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

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

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

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

   a) $\mathcal{B}i(n, p)$. *Sugerencia*: ¿Qué distribución tiene la suma de n variables aleatorias independientes $Be(p)$?

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

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

   d) $\mathcal{P}(\lambda)$. *Sugerencia*: Para hallar $\mathrm{Var}(X)$ calcular primero $\mathbb{E}\left(X\left(X-1\right)\right)$.

   e) $\Gamma(\alpha, \lambda)$. *Sugerencia*: Si $X \sim \Gamma(\alpha, \lambda)$ entonces $\int_{-\infty}^{+\infty} f_X\left(x\right)dx = 1$ para **cualquier** α y λ positivos.

   f) $\varepsilon(\lambda)$.

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

   h) $\mathcal{N}(\mu, \sigma^2)$.

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

   j) $\beta(a, b)$. *Sugerencia*: Si $X \sim \beta(a, b)$ entonces $\int_{-\infty}^{+\infty} f_X\left(x\right)dx = 1$ para **cualquier** a y b positivos.

5. Sean $X_1, \ldots, X_n$ variables aleatorias positivas independientes e idénticamentes distribuídas. Probar que

$$\mathbb{E}\left(\frac{X_1 + \cdots + X_k}{X_1 + \cdots + X_n}\right) = \frac{k}{n}.$$

   Sugerencia: Probar que $\dfrac{X_i}{X_1 + \cdots + X_n} \sim \dfrac{X_j}{X_1 + \cdots + X_n}$ para todo par $1 \leq i, j \leq n$.

6. Se distribuyen al azar N bolillas indistinguibles en m urnas. Sean X el número de urnas vacías, Y el número de urnas que contienen exactamente una bolilla y Z el número de urnas que contienen dos o más bolillas.

   a) Hallar $\mathbb{E}(X)$.
   Sugerencia: Sea

$$X_i = \begin{cases} 1 & \text{si la } i\text{-ésima urna está vacía} \\ 0 & \text{en caso contrario.} \end{cases}$$

   Verificar que $X = \sum_{i=1}^{m} X_i$.

   b) Hallar $\mathbb{E}(Y)$ y $\mathbb{E}(Z)$.

c) Un centro cultural dispone de m cuentas de correo electrónico para comunicarse con el público. Durante un día en particular, N personas envían sus inquietudes vía e-mail al centro cultural, eligiendo una cuenta al azar para hacerlo. Hallar la esperanza del número de cuentas de correo que no son usadas durante dicho día.

7. Muestreo estratificado
Se quiere saber cuántos habitantes viven en una cierta ciudad. Se sabe que dicha ciudad tiene n manzanas, de las cuales n_j tienen x_j habitantes cada una ($n_1 + n_2 + \cdots = n$). Sea $m = \sum_j \frac{n_j x_j}{n}$ el número medio de habitantes por manzana y sea $a^2 = \frac{1}{n}\sum_j n_j x_j^2 - m^2$. Para averiguarlo se sortean al azar r manzanas para encuestarlas. Un encuestador es enviado a cada una de las manzanas sorteadas para contar la cantidad de habitantes que viven en ellas. Definamos las variables aleatorias

$$Z_i = \text{cantidad de habitantes que viven en la } i\text{-ésima manzana sorteada}$$
$$Y = \text{cantidad de personas encuestadas.}$$

a) Hallar $\mathbb{E}(Z_i)$ y $\mathrm{Var}(Z_i)$.

b) Mostrar que

$$\mathbb{E}(Y) = mr$$
$$\mathrm{Var}(Y) = \frac{a^2 r(n-r)}{n-1}.$$

8. Sean $X_1, \ldots, X_n$ variables aleatorias independientes tales que

$$P(X_k \leq t) = \begin{cases} 0 & \text{si} \quad t < 0 \\ t^k & \text{si} \quad 0 \leq t \leq 1 \\ 1 & \text{si} \quad t > 1 \end{cases} \qquad k = 1, \ldots, n$$

a) Sea $Y = \max(X_1, \ldots, X_n)$. Hallar F_Y, f_Y y $\mathbb{E}(Y)$.

b) Hallar $\mathbb{E}(\prod_{i=1}^n X_i)$.

