5}{24} + \frac{13}{24} = \frac{3}{4}. \end{aligned}$$

(2)(a). Se eligen los ocho grupos que aportan un tercer clasificado. En cada uno de ellos se eligen tres de sus cuatro equipos y, en cada grupo restante, dos de cuatro. El número de configuraciones es

$$\binom{12}{8} \binom{4}{3}^8 \binom{4}{2}^4.$$

(2)(b). Cada uno de los 16 partidos tiene dos resultados posibles y los cruces están fijados. Por el principio multiplicativo, existen 2^{16} configuraciones.

Capítulo 8

Propiedades de la Esperanza

8.1. Esperanza condicional, sumas y transformadas

Resumen

Esperanza condicional.

Condicionar en $Y = y$ reemplaza la ley de X por su ley condicional:

$$\mathbb{E}[g(X) | Y = y] = \sum_x g(x) p_{X|Y}(x | y),$$

$$\mathbb{E}[g(X) | Y = y] = \int_{\mathbb{R}} g(x) f_{X|Y}(x | y) dx,$$

en los casos discreto y continuo, respectivamente. Como variable aleatoria, $\mathbb{E}[X | Y]$ es una función de Y .

- **Linealidad.** $\mathbb{E}[aX + bZ | Y] = a\mathbb{E}[X | Y] + b\mathbb{E}[Z | Y]$.
- **Información conocida.** $\mathbb{E}[h(Y)X | Y] = h(Y)\mathbb{E}[X | Y]$ y $\mathbb{E}[h(Y) | Y] = h(Y)$.
- **Monotonía.** Si $X \leq Z$ casi seguramente, entonces $\mathbb{E}[X | Y] \leq \mathbb{E}[Z | Y]$ casi seguramente.
- **Independencia.** Si X e Y son independientes, $\mathbb{E}[X | Y] = \mathbb{E}[X]$.

Varianza condicional.

$$\begin{aligned}\text{Var}(X | Y) &= \mathbb{E}[(X - \mathbb{E}[X | Y])^2 | Y] \\ &= \mathbb{E}[X^2 | Y] - \mathbb{E}[X | Y]^2.\end{aligned}$$

Leyes totales.

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X | Y]],$$

$$\text{Var}(X) = \mathbb{E}[\text{Var}(X | Y)] + \text{Var}(\mathbb{E}[X | Y]).$$

La varianza total separa la variabilidad interna de la variabilidad entre medias condicionales.

Sumas de variables aleatorias.

La linealidad no requiere independencia:

$$\mathbb{E}\left[\sum_i X_i\right] = \sum_i \mathbb{E}[X_i].$$

Para varianzas finitas,

$$\text{Var}\left(\sum_i X_i\right) = \sum_i \text{Var}(X_i) + 2 \sum_{i < j} \text{Cov}(X_i, X_j).$$

Convolución.

Si X e Y son independientes, la ley de $X + Y$ se obtiene mediante

$$p_{X+Y}(z) = \sum_x p_X(x) p_Y(z - x),$$

$$f_{X+Y}(z) = \int_{\mathbb{R}} f_X(x) f_Y(z - x) dx,$$

en los casos discreto y continuo, respectivamente. El soporte de la suma debe determinarse antes de fijar los límites.

Sumas aleatorias.

Sea N una variable entera no negativa, independiente de las variables i.i.d. $X_1, X_2, \dots$, y defina $S_N = \sum_{i=1}^N X_i$, con $S_0 = 0$. Entonces

$$\mathbb{E}[S_N] = \mathbb{E}[N] \mathbb{E}[X_1],$$

$$\text{Var}(S_N) = \mathbb{E}[N] \text{Var}(X_1) + \text{Var}(N) \mathbb{E}[X_1]^2.$$

Se obtienen condicionando en N ; sin independencia deben usarse las leyes condicionales correspondientes.

FGM y función característica.

$$M_X(t) = \mathbb{E}[e^{tX}], \quad \varphi_X(t) = \mathbb{E}[e^{itX}].$$

- **Existencia.** La función característica siempre existe; la FGM puede no ser finita fuera de $t = 0$.
- **Momentos.** Si M_X es finita en un entorno de cero, $M_X^{(k)}(0) = \mathbb{E}[X^k]$.
- **Sumas independientes.** Si $X_1, \dots, X_n$ son independientes, $M_{\sum_i X_i}(t) = \prod_i M_{X_i}(t)$ y $\varphi_{\sum_i X_i}(t) = \prod_i \varphi_{X_i}(t)$.
- **Unicidad.** Cuando existe en un entorno de cero, la FGM determina la distribución; la función característica siempre la determina.

8.1.1. Preguntas

Pregunta 130.

Sea p un número escogido al azar en $[0, 1]$. Se lanza una moneda con probabilidad de cara igual a p , n veces independientemente. Se define Y como la variable aleatoria igual al número de caras en los lanzamientos.

(a) Averigüe si Y es una variable aleatoria discreta o continua.

(b) Calcule $\mathbb{E}[Y]$ y $\text{Var}(Y)$.

Solución

Sea $U \sim \text{Unif}(0, 1)$ el valor escogido para la probabilidad de cara. Esta es la interpretación uniforme de escoger p al azar en $[0, 1]$. Condicionalmente a $U = u$,

$$Y \mid (U = u) \sim \text{Binomial}(n, u).$$

(a). Aunque U es continua, Y cuenta el número de caras y solo puede tomar valores en $\{0, 1, \dots, n\}$; por tanto, Y es discreta. Para identificar su distribución marginal, fijemos $k \in \{0, \dots, n\}$. Por la ley de probabilidades totales,

$$\begin{aligned} \mathbb{P}(Y = k) &= \int_0^1 \mathbb{P}(Y = k \mid U = u) du \\ &= \binom{n}{k} \int_0^1 u^k (1 - u)^{n-k} du \\ &= \binom{n}{k} B(k + 1, n - k + 1). \end{aligned}$$

La integral es la función beta evaluada en enteros. Como

$$B(k + 1, n - k + 1) = \frac{k!(n - k)!}{(n + 1)!},$$

se obtiene

$$\mathbb{P}(Y = k) = \frac{1}{n + 1}.$$

Así, Y es uniforme sobre $\{0, 1, \dots, n\}$; en particular, no es binomial marginalmente.

(b). Los momentos condicionales de una variable binomial son

$$\mathbb{E}[Y \mid U] = nU, \quad \text{Var}(Y \mid U) = nU(1 - U).$$

Como $U \sim \text{Unif}(0, 1)$,

$$\mathbb{E}[U] = \frac{1}{2}, \quad \mathbb{E}[U(1 - U)] = \frac{1}{6}, \quad \text{Var}(U) = \frac{1}{12}.$$

La ley de la esperanza total entrega

$$\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y \mid U]] = n\mathbb{E}[U] = \frac{n}{2}.$$

La ley de la varianza total separa la variabilidad de los lanzamientos y la del parámetro escogido:

$$\begin{aligned} \text{Var}(Y) &= \mathbb{E}[\text{Var}(Y \mid U)] + \text{Var}(\mathbb{E}[Y \mid U]) \\ &= n\mathbb{E}[U(1 - U)] + n^2 \text{Var}(U) \\ &= \frac{n}{6} + \frac{n^2}{12} = \frac{n(n + 2)}{12}. \end{aligned}$$

Pregunta 131.

Se tiene una barra de largo L a la cual se le quieren hacer dos cortes al azar. Sea $X \sim \text{Unif}(0, L)$, la v.a. que representa el lugar del primer corte, e $Y | X = x \sim \text{Unif}(x, L)$, v.a. que representa el lugar del segundo corte entre el primero y el extremo L de la barra.

- a) Obtenga la función de densidad de Y .
b) Calcule $\mathbb{E}[Y]$.

Solución

(a). Condicionalmente a $X = x$, el segundo corte es uniforme en $[x, L]$. Por lo tanto,

$$f_{X,Y}(x, y) = \begin{cases} \frac{1}{L(L-x)}, & 0 \leq x < L, \ x \leq y \leq L, \\ 0, & \text{en otro caso.} \end{cases}$$

Para $0 \leq y < L$,

$$f_Y(y) = \int_0^y \frac{1}{L(L-x)} dx = \frac{1}{L} \ln\left(\frac{L}{L-y}\right).$$

Fuera de ese intervalo, $f_Y(y) = 0$.

(b). Como $Y | X \sim \text{Unif}(X, L)$, $\mathbb{E}[Y | X] = (X + L)/2$. Así,

$$\mathbb{E}[Y] = \mathbb{E}\left[\frac{X+L}{2}\right] = \frac{\mathbb{E}[X] + L}{2} = \frac{3L}{4}.$$

Pregunta 132.

Una empresa vende combustible, donde el precio es una v.a. $Y \sim \text{Unif}(0, 1)$. Hay un solo cliente, el que comprará una cantidad $X \sim \text{Exp}(\lambda)$. Determine el ingreso esperado de la empresa sabiendo que condicional en $Y = y$, X distribuye exponencial de parámetro λy .

Solución

Sea $R = YX$ el ingreso. El enunciado entrega simultáneamente la ley marginal $X \sim \text{Exp}(\lambda)$ y la ley condicional $X | Y = y \sim \text{Exp}(\lambda y)$; ambas especificaciones no son compatibles cuando $Y \sim \text{Unif}(0, 1)$. Interpretamos que la ley condicional, que describe explícitamente cómo depende la compra del precio, es la que define el modelo. Para $y > 0$, esta ley implica $\mathbb{E}[X | Y = y] = 1/(\lambda y)$. Por la ley de la esperanza total,

$$\begin{aligned} \mathbb{E}[R] &= \mathbb{E}[\mathbb{E}[YX | Y]] = \mathbb{E}[Y\mathbb{E}[X | Y]] \\ &= \mathbb{E}\left[Y \frac{1}{\lambda Y}\right] = \frac{1}{\lambda}. \end{aligned}$$

El valor de la ley condicional en $Y = 0$ no afecta el cálculo porque $\mathbb{P}(Y = 0) = 0$.

Si, en cambio, se descartara la ley condicional y se supusiera que $X \sim \text{Exp}(\lambda)$ es independiente de Y , se obtendría $\mathbb{E}[R] = 1/(2\lambda)$; la ley marginal por sí sola no determina el ingreso esperado.

Pregunta 133.

Usted acaba de ser contratado en un centro de datos de una empresa muy importante, su tarea es ver cuanto tarda el sistema en llegar al máximo de su capacidad, según lo que le cuenta su jefe, para que esto ocurra los dos servidores que distribuyen deben llegar a su límite. El tiempo que toma que esto suceda distribuye $\text{Exp}(\lambda)$ para cada uno de los servidores.

Suponga que los servidores no trabajan al mismo tiempo, es decir, cuando el servidor 1 llega a su límite, en ese punto recién empieza a funcionar el servidor 2.

- Calcule la FGM para el sistema completo
- Calcule la distribución que sigue el sistema completo
- ¿Cuál es el tiempo esperado en que los servidores llegan a su capacidad máxima?

Solución

Sean X_1 y X_2 los tiempos que tardan los servidores 1 y 2 en alcanzar su límite. Que funcionen sucesivamente determina que el tiempo total sea $Z = X_1 + X_2$. Para identificar la distribución de Z , interpretaremos, como es usual en este ejercicio, que X_1 y X_2 son independientes y que ambos distribuyen $\text{Exp}(\lambda)$. La operación sucesiva, por sí sola, no implica independencia; sin esta hipótesis, las distribuciones marginales no bastarían para determinar la ley de Z .

(a). Para $t < \lambda$, la FGM de cada tiempo es

$$\begin{aligned} M_{X_i}(t) &= \mathbb{E}[e^{tX_i}] \\ &= \int_0^\infty e^{tx} \lambda e^{-\lambda x} dx \\ &= \lambda \int_0^\infty e^{-(\lambda-t)x} dx = \frac{\lambda}{\lambda-t}. \end{aligned}$$

Por independencia, la FGM de la suma es el producto:

$$M_Z(t) = M_{X_1}(t)M_{X_2}(t) = \left(\frac{\lambda}{\lambda-t}\right)^2, \quad t < \lambda.$$

(b). Para $z \geq 0$, la densidad de la suma se obtiene mediante convolución:

$$\begin{aligned} f_Z(z) &= \int_0^z f_{X_1}(x)f_{X_2}(z-x) dx \\ &= \int_0^z \lambda e^{-\lambda x} \lambda e^{-\lambda(z-x)} dx \\ &= \lambda^2 e^{-\lambda z} \int_0^z dx \\ &= \lambda^2 z e^{-\lambda z}. \end{aligned}$$

Para $z < 0$, $f_Z(z) = 0$. Esta es la densidad de una distribución Gamma con forma 2 y tasa λ :

$$Z \sim \text{Gamma}(2, \lambda).$$

(c). La linealidad de la esperanza permite sumar los tiempos esperados:

$$\mathbb{E}[Z] = \mathbb{E}[X_1] + \mathbb{E}[X_2] = \frac{1}{\lambda} + \frac{1}{\lambda} = \frac{2}{\lambda}.$$

Este resultado también se obtiene derivando $M_Z(t)$ y evaluando en $t = 0$.

Pregunta 134.

Sean $X_i \forall i \in \{1, \dots, n\}$ v.a independientes y $Z = \sum_{i=1}^n X_i$. Utilizando FGM:

- Demuestre que si $X_i \sim \text{Poi}(\lambda_i)$, $Z \sim \text{Poi}(\sum_{i=1}^n \lambda_i)$
- Demuestre que si $X_i \sim \text{Bernoulli}(p)$, $Z \sim \text{Bin}(n, p)$

Solución

Sea $M_X(s) = \mathbb{E}[e^{sX}]$ la función generadora de momentos. **(a).** Si $X_i \sim \text{Poi}(\lambda_i)$, entonces

$$M_{X_i}(s) = \exp\{\lambda_i(e^s - 1)\}.$$

Por independencia,

$$M_Z(s) = \prod_{i=1}^n M_{X_i}(s) = \exp\left\{\left(\sum_{i=1}^n \lambda_i\right)(e^s - 1)\right\}.$$

Esta es la FGM de una variable $\text{Poi}(\sum_{i=1}^n \lambda_i)$.

(b). Si $X_i \sim \text{Bernoulli}(p)$, entonces $M_{X_i}(s) = 1 - p + pe^s$. Por independencia,

$$M_Z(s) = (1 - p + pe^s)^n,$$

que corresponde a $Z \sim \text{Bin}(n, p)$.

Pregunta 135.

Un encargado del Transantiago quiere medir el tiempo que se demoran en pasar los buses en un determinado paradero. Para ello, el encargado comienza a medir el tiempo con un cronómetro justo en el momento en que un bus sale del paradero, y detiene el cronómetro justo en el instante en que llega otro bus. El tiempo real entre ambos buses es una variable aleatoria T que distribuye Exponencial de parámetro $\lambda = 0,1$ [minutos] $^{-1}$, pero debido a errores de medición del encargado y a problemas de calibración del cronómetro, se tiene que, dado el valor de T , el tiempo medido por el encargado es una variable X que distribuye como una normal de media T y varianza 1.

- Calcule el valor esperado de X .
- Calcule la varianza de X .
- Calcule la covarianza entre X y T . *Hint: Puede servirle que $\mathbb{E}[XT] = \mathbb{E}[\mathbb{E}[XT | T]] = \mathbb{E}[\mathbb{E}[X | T] \cdot T]$*
- Use la desigualdad de Chebyshev para demostrar que la probabilidad de que *la magnitud del error de medición sea mayor o igual a 2 minutos* es menor o igual a 0.25. *Hint: Defina la variable $Z = T - X$.*

Solución

Como $T \sim \text{Exp}(0,1)$ y $X | T \sim \mathcal{N}(T, 1)$,

$$\mathbb{E}[T] = 10, \quad \text{Var}(T) = 100, \quad \mathbb{E}[X | T] = T, \quad \text{Var}(X | T) = 1.$$

(a). Por la ley de la esperanza total,

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X | T]] = \mathbb{E}[T] = 10.$$

(b). La ley de la varianza total entrega

$$\text{Var}(X) = \mathbb{E}[\text{Var}(X | T)] + \text{Var}(\mathbb{E}[X | T]) = 1 + 100 = 101.$$

(c). Como T queda fija al condicionar,

$$\mathbb{E}[XT] = \mathbb{E}[T\mathbb{E}[X | T]] = \mathbb{E}[T^2] = 200,$$

$$\text{Cov}(X, T) = \mathbb{E}[XT] - \mathbb{E}[X]\mathbb{E}[T] = 200 - 100 = 100.$$

(d). Para $Z = T - X$,

$$\mathbb{E}[Z] = 0, \quad \text{Var}(Z) = 100 + 101 - 2(100) = 1.$$

La magnitud del error es $|Z|$. Por Chebyshev,

$$\mathbb{P}(|T - X| \geq 2) = \mathbb{P}(|Z| \geq 2) \leq \frac{\text{Var}(Z)}{2^2} = \frac{1}{4}.$$

Las varianzas y la covarianza se expresan en minutos cuadrados.

