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 $\int_{-\infty}^{+\infty} f_X(x)\, dx = 1$ 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 $\int_{-\infty}^{+\infty} f_X(x)\, dx = 1$, 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

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

   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 ($n_1 + n_2 + \cdots = n$ y $x_i \neq x_j$ si $i \neq j$). Sea $m = \sum_j \dfrac{n_j x_j}{n}$ el número medio de habitantes por manzana y sea $a^2 = \dfrac{1}{n} \sum_j n_j x_j^2 - m^2$. Para averiguarlo se sortean al azar r manzanas para encuestarlas. Un encuestador es enviado a cada una de las manzanas sorteadas para contar la cantidad de habitantes que viven en ellas. Definamos las variables aleatorias

$$\begin{aligned} Z_i &= \text{cantidad de habitantes que viven en la iésima manzana sorteada} \\ Y &= \text{cantidad de personas encuestadas.} \end{aligned}$$

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

b) Mostrar que

$$\mathbb{E}(Y) = mr$$

$$\operatorname{Var}(Y) = \frac{a^2 r(n-r)}{n-1}.$$

8. Sean $X_1, \ldots, X_n$ v.a. 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, $\mathbb{E}(Y)$.

b) Hallar $\mathbb{E}(X_1 \cdots X_n)$.

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

b) Calcular $\mathbb{E}(Z)$ y $\operatorname{Var}(Z)$ 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 $\mathbb{E}(X)$, $\mathbb{E}(Y)$, $\operatorname{Var}(X)$, $\operatorname{Var}(Y)$, $\operatorname{Cov}(X, Y)$, $\operatorname{Var}(X - Y)$ y $\rho(X, Y)$.

11. El problema del coleccionista de cupones

Un hombre colecciona cupones de un álbum compuesto por N cupones distintos. Dicho hombre compra sus cupones de a uno eligiendo en cada oportunidad uno al azar entre los N distintos del álbum.

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

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

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

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

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

a) Probar que:

$$P(X_i = 1, X_j = 1) = \frac{D(D-1)}{N(N-1)} \qquad \text{para } i \neq j$$

$$P(X_i = 1) = \frac{D}{N}$$

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

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

Sugerencia: Usar que $X = \sum_{i=1}^{n} X_i$. Notar que $X \sim \mathcal{H}(N, D, n)$.

13. Sea $(X_1, \ldots, X_k) \sim \mathcal{M}(p_1, \ldots, p_k, n)$, $n > 2$.

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

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

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. Sea X una variable aleatoria con distribución simétrica respecto de $m \in \mathbb{R}$.

a) Probar que si X es una variable aleatoria absolutamente continua y simétrica respecto de m tal que $\mathbb{E}(|X|) < \infty$ entonces $\mathbb{E}(X) = m$.

b) Una variable aleatoria X tiene distribución logística si es absolutamente continua con densidad

$$f(x) = \frac{e^{-x}}{(1 + e^{-x})^2}.$$

i. Probar que X tiene distribución simétrica respecto de 0. Hallar $\mathbb{E}(X)$.

ii. Hallar la densidad de $Y = e^X$. Verificar que $Y \sim F_{2,2}$, es decir, Y tiene distribución F de Snedecor con 2 grados de libertad en el numerador y 2 en el denominador. ¿Tiene esperanza finita?

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. Sean Z_1 y Z_2 variables aleatorias independientes con distribución $\mathrm{N}(0,1)$. Sea $A = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$ tal que $\det A \neq 0$ y $c = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$. Consideremos el vector aleatorio $\mathbf{Y} = (Y_1, Y_2)$ definido por

$$\mathbf{Y} = A \cdot \mathbf{Z} + c. \tag{1}$$

a) Hallar la distribución conjunta del vector $\mathbf{Y}$[1].

b) Apelando al Ejercicio 19 de la Práctica 5, ¿qué distribución tienen Y_1 e Y_2? Calcular sus esperanzas y varianzas.

Sugerencia: Utilizar el Ejercicio 4 h) de esta Práctica.

c) 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 $|\rho(Y_1, Y_2)| < 1$.

d) Calcular los parámetros de la distribución de $\mathbf{Y}$. Verificar que $\mu = c$ y $\Sigma = A \cdot A^t$.

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

f) Mostrar que la función de densidad conjunta de $\mathbf{Y}$ hallada en *a*) puede escribirse como

$$f_{Y_1 Y_2}(y_1, y_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{y_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{y_2-\mu_2}{\sigma_2}\right)^2 - 2\rho\left(\frac{y_1-\mu_1}{\sigma_1}\right)\left(\frac{y_2-\mu_2}{\sigma_2}\right)\right]\right\}} \tag{2}$$

donde

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \qquad \text{y} \qquad \Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}.$$

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 $f_{\mathbf{Y}}$?

g) Probar que si $\mathbf{X}$ es un vector aleatorio con densidad dada por (2) entonces tiene distribución normal multivariada 2-dimensional, es decir, probar que existen $c \in \mathbb{R}^2$ y $A \in \mathbb{R}^{2\times 2}$ con $\det A \neq 0$ tales que $\mathbf{X} = A \cdot \mathbf{Z} + c$ donde $\mathbf{Z}$ es un vector aleatorio de marginales independientes con distribución $\mathrm{N}(0,1)$. Exhibir A y c.

[1] Esta distribución se conoce como distribución normal bivariada o multivariada 2-dimensional de parámetros μ y Σ y se denota por $\mathrm{N}_2(\mu, \Sigma)$ donde μ es el vector de medias definido por $\mu_i = \mathbb{E}(Y_i)$ y Σ es la matriz de covarianzas dada por $\Sigma_{i,j} = \mathrm{Cov}(Y_i, Y_j)$.

h) A partir del ítem anterior deducir que si $\mathbf{X} = (X_1, X_2)$ es un vector aleatorio con distribución normal multivariada entonces X_1 y X_2 son independientes si y sólo si $\mathrm{Cov}(Z_1, Z_2) = 0$.

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

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 $Y \sim \mathrm{N}(0,1)$.

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 13 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*) Probar la desigualdad de **Cauchy-Schwartz**:

$$(\mathbb{E}(XY))^2 \leq \mathbb{E}\left(X^2\right)\mathbb{E}\left(Y^2\right). \tag{3}$$

Sugerencia: Excepto cuando $Y = -tX$, en cuyo caso la desigualdad anterior se convierte en igualdad, vale que para todo $t \in \mathbb{R}$ se verifica

$$0 < \mathbb{E}\left((tX+Y)^2\right) = \mathbb{E}\left(t^2X^2 + 2tXY + Y^2\right). \tag{4}$$

Observar que (4) define un polinomio en t de grado dos sin raíces reales. Concluir que el discriminante de la ecuación cuadrática correspondiente es negativo y obtener de allí la desigualdad en cuestión.

b) Deducir (o probar del mismo modo que en el ítem anterior pero a partir del cálculo de $\mathrm{Var}(tX+Y)$) la siguiente desigualdad:

$$(\mathrm{Cov}(X,Y))^2 \leq \mathrm{Var}(X)\,\mathrm{Var}(Y). \tag{5}$$

¿Qué relación deben cumplir X e Y para que valga la igualdad en (5)?

21. *a*) Sea X una variable aleatoria discreta con $R_X \subseteq \mathbb{N}_0$. Probar que

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

[2] Puede probarse, aunque no lo veremos en esta materia, que ambas afirmaciones son equivalentes si se asume de antemano que $\mathbf{X}$ es absolutamente continuo. Esta 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) 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 $[|X|]$.