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

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

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

a) Hallar $\mathbb{E}(X)$, $\mathbb{E}(Y)$, $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$, $\mathrm{Cov}(X,Y)$, $\mathrm{Var}(X-Y)$ y $\rho(X,Y)$.

b) Calcular $\mathbb{E}(Z)$ y $\mathrm{Var}(Z)$ donde $Z = -4X + 1$.

10. Sea (X, Y) un vector aleatorio con función de densidad conjunta

$$f_{XY}(x,y) = \begin{cases} \frac{6}{7}(3y + x) & 0 \leq x \leq y \leq 1 \\ 0 & \text{en caso contrario} \end{cases}$$

Hallar $\mathbb{E}(X)$, $\mathbb{E}(Y)$, $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$, $\mathrm{Cov}(X,Y)$, $\mathrm{Var}(X-Y)$ y $\rho(X,Y)$.

11. Sea $(X_1, \ldots, X_k)$ un vector aleatorio con distribución $\mathcal{M}(p_1, \ldots, p_k, n)$ para $n > 2$.

a) Hallar $\mathbb{E}(X_i)$, $\mathrm{Var}(X_i)$ y $\mathrm{Cov}(X_i, X_j)$. Interpretar.

b) Hallar el mejor predictor lineal de X_1 basado en $X_2 + X_3$ y el error de predicción.

12. Dada una urna con N bolillas de las cuales D son blancas y $N - D$ son negras, se extraen n sin reposición. Sean

$$X = \text{número de bolillas blancas extraídas}$$

$$X_i = \begin{cases} 1 & \text{si la i-ésima bolilla extraída es blanca} \\ 0 & \text{si la i-ésima bolilla extraída es negra.} \end{cases}$$

Observar que $X \sim \mathcal{H}(N, D, n)$.

a) Mostrar que $P(X_i = 1) = \frac{D}{N}$ para todo $i = 1, \ldots, n$ y que para $i \neq j$ se tiene

$$P(X_i = 1, X_j = 1) = \frac{D(D-1)}{N(N-1)}.$$

Determinar la distribución conjunta del vector (X_i, X_j).

b) Calcular $\mathbb{E}(X_i)$, $\mathrm{Var}(X_i)$.

c) Calcular $\mathrm{Cov}(X_i, X_j)$ para $i \neq j$.

d) Hallar $\mathbb{E}(X)$. Verificar que

$$\mathrm{Var}(X) = \frac{(N-n)nD(N-D)}{(N-1)N^2}$$

13. El problema del coleccionista de cupones

Un hombre colecciona cupones de un álbum compuesto por n cupones distintos. El hombre adquiere sus cupones comprando uno por día en el kiosko de la esquina de su casa y, cada vez que adquiere uno, éste tiene igual probabilidad de ser cualquiera de los n que componen el álbum.

a) Hallar la esperanza del número de cupones diferentes que hay en un conjunto de k cupones.

b) Hallar el número esperado de cupones que es necesario juntar para completar el álbum.

14. *a*) Mostrar con un ejemplo que $\mathrm{Cov}(X, Y) = 0$ no implica que X e Y sean independientes.

Sugerencia: Puede considerar un vector aleatorio con densidad uniforme en el pentágono de vértices $(-1, 1)$, $(1, 1)$, $(-1, 0)$, $(0, -1)$ y $(1, 0)$.

b) Sean X e Y dos variables aleatorias que toman sólo dos valores cada una. Probar que si $\mathrm{Cov}(X, Y) = 0$ entonces X e Y son independientes.

Sugerencia: Considerar primero el caso en que los dos valores que toma cada variable son 0 y 1.

15. Probar que si X es una variable aleatoria con distribución simétrica respecto de $\theta \in \mathbb{R}$ y $\mathbb{E}(|X|) < \infty$ entonces $\mathbb{E}(X) = \theta$.