8.1.2. Preguntas de Evaluación

Pregunta 136. (Control 3 - Primavera 2022, P2)

- (a) Sirius Black está escapando de Azkabán cuando llega a una sala donde tiene que escoger entre tres puertas para lograr escapar. La primera puerta lo lleva de regreso a la sala donde escogió la puerta y le toma dos días de viaje (es decir, luego de dos días de viaje, regresa al mismo lugar de partida). La segunda puerta también lo lleva de regreso a la sala donde escogió la puerta, pero ahora le toma cuatro días de viaje. Por último, la tercera puerta le permite escapar después de un día de viaje. Si se asume que Sirius siempre escoge las puertas uno, dos, y tres con probabilidad 0.5, 0.3 y 0.2, respectivamente, ¿Cuál es el número esperado de días hasta que Sirius escape y llegue a la libertad? Note que en caso de que Sirius Black escoja una puerta que lo devuelve al mismo lugar, él vuelve a escoger una puerta (con las mismas probabilidades descritas). Así sucesivamente hasta lograr escapar. Asuma que las elecciones de las puertas en los diferentes intentos son independientes.

Hint 1: Podría serle útil definir una VA para el tiempo que le toma en salir de la sala, y otra VA sobre la puerta que escoge.

- (b) Usted es un acérrimo/a fan de Mario Bros y está a punto de finalizar el juego cuando le quedan dos niveles por superar. En el primer nivel, en cada intento usted lo supera con probabilidad p , mientras que en el segundo nivel, en cada intento usted lo supera con probabilidad q . Asuma que la probabilidad de pasar una etapa en cada intento es independiente del resto de los intentos. ¿Cuál es la probabilidad de que supere los dos niveles en exactamente 10 intentos?

Hint: Puede definir dos VA, una que indique el número de intentos usados para pasar la primera de las dos etapas mencionadas, y otra VA que indique el número de intentos realizados para pasar la segunda etapa.

Nota: No es necesario que simplifique expresiones matemáticas, vale decir, puede dejar expresado sumas, multiplicaciones, integrales, etc. (Note que la fórmula que entregue debe ser una fórmula con valores numéricos, no respuestas del estilo $\mathbb{P}(X = 10)$.)

Solución

(a). Sea D el número de días hasta escapar. Condicionando en la primera puerta escogida y usando que, al regresar a la sala, el problema vuelve a comenzar,

$$\mathbb{E}[D] = \frac{1}{2}(2 + \mathbb{E}[D]) + \frac{3}{10}(4 + \mathbb{E}[D]) + \frac{1}{5}.$$

Así, $(1/5)\mathbb{E}[D] = 12/5$ y $\mathbb{E}[D] = 12$ días.

(b). Sean $X \sim \text{Geom}(p)$ y $Y \sim \text{Geom}(q)$ los intentos requeridos en cada nivel. Por independencia,

$$\begin{aligned}\mathbb{P}(X + Y = 10) &= \sum_{y=1}^9 \mathbb{P}(X = 10 - y) \mathbb{P}(Y = y) \\ &= \sum_{y=1}^9 p(1-p)^{9-y} q(1-q)^{y-1}.\end{aligned}$$

Pregunta 137. (Control 3 - Otoño 2023, P3)

- (a) Comente la veracidad o falsedad de las siguientes afirmaciones. Justifique clara y brevemente su respuesta.
- (I) Al igual que sucede con la esperanza, la varianza de la suma de variables aleatorias es igual a la suma de las varianzas individuales de dichas variables aleatorias.
 - (II) Toda variable aleatoria posee función generadora de momento (FGM), la cual existe para todos los valores del rango de la VA.
 - (III) Sea $X \sim \text{Uniforme}(a, b)$ e $Y \sim \text{Uniforme}(c, d)$ variables aleatorias independientes, luego $X + Y$ sigue una distribución $\text{Uniforme}(a + c, b + d)$.
- (b) La empresa *GameWorld*, reconocida a nivel mundial en el mercado de los videojuegos, lanzó su nuevo videojuego de realidad virtual llamado *Nebulosa*. Con el objetivo de evaluar el rendimiento en las ventas de este juego, la empresa ha decidido registrar las ventas correspondientes a su tienda ubicada en Santiago.

Se ha establecido que cada cliente puede adquirir solo una unidad de *Nebulosa* y toma su decisión de manera independiente del resto. Según los estudios realizados, se estima que la probabilidad de que un cliente compre una unidad de este juego es p . Además, la reposición del producto es instantánea en caso de agotarse, por lo que ningún cliente se quedará sin poder comprarlo.

- (I) Identifique la distribución del total de unidades de *Nebulosa* que se venden durante el día si asisten n clientes ese día a la tienda.
- (II) El día sábado llegan m clientes a la tienda, mientras que el domingo llegan l clientes. Asuma que para cada día la venta de videojuegos es independiente y que la probabilidad p se mantiene constante. Sea X las unidades del videojuego vendidas el día sábado e Y las unidades vendidas el día domingo. Demuestre que el total de unidades de *Nebulosa* vendidas durante el fin de semana, representada por $X + Y$, sigue una distribución conocida; en particular, indique el nombre de dicha distribución y encuentre su(s) parámetro(s).

Nota: No se considera como demostración utilizar propiedades de la suma de variables aleatorias conocidas.

Solución

(a)(i). La afirmación es falsa en general. Para variables con segundos momentos finitos,

$$\text{Var} \left(\sum_{i=1}^n X_i \right) = \sum_{i=1}^n \text{Var}(X_i) + 2 \sum_{i < j} \text{Cov}(X_i, X_j).$$

La igualdad propuesta queda garantizada, por ejemplo, cuando las variables son independientes o, más generalmente, cuando son incorrelacionadas dos a dos.

(a)(ii). La afirmación es falsa. La FGM se define mediante $M_X(s) = \mathbb{E}[e^{sX}]$ y debe ser finita en un intervalo abierto que contenga a 0. Por ejemplo, una variable con distribución de Cauchy no posee FGM, pues $\mathbb{E}[e^{sX}]$ diverge para todo $s \neq 0$.

(a)(iii). La afirmación es falsa. Basta considerar X e Y independientes y uniformes en $[0, 1]$. La convolución entrega

$$f_{X+Y}(s) = \begin{cases} s, & 0 \leq s \leq 1, \\ 2 - s, & 1 < s \leq 2, \\ 0, & \text{en otro caso,} \end{cases}$$

que es una densidad triangular, no uniforme.

(b)(i). Si N es el total de unidades vendidas durante el día, cada cliente genera un ensayo Bernoulli independiente con probabilidad de compra p . Por tanto,

$$N \sim \text{Binomial}(n, p).$$

(b)(ii). Las ventas de cada día satisfacen $X \sim \text{Binomial}(m, p)$ e $Y \sim \text{Binomial}(l, p)$. Como X e Y son independientes, la FGM de la suma es

$$\begin{aligned} M_{X+Y}(s) &= M_X(s)M_Y(s) \\ &= (1 - p + pe^s)^m (1 - p + pe^s)^l \\ &= (1 - p + pe^s)^{m+l}. \end{aligned}$$

Esta es la FGM de una variable $\text{Binomial}(m + l, p)$, de modo que $X + Y \sim \text{Binomial}(m + l, p)$.

Pregunta 138. (Control 3 - Otoño 2024, P2)

El Parque Eólico de La Serena, localizado en la región de Coquimbo en Chile, desempeña un rol esencial en la generación de energía renovable para la localidad. El complejo incluye k aerogeneradores que transforman la energía cinética del viento en electricidad. Cada aerogenerador, independiente del resto, tiene una probabilidad p de estar operativo en un momento dado, mientras que la probabilidad de estar inactivo por mantenimiento o fallos técnicos es $1 - p$.

Los aerogeneradores funcionan capturando la corriente de aire, lo que a su vez hace girar sus aspas y genera energía. La cantidad de energía A_i producida por el i -ésimo aerogenerador, cuando está operativo, sigue una distribución Normal con media μ y desviación estándar σ . Por otra parte, en caso de que un aerogenerador esté inactivo, no producirá energía. La producción de energía de cada unidad es independiente de las demás.

- Calcule el valor esperado de la energía total producida cuando hay exactamente n aerogeneradores que están funcionando.
- Determine la esperanza de la energía total producida por el parque eólico.
- Calcule la varianza de la energía total producida.

Solución

Sea X el número de aerogeneradores operativos y A la energía total producida.

(a). Condicional en $X = n$, la energía total es la suma de n producciones independientes, cada una con esperanza μ . Luego,

$$\mathbb{E}[A \mid X = n] = n\mu.$$

(b). Como cada aerogenerador está operativo independientemente con probabilidad p , se tiene

$X \sim \text{Binomial}(k, p)$. La ley de la esperanza total da

$$\mathbb{E}[A] = \mathbb{E}[\mathbb{E}[A | X]] = \mathbb{E}[\mu X] = kp\mu.$$

(c). Dado X , la suma de X producciones independientes tiene varianza $X\sigma^2$. Usando la ley de la varianza total,

$$\begin{aligned}\text{Var}(A) &= \mathbb{E}[\text{Var}(A | X)] + \text{Var}(\mathbb{E}[A | X]) \\ &= \mathbb{E}[X\sigma^2] + \text{Var}(\mu X) \\ &= kp\sigma^2 + kp(1-p)\mu^2.\end{aligned}$$

Pregunta 139. (Control 3 - Primavera 2024, P2)

- (a) La aplicación Pedidos Ya utiliza un sistema de calificación para los pedidos, en una escala de valores enteros entre 0 y 5 estrellas (contando el 0 y el 5). Se ha observado que esta calificación, o *rating*, está influenciada por el tiempo de entrega: a mayor tiempo de espera, es más probable que los usuarios otorguen una menor cantidad de estrellas en su evaluación.

Para modelar esta situación, pueden definirse dos variables aleatorias. Sea T la variable aleatoria que representa el tiempo de entrega de un pedido, distribuida de forma Exponencial con un parámetro de tasa λ , y R la variable aleatoria que representa la calificación del pedido. En particular, la distribución del rating sigue una distribución Binomial con 5 intentos y probabilidad de éxito e^{-T} , donde T es la VA del tiempo de entrega. Calcule el rating esperado según la situación descrita.

- (b) Una planta química utiliza dos escalas de temperatura para monitorear el agua en sus experimentos. La temperatura del agua, que puede ser modelada como una VA, suele medirse en grados Celsius, representada con la variable C . Sin embargo, algunos reportes se realizan en grados Fahrenheit, F , tal que $F = 1.8C + 32$. Además, la planta mide otra variable aleatoria, X , que representa el nivel de actividad de una reacción química, que depende de la temperatura del agua. A lo largo de los experimentos, se ha observado y registrado una correlación entre X y C . El jefe de la planta necesita reportar la correlación entre X y F .

Ante esto, un practicante plantea la hipótesis de que la correlación entre X y C es igual a la correlación entre X y F , pues de lo contrario los experimentos medidos en Fahrenheit deberán repetirse. Se le ha contratado a usted para verificar la veracidad de esta hipótesis: demuestre su validez o, en caso contrario, justifique por qué esto no necesariamente se cumple.

Solución

(a). Condicional en T , el rating satisface $R | T \sim \text{Binomial}(5, e^{-T})$, por lo que $\mathbb{E}[R | T] = 5e^{-T}$. La ley de la esperanza total y la densidad de $T \sim \text{Exponencial}(\lambda)$ dan

$$\begin{aligned}\mathbb{E}[R] &= \mathbb{E}[\mathbb{E}[R | T]] = 5\mathbb{E}[e^{-T}] \\ &= 5 \int_0^\infty e^{-t} \lambda e^{-\lambda t} dt = \frac{5\lambda}{\lambda + 1}.\end{aligned}$$

(b). Suponiendo que las varianzas de X y C son positivas, las correlaciones están bien definidas. Como $F = 1.8C + 32$,

$$\text{Cov}(X, F) = 1.8 \text{Cov}(X, C), \quad \sigma_F = \sqrt{\text{Var}(F)} = 1.8\sigma_C.$$

Por tanto,

$$\text{Corr}(X, F) = \frac{1,8 \text{Cov}(X, C)}{\sigma_X(1,8\sigma_C)} = \text{Corr}(X, C),$$

de modo que la hipótesis es válida.

Pregunta 140. (Control 3 - Otoño 2025, P3)

- (a) Un distribuidor debe retirar un pedido en una tienda para luego entregarlo al domicilio del cliente. El proceso ocurre en dos etapas:

- **Etapla 1:** Al final del día repartidor va a la tienda a retirar el pedido. La tienda puede o no tener stock disponible, y la probabilidad de que haya stock en el momento del retiro es p . Se considera que el retiro es exitoso cuando el repartidor llega a la tienda y hay stock, en dicho caso pasa a la siguiente etapa. En caso contrario (de no ser exitoso) volverá a intentar retirar el pedido al día siguiente. La probabilidad de tener un retiro exitoso es igual para diferentes días, e independiente entre estos.
- **Etapla 2:** Al inicio del día intenta de entregar el pedido al cliente. La probabilidad de que el cliente esté presente para recibir el pedido ese día es q . En caso de que la entrega no sea exitosa el repartidor vuelve a intentarlo al día siguiente. La probabilidad de tener una entrega exitosa es igual para diferentes días, e independiente entre estos.

Ejemplo: El día 1 el repartidor intenta retirar el producto y es un retiro no exitoso, luego lo intenta el día 2 y es exitoso. Entonces, el día 3 intenta ir a dejarlo, pero no es exitoso, luego lo vuelve a intentar el día 4 y lo consigue.

Se asume que el 24 de junio (hoy) el repartidor hizo el primer intento de retirar el pedido en la tienda.

Calcule la probabilidad de que el pedido sea entregado exitosamente **exactamente** el 1 de julio. (**Recuerde que junio tiene 30 días**).

- (b) Considera el siguiente juego: se lanza un solo dado balanceado. Si el resultado es **3**, entonces el juego termina y ganas \$0. Si el resultado no es 3, entonces tienes dos opciones: (i) detener el juego y recibir un monto igual al número obtenido en ese lanzamiento (por ejemplo, si sale 5, ganas \$5), o (ii) volver a comenzar el juego.

Supongamos que un jugador decide **seguir la estrategia de detenerse la primera vez que aparece un número mayor o igual a 5**, ¿cuál es su ganancias esperada bajo esta estrategia?

Hint: Le podría ser útil considerar la VA del número obtenido en el primer lanzamiento.

Solución

(a). Sea X el número de intento en que se logra retirar el pedido y Y el número de intento en que se logra entregarlo una vez retirado. Suponiendo, además, que ambas secuencias de intentos son independientes entre sí,

$$X \sim \text{Geom}(p), \quad Y \sim \text{Geom}(q),$$

ambas con soporte $\{1, 2, \dots\}$. Si el retiro ocurre en el día X , el primer intento de entrega ocurre al día siguiente, por lo que la entrega exitosa sucede en el día $X + Y$. Contando el 24 de junio

como día 1, el 1 de julio es el día 8. En consecuencia,

$$\begin{aligned}\mathbb{P}(X + Y = 8) &= \sum_{k=1}^7 \mathbb{P}(X = k) \mathbb{P}(Y = 8 - k) \\ &= \sum_{k=1}^7 (1-p)^{k-1} p (1-q)^{7-k} q.\end{aligned}$$

(b). Sea G la ganancia final. Al condicionar en el primer lanzamiento, los resultados 1, 2 y 4 reinician el juego; el resultado 3 entrega cero, y los resultados 5 y 6 entregan su valor. Por tanto,

$$\mathbb{E}[G] = \frac{3}{6} \mathbb{E}[G] + \frac{1}{6} (0 + 5 + 6).$$

Despejando se obtiene

$$\mathbb{E}[G] = \frac{11}{3}.$$

Pregunta 141. (Control 3 - Primavera 2025, P3)

Un equipo de vóleibol tiene un baúl con **10 pelotas**, de las cuales **4 están infladas** y **6 están desinfladas**. El preparador saca una **muestra aleatoria de 5 pelotas** e infla las pelotas desinfladas de dicha muestra, y luego las devuelve al baúl. Para el partido, se saca una muestra aleatoria de **3 pelotas** del baúl.

- (i) Calcule la esperanza del número de pelotas que infla el preparador. *Hint: Podría serle de utilidad reconocer alguna de las distribuciones vistas en el curso.*
- (ii) Determine la distribución (nombre y parámetros si es conocida, o su PMF/PDF en caso contrario) del número de pelotas infladas que se tiene para el partido, dado que el preparador ha inflado t pelotas. Calcule la esperanza condicional del número de pelotas infladas que se tiene para el partido, dado que el preparador ha inflado t pelotas.
- (iii) Calcule la esperanza del número de pelotas infladas seleccionadas para el partido.
- (iv) Uno de los jugadores del equipo afirma que para calcular la varianza de las pelotas infladas a tener en el partido, se podría usar la Ley de la Esperanza Total. Indique la veracidad de dicha afirmación. Si es que es verdadera, entonces explique cómo se haría, sea claro en las variables aleatorias y explique cada término de cada fórmula que utilice. Si cree que es falsa la afirmación, justifique claramente.

Solución

(i). Sea X el número de pelotas desinfladas en la muestra inicial de cinco. Como la muestra se extrae sin reemplazo desde diez pelotas, seis de ellas desinfladas,

$$X \sim \text{Hipergeom}(10, 6, 5).$$

El preparador infla exactamente esas X pelotas. Por la esperanza de la distribución hipergeométrica,

$$\mathbb{E}[X] = 5 \frac{6}{10} = 3.$$

(ii). Sea Y el número de pelotas infladas en la muestra de tres para el partido. Condicionado a $X = t$, en el baúl hay $4 + t$ pelotas infladas entre las diez. Por tanto, para $t \in \{1, 2, 3, 4, 5\}$,

$$Y \mid (X = t) \sim \text{Hipergeom}(10, 4 + t, 3),$$

con función de masa

$$\mathbb{P}(Y = y \mid X = t) = \frac{\binom{4+t}{y} \binom{6-t}{3-y}}{\binom{10}{3}}$$

para los valores enteros y que hacen no nulos ambos coeficientes binomiales. En particular,

$$\mathbb{E}[Y \mid X = t] = 3 \frac{4+t}{10}.$$

(iii). La ley de la esperanza total permite promediar la esperanza condicional anterior sobre los posibles valores de X :

$$\begin{aligned} \mathbb{E}[Y] &= \mathbb{E}[\mathbb{E}[Y \mid X]] \\ &= \mathbb{E}\left[3 \frac{4+X}{10}\right] \\ &= \frac{3}{10}(4 + \mathbb{E}[X]) \\ &= \frac{21}{10}. \end{aligned}$$

(iv). La afirmación es correcta siempre que la esperanza total se aplique tanto a Y como a Y^2 . Equivalentemente, puede usarse la ley de la varianza total:

$$\text{Var}(Y) = \mathbb{E}[\text{Var}(Y \mid X)] + \text{Var}(\mathbb{E}[Y \mid X]).$$

El primer término mide la variabilidad de la muestra final cuando se fija cuántas pelotas infló el preparador; el segundo mide la variabilidad producida por X . Como

$$\mathbb{E}[Y \mid X] = \frac{3}{10}(4 + X), \quad \text{Var}(Y \mid X) = \frac{7}{300}(4 + X)(6 - X),$$

y $\mathbb{E}[X] = 3$, $\text{Var}(X) = 2/3$, se obtiene

$$\mathbb{E}[\text{Var}(Y \mid X)] = \frac{427}{900}, \quad \text{Var}(\mathbb{E}[Y \mid X]) = \frac{3}{50}.$$

Por tanto,

$$\text{Var}(Y) = \frac{427}{900} + \frac{3}{50} = \frac{481}{900}.$$

Pregunta 142. (Control 3 - Otoño 2026, P2)

Toñito ha decidido viajar al sur de Chile para pescar en el Lago Probabilístico. El tiempo (en horas) que Toñito decide pasar en el lago es una variable aleatoria que sigue una distribución Uniforme en el intervalo $[1, 3]$.

A lo largo de años de observación, los lugareños han notado que el número de peces capturados sigue una distribución de Poisson cuya tasa es igual al doble del tiempo que la persona ha estado en el lago.

- Determine el número esperado de peces capturados por Toñito.
- Uno de los pescadores del lago le comenta a Toñito: “Como el número de peces capturados, dado un tiempo de permanencia en el lago, sigue una distribución Poisson, y en una distribución Poisson la esperanza es igual a la varianza, entonces la varianza de peces capturados debe ser igual a la esperanza de peces capturados.”.

Evalúe la veracidad de esta afirmación. Demuestre matemáticamente y además explique intuitivamente su respuesta.

Solución

(a). Sea T el tiempo de permanencia y N el número de peces. El modelo establece que

$$N \mid (T = t) \sim \text{Poisson}(2t),$$

por lo que $\mathbb{E}[N \mid T] = 2T$. La ley de la esperanza total entrega

$$\mathbb{E}[N] = \mathbb{E}[\mathbb{E}[N \mid T]] = 2\mathbb{E}[T].$$

Como $T \sim \text{Uniforme}(1, 3)$, $\mathbb{E}[T] = (1 + 3)/2 = 2$, y así

$$\mathbb{E}[N] = 4.$$

(b). La afirmación es falsa porque la igualdad entre media y varianza se cumple después de fijar T , pero la tasa $2T$ también es aleatoria. La ley de la varianza total separa ambas fuentes de variabilidad:

$$\text{Var}(N) = \mathbb{E}[\text{Var}(N \mid T)] + \text{Var}(\mathbb{E}[N \mid T]).$$

Como una variable Poisson tiene media y varianza iguales a su tasa,

$$\text{Var}(N \mid T) = 2T, \quad \mathbb{E}[N \mid T] = 2T.$$

Además,

$$\text{Var}(T) = \frac{(3 - 1)^2}{12} = \frac{1}{3}.$$

Por consiguiente,

$$\begin{aligned} \text{Var}(N) &= \mathbb{E}[2T] + \text{Var}(2T) \\ &= 4 + 4 \left(\frac{1}{3} \right) \\ &= \frac{16}{3}. \end{aligned}$$

La diferencia $\text{Var}(N) - \mathbb{E}[N] = 4/3$ proviene de que distintos tiempos de permanencia generan distintas tasas de captura. Intuitivamente, al azar propio del conteo Poisson se suma la incertidumbre sobre cuánto tiempo pesca Toño.

Pregunta 143. (Examen - Primavera 2021, P2)

En una elección de curso hay dos candidatos, A y B , y un número impar N de votantes. Todos votan por uno de los candidatos. Salvo que se indique lo contrario, los votos son independientes. Cada inciso describe un escenario distinto.

- Se han contado K votos y A_1 favorecen a A , donde $0 \leq A_1 \leq K \leq N$. Cada voto restante favorece a A con probabilidad $p \in [0, 1]$. Calcule la probabilidad de que A gane.
- El curso tiene H hombres y $M = N - H$ mujeres. Cada hombre vota por A con probabilidad p_H y cada mujer con probabilidad p_M . Calcule la probabilidad de que A gane exactamente por un voto.
- En el escenario del inciso anterior, calcule el número esperado de votos de A .
- En otro curso hay N_d votantes. El primer día votan K_d personas, donde K_d es un entero entre 1 y $N_d - 1$, cada una por A con probabilidad $p_d \in [0, 1]$. Si F es la fracción de votos obtenida por A ese día, cada una de las $N_d - K_d$ personas del segundo día vota por A , condicional en F , con probabilidad $\lambda p_d + (1 - \lambda)F$, donde $\lambda \in [0, 1]$. Los votos son condicionalmente independientes dentro de cada día. Calcule el número esperado de votos de A .

Solución

Sea $r = (N + 1)/2$ el número mínimo de votos que necesita A para ganar, pues N es impar. **(a).** Sea Z el número de votos para A entre los $N - K$ votos restantes. Entonces

$$Z \sim \text{Binomial}(N - K, p).$$

El candidato A gana si $A_1 + Z \geq r$. Por tanto,

$$\mathbb{P}(A \text{ gana}) = \sum_{j=\max\{0, r-A_1\}}^{N-K} \binom{N-K}{j} p^j (1-p)^{N-K-j},$$

con valor cero cuando el límite inferior supera $N - K$. La expresión se interpreta con la convención binomial usual en los casos $p = 0$ o $p = 1$.

(b). Sean $X \sim \text{Binomial}(H, p_H)$ e $Y \sim \text{Binomial}(M, p_M)$ los votos masculinos y femeninos para A . Los dos conteos son independientes. Ganar exactamente por un voto equivale a obtener r votos en total. Al condicionar en el número y de votos femeninos,

$$\begin{aligned} \mathbb{P}(X + Y = r) &= \sum_{y=\max\{0, r-H\}}^{\min\{M, r\}} \binom{H}{r-y} p_H^{r-y} (1-p_H)^{H-r+y} \\ &\quad \cdot \binom{M}{y} p_M^y (1-p_M)^{M-y}. \end{aligned}$$

Los límites garantizan simultáneamente $0 \leq y \leq M$ y $0 \leq r - y \leq H$.

(c). Por linealidad de la esperanza,

$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y] = Hp_H + Mp_M.$$

(d). Sea X_1 el número de votos para A el primer día y sea X_2 el del segundo. Entonces

$$X_1 \sim \text{Binomial}(K_d, p_d), \quad F = \frac{X_1}{K_d}.$$

Condicionado en X_1 , la probabilidad de voto por A el segundo día es $\lambda p_d + (1 - \lambda)X_1/K_d$. Por ello,

$$\mathbb{E}[X_2 | X_1] = (N_d - K_d) \left(\lambda p_d + (1 - \lambda) \frac{X_1}{K_d} \right).$$

La ley de la esperanza iterada y $\mathbb{E}[X_1] = K_d p_d$ dan

$$\begin{aligned} \mathbb{E}[X_1 + X_2] &= K_d p_d + (N_d - K_d) \left(\lambda p_d + (1 - \lambda) \frac{\mathbb{E}[X_1]}{K_d} \right) \\ &= K_d p_d + (N_d - K_d) p_d \\ &= N_d p_d. \end{aligned}$$

Aunque los votos del segundo día dependen de la fracción observada el primero, esa dependencia no modifica el número total esperado de votos para A .

Pregunta 144. (Examen - Otoño 2022, P2)

El local *Lucky Cat* es muy famoso por vender papas fritas a un bajo precio. Para asegurar que los locales que venden comida cumplen con las medidas de higiene y salubridad, la municipalidad dispone del *Organismo Municipal Fiscalizador de Salubridad* (OMFS), cuya función es verificar si dichos locales, tales como *Lucky Cat*, cumplen las medidas mínimas de salubridad. Específicamente, al inicio de **cada hora del día, todos los días de la semana**, el OMFS inspecciona si el local está limpio. Por este motivo, *Lucky Cat* ha contratado a don Fermín, quien está encargado de hacer el aseo en el local en el último minuto **de cada hora**, lo cual le toma exactamente un minuto. A modo de ejemplo, si don Fermín hace el aseo a las 16:59 y luego el OMFS inspecciona el local a las 17:00, el OMFS aprueba el funcionamiento

del local y este puede seguir funcionando (al menos por una hora más). El jefe del local se da cuenta de que, cada vez que don Fermín tiene que realizar el aseo, es decir, al final de cada hora, hay una probabilidad q de que olvide su labor y, por ende, no haga el aseo cuando le corresponde. Además, dichas probabilidades son independientes entre diferentes horas. Nuevamente, a modo de ejemplo, si a don Fermín se le olvida hacer el aseo a las 14:59, cuando el OMFS visite el local a las 15:00, este se clausurará y no podrá seguir funcionando.

Cada hora, *Lucky Cat* recibe una demanda en unidades de *porciones de papas fritas*. La demanda en la i -ésima hora se puede descomponer en dos partes: una demanda basal X_0 , que sigue una distribución Normal con media μ y varianza σ_0^2 , y una demanda específica X_i , que sigue una distribución Normal con media 0 y varianza σ_1^2 . La demanda en la i -ésima hora es entonces $X_0 + X_i$. La demanda basal y las demandas específicas son todas independientes, es decir, las variables aleatorias $X_0, X_1, X_2, \dots$ son todas independientes.

Además de lo mencionado, haga los siguientes supuestos:

- Una vez clausurado el local, este no vuelve a abrir.
 - El local abre a las 10:00 del día viernes 15 de julio de 2022 y no cierra nunca por voluntad propia, sino que solo cierra si es clausurado por el OMFS.
 - El local se ensucia en toda hora de funcionamiento, independiente del nivel de la demanda durante dicha hora, de tal manera que, si no se hace el aseo al final de la hora respectiva, el local será clausurado por el OMFS.
 - El local parte limpio; por ende, cuando el OMFS visita el local en el primer minuto de operación, i.e., a las 10:00 del día viernes 15 de julio de 2022, el local es aprobado para su posterior funcionamiento de al menos la siguiente hora.
- (a) Considere que el local atiende **exactamente** durante seis horas. Determine la distribución de probabilidad que sigue la suma de las demandas durante este lapso. Si la distribución de probabilidad es conocida, es decir, fue vista en clases, especifique sus parámetros.
- (b) Encuentre la correlación entre las demandas de dos horas diferentes en las que el local está funcionando.
- (c) Calcule la esperanza de la **demanda total** por *porciones de papas fritas*, es decir, la demanda durante todas las horas de operación.
- (d) Suponga que el local parte sus operaciones con un stock de s *porciones de papas fritas*. No hay reposición del stock una vez que empieza a funcionar el local. Se desea encontrar una cota superior para la probabilidad de que la **demanda total** (i.e., la demanda durante todas las horas de operación) sea superior al stock inicial. Para ello, calcule la varianza de la **demanda total** usando la *Ley de la Varianza Total*. Luego, obtenga la cota mencionada usando la desigualdad de Chebyshev. Haga el desarrollo algebraico y, finalmente, una vez que obtenga el valor de la cota, reemplace los parámetros del problema por $\mu = 52$, $\sigma_0 = 12$, $\sigma_1 = 8$, $q = 0,05$ y $s = 3600$, para así obtener el valor numérico de la cota.

Note que en las partes (c) y (d), a diferencia de la parte (a), el número de horas que opera el local podría variar.