16. **Esquema de Polya**. De un bolillero que contiene B bolillas blancas y R rojas se extrae una bolilla al azar y se la devuelve al bolillero junto con otras c bolillas del mismo color. Se repite este procedimiento sucesivamente comenzando en cada nuevo paso con la composición del bolillero resultante del paso anterior. Sean

$$X_i = \begin{cases} 1 & \text{si la } i\text{-ésima bolilla extraida es roja} \\ 0 & \text{si la } i\text{-ésima bolilla extraida es blanca} \end{cases}$$

a) ¿Qué distribución tienen las variables aleatorias X_i?

b) Hallar $\mathbb{E}(X_i)$, $\mathrm{Var}(X_i)$, $\mathrm{Cov}(X_i, X_j)$ y $\rho(X_i, X_j)$ para $i < j$. ¿Son independientes las variables X_i? Contestar a partir las cantidades calculadas.

c) Sea $S_j = \sum_{i=1}^{j} X_i$ el número de bolillas rojas extraídas luego de j extracciones. Hallar $\mathbb{E}(S_j)$, $\mathrm{Var}(S_j)$, $\mathrm{Cov}(S_i, S_j)$ y $\rho(S_i, S_j)$ para $i < j$.

17. Sea $\mathbf{Z} = (Z_1, Z_2)$ un vector aleatorio de marginales Z_1 y Z_2 independientes y con distribución $\mathrm{N}(0, 1)$. Dada $A \in \mathbb{R}^{2\times 2}$ inversible y $\mathbf{c} \in \mathbb{R}^2$, sea $\mathbf{Y} = (Y_1, Y_2)$ el vector aleatorio definido por $\mathbf{Y} = A \cdot \mathbf{Z} + \mathbf{c}$.

a) Calcular $\mathbb{E}(Y_i)$ y $\mathrm{Var}(Y_i)$ para todo i. Observar que, por el ejercicio 18 de la Práctica 5, cada Y_i tiene distribución normal.

b) Hallar $\mathrm{Cov}(Y_1, Y_2)$ y $\rho(Y_1, Y_2)$ a partir del cálculo de $\mathrm{Cov}(Z_1, Z_2)$. Observar que $|\rho(Y_1, Y_2)| < 1$.

c) Verificar que si para $\mathbf{Y}$ definimos su vector de medias $\mu \in \mathbb{R}^2$ por $\mu_i = \mathbb{E}(Y_i)$ y su matriz de covarianzas $\Sigma \in \mathbb{R}^{2\times 2}$ como $\Sigma_{ij} = \mathrm{Cov}(Y_i, Y_j)$ entonces $\mu = \mathbf{c}$ y $\Sigma = A \cdot A^t$.

d) Concluir a partir del ítem anterior que Σ es inversible y que tanto ella como su inversa son matrices simétricas y definidas positivas.

e) Mostrar que la función de densidad conjunta de $\mathbf{Y}$ viene dada por

$$f_{Y_1 Y_2}(y_1, y_2) = \frac{1}{2\pi\sigma_{Y_1}\sigma_{Y_2}\sqrt{1-\rho^2}} e^{\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{y_1-\mu_{Y_1}}{\sigma_{Y_1}}\right)^2 + \left(\frac{y_2-\mu_{Y_2}}{\sigma_{Y_2}}\right)^2 - 2\rho\left(\frac{y_1-\mu_{Y_1}}{\sigma_{Y_1}}\right)\left(\frac{y_2-\mu_{Y_2}}{\sigma_{Y_2}}\right)\right]\right\}} \tag{1}$$

donde notamos $\mu_{Y_i} = \mathbb{E}(Y_i)$, $\sigma^2_{Y_i} = \mathrm{Var}(Y_i)$ y $\rho = \rho(Y_1, Y_2)$.[1] Verificar que la densidad también puede escribirse en notación matricial como

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{1}{2\pi\left[\det(\Sigma)\right]^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(\mathbf{y}-\mu)^t \Sigma^{-1}(\mathbf{y}-\mu)\right\}.$$

¿Qué tipo de curvas de nivel tiene $f_{\mathbf{Y}}$?

f) Deducir que Y_1 e Y_2 son independientes si y sólo si $\mathrm{Cov}(Y_1, Y_2) = 0$.