Solución

(a). Denotemos por $Y_i = X_0 + X_i$ la demanda de la hora i . Si $Z_6 = \sum_{i=1}^6 Y_i$, entonces

$$Z_6 = 6X_0 + \sum_{i=1}^6 X_i.$$

Al ser una combinación lineal de normales independientes,

$$Z_6 \sim \mathcal{N}(6\mu, 36\sigma_0^2 + 6\sigma_1^2).$$

(b). Para $i \neq j$,

$$\text{Cov}(Y_i, Y_j) = \text{Var}(X_0) = \sigma_0^2, \quad \text{Var}(Y_i) = \sigma_0^2 + \sigma_1^2.$$

Por tanto,

$$\text{Corr}(Y_i, Y_j) = \frac{\sigma_0^2}{\sigma_0^2 + \sigma_1^2}.$$

(c). Sea N el número de horas de operación. Bajo la independencia implícita entre los olvidos de don Fermín y las demandas, N es independiente de $X_0, X_1, X_2, \dots$. Con la convención geométrica sobre $\{1, 2, \dots\}$, $N \sim \text{Geom}(q)$ y $\mathbb{E}[N] = 1/q$. Si $Z = \sum_{i=1}^N Y_i$, entonces $\mathbb{E}[Z | N] = N\mu$, de donde

$$\mathbb{E}[Z] = \mathbb{E}[\mathbb{E}[Z | N]] = \mu \mathbb{E}[N] = \frac{\mu}{q}.$$

(d). Condicionalmente a N ,

$$Z = NX_0 + \sum_{i=1}^N X_i, \quad \text{Var}(Z | N) = N^2\sigma_0^2 + N\sigma_1^2.$$

Como $\text{Var}(N) = (1-q)/q^2$ y $\mathbb{E}[N^2] = (2-q)/q^2$, la varianza total da

$$\text{Var}(Z) = \frac{(2-q)\sigma_0^2}{q^2} + \frac{\sigma_1^2}{q} + \frac{(1-q)\mu^2}{q^2}.$$

Para $s > \mu/q$, Chebyshev implica

$$\mathbb{P}(Z \geq s) \leq \frac{\text{Var}(Z)}{(s - \mu/q)^2}.$$

Con los valores entregados,

$$\text{Var}(Z) = 1\,141\,120, \quad \mathbb{P}(Z \geq 3600) \leq \frac{1\,141\,120}{2560^2} \approx 0,174.$$

Pregunta 145. (Examen - Otoño 2023, P2)

Vístete es una empresa que vende ropa usada por internet. Cada día, la empresa vende exactamente veinte prendas, cada una a un cliente diferente. Una vez que un cliente selecciona y paga su compra en la página web de *Vístete*, se le hace el despacho de su producto al día siguiente. Por ende, cada día la empresa debe despachar las prendas a cada uno de los 20 clientes que compraron el día anterior. Cada día, *Vístete* subcontrata a una de las dos empresas que hay para realizar el envío de los pedidos: *Transporte A* y *Transporte B*. Específicamente, al final de cada día, *Vístete* llama a *Transporte A* por teléfono para coordinar el transporte de los 20 productos que debe despachar al día siguiente. El 70 % de las veces, *Transporte A* acepta hacer el despacho. En caso contrario, el despacho lo hace *Transporte B*, la cual siempre tiene disponibilidad. La razón por la cual *Vístete* opta por dar prioridad en los envíos a *Transporte A* es por sus mejores niveles de servicio. De estadísticas previas, se sabe que cada producto despachado por *Transporte A* llega exitosamente a destino (i.e., al cliente) con probabilidad 0,9, independiente entre clientes, mientras que dicha probabilidad es 0,6 en el caso de *Transporte B*, también independiente entre clientes. Las razones por las que un producto no llega a destino son varias, como, por ejemplo, que no confirman bien la dirección o no informan al cliente la hora adecuada de entrega.

Conteste las siguientes preguntas indicando todo el desarrollo. En aquellas preguntas en donde la respuesta es de tipo numérica, basta con que entregue una expresión matemática con operaciones (suma, resta, multiplicación, exponente natural, logaritmo, factorial, coeficiente

binomial, etc.) de valores numéricos.

- (a) Suponga que se sabe que un día el despacho lo hizo *Transporte B*. ¿Qué distribución sigue el número de paquetes (es decir, prendas) que exitosamente reciben los clientes ese día? Si es una distribución conocida, indique su nombre y el valor de su(s) parámetro(s) solamente; en caso contrario, entregue su rango y PMF.
- (b) Un día se han despachado exitosamente 12 paquetes de los 20. ¿Cuál es la probabilidad de que el despacho se haya hecho con *Transporte A*?

Este párrafo aplica para todas las preguntas que siguen a continuación. Considere un día específico. Sea W la VA que vale 1 si en dicho día se hace el despacho con *Transporte A*, y 0 si se hace con *Transporte B*. Además, sea Z la VA que toma el valor de la probabilidad de que un despacho a un cliente en un día cualquiera se haga exitosamente, es decir, $Z = 0,6 + 0,3W$. Sea X el número de paquetes entregados exitosamente en un día cualquiera.

- (c) ¿Qué distribución sigue W ? Si es una distribución conocida, indique su nombre y el valor de su(s) parámetro(s) solamente; en caso contrario, entregue su rango y PMF.
- (d) ¿Qué distribución sigue Z ? Si es una distribución conocida, indique su nombre y el valor de su(s) parámetro(s) solamente; en caso contrario, entregue su rango y PMF.
- (e) Encuentre la esperanza y la varianza de Z .
- (f) Encuentre la distribución de $X | Z$, donde X es el número de paquetes entregados en un día cualquiera. Si es una distribución conocida, indique su nombre y el valor de su(s) parámetro(s) solamente; en caso contrario, entregue su rango y PMF.
- (g) Encuentre la varianza del número de paquetes que se reparte en un día. Puede dejar su respuesta en función del primer y segundo momento de Z . Puede ser útil usar la Ley de la Varianza Total:

$$\text{Var}(X) = \mathbb{E}_Z[\text{Var}_X(X | Z)] + \text{Var}_Z(\mathbb{E}_X[X | Z]).$$

Solución

(a). Condicionalmente a que despacha B, $X \sim \text{Binomial}(20, 0,6)$.

(b). Definamos los eventos

- A : el despacho lo realiza Transporte A; ■ B : el despacho lo realiza Transporte B.

El Teorema de Bayes, con la ley de probabilidades totales en el denominador, entrega

$$\begin{aligned} \mathbb{P}(A | X = 12) &= \frac{\mathbb{P}(X = 12 | A)\mathbb{P}(A)}{\mathbb{P}(X = 12 | A)\mathbb{P}(A) + \mathbb{P}(X = 12 | B)\mathbb{P}(B)} \\ &= \frac{\binom{20}{12}0,9^{12}0,1^8(0,7)}{\binom{20}{12}0,9^{12}0,1^8(0,7) + \binom{20}{12}0,6^{12}0,4^8(0,3)}. \end{aligned}$$

(c). $W \sim \text{Bernoulli}(0,7)$.

(d). El rango de Z es $\{0,6, 0,9\}$, con

$$p_Z(0,6) = 0,3, \quad p_Z(0,9) = 0,7.$$

(e). Usando la transformación afín,

$$\mathbb{E}[Z] = 0,6 + 0,3(0,7) = 0,81, \quad \text{Var}(Z) = 0,3^2(0,7)(0,3) = 0,0189.$$

(f). Condicionalmente a Z , los veinte resultados son Bernoulli independientes con parámetro Z ;

por tanto,

$$X \mid Z \sim \text{Binomial}(20, Z).$$

(g). La varianza total da

$$\begin{aligned}\text{Var}(X) &= \mathbb{E}[20Z(1-Z)] + \text{Var}(20Z) \\ &= 20\mathbb{E}[Z] + 380\mathbb{E}[Z^2] - 400\mathbb{E}[Z]^2.\end{aligned}$$

Como $\mathbb{E}[Z^2] = \text{Var}(Z) + \mathbb{E}[Z]^2 = 0,675$, resulta

$$\text{Var}(X) = 10,26.$$

Pregunta 146. (Examen - Otoño 2025, P1)

Responda las siguientes preguntas.

- (a) Denisse tiene un álbum de Snoopy, el cual se completa con un total de 400 láminas distintas. Ella ya ha reunido 100 de ellas. Compra un sobre que contiene 10 láminas distintas entre sí, donde cada lámina del álbum tiene la misma probabilidad de estar en un sobre. ¿Cuál es el valor esperado de láminas nuevas que obtendrá en ese sobre?
- (b) En un vivero se comercializan dos tipos de plantas ornamentales: lavanda y romero. Las alturas de estas plantas siguen distribuciones normales con medias de 25 cm y 35 cm, y desviaciones estándar de 5 cm y 7 cm, respectivamente. El vivero dispone de 20 ejemplares de lavanda y 80 de romero.
- Don Ambrosio visita el vivero y compra una planta elegida al azar. Al medirla, observa que tiene una altura menor o igual que 28 cm, aunque desconoce la altura exacta. Él sostiene que se trata de una lavanda. ¿Cuál es la probabilidad de que esté en lo cierto?
- (c) Jack es un aficionado jugador de béisbol, el cual batea la pelota con un ángulo de elevación Θ uniforme entre 0 y $\pi/4$ radianes, es decir, entre 0 y 45 grados. La distancia horizontal a la que cae la pelota luego de ser bateada es igual a

$$\frac{v^2 \sin(2\Theta)}{g},$$

donde v es la velocidad a la cual se batea la pelota, cuyo valor es fijo, Θ es el ángulo de elevación y g es la aceleración de gravedad. Jack afirma que la esperanza de la distancia horizontal que recorre la pelota es $v^2 \sin(2 \cdot \pi/8)/g$. Comente sobre la veracidad de dicha afirmación.

Solución

(a). Si X es el número de láminas nuevas, entonces

$$X \sim \text{Hipergeom}(400, 300, 10).$$

La media de esta distribución es

$$\mathbb{E}[X] = n \frac{K}{N} = 10 \frac{300}{400} = 7,5.$$

(b). Definamos los eventos

- L : la planta escogida es una lavanda;
- A : la planta escogida mide a lo más 28 cm.

Entonces

$$\mathbb{P}(L) = 0,2, \quad \mathbb{P}(A \mid L) = \Phi(0,6), \quad \mathbb{P}(A \mid L^c) = \Phi(-1).$$

El Teorema de Bayes, con la ley de probabilidades totales en el denominador, entrega

$$\begin{aligned}\mathbb{P}(L | A) &= \frac{\mathbb{P}(A | L)\mathbb{P}(L)}{\mathbb{P}(A | L)\mathbb{P}(L) + \mathbb{P}(A | L^c)\mathbb{P}(L^c)} \\ &= \frac{0,2\Phi(0,6)}{0,2\Phi(0,6) + 0,8\Phi(-1)} \approx 0,533.\end{aligned}$$

(c). La afirmación es falsa: la esperanza no atraviesa una función no lineal. Como la densidad de Θ es $4/\pi$,

$$\mathbb{E}\left[\frac{v^2 \sin(2\Theta)}{g}\right] = \frac{4v^2}{\pi g} \int_0^{\pi/4} \sin(2\theta) d\theta = \frac{2v^2}{\pi g}.$$

En cambio, la cantidad propuesta por Jack es

$$\frac{v^2 \sin(\pi/4)}{g} = \frac{v^2}{\sqrt{2}g},$$

que es estrictamente mayor. Equivalentemente, Jensen aplica porque $\theta \mapsto \sin(2\theta)$ es estrictamente cóncava en el intervalo.

Pregunta 147. (Examen - Primavera 2025, P1)

- Doña Anastacia tiene un emprendimiento, el cual consiste en vender sacos de comida de 18 kg para mascotas. No obstante, el peso real de un saco suele diferir de dicho valor. Específicamente, la diferencia entre el peso real de un saco con respecto al peso que debiese tener (18 kg); se puede modelar como una normal estándar. Para un saco que se produce, ¿cómo distribuye el cuadrado de dicha diferencia? Si es una distribución conocida (vista en clases), indique su nombre y valor(es) de su(s) parámetro(s), en caso contrario, entregue su PMF o PDF. Debe entregar un desarrollo justificado para su respuesta. *Hint: Notar que la función cuadrática no es monótona en todo su dominio.*
- En una elección con dos candidatos, A y B , se tiene una mesa en donde votarán 40 mujeres y 60 hombres. Se sabe que la probabilidad de que un hombre vote por el candidato A es 55 %, y por el candidato B 45 %. En el caso de las mujeres, 72 % declara preferir al candidato A y un 28 % al candidato B . Conteste las siguientes preguntas con respecto a la mesa mencionada.
 - ¿Cuál es el número esperado de votos que tendrá el candidato A ?
 - ¿Cuál es la probabilidad de que el candidato A obtenga exactamente 80 votos?

Solución

(1). Sea $X \sim \mathcal{N}(0, 1)$ la diferencia entre el peso real y 18 kg, y defina $Y = X^2$. Como $Y \geq 0$, para $y < 0$ se tiene $F_Y(y) = 0$. Para $y \geq 0$,

$$\begin{aligned}F_Y(y) &= \mathbb{P}(X^2 \leq y) \\ &= \mathbb{P}(-\sqrt{y} \leq X \leq \sqrt{y}) \\ &= \Phi(\sqrt{y}) - \Phi(-\sqrt{y}).\end{aligned}$$

Al derivar para $y > 0$ y usar la simetría $\phi(-x) = \phi(x)$,

$$\begin{aligned}f_Y(y) &= \frac{\phi(\sqrt{y}) + \phi(-\sqrt{y})}{2\sqrt{y}} \\ &= \frac{1}{\sqrt{2\pi y}} e^{-y/2}.\end{aligned}$$

Por tanto,

$$f_Y(y) = \begin{cases} \frac{1}{\sqrt{2\pi y}} e^{-y/2}, & y > 0, \\ 0, & y \leq 0. \end{cases}$$

Esta es la densidad de una distribución chi-cuadrado con un grado de libertad: $Y \sim \chi_1^2$.

(2)(a). Sea M el número de votos para A entre las 40 mujeres y H el número correspondiente entre los 60 hombres. Si $V = M + H$, la linealidad de la esperanza, que no requiere independencia, da

$$\mathbb{E}[V] = 40(0,72) + 60(0,55) = 61,8.$$

(2)(b). El enunciado sólo entrega las probabilidades marginales de voto y no declara independencia entre votantes. Por ello, la probabilidad solicitada no queda determinada de manera única. Bajo el supuesto adicional de que todas las decisiones son independientes,

$$M \sim \text{Binomial}(40, 0,72), \quad H \sim \text{Binomial}(60, 0,55),$$

y ambos conteos son independientes. La función de masa de $V = M + H$ se obtiene entonces por convolución. Para $V = 80$, si $M = k$, las restricciones $0 \leq k \leq 40$ y $0 \leq 80 - k \leq 60$ implican $20 \leq k \leq 40$. Por ello,

$$\mathbb{P}(V = 80) = \sum_{k=20}^{40} \binom{40}{k} (0,72)^k (0,28)^{40-k} \binom{60}{80-k} (0,55)^{80-k} (0,45)^{k-20}.$$

Capítulo 9

Convergencia y Teoremas Límite

9.1. Desigualdades, convergencia y teoremas límite

Resumen

Desigualdades de probabilidad.

- **Boole.** Para una familia numerable de eventos,

$$\mathbb{P}\left(\bigcup_i A_i\right) \leq \sum_i \mathbb{P}(A_i).$$

- **Markov.** Si $X \geq 0$ y $a > 0$,

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[X]}{a}.$$

- **Chebyshev.** Si $\mathbb{E}[X] = \mu$, $\text{Var}(X) = \sigma^2$ y $\varepsilon > 0$,

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

- **Chernoff.** Cuando la FGM es finita para los valores considerados,

$$\mathbb{P}(X \geq a) \leq \inf_{t>0} e^{-ta} M_X(t),$$

$$\mathbb{P}(X \leq a) \leq \inf_{t<0} e^{-ta} M_X(t).$$

Markov exige una variable no negativa; Chebyshev se obtiene aplicándola a $(X - \mu)^2$.

Cauchy-Schwarz y Jensen.

Si $\mathbb{E}[X^2], \mathbb{E}[Y^2] < \infty$ y g es convexa,

$$|\mathbb{E}[XY]| \leq \sqrt{\mathbb{E}[X^2]\mathbb{E}[Y^2]},$$

$$g(\mathbb{E}[X]) \leq \mathbb{E}[g(X)].$$

Consecuencias y casos de igualdad:

- **Correlación.** Cauchy-Schwarz implica $|\text{Corr}(X, Y)| \leq 1$; hay igualdad cuando las variables centradas son proporcionales casi seguramente.
- **Jensen afín.** Si g es afín, Jensen se cumple con igualdad. Si g es estrictamente convexa, la igualdad exige que X sea constante casi seguramente.

Modos de convergencia.

- **Casi segura.** $\mathbb{P}(\{\omega : X_n(\omega) \rightarrow X(\omega)\}) = 1$.

- **En probabilidad.** Para todo $\varepsilon > 0$, $\mathbb{P}(|X_n - X| > \varepsilon) \rightarrow 0$.

- **En distribución.** $F_{X_n}(x) \rightarrow F_X(x)$ en cada punto de continuidad de F_X .

Estas nociones satisfacen

$$X_n \xrightarrow{\text{c.s.}} X \implies X_n \xrightarrow{\mathbb{P}} X \implies X_n \xrightarrow{d} X,$$

sin recíprocas en general. Si el límite es una constante c , entonces $X_n \xrightarrow{d} c$ también implica $X_n \xrightarrow{\mathbb{P}} c$. Para variables enteras no negativas, la convergencia en distribución equivale a la convergencia de $\mathbb{P}(X_n = k)$ para cada entero $k \geq 0$.

Leyes de los grandes números.

Para variables i.i.d. con $\mathbb{E}[X_i] = \mu$ y $\text{Var}(X_i) = \sigma^2 < \infty$, la media $\bar{X}_n = n^{-1} \sum_{i=1}^n X_i$ satisface:

- **Ley débil.** $\bar{X}_n \xrightarrow{\mathbb{P}} \mu$; en particular, Chebyshev da

$$\mathbb{P}(|\bar{X}_n - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2}.$$

- **Ley fuerte.** $\bar{X}_n \xrightarrow{\text{c.s.}} \mu$.

Teorema Central del Límite.

Si $X_1, X_2, \dots$ son i.i.d., con media μ y varianza $0 < \sigma^2 < \infty$, entonces

$$\frac{\sum_{i=1}^n X_i - n\mu}{\sigma\sqrt{n}} = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{d} \mathcal{N}(0, 1).$$

Por ello, para n grande,

$$\sum_{i=1}^n X_i \approx \mathcal{N}(n\mu, n\sigma^2), \quad \bar{X}_n \approx \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right).$$

La independencia es una hipótesis esencial de esta versión. Si la suma es entera, la corrección de continuidad aproxima

$$\mathbb{P}\left(\sum_{i=1}^n X_i \leq k\right) \approx \Phi\left(\frac{k + 1/2 - n\mu}{\sigma\sqrt{n}}\right).$$

9.1.1. Preguntas

Pregunta 148.

Una fábrica produce un 5 % de artículos defectuosos. ¿De qué tamaño se debe tomar una muestra aleatoria de artículos de manera que la proporción de defectuosos en la muestra sea inferior al 6 % con probabilidad al menos 0,99? Indicación: Aplique TCL para aproximar la probabilidad y despeje n de la desigualdad resultante.

Solución

Sea X_i el indicador de que el artículo i es defectuoso. Entonces $X_i \sim \text{Bernoulli}(0,05)$, con media 0,05 y varianza 0,0475. Para $\bar{X}_n = n^{-1} \sum_{i=1}^n X_i$, el Teorema Central del Límite entrega

$$\mathbb{P}(\bar{X}_n < 0,06) \approx \Phi\left(\sqrt{\frac{n}{475}}\right).$$

Por lo tanto, se requiere

$$n \geq 475(\Phi^{-1}(0,99))^2 \approx 2570,65,$$

y el menor tamaño entero bajo esta aproximación es $n = 2571$.

Pregunta 149.

Se dispone de dos monedas, cada una de las cuales se lanza n veces. Suponga que los lanzamientos individuales son independientes (para una misma moneda y entre ellas) y que las probabilidades de cara respectivas son p_1, p_2 . Suponga que n es grande y aplique el teorema central del límite para calcular aproximadamente la probabilidad de que la primera moneda muestre un r % (con $r > 0$) más de caras que la segunda. Dejar la respuesta en términos de la función Φ (función de probabilidad acumulada de la distribución normal $\mathcal{N}(0,1)$).

Solución

Sea $t = 1 + r/100$, sean $N_1 = \sum_{i=1}^n X_i$ y $N_2 = \sum_{i=1}^n Y_i$ los números de caras, y definamos $Z_i = X_i - tY_i$, donde X_i e Y_i son los indicadores de cara en el lanzamiento i de las monedas 1 y 2. Entonces

$$\mu := \mathbb{E}[Z_i] = p_1 - tp_2, \quad \sigma^2 := \text{Var}(Z_i) = p_1(1 - p_1) + t^2 p_2(1 - p_2).$$

Si $\sigma^2 > 0$, el Teorema Central del Límite entrega

$$\mathbb{P}(N_1 > tN_2) \approx \Phi\left(\frac{\sqrt{n}\mu}{\sigma}\right).$$

Es decir,

$$\mathbb{P}\left(N_1 > \left(1 + \frac{r}{100}\right) N_2\right) \approx \Phi\left(\frac{\sqrt{n}[p_1 - (1 + r/100)p_2]}{\sqrt{p_1(1 - p_1) + (1 + r/100)^2 p_2(1 - p_2)}}\right).$$

Si $\sigma^2 = 0$, ambos conteos son deterministas. En ese caso, la probabilidad solicitada vale 1 únicamente cuando $(p_1, p_2) = (1, 0)$ y vale 0 en los demás casos degenerados.

Pregunta 150.

Un examen está conformado por 60 preguntas verdadero y falso. Estas no requieren justificación, sino que solo indicar si son verdadero o falso. Cada pregunta correcta vale un punto, mientras que cada pregunta incorrecta **resta un punto**. Una pregunta omitida no suma ni resta puntos. El examen se considera aprobado si el puntaje total es mayor o igual a 45 puntos. Pedro tomará el examen que se menciona. Considere que de las 60 preguntas, Pedro sabe la respuesta de solo 18 preguntas, las cuales contestará de manera correcta. Para las 42 preguntas restantes, Pedro las responderá al azar, de modo que cada pregunta estará contestada de manera correcta con probabilidad $1/2$, o incorrecta con probabilidad complementaria. Asuma independencia en el resultado de las 42 preguntas contestadas al azar. Encuentre una aproximación del valor de la probabilidad de aprobar el examen. Note que se pide un valor aproximado de la probabilidad, y no una expresión matemática de dicha probabilidad.

Solución

Sea X_i el puntaje de la respuesta i contestada al azar, con valores 1 y -1 equiprobables, y sea $S = \sum_{i=1}^{42} X_i$. Pedro aprueba cuando $18 + S \geq 45$; como S sólo toma valores pares, esto equivale a $S \geq 28$. Con corrección por continuidad y el Teorema Central del Límite,

$$\mathbb{P}(S \geq 28) \approx 1 - \Phi\left(\frac{27}{\sqrt{42}}\right) \approx 1,55 \times 10^{-5}.$$

Pregunta 151.

- a) En un evento hay 500 invitados. De ellos, hay 300 que están muy *entusiasmados* con ir al evento, y 200 que son *indiferentes*. En particular, para cada invitado *entusiasmado* se considera la variable X_i que vale 1 si es que el invitado i va al evento y 0 si no, y se tiene que $\mathbb{P}(X_i = 1) = 0,8$ para todo $i \in \{1, \dots, 300\}$. Para cada invitado *indiferente*, se considera la variable Y_j que vale 1 si es que el invitado j va al evento y 0 si no, y se tiene que $\mathbb{P}(Y_j = 1) = 0,5$ para todo $j \in \{1, \dots, 200\}$. Considere que todos los invitados son independientes entre sí. La organización decide realizar el evento sólo si la probabilidad de *que vayan menos de 400 personas* es menor a 0,1. Ocupe la desigualdad de Markov para aconsejar a la organización sobre la realización del evento.
- b) Una fábrica de botellas de vidrio debe generar 300 unidades en una jornada laboral de 6 horas. El tiempo de fabricación de cada botella i sigue una distribución exponencial de parámetro $\lambda_i = 0,875$ [botellas/minuto]. Asuma que los tiempos de fabricación de cada botella son independientes. Usando el TCL, estime la probabilidad de que se logren fabricar las 300 unidades en la jornada laboral.

Solución

(a). Si N es la asistencia total,

$$\mathbb{E}[N] = 300(0,8) + 200(0,5) = 340.$$

Markov implica $\mathbb{P}(N \geq 400) \leq 0,85$, de modo que $\mathbb{P}(N < 400) \geq 0,15 > 0,1$. El criterio no se satisface y el evento no debería realizarse.

(b). Para los tiempos $T_i \sim \text{Exp}(0,875)$, $\mathbb{E}[T_i] = \sigma_{T_i} = 8/7$. Como la jornada dura 360 minutos, el Teorema Central del Límite entrega

$$\mathbb{P}\left(\sum_{i=1}^{300} T_i \leq 360\right) \approx \Phi\left(\frac{\sqrt{300}}{20}\right) \approx 0,807.$$

Pregunta 152.

Una astrónoma quiere conocer la distancia d (en años luz) de la Tierra a una estrella lejana. Para medir esta distancia dispone de un instrumento adecuado, pero debido a las variaciones de las condiciones atmosféricas, el valor obtenido en cada medición no corresponde a la distancia exacta, sino a una v.a con esperanza d y varianza 4. Por esta razón, la astrónoma planea tomar n mediciones y utilizar su promedio como el valor estimado de la distancia real. Asumiendo que las mediciones son independientes e idénticamente distribuidas. ¿Cuál es el mínimo valor de n tal que con probabilidad de al menos 95 % su estimación tenga un error de a lo más $\pm 0,5$ años luz? Obtenga un resultado usando la desigualdad de Chebyshev y otro aproximando por TCL.

Solución

Para $\bar{X}_n = n^{-1} \sum_{i=1}^n X_i$,

$$\mathbb{E}[\bar{X}_n] = d, \quad \text{Var}(\bar{X}_n) = \frac{4}{n}.$$

El Teorema Central del Límite aproxima

$$\mathbb{P}(|\bar{X}_n - d| \leq 0,5) \approx 2\Phi\left(\frac{\sqrt{n}}{4}\right) - 1,$$

por lo que basta $n \geq (4\Phi^{-1}(0,975))^2 \approx 61,47$: el menor entero es 62. En cambio, Chebyshev exige $16/n \leq 0,05$, y por ello entrega la garantía conservadora $n = 320$.

Pregunta 153.

En Economía se dice que un agente, con función de utilidad U , frente a una v.a. X es:

- **Averso al Riesgo** ssi $U(\mathbb{E}[X]) > \mathbb{E}[U(X)]$.
- **Favorable al Riesgo** ssi $U(\mathbb{E}[X]) < \mathbb{E}[U(X)]$.
- **Neutral al Riesgo** ssi $U(\mathbb{E}[X]) = \mathbb{E}[U(X)]$.

Si $X \sim \text{Unif}(0, 1)$, indique el tipo de agente las siguientes funciones de utilidad:

- a) $U(t) = t^2$.
- b) $U(t) = \ln(t)$.
- c) $U(t) = a + bt$

Solución

Como $X \sim \text{Unif}(0, 1)$, se tiene $\mathbb{E}[X] = 1/2$. **(a).** Para $U(t) = t^2$,

$$U(\mathbb{E}[X]) = \frac{1}{4}, \quad \mathbb{E}[U(X)] = \int_0^1 x^2 dx = \frac{1}{3}.$$

Como $U(\mathbb{E}[X]) < \mathbb{E}[U(X)]$, el agente es favorable al riesgo según la definición del enunciado.

(b). Para $U(t) = \ln t$,

$$U(\mathbb{E}[X]) = \ln(1/2) = -\ln 2, \quad \mathbb{E}[U(X)] = \int_0^1 \ln x dx = -1.$$

Como $-\ln 2 > -1$, el agente es averso al riesgo.

(c). Para $U(t) = a + bt$,

$$U(\mathbb{E}[X]) = a + \frac{b}{2} = \mathbb{E}[a + bX].$$

El agente es neutral al riesgo.

Pregunta 154.

- a) Demuestre la desigualdad de Jensen para una función convexa $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ y una variable aleatoria discreta X . Nombre ejemplos de funciones validas para este caso.
- b) Utilizando la desigualdad de Cauchy-Schwarz, demuestre que para cualquier variable aleatoria X e Y :

$$|\text{Corr}(X, Y)| \leq 1.$$

Además, muestre que $|\text{Corr}(X, Y)| = 1$ si y sólo si $Y = aX + b$ para $a, b \in \mathbb{R}$.

Solución

(a). Si X toma los valores $x_1, \dots, x_m \in \mathbb{R}^n$ con probabilidades $p_1, \dots, p_m$, la convexidad de ϕ entrega

$$\phi(\mathbb{E}[X]) = \phi\left(\sum_{i=1}^m p_i x_i\right) \leq \sum_{i=1}^m p_i \phi(x_i) = \mathbb{E}[\phi(X)].$$

Para soporte numerable, defina $X_m = X$ en sus primeros m valores y $X_m = 0$ en el resto. Si X y $\phi(X)$ son integrables, entonces $\mathbb{E}[X_m] \rightarrow \mathbb{E}[X]$ y $\mathbb{E}[\phi(X_m)] \rightarrow \mathbb{E}[\phi(X)]$; la continuidad de ϕ permite tomar límite en la desigualdad finita. Son ejemplos convexos las funciones afines, $x \mapsto \|x\|_2^2$, $x \mapsto e^{a^\top x}$ y, en una dimensión, $x \mapsto -\log x$ sobre $(0, \infty)$.

(b). Supongamos varianzas finitas y positivas, y definamos $U = (X - \mathbb{E}[X])/\sigma_X$ y $V = (Y - \mathbb{E}[Y])/\sigma_Y$. Por Cauchy-Schwarz,

$$|\text{Corr}(X, Y)| = |\mathbb{E}[UV]| \leq \sqrt{\mathbb{E}[U^2]\mathbb{E}[V^2]} = 1.$$

La igualdad en Cauchy-Schwarz ocurre si y sólo si $V = cU$ casi seguramente. Como $\mathbb{E}[U^2] = \mathbb{E}[V^2] = 1$, se tiene $c \in \{-1, 1\}$. Al despejar, esto equivale a $Y = aX + b$ casi seguramente, donde $a = c\sigma_Y/\sigma_X \neq 0$ y $b = \mathbb{E}[Y] - a\mathbb{E}[X]$. Recíprocamente, toda relación afín de esta forma produce $|\text{Corr}(X, Y)| = 1$, con signo igual al de a .

9.1.2. Preguntas de Evaluación

Pregunta 155. (Examen - Otoño 2021, P2)

Don Eustacio tiene un camión para hacer despachos. Cada día se tiene una carga, X , que debe ser despachada dentro de ese mismo día. La carga (en *toneladas*) que debe repartir el camión en un día sigue una distribución Uniforme continua entre 1 y 2. Por otra parte, se tiene que el tiempo, Y , (en *horas*) que tarda el camión en repartir toda la carga de un día sigue una distribución exponencial con parámetro $X^{-1/2}$. Por lo tanto, dado el tamaño de la carga de un día X , tenemos que el promedio del tiempo de despacho es $X^{1/2}$. La distribución de probabilidad de la carga de un día es independiente de la distribución de probabilidad de la carga de los otros días.

Don Eustacio tiene un chofer que trabaja 8 *horas* al día. Por lo tanto, si es que el tiempo de despacho de un día es mayor a 8 *horas*, entonces no se puede repartir toda la carga de ese día. La carga que no se alcanza a repartir un día se pierde, es decir, no se guarda para el día siguiente.

Don Eustacio hace la siguiente afirmación: «Aplicando sobre la VA Y (i.e., el tiempo de despacho de la carga de un día) una de las desigualdades vistas en el curso IN3141 que usa la varianza, se puede concluir que la probabilidad de que no se reparta toda la carga de un día es menor a 0,03».

Determine la veracidad de la afirmación anterior. Note que establecer la veracidad de la afirmación debe restringirse a la metodología a la cual se remite en la afirmación, y no a computar el valor exacto de dicha probabilidad. Sea claro en su desarrollo en cómo llega a su respuesta.

Hint: No existe fórmula cerrada para la PDF marginal de la VA Y , i.e., del tiempo en repartir la carga de un día, de modo que evite obtener una expresión en fórmula cerrada para esto. Esto no implica que no se pueda resolver la pregunta haciendo uso de la PDF marginal de la VA Y . Por otra parte, note que existen resultados vistos en clases sobre la esperanza y varianza que le ahorrarían hacer cálculos con distribuciones marginales o conjuntas.

Hint 2: Si X es una VA y $h: \mathbb{R} \rightarrow \mathbb{R}$ es una función, entonces $\text{Var}(h(X)) = \mathbb{E}[h(X)^2] - \mathbb{E}[h(X)]^2$.

Solución

Como $Y | X = x \sim \text{Exp}(x^{-1/2})$, se tiene $\mathbb{E}[Y | X] = \sqrt{X}$ y $\text{Var}(Y | X) = X$. Por esperanza y varianza totales,

$$m := \mathbb{E}[Y] = \mathbb{E}[\sqrt{X}] = \int_1^2 \sqrt{x} dx = \frac{2}{3}(2^{3/2} - 1) \approx 1,21895,$$

$$\text{Var}(Y) = \mathbb{E}[X] + \text{Var}(\sqrt{X}) = 2\mathbb{E}[X] - m^2 = 3 - m^2 \approx 1,51416.$$

Puesto que $8 > m$, la desigualdad de Chebyshev entrega

$$\begin{aligned} \mathbb{P}(Y \geq 8) &\leq \mathbb{P}(|Y - m| \geq 8 - m) \leq \frac{\text{Var}(Y)}{(8 - m)^2} \\ &\approx \frac{1,51416}{(8 - 1,21895)^2} \approx 0,03293. \end{aligned}$$

Esta cota es mayor que 0,03. Por consiguiente, el procedimiento indicado no permite concluir la desigualdad afirmada y, en el sentido preciso solicitado, la afirmación es falsa.