g) Sea $\mathbf{X}$ un vector aleatorio continuo con densidad conjunta como en (1). Probar que existen $\mathbf{c} \in \mathbb{R}^2$ y $A \in \mathbb{R}^{2\times 2}$ con $\det A \neq 0$ tales que $\mathbf{X} = A \cdot \mathbf{Z} + \mathbf{c}$ para cierto vector aleatorio $\mathbf{Z}$ de marginales independientes con distribución $\mathrm{N}(0, 1)$.[2]

h) Sean $\mathbf{X}$ un vector aleatorio con distribución $\mathrm{N}_2(\mu, \Sigma)$, $\mathbf{d} \in \mathbb{R}^2$ y $B \in \mathbb{R}^{2\times 2}$ con $\det B \neq 0$. Probar que $B \cdot \mathbf{X} + \mathbf{d} \sim \mathrm{N}_2(B \cdot \mu + \mathbf{d}, B \cdot \Sigma \cdot B^t)$.

i) Probar que si $\mathbf{X}$ es un vector con distribución normal multivariada 2-dimensional entonces $\mathbf{v} \cdot \mathbf{X}$ tiene distribución normal para todo $\mathbf{v} \in \mathbb{R}^2 \setminus \{\mathbf{0}\}$. [3]

18. Sean X y W variables aleatorias independientes tales que $X \sim \mathrm{N}(0, 1)$ y $W \sim Be(\frac{1}{2})$. Definimos

$$Y = \begin{cases} X & \text{si } W = 1 \\ -X & \text{si } W = 0 \end{cases}$$

a) Probar que Y es una variable aleatoria con distribución $\mathrm{N}(0, 1)$.

[1] Esta distribución se conoce como distribución normal bivariada o multivariada 2-dimensional de parámetros μ y Σ y se denota por $\mathrm{N}_2(\mu, \Sigma)$. Observar que Σ debe ser siempre una matriz simétrica y definida positiva.

[2] Observar que esto da una caracterización de la distribución normal multivariada: un vector $\mathbf{X}$ tiene distribución normal multivariada si y sólo si es una transformación afín de un vector $\mathbf{Z}$ de coordenadas independientes con distribución normal estándar.

[3] Aunque no contamos con las herramientas necesarias para probarlo, esta implicación resulta una equivalencia. Esta nueva caracterización de la distribución normal multivariada en $\mathbb{R}^2$ es la que se adopta a la hora de definir dicha distribución sobre espacios más generales.

b) ¿Son X e Y independientes?

c) ¿Son Y y W independientes?

d) Mostrar que $\mathrm{Cov}\,(X, Y) = 0$.

e) Deducir que el vector aleatorio (X, Y) no tiene distribución normal multivariada a pesar de tener marginales con distribución normal. ¿Contradice esto los resultados del ejercicio anterior?¿Por qué?

19. Sean $U_1, \ldots, U_n$ variables aleatorias independientes con distribución $\mathcal{U}\,[0, 1]$ y consideremos sus estadísticos de orden $U^{(1)}, \ldots, U^{(n)}$.

a) Hallar $\mathbb{E}\left(U^{(i)}\right)$. ¿Qué relación guardan las esperanzas entre sí?

Sugerencia: Recordar la distribución de $U^{(i)}$ calculada en el Ejercicio 12 de la Práctica 5.

b) Calcular $\mathrm{Var}\left(U^{(i)}\right)$.

c) ¿Para qué valor de i se minimiza la varianza? ¿Para cuál se maximiza?

20. *a*) Sea X una variable aleatoria discreta con $\mathcal{R}_X \subseteq \mathbb{N}_0$. Probar que

$$\mathbb{E}(X) = \sum_{n=0}^{\infty} P(X > n) = \sum_{n=1}^{\infty} P(X \geq n).$$

b) Sea X una variable aleatoria arbitraria. Probar que

$$\sum_{n=1}^{\infty} P(|X| \geq n) \leq \mathbb{E}(|X|) \leq 1 + \sum_{n=1}^{\infty} P(|X| \geq n)$$

Concluir que

$$\mathbb{E}(|X|) < \infty \Longleftrightarrow \sum_{n=1}^{\infty} P(|X| \geq n) < \infty.$$

Sugerencia: Considerar la variable aleatoria discreta $[|X|]$.