Pregunta 156. (Examen - Otoño 2021, P3)

- Sea $n \geq 1$, $p \in (0, 1)$ y $X \sim \text{Binomial}(n, p)$. Para cada $k \geq 1$, defina $X_k = n - X$. Determine si $(X_k)_{k \geq 1}$ converge en distribución o en probabilidad a X . Identifique además el límite natural de la sucesión.
- En una autopista pasan exactamente 100 vehículos al día. Durante 30 días, sea $X_i \sim \text{Binomial}(100, p)$ el número de vehículos contabilizados por TAG en el día i , donde $p \in (0, 1)$ y las variables $X_1, \dots, X_{30}$ son independientes. Calcule la probabilidad de que durante el período se contabilicen más de 920 vehículos y entregue también una aproximación mediante el Teorema Central del Límite.

Solución

(1). La sucesión no depende de k : $X_k = n - X$ para todo k . Por ello,

$$X_k \longrightarrow n - X \quad \text{casi seguramente,}$$

y, en consecuencia, también converge a $n - X$ en probabilidad y en distribución. Además,

$$n - X \sim \text{Binomial}(n, 1 - p).$$

Así, X_k converge en distribución a X si y solo si $p = 1/2$, pues solo entonces las distribuciones $\text{Binomial}(n, 1 - p)$ y $\text{Binomial}(n, p)$ coinciden.

En cambio, X_k no converge en probabilidad a X para ningún $p \in (0, 1)$. Como la sucesión es constante en k , tal convergencia exigiría $n - X = X$ casi seguramente. Esto no ocurre porque una binomial con $p \in (0, 1)$ no está concentrada en el único valor $n/2$.

(2). La suma de variables binomiales independientes con el mismo parámetro p vuelve a ser binomial. Por tanto,

$$S = \sum_{i=1}^{30} X_i \sim \text{Binomial}(3000, p),$$

y la probabilidad exacta es

$$\mathbb{P}(S > 920) = \sum_{s=921}^{3000} \binom{3000}{s} p^s (1-p)^{3000-s}.$$

Como $\mathbb{E}[S] = 3000p$ y $\text{Var}(S) = 3000p(1-p)$, el Teorema Central del Límite, con corrección de continuidad, sugiere

$$\mathbb{P}(S > 920) \approx 1 - \Phi\left(\frac{920,5 - 3000p}{\sqrt{3000p(1-p)}}\right),$$

donde Φ es la CDF de una Normal estándar. Esta aproximación es adecuada cuando tanto $3000p$ como $3000(1-p)$ son suficientemente grandes; la suma binomial anterior sigue siendo la expresión exacta para todo $p \in (0, 1)$.

Pregunta 157. (Examen - Primavera 2021, P3)

Una estación de metro posee un ascensor con capacidad de peso C kilos. La jefa de operaciones limita mediante un cartel el número de personas que pueden estar simultáneamente dentro del ascensor; el límite actualmente indicado es n_0 . Suponga que los pesos de las personas son independientes y siguen una distribución uniforme continua con parámetros a y b , donde $0 < a < b$. El ascensor se usa siempre con exactamente el número de personas indicado en el cartel. Se desea que, cada vez que se use, la probabilidad de exceder la capacidad de peso sea a lo más α , con $0 < \alpha < 1/2$.

La jefa contrata a dos practicantes. El primero usa la desigualdad de Markov y el segundo usa el Teorema Central del Límite. Determine el máximo límite entero de personas que propone cada practicante y establezca, en cada caso, si su resultado obliga a reducir el límite actual n_0 . Incluya los cálculos y razonamientos utilizados.

Solución

Sea $S_n = \sum_{i=1}^n X_i$, donde los X_i son los pesos. Denotemos

$$\mu = \frac{a+b}{2}, \quad \sigma = \frac{b-a}{\sqrt{12}}.$$

Por Markov,

$$\mathbb{P}(S_n \geq C) \leq \frac{\mathbb{E}[S_n]}{C} = \frac{n\mu}{C}.$$

La cota garantiza el nivel requerido cuando $n\mu/C \leq \alpha$; por tanto,

$$n_M = \left\lfloor \frac{\alpha C}{\mu} \right\rfloor.$$

Este método obliga a reducir el límite exactamente cuando $n_M < n_0$.

Para el segundo practicante, el TCL aproxima S_n por una normal de media $n\mu$ y varianza $n\sigma^2$.

Si $z_{1-\alpha} = \Phi^{-1}(1-\alpha)$, la condición aproximada es

$$1 - \Phi\left(\frac{C - n\mu}{\sigma\sqrt{n}}\right) \leq \alpha \iff \mu x^2 + z_{1-\alpha}\sigma x - C \leq 0, \quad x = \sqrt{n}.$$

La raíz positiva de la cuadrática es

$$x_+ = \frac{-z_{1-\alpha}\sigma + \sqrt{z_{1-\alpha}^2\sigma^2 + 4\mu C}}{2\mu},$$

de modo que el límite propuesto es

$$n_{\text{TCL}} = \lfloor x_+^2 \rfloor.$$

La segunda metodología obliga a reducir el límite si y solo si $n_{\text{TCL}} < n_0$.

Pregunta 158. (Examen alternativo - Primavera 2021, P3)

El poblado de *Rincones con Agua* no posee red de agua para su población, por lo que cuenta con un estanque con 3200 litros de agua para ser usados cada semana para su población de 106 habitantes. Cada habitante tiene un balde con capacidad de 5 litros, el cual se lleva al estanque de agua para rellenar. El consumo de agua semanal de cada persona podría ser diferente. En particular, considere que cada persona acude a rellenar su balde en la semana un número entero **entre** 4 y 8 veces con igual probabilidad en cada caso (ojo, notar que el número de veces que cada persona llena su balde en la semana será de 4, 5, 6, 7, u 8). Cada vez que los habitantes se quedan sin agua (es decir, que acuden al estanque a rellenar su balde pero esto no es posible dado que no queda agua en el estanque), la municipalidad se encarga de hacer un reabastecimiento con una empresa contratista que posee agua suficiente como para cubrir la falta de agua de la respectiva semana. La municipalidad de *Rincones con Agua* tiene que pagar un monto de \$1200 por cada semana que se acude a la empresa contratista.

Asuma que: (i) los habitantes no poseen otra fuente de agua que la del estanque que se menciona, o la empresa contratista que funciona una vez que el estanque se queda sin agua, (ii) cada vez que un habitante rellena el balde, el balde está vacío inicialmente y es llenado con agua hasta su capacidad máxima de 5 litros, (iii) al inicio de la semana todos los habitantes tienen su balde vacío, mientras que el estanque esta lleno pues se llena al inicio de cada semana.

Otro pueblo vecino ha ofrecido a *Rincones con Agua* instalar un estanque con una capacidad de 200 litros por semana. Al igual que el estanque original, este estanque se rellenaría al inicio de cada semana. (De modo que si es que se tiene este estanque, se tendrían en total 3400 litros disponibles a la semana.) Dado que dicho pueblo vecino incurriría en ciertos costos de abastecimiento, el precio semanal de arriendo de dicho estanque es de \$250 (costo que, en caso de tomar el arriendo, lo debe pagar la municipalidad de *Rincones con Agua*). Notar que el costo de arriendo es independiente de cuánta agua se use del estanque. En caso de tomar el arriendo, si es que en una semana la demanda de agua de los habitantes es superior a la capacidad de ambos estanques, se debe entonces incurrir a lo indicado al final del primer párrafo. Tomando en consideración los costos asociados (y no aspectos no-monetarios) ¿sería conveniente tomar el contrato de arriendo del estanque?

Solución

Para $i \in \{1, \dots, 106\}$, sea Y_i el número de veces que la persona i llena su balde durante una semana y sea $X_i = 5Y_i$ su consumo, medido en litros. Suponiendo que los consumos individuales son independientes, las variables Y_i son i.i.d. y uniformes sobre $\{4, 5, 6, 7, 8\}$. Sus dos primeros momentos son

$$\begin{aligned}\mathbb{E}[Y_i] &= \frac{4 + 5 + 6 + 7 + 8}{5} = 6, \\ \mathbb{E}[Y_i^2] &= \frac{4^2 + 5^2 + 6^2 + 7^2 + 8^2}{5} = 38,\end{aligned}$$

por lo que $\text{Var}(Y_i) = 38 - 6^2 = 2$. Al escalar por la capacidad del balde,

$$\mathbb{E}[X_i] = 5\mathbb{E}[Y_i] = 30, \quad \text{Var}(X_i) = 5^2 \text{Var}(Y_i) = 50.$$

Sea $W = \sum_{i=1}^{106} X_i$ la demanda semanal total. Por independencia,

$$\mathbb{E}[W] = 106 \cdot 30 = 3180, \quad \text{Var}(W) = 106 \cdot 50 = 5300.$$

Si C_0 y C_1 son los costos semanales sin y con arriendo, respectivamente, entonces

$$C_0 = 1200 \mathbf{1}_{\{W > 3200\}}, \quad C_1 = 250 + 1200 \mathbf{1}_{\{W > 3400\}}.$$

El Teorema Central del Límite aproxima $(W - 3180)/\sqrt{5300}$ mediante una variable $Z \sim \mathcal{N}(0, 1)$. Como W toma valores en $5\mathbb{Z}$, los eventos $W > 3200$ y $W > 3400$ equivalen a $W \geq 3205$ y $W \geq 3405$, respectivamente. Usando la corrección de continuidad en los puntos medios 3202,5 y 3402,5, el costo esperado sin arriendo es

$$\begin{aligned} \mathbb{E}[C_0] &= 1200 \mathbb{P}(W > 3200) \\ &\approx 1200 \left[1 - \Phi\left(\frac{3202,5 - 3180}{\sqrt{5300}}\right) \right] \\ &= 1200 [1 - \Phi(0,3091)] \approx 454,36. \end{aligned}$$

Con el estanque adicional, el costo esperado es

$$\begin{aligned} \mathbb{E}[C_1] &= 250 + 1200 \mathbb{P}(W > 3400) \\ &\approx 250 + 1200 \left[1 - \Phi\left(\frac{3402,5 - 3180}{\sqrt{5300}}\right) \right] \\ &= 250 + 1200 [1 - \Phi(3,0563)] \approx 251,34. \end{aligned}$$

Como el costo esperado con arriendo es menor, conviene aceptar el contrato bajo el criterio de costo esperado.

Pregunta 159. (Examen - Otoño 2022, P3)

Usted está a cargo de un servicio de entregas y está intentando reorganizar la distribución de la carga entre los repartidores. Un repartidor debe entregar 20 paquetes en un día. El tiempo que se demora lo podemos separar en dos partes: el tiempo que tarda en llegar al lugar de cada entrega y el tiempo que se demora en la entrega misma. En la entrega misma de cada paquete se demora un tiempo aleatorio independiente del resto: se demora 3 minutos con probabilidad $1/2$ y 7 minutos con probabilidad $1/2$. El recorrido del repartidor comienza en la bodega y termina cuando entrega el último paquete, y en cada traslado se demora exactamente 10 minutos (no aleatorios). Es decir, desde la bodega hasta el lugar de entrega del primer paquete se demora 10 minutos en trasladarse, y entre cada par de lugares de entrega de paquetes se demora 10 minutos en trasladarse.

- Sea T el tiempo total que se demora el repartidor en completar las 20 entregas. Calcule $\mathbb{E}[T]$ y $\text{Var}(T)$.
- Normalice T , es decir, encuentre $a, b \in \mathbb{R}$ tales que, si definimos $Z := a(T - b)$, se cumpla que $\mathbb{E}[Z] = 0$ y $\text{Var}(Z) = 1$. ¿Qué distribución continua sigue aproximadamente la VA Z ? Justifique su respuesta.
- Usted está considerando limitar el tiempo total del recorrido, pero debe tomar en cuenta cómo eso afectaría las entregas. Usando la aproximación del ítem anterior, calcule el tiempo necesario para que el repartidor alcance a entregar todos los paquetes con probabilidad 0,95. Puede usar el hecho de que $\Phi(1,64) \approx 0,95$, donde Φ es la CDF de una normal estándar.

- (d) Si quisiéramos, en lugar de limitar el tiempo, aumentar el número de paquetes que se reparte cada día, ¿cuántos paquetes le podemos asignar al repartidor en un día como máximo si queremos asegurar que con probabilidad al menos 0,95 alcance a terminar en una jornada de 8 horas?

Solución

(a). Sea X_i el tiempo de la i -ésima entrega. Entonces

$$\mathbb{E}[X_i] = 5, \quad \text{Var}(X_i) = \frac{1}{2}(3-5)^2 + \frac{1}{2}(7-5)^2 = 4.$$

Como $T = 200 + \sum_{i=1}^{20} X_i$ y los X_i son independientes,

$$\mathbb{E}[T] = 200 + 20 \cdot 5 = 300, \quad \text{Var}(T) = 20 \cdot 4 = 80.$$

(b). Basta tomar $b = 300$ y $a = 1/\sqrt{80}$, de modo que

$$Z = \frac{T - 300}{\sqrt{80}}.$$

Esta es una suma estandarizada de variables i.i.d.; por el TCL, Z se aproxima por una $\mathcal{N}(0, 1)$.

(c). Como $\mathbb{P}(Z \leq 1,64) \approx 0,95$,

$$T \leq 300 + 1,64\sqrt{80} \approx 314,67.$$

Así, la aproximación normal sugiere un límite de 314,67 minutos. En el modelo exacto, sin embargo,

$$T = 260 + 4K, \quad K \sim \text{Binomial}\left(20, \frac{1}{2}\right),$$

por lo que los tiempos posibles están separados por cuatro minutos. En particular,

$$\mathbb{P}(T \leq 312) = \mathbb{P}(K \leq 13) \approx 0,9423,$$

$$\mathbb{P}(T \leq 316) = \mathbb{P}(K \leq 14) \approx 0,9793.$$

Como T solo toma valores en esa retícula, el menor límite que alcanza una probabilidad de al menos 0,95 es 316 minutos.

(d). Con n paquetes,

$$\mathbb{E}[T_n] = 15n, \quad \text{Var}(T_n) = 4n.$$

El cuantil aproximado de nivel 0,95 es $15n + 1,64(2\sqrt{n})$. Se exige

$$15n + 3,28\sqrt{n} \leq 480.$$

La expresión vale aproximadamente 467,97 para $n = 30$ y 483,26 para $n = 31$. Así, el máximo es $n = 30$.

Pregunta 160. (Examen - Primavera 2022, P3)

- a) La empresa de chalecos *Abrigaditto* toma la decisión de arrendar una máquina para tejer chalecos para potenciar la producción de la temporada de invierno 2023. Esta máquina tarda un tiempo $T \sim \text{Unif}[25, 45]$ minutos en producir un chaleco. Se desean producir 200 chalecos. El arriendo de la máquina se hace por un tiempo t [minutos] definido por el arrendatario, i.e., *Abrigaditto*. La estructura de costo por el arriendo es la siguiente: al solicitar la máquina por t minutos, se paga un costo de $\$c_1 \cdot t$ (este costo **siempre se paga**, independiente si la máquina se usa un tiempo menor a t), y en caso de exceder el tiempo acordado (t) se paga adicionalmente una multa de $\$c_2$ (independiente de cuánto tiempo extra se use la máquina).

- (i) Encuentre una aproximación para la probabilidad de que la máquina tarde más de t minutos en tejer los 200 chalecos utilizando el Teorema Central del Límite. *Observación:* El resultado debe quedar expresado sólo en términos de t y funciones

conocidas.

- (ii) Suponga que la empresa solo arrienda la máquina por 7100 minutos, bajo la misma estructura de costos mencionada antes. Sea $c_1 = \$1$ y $c_2 = \$10\,000$, calcule el costo esperado de arrendar la máquina por $t = 7100$ minutos.
- b) En un casino se encuentra el juego de la ruleta: donde uno apuesta un dólar por un número del 1 al 10 y, en caso de acertar, recibe un premio de \$25 dólares. Un gran jugador, Diego, es fanático del número 7, y decide apostar un dólar a dicho número hasta que éste aparezca. Notemos que, cada vez que Diego juega, o bien pierde un dólar, o gana \$25 dólares. Usando la desigualdad de Chebyshev, acote la probabilidad de que Diego termine con utilidad positiva, i.e., utilidad mayor a cero.
Hint 1: Podría serle de utilidad escribir la VA de la utilidad neta, computar su esperanza y varianza.
Hint 2: Note que según lo expuesto en el enunciado, para que Diego termine con utilidad positiva, debe jugar menos de 26 veces.
- c) Se jugará un campeonato de ajedrez a un solo partido entre dos jugadores. Ambos tienen prácticamente el mismo nivel, sin embargo, quien juegue con las piezas blancas partirá jugando, de modo que se podría afirmar que tiene más probabilidades de ganar. Considere que el rendimiento en el tablero del jugador con piezas blancas se puede modelar como una VA con distribución $\text{Unif}[4, 6]$, mientras que el rendimiento del jugador con negras puede modelarse como una VA $\text{Unif}[3, 6]$. Gana el partido el jugador que tiene mayor rendimiento. Calcule la probabilidad de que la partida la gane el jugador que juega con piezas negras. Asuma que las VA asociadas al rendimiento de ambos jugadores son independientes.

Solución

(a)(i). Bajo el supuesto adicional de que los tiempos $T_1, \dots, T_{200}$ son independientes e idénticamente distribuidos como $\text{Unif}[25, 45]$, sea $S = \sum_{i=1}^{200} T_i$. Entonces $\mathbb{E}[T_i] = 35$ y $\text{Var}(T_i) = 100/3$, de modo que

$$\mathbb{P}(S > t) \approx 1 - \Phi\left(\sqrt{6}\left(\frac{t}{200} - 35\right)\right).$$

(a)(ii). Para $t = 7100$, $c_1 = 1$ y $c_2 = 10000$,

$$\mathbb{E}[C(S)] \approx 7100 + 10000 \left[1 - \Phi\left(\frac{\sqrt{6}}{2}\right)\right] \approx 8203,4.$$

(b). Si $T \sim \text{Geom}(0,1)$ es el número de jugadas hasta obtener el primer 7, entonces $\mathbb{E}[T] = 10$ y $\text{Var}(T) = 90$. Como $\{T \geq 26\} \subseteq \{|T - 10| \geq 16\}$, Chebyshev entrega

$$\begin{aligned} \mathbb{P}(T < 26) &= 1 - \mathbb{P}(T \geq 26) \\ &\geq 1 - \mathbb{P}(|T - 10| \geq 16) \\ &\geq 1 - \frac{90}{16^2} = \frac{83}{128} \approx 0,6484. \end{aligned}$$

(c). Para rendimientos independientes $B \sim \text{Unif}[4, 6]$ y $N \sim \text{Unif}[3, 6]$,

$$\mathbb{P}(N > B) = \int_4^6 \int_4^n \frac{1}{2} \frac{1}{3} db dn = \frac{1}{3}.$$

Pregunta 161. (Examen - Otoño 2023, P1)

- (a) Una persona que va a viajar en tren hacia la playa, decide contratar un servicio de transporte de pasajeros que lo lleve en un vehículo desde su casa hasta el terminal de buses. El vehículo del servicio de transporte sale desde una central que está a 5 cuadras¹ de la casa del cliente, mientras que el terminal de buses está a 25 cuadras de la casa del cliente. Debido al tráfico y a los semáforos, el tiempo que el vehículo demora en recorrer una cuadra distribuye uniforme entre 0,5 y 2,5 minutos. Asuma que el pasajero se sube inmediatamente al auto cuando pasan a buscarlo a su casa, y que el tiempo que le toma al vehículo recorrer cualquier par de cuadras son independientes entre sí. Considerando que el vehículo sale desde la central a las 10:00, y que la persona debe estar en el terminal a las 10 : 50, estime la probabilidad de que el pasajero logre llegar a tiempo.

Hint: Considere que $n = 30$ es lo suficientemente grande para poder realizar aproximaciones ocupando los resultados de convergencia vistos en clases.

- (b) Una productora de eventos está planeando hacer un festival de música. El costo total de organizar el festival es de 1000 millones de pesos, y duraría 4 días: Jueves, Viernes, Sábado, y Domingo. Desde la directiva de la productora deciden que harán el festival sólo si *los ingresos totales son mayores o iguales al costo de organizar el festival* con probabilidad mayor a 0,7. Se sabe que en general los días Sábado y Domingo va más gente que los días Jueves y Viernes. Lo anterior se ve reflejado en que para cada día, los ingresos diarios (en millones de pesos) distribuyen Exponencial, tal que el parámetro de dicha distribución Exponencial depende del día:

$$\lambda_{\text{Jueves}} = \lambda_{\text{Viernes}} = 0,01 \text{ [millones de pesos]}^{-1}$$

$$\lambda_{\text{Sábado}} = \lambda_{\text{Domingo}} = 0,005 \text{ [millones de pesos]}^{-1}$$

Además, se sabe que los ingresos entre días diferentes son independientes.

- (i) Calcule el valor esperado y la varianza de los *ingresos totales*.
- (ii) Use la desigualdad de Markov para determinar si es posible concluir si la productora debería o no realizar el festival.
- (c) Un grupo de inversionistas debe analizar si comprar acciones de dos empresas. La *ganancia* (en millones de pesos) de las acciones de *Empresa 1* sigue una distribución normal con media 5 y varianza 2, mientras que la *ganancia* (en millones de pesos) de las acciones de *Empresa 2* distribuyen uniformemente entre -3 y 2 . Además, la correlación entre las ganancias de las acciones de ambas empresas tiene un valor de $\rho = 0,5$. Note que las *ganancias* de las acciones de ambas empresas pueden ser negativas, lo que significaría que el grupo de inversionistas tuvo pérdidas. Los inversionistas concluyen que invertirán comprando acciones de ambas empresas sólo si la *ganancia total esperada es positiva* y si la *diferencia absoluta entre la ganancia total y su valor esperado es menor a 2 desviaciones estándar* con probabilidad mayor a 0,6 (considerando la *ganancia total esperada* como la ganancia obtenida en conjunto por las acciones de ambas empresas).
- (i) Calcule el valor esperado y la varianza de la *ganancia esperada*.
- (ii) Use la desigualdad de Chebyshev para concluir si los inversionistas deberían o no comprar las acciones de ambas empresas.

¹Al hacer referencia a una *cuadra*, nos referimos simplemente a una métrica de distancia.

Solución

(a). Si $T = \sum_{i=1}^{30} U_i$ y $U_i \sim \text{Unif}[0,5, 2,5]$, entonces $\mathbb{E}[U_i] = 1,5$ y $\text{Var}(U_i) = 1/3$. Por el Teorema Central del Límite,

$$\mathbb{P}(T \leq 50) \approx \Phi\left(\frac{5}{\sqrt{10}}\right) \approx 0,943.$$

(b)(i). Si X es la suma de los cuatro ingresos diarios, la independencia entrega

$$\mathbb{E}[X] = 2 \frac{1}{0,01} + 2 \frac{1}{0,005} = 600,$$

$$\text{Var}(X) = 2 \frac{1}{0,01^2} + 2 \frac{1}{0,005^2} = 100000.$$

(b)(ii). Como $X \geq 0$, Markov implica $\mathbb{P}(X \geq 1000) \leq 600/1000 = 0,6 < 0,7$; según el criterio del enunciado, no debería realizarse el festival.

(c)(i). Sean G_1 y G_2 las ganancias de ambas empresas y $Z = G_1 + G_2$. Como $\text{Var}(G_2) = 25/12$ y $\text{Cov}(G_1, G_2) = 0,5\sqrt{2}\sqrt{25/12} = 5/(2\sqrt{6})$,

$$\mathbb{E}[Z] = 5 - \frac{1}{2} = \frac{9}{2}, \quad \text{Var}(Z) = 2 + \frac{25}{12} + \frac{5}{\sqrt{6}} \approx 6,125.$$

(c)(ii). Chebyshev garantiza $\mathbb{P}(|Z - \mathbb{E}[Z]| < 2\sigma_Z) \geq 3/4 > 0,6$ y, además, $\mathbb{E}[Z] > 0$. Por ello se satisfacen ambas condiciones de inversión.

Pregunta 162. (Examen - Primavera 2023, P3)

Un barco carguero lleva *contenedores* de un puerto a otro, su capacidad es de 5 espacios distribuidos a lo ancho del barco donde puede ser asignada la carga. Dichos espacios se ilustran en la Figura 9.1, donde el punto “C” indica el centro del barco, mientras que “A” indica la posición de la izquierda del barco, y “E” indica la posición de la derecha del barco. La distancia entre dos posiciones consecutivas de la Figura 9.1 es de 10 [m]. Suponga que existen dos tipos de carga que se puede llevar: la carga pesada de peso M [kg] y la carga liviana de peso 0 [kg] (despreciable). Dependiendo de como se distribuye la carga, el barco puede desequilibrarse según el torque que pueda ejercer el conjunto de cargas sobre el barco. Considere que el torque es con respecto al punto “C” indicado en la Figura 9.1. Asuma que la cantidad de *contenedores* con carga pesada es 2 y con carga liviana 3 (fijas).

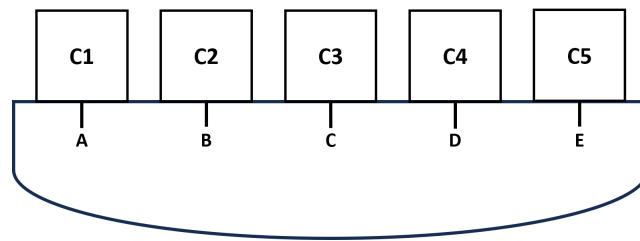


Figura 9.1: Distribución de cargas visto desde el frente del barco

a) Encuentre la función de masa de probabilidad del torque.

Hint: El toque se define como la suma de la fuerza de las masas por el largo del brazo

$$\vec{t} = \vec{r} \cdot F$$

donde $\vec{r}$ es el brazo y F es la fuerza. Note que la fuerza es igual a la masa del objeto por la aceleración de gravedad, g . Por ejemplo, si hubiese una única carga de a [kg] en el espacio

“A”, entonces el toque sería de $20 \cdot a \cdot g$. Por otra parte si hubiese una carga de única carga de b [kg] en el espacio “D”, entonces el toque sería de $-10 \cdot b \cdot g$ (note que es negativo el torque en este caso). Si se instala una única carga en “C”, entonces el torque sería cero.

- b) Encuentre la esperanza y la varianza del torque.
- c) El barco se hunde cuando el desequilibrio de masas es tal que el valor absoluto del torque es mayor a $25Mg$. Encuentre una cota, usando la desigualdad de Chebyshev, para la probabilidad de que el barco llegue a hundirse.

Solución

El enunciado no especifica el mecanismo de asignación de las cargas a los cinco espacios, por lo que la función de masa del torque no queda determinada sin una hipótesis adicional. En lo que sigue suponemos que cada uno de los $\binom{5}{2}$ pares de posiciones para las dos cargas pesadas es igualmente probable.

(a). Las posiciones de los dos contenedores pesados forman uno de los $\binom{5}{2} = 10$ pares equiprobables. Al sumar sus brazos de palanca, medidos en unidades de 10 metros,

$$p_T(t) = \begin{cases} 1/10, & t \in \{-30Mg, -20Mg, 20Mg, 30Mg\}, \\ 1/5, & t \in \{-10Mg, 0, 10Mg\}, \\ 0, & \text{en otro caso.} \end{cases}$$

(b). Por simetría, $\mathbb{E}[T] = 0$. Usando la función de masa anterior,

$$\text{Var}(T) = \mathbb{E}[T^2] = \frac{(10Mg)^2}{10} (2 \cdot 3^2 + 2 \cdot 2^2 + 4 \cdot 1^2) = 300(Mg)^2.$$

(c). La desigualdad de Chebyshev entrega

$$\mathbb{P}(|T| > 25Mg) \leq \frac{300}{25^2} = 0,48.$$

La probabilidad exacta, aunque no es necesaria para la cota, es $\mathbb{P}(T \in \{-30Mg, 30Mg\}) = 0,2$.

Pregunta 163. (Examen - Otoño 2024, P3)

Responda las siguientes preguntas.

- (a) El número de clientes que visitan una tienda durante un día es una variable aleatoria X con $\mathbb{E}[X] = 100$ y $\text{Var}(X) = 225$. Usando la desigualdad de Chebyshev, encuentre una cota superior para

$$\mathbb{P}(X \leq 80 \cup X \geq 120).$$

- (b) En el automovilismo internacional, las competiciones se organizan en circuitos de carreras donde compiten varias categorías, como la Fórmula 1, Fórmula 2, Fórmula E y otras. El número de incidentes en un determinado circuito se modela mediante una variable aleatoria cuya distribución es Poisson con parámetro igual a 10 accidentes. A lo largo de un año se realizan carreras en 22 circuitos diferentes; suponga que el número de incidentes en los circuitos es independiente entre circuitos. Utilice el Teorema Central del Límite para aproximar la probabilidad de que haya más de 250 incidentes en un determinado año.

Solución

(a). El evento pedido es $\{|X - 100| \geq 20\}$. Chebyshev entrega

$$\mathbb{P}(X \leq 80 \cup X \geq 120) = \mathbb{P}(|X - 100| \geq 20) \leq \frac{225}{20^2} = \frac{9}{16}.$$

(b). Si $S = \sum_{i=1}^{22} X_i$, entonces $\mathbb{E}[S] = 220$ y $\text{Var}(S) = 220$. Como S toma valores enteros, el evento $\{S > 250\}$ equivale a $\{S \geq 251\}$. Aplicando el TCL con corrección de continuidad,

$$\mathbb{P}(S > 250) \approx 1 - \Phi\left(\frac{250,5 - 220}{\sqrt{220}}\right) = 1 - \Phi(2,0563) \approx 0,0199.$$

Pregunta 164. (Examen - Primavera 2024, P3)

- (a) Demuestre la Ley Débil de los Grandes Números: Sea $X_1, X_2, \dots, X_n$ VAs i.i.d con $\mathbb{E}[X_i] = \mu < \infty$, entonces $\forall \epsilon > 0$, $\bar{X}_n \xrightarrow{p} \mu$ o equivalentemente:

$$\lim_{n \rightarrow \infty} \mathbb{P}(|\bar{X}_n - \mu| \geq \epsilon) = 0$$

Recuerde que $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$. Asuma, para la demostración, que $\text{Var}(X_i) = \sigma^2 < \infty$.
Hint: Le puede ser útil la materia vista post Control 3.

- (b) Un estudiante usa dos medios de transporte en su trayecto para llegar a su casa desde la universidad. El primer trayecto es en auto; el segundo, caminando.

- i) Considere que la distribución conjunta del tiempo de viaje en auto y caminando se pueden representar con el vector Normal Multivariado

$$\mathcal{N}\left(\begin{pmatrix} 25 \\ 20 \end{pmatrix}, \begin{pmatrix} 16 & 0.8 \\ 0.8 & 4 \end{pmatrix}\right).$$

Calcule la probabilidad de que el estudiante se demore 40 minutos o menos en llegar a su casa desde la universidad. Compute el valor exacto.

- ii) Ahora, asuma que el tiempo de viaje en auto distribuye $\text{Exp}(0.05)$, y el tiempo de caminata distribuye $\text{Uniforme}[15, 20]$. Asuma que ambas VA son independientes. Calcule la probabilidad de que el tiempo total de transporte sea mayor a 30 minutos.

Solución

(a). La independencia y la idéntica distribución implican

$$\begin{aligned} \mathbb{E}[\bar{X}_n] &= \frac{1}{n} \sum_{i=1}^n \mathbb{E}[X_i] = \mu, \\ \text{Var}(\bar{X}_n) &= \frac{1}{n^2} \sum_{i=1}^n \text{Var}(X_i) = \frac{\sigma^2}{n}. \end{aligned}$$

Para todo $\epsilon > 0$, la desigualdad de Chebyshev da

$$0 \leq \mathbb{P}(|\bar{X}_n - \mu| \geq \epsilon) \leq \frac{\text{Var}(\bar{X}_n)}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2}.$$

El extremo derecho converge a cero cuando $n \rightarrow \infty$. Por el teorema del sándwich,

$$\lim_{n \rightarrow \infty} \mathbb{P}(|\bar{X}_n - \mu| \geq \epsilon) = 0,$$

que es precisamente $\bar{X}_n \xrightarrow{p} \mu$.

(b)(i). Sean X el tiempo en auto y Y el tiempo caminando. Como (X, Y) es un vector Normal multivariado, la suma $T = X + Y$ es Normal. A partir del vector de medias y la matriz de

covarianza,

$$\begin{aligned}\mathbb{E}[T] &= 25 + 20 = 45, \\ \text{Var}(T) &= 16 + 4 + 2(0,8) = 21,6.\end{aligned}$$

Por tanto,

$$T \sim \mathcal{N}(45, 21,6),$$

y

$$\mathbb{P}(T \leq 40) = \Phi\left(\frac{40 - 45}{\sqrt{21,6}}\right) = \Phi\left(-\frac{5}{\sqrt{21,6}}\right) \approx 0,1410.$$

(b)(ii). Ahora $X \sim \text{Exponencial}(0,05)$ y $Y \sim \text{Uniforme}[15, 20]$. Por independencia, conviene condicionar respecto de Y :

$$\begin{aligned}\mathbb{P}(X + Y > 30) &= \int_{15}^{20} \mathbb{P}(X > 30 - y \mid Y = y) f_Y(y) dy \\ &= \frac{1}{5} \int_{15}^{20} e^{-0,05(30-y)} dy \\ &= 4 \left(e^{-1/2} - e^{-3/4} \right).\end{aligned}$$

Pregunta 165. (Examen - Otoño 2025, P3)

Responda las siguientes preguntas.

- (a) Agustín visita un casino a probar una ruleta con 100 números equiprobables. En cada jugada apuesta \$10 a un número. Si acierta, recibe un pago de a veces su apuesta, es decir, gana \$10 a ; si no acierta, pierde su apuesta. Agustín realiza 1100 apuestas. Sea X_i la ganancia obtenida en la i -ésima apuesta.

El casino desea que la probabilidad de que Agustín obtenga una ganancia promedio mayor que cero sea 0,0025. Usando el Teorema Central del Límite, encuentre el mínimo valor de a que cumple esta condición. Puede usar que $X_i = c_1 + c_2 Z_i$, donde Z_i distribuye Bernoulli.

- (b) Un centro de salud llama, de forma aleatoria y anónima, a personas vacunadas para detectar efectos adversos leves. Una de cada cuatro personas reporta algún efecto y las llamadas son independientes. Se llama hasta encontrar a la primera persona que confirme un efecto adverso. Utilizando alguna desigualdad vista en el curso, entregue una cota inferior para la probabilidad de que el número de llamadas necesarias sea menor que 7.

Solución

(a). Sea $Z_i \sim \text{Bernoulli}(1/100)$. Entonces

$$X_i = -10 + 10(a + 1)Z_i,$$

y, por las propiedades de una transformación afín,

$$\mu := \mathbb{E}[X_i] = \frac{a - 99}{10}, \quad \sigma^2 := \text{Var}(X_i) = \frac{99(a + 1)^2}{100}.$$

Para $n = 1100$, el TCL aproxima

$$\mathbb{P}(\bar{X} > 0) \approx 1 - \Phi\left(\frac{-\mu}{\sigma/\sqrt{n}}\right).$$

Sea $z = \Phi^{-1}(0,9975) \approx 2,807$. La condición pedida equivale a

$$\frac{(99-a)\sqrt{1100}}{\sqrt{99}(a+1)} = z.$$

Si $A = \sqrt{1100/99}$, al despejar se obtiene

$$a = \frac{99A - z}{A + z} \approx 53,29.$$

Este es el valor mínimo que entrega el sustituto normal sin corrección de continuidad. No constituye una garantía exacta para el modelo binomial, cuya probabilidad es una función escalonada de a .

(b). Sea $X \sim \text{Geom}(1/4)$ sobre $\{1, 2, \dots\}$. Entonces $\mathbb{E}[X] = 4$. Markov implica

$$\mathbb{P}(X \geq 7) \leq \frac{\mathbb{E}[X]}{7} = \frac{4}{7}.$$

Tomando complementos,

$$\mathbb{P}(X < 7) \geq 1 - \frac{4}{7} = \frac{3}{7}.$$

Pregunta 166. (Examen - Primavera 2025, P3)

Usted y 35 amigos están dando un paseo en bicicleta. Cada uno pedalea una distancia que sigue una distribución exponencial independiente, con igual parámetro, tal que la esperanza de la distancia que recorre un ciclista es de 2 kilómetros.

- (a) Usted apuesta a que la distancia promedio recorrida por el grupo sea mayor que 2,5 kilómetros, por lo que paga necesariamente un monto $\$c$ y recibe un monto $\$r$ si ocurre el evento. Usando el Teorema Central del Límite, ¿cuánto debe valer r para que decida apostar, es decir, para que su esperanza de ganancia sea no negativa?
- (b) ¿Cuántos ciclistas se necesitan para garantizar, por la desigualdad de Chebyshev, que la probabilidad de que el promedio de la distancia recorrida sea mayor que 2,4 km sea menor o igual que 0,1?

Suponga ahora que hay solo dos corredores: uno experimentado y otro novato. La distancia recorrida por cada uno distribuye exponencialmente, de manera independiente, con tasas diferentes. Defina claramente la notación usada para cada tasa e indique, con justificación, qué tasa debería ser mayor. Además, determine la probabilidad de que el corredor novato recorra más kilómetros que el experimentado.

Solución

(a). Cada distancia X_i tiene media 2 y varianza 4. Con $n = 36$, el TCL entrega

$$\mathbb{P}(\bar{X} > 2,5) \approx \mathbb{P}\left(Z > \frac{2,5 - 2}{2/\sqrt{36}}\right) = 1 - \Phi(1,5) \approx 0,0668.$$

La esperanza de la ganancia es $r \mathbb{P}(\bar{X} > 2,5) - c$. Por tanto, debe cumplirse

$$r \geq \frac{c}{0,0668} \approx 14,97c.$$

(b). Para n ciclistas, $\text{Var}(\bar{X}) = 4/n$. Entonces

$$\mathbb{P}(\bar{X} > 2,4) \leq \mathbb{P}(|\bar{X} - 2| \geq 0,4) \leq \frac{4/n}{0,4^2} = \frac{25}{n}.$$

Para que esta cota sea a lo más 0,1, se requiere $n \geq 250$.

Dos corredores. Sea $E \sim \text{Exp}(\lambda_E)$ la distancia del experimentado y $N \sim \text{Exp}(\lambda_N)$ la del novato. Si el experimentado recorre más en promedio, entonces

$$\frac{1}{\lambda_E} > \frac{1}{\lambda_N}, \quad \text{es decir,} \quad \lambda_E < \lambda_N.$$

Por independencia,

$$\mathbb{P}(N > E) = \int_0^\infty \mathbb{P}(N > x) f_E(x) dx = \int_0^\infty e^{-\lambda_N x} \lambda_E e^{-\lambda_E x} dx = \frac{\lambda_E}{\lambda_E + \lambda_N}.$$

Pregunta 167. (Examen v2 - Primavera 2025, P3)

1. Se lanza repetidamente un dado hasta que la suma total de todos los lanzamientos excede 300. Aproximar la probabilidad de que sean necesarios al menos 80 lanzamientos.
2. Suponga que X es una variable aleatoria con media y varianza ambas iguales a 20. ¿Qué se puede decir acerca de $\mathbb{P}(0 < X < 40)$?
3. La densidad conjunta de X e Y está dada por

$$f(x, y) = \frac{e^{-y}}{y}, \quad 0 < x < y, \quad 0 < y < \infty.$$

Calcule $\mathbb{E}[X^3 \mid Y = y]$.

Solución

(1). Sea X_i el resultado del lanzamiento i y defina $S_n = \sum_{i=1}^n X_i$. Para un dado balanceado,

$$\mathbb{E}[X_i] = \frac{1 + 2 + 3 + 4 + 5 + 6}{6} = \frac{7}{2},$$

$$\mathbb{E}[X_i^2] = \frac{1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2}{6} = \frac{91}{6},$$

$$\text{Var}(X_i) = \mathbb{E}[X_i^2] - \mathbb{E}[X_i]^2 = \frac{35}{12}.$$

Si $N = \min\{n : S_n > 300\}$ es el número de lanzamientos requeridos, los incrementos positivos implican la identidad de eventos

$$\{N \geq 80\} = \{S_{79} \leq 300\}.$$

Por el Teorema Central del Límite,

$$S_{79} \approx \mathcal{N}\left(79 \frac{7}{2}, 79 \frac{35}{12}\right) = \mathcal{N}\left(276,5, \frac{2765}{12}\right).$$

Como S_{79} es entera, incorporamos la corrección de continuidad en 300,5. Así,

$$\begin{aligned} \mathbb{P}(N \geq 80) &= \mathbb{P}(S_{79} \leq 300) \\ &\approx \Phi\left(\frac{300,5 - 276,5}{\sqrt{2765/12}}\right) \\ &= \Phi(1,5811) \approx 0,9431. \end{aligned}$$

(2). Como $\mathbb{E}[X] = 20$ y $\text{Var}(X) = 20$, el evento de interés puede escribirse como

$$\{0 < X < 40\} = \{|X - 20| < 20\}.$$

La desigualdad de Chebyshev entrega

$$\begin{aligned} \mathbb{P}(0 < X < 40) &= 1 - \mathbb{P}(|X - 20| \geq 20) \\ &\geq 1 - \frac{\text{Var}(X)}{20^2} = 1 - \frac{20}{400} = 0,95. \end{aligned}$$

(3). Para $y > 0$, la densidad marginal de Y es

$$f_Y(y) = \int_0^y f(x, y) dx = \int_0^y \frac{e^{-y}}{y} dx = e^{-y}.$$

Por tanto, la densidad condicional de X dado $Y = y$ es

$$f_{X|Y}(x | y) = \frac{f(x, y)}{f_Y(y)} = \frac{1}{y}, \quad 0 < x < y.$$

Al integrar respecto de esta densidad,

$$\begin{aligned} \mathbb{E}[X^3 | Y = y] &= \int_0^y x^3 f_{X|Y}(x | y) dx \\ &= \frac{1}{y} \int_0^y x^3 dx = \frac{y^3}{4}. \end{aligned}$$

Pregunta 168. (Examen - Otoño 2026, P3)

- Una empresa tecnológica ofrece un servicio de inteligencia artificial en la nube. Han determinado que el tiempo X , medido en segundos, que tarda su servidor principal en procesar una consulta compleja sigue una distribución Exponencial con una tasa de λ consultas por segundo.

Para garantizar la calidad del servicio, el equipo de ingeniería está preocupado por los “tiempos de espera extremos”. Han definido que si una consulta tarda **cinco** veces el tiempo esperado, el sistema lanzará un error de *Timeout*. A usted se le ha pedido encontrar una cota superior para la probabilidad de que ocurra este tipo de error, usando:

- La desigualdad de Markov.
 - La desigualdad de Chebyshev.
 - La cota de Chernoff, encontrando la mejor cota posible.
- El Instituto Meteorológico Nacional está realizando un estudio sobre los patrones de temperatura durante la temporada de verano en la ciudad. Para ello, instala estaciones de monitoreo que registran la temperatura máxima diaria.

De acuerdo con el historial, se asume que las temperaturas máximas diarias son independientes e idénticamente distribuidas, con temperatura media de 32°C y desviación estándar de 3°C .

El equipo quiere determinar cuántos días debe registrar para que el promedio de las temperaturas observadas difiera de la media histórica de 32°C en a lo más 1°C con una probabilidad de al menos 95 %. Ayude al equipo a determinar este número mínimo de días, utilizando la aproximación del Teorema Central del Límite.

- Un amigo afirma que el Teorema Central del Límite establece que el promedio de un gran número de variables aleatorias con la misma distribución, media finita y varianza finita tiene, aproximadamente, una distribución normal. Determine si esta afirmación es verdadera o falsa. Si es verdadera, justifique su respuesta; si es falsa, proporcione un contraejemplo.

Solución

(1)(a). Como $X \sim \text{Exponencial}(\lambda)$, se tiene $\mathbb{E}[X] = 1/\lambda$, y el evento de *timeout* es $\{X \geq 5/\lambda\}$.

La variable X es no negativa, por lo que la desigualdad de Markov entrega

$$\mathbb{P}\left(X \geq \frac{5}{\lambda}\right) \leq \frac{\mathbb{E}[X]}{5/\lambda} = \frac{1}{5}.$$

(1)(b). Además, $\text{Var}(X) = 1/\lambda^2$. La inclusión de eventos

$$\left\{X \geq \frac{5}{\lambda}\right\} \subseteq \left\{\left|X - \frac{1}{\lambda}\right| \geq \frac{4}{\lambda}\right\}$$

permite aplicar la desigualdad de Chebyshev:

$$\begin{aligned}\mathbb{P}\left(X \geq \frac{5}{\lambda}\right) &\leq \mathbb{P}\left(\left|X - \frac{1}{\lambda}\right| \geq \frac{4}{\lambda}\right) \\ &\leq \frac{1/\lambda^2}{(4/\lambda)^2} = \frac{1}{16}.\end{aligned}$$

(1)(c). La función generadora de momentos de X es

$$M_X(s) = \frac{\lambda}{\lambda - s}, \quad s < \lambda.$$

Para cualquier $s \in (0, \lambda)$, la cota de Chernoff da

$$\mathbb{P}\left(X \geq \frac{5}{\lambda}\right) \leq e^{-5s/\lambda} M_X(s) = \frac{\lambda}{\lambda - s} e^{-5s/\lambda} =: g(s).$$

Para minimizarla, basta minimizar su logaritmo. En efecto,

$$\frac{d}{ds} \log g(s) = \frac{1}{\lambda - s} - \frac{5}{\lambda}, \quad \frac{d^2}{ds^2} \log g(s) = \frac{1}{(\lambda - s)^2} > 0.$$

El único punto crítico es $s^* = 4\lambda/5$, que por la segunda derivada es el minimizador. Así, la mejor cota de esta familia es

$$\mathbb{P}\left(X \geq \frac{5}{\lambda}\right) \leq g(s^*) = 5e^{-4} \approx 0,0916.$$

En este ejemplo la cota de Chernoff es menos ajustada que la de Chebyshev; las distintas desigualdades no tienen por qué quedar ordenadas. Como referencia, la probabilidad exacta es e^{-5} , pero no se necesita para obtener ninguna de las cotas pedidas.

(2). Sean $X_1, \dots, X_n$ las temperaturas máximas diarias y $\bar{X}_n = n^{-1} \sum_{i=1}^n X_i$. Por independencia y distribución idéntica, el teorema central del límite aproxima

$$\frac{\bar{X}_n - 32}{3/\sqrt{n}} \quad \text{por una variable } Z \sim \text{Normal}(0, 1).$$

Por simetría de la normal estándar,

$$\begin{aligned}\mathbb{P}(|\bar{X}_n - 32| \leq 1) &\approx \mathbb{P}\left(-\frac{\sqrt{n}}{3} \leq Z \leq \frac{\sqrt{n}}{3}\right) \\ &= 2\Phi\left(\frac{\sqrt{n}}{3}\right) - 1.\end{aligned}$$

Para que esta aproximación sea al menos 0,95, se requiere

$$\begin{aligned}2\Phi\left(\frac{\sqrt{n}}{3}\right) - 1 &\geq 0,95 \\ \iff \frac{\sqrt{n}}{3} &\geq \Phi^{-1}(0,975) \approx 1,96 \\ \iff n &\geq (3 \cdot 1,96)^2 = 34,5744.\end{aligned}$$

El menor número entero que satisface el criterio aproximado es $n = 35$ días.

(3). La afirmación es falsa porque omite una hipótesis de independencia. Sea $B \sim \text{Bernoulli}(1/2)$ y defina $X_i = B$ para todo $i \geq 1$. Todas las variables X_i tienen la misma distribución, media y

varianza finitas, pero son perfectamente dependientes. Para todo n ,

$$\overline{X}_n = \frac{1}{n} \sum_{i=1}^n X_i = B,$$

de modo que el promedio conserva una distribución Bernoulli y no se aproxima a una normal. La versión usual del teorema central del límite requiere, entre otras condiciones posibles, que las variables sean independientes e idénticamente distribuidas.