PRÁCTICA 5: VECTORES ALEATORIOS

"Conviene que todos los ciudadanos entren en contacto con la verdadera matemática, que es método, arte y ciencia, muy distinta de la calculatoria, que es técnica y rutina."
LUIS ANTONIO SANTALÓ

Ejercicio 1. Una urna contiene 4 bolitas blancas y 6 negras. Se extraen 3 bolitas sin reposición y se definen las siguientes variables aleatorias:

$$\begin{aligned} X &= \begin{cases} 1 & \text{si el número de bolitas blancas extraídas es par} \\ 0 & \text{si el número de bolitas blancas extraídas es impar} \end{cases} \\ Y &= \text{Número de bolillas negras extraídas.} \end{aligned}$$

a) Hallar p_{XY} y F_{XY}.

b) Hallar p_X y p_Y. Determinar si X e Y son independientes.

Ejercicio 2. Sean X e Y variables aleatorias independientes. Probar las siguientes afirmaciones:

a) $X \sim \mathcal{B}i(n, p)$, $Y \sim \mathcal{B}i(m, p) \Rightarrow X + Y \sim \mathcal{B}i(n + m, p)$.

b) $X \sim \mathcal{P}(\lambda)$, $Y \sim \mathcal{P}(\mu) \Rightarrow X + Y \sim \mathcal{P}(\lambda + \mu)$.

c) $X \sim \mathcal{G}(p)$, $Y \sim \mathcal{G}(p) \Rightarrow X + Y \sim \mathcal{BN}(2, p)$.

Ejercicio 3. Se sabe que en la provincia de Salta la proporción de hombres de ojos azules es 20%, de ojos verdes es 5%, de ojos negros es 10% y otro color de ojos es 65%. Josefina decide viajar de la capital salteña a una ciudad a 200 km. donde se realizará un congreso médico sobre alcoholismo. Para ello debe tomar dos colectivos en los que viajan sólo salteños. Para llevar a cabo una prueba decide tomar una copa de jerez con cada hombre de ojos verdes o azules que encuentre en su viaje. Como su belleza es irresistible, todos los hombres aceptan su invitación. En el primer colectivo viajan 10 hombres de los cuales ninguno transborda al siguiente. En el segundo hay 8 hombres.

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

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

Ejercicio 4. El 10% de la población fuma cigarrillos negros, el 35% fuma cigarrillos rubios, el 3% fuma pipa y el resto no fuma. Se realiza una encuesta a 35 personas y con los resultados obtenidos se definen las variables aleatorias

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

a) Hallar la probabilidad puntual del vector aleatorio (Y_1, Y_2, Y_3, Y_4).

b) Hallar la probabilidad puntual del vector aleatorio $(Y_1, Y_2 + Y_3, Y_4)$.

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

Ejercicio 5.

a) El número de personas que ingresa por día a un cierto banco es una variable aleatoria de Poisson de parámetro λ. Cada persona que entra tiene probabilidad p de ser hombre y $1-p$ de ser mujer. Mostrar que el número hombres y de mujeres que ingresan por día a dicho banco son variables aleatorias independientes con distribución Poisson de parámetros λp y $\lambda(1-p)$ respectivamente.

b) Recíprocamente, supongamos que el número de hombres y de mujeres que ingresan al banco por día son variables aleatorias de Poisson independientes, ¿Cuál es la distribución del número total de personas que ingresan al banco por día?

c) El número de hombres y de mujeres que ingresan por día a un cierto banco son variables aleatorias independientes con distribución Poisson de parámetros 2 y 2.5 respectivamente. Sabiendo que en un determinado día ingresaron únicamente dos personas al banco, calcular la probabilidad de que hayan sido un hombre y una mujer.

Ejercicio 6. Sea (X, Y) un vector aleatorio absolutamente continuo con función de densidad conjunta

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

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

Ejercicio 7. Tres integrantes de un equipo de salto en largo participan en un campeonato. Se asigna al equipo la mejor de las tres distancias obtenidas. La distancia (en metros) que salta el atleta A tiene distribución uniforme en el intervalo $[7, 9]$ mientras que la distancia del salto del atleta B tiene distribución absolutamente continua con función de densidad

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

La distancia que salta el atleta C tiene otra distribución absolutamente continua cuya densidad es

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

a) Hallar la distribución[1] de la distancia asignada al equipo. Dar tanto la función de distribución como la función de densidad.

b) Encontrar la probabilidad de que la distancia sea mayor que 8.2.

[1]Cuando se pide la distribución de una variable aleatoria se puede contestar dando cualquier objeto que la caracterice, como la función de probabilidad puntual en el caso discreto, la función de densidad en el caso continuo, la función de distribución acumulada, o, en el caso en el que la distribución pertenezca a una familia conocida, basta dar el nombre y los parámetros en cuestión.

Ejercicio 8.

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

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

$$X^{(n)} = \max_{1 \leq i \leq n} X_i.$$

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

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

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

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

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

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

Ejercicio 9. Sean $X_1, \dots, X_n$ variables aleatorias independientes con distribución exponencial de parámetros $\alpha_1, \dots, \alpha_n$ respectivamente.

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

b) Probar que

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

c) Calcular $P\left(\min_{1 \leq i \leq n} Y_i \leq 2\right)$, donde para cada $i = 1, ..., n$ se define $Y_i = [X_i] + 1$. ¿Qué distribución tiene Y_i?

Ejercicio 10. Dos servidores A y B procesan trabajos a medida que van llegando. El tiempo que tarda el servidor A en procesar un trabajo es una variable aleatoria $X \sim \mathcal{E}(\lambda_1)$, mientras que el tiempo que tarda el servidor B es una variable aleatoria $Y \sim \mathcal{E}(\lambda_2)$. Ambos servidores actúan en forma independiente.

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

b) Supongamos que tres trabajos llegan simultáneamente. Uno es atendido por A, otro por B y el tercero queda esperando a que uno de los servidores se libere. Si $\lambda_1 = \lambda_2$, pruebe que la probabilidad de que el último trabajo en ser atendido sea el último en ser completado es 0.5.

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

Ejercicio 11.

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

b) Sean X e Y variables aleatorias independientes tales que $X \sim \Gamma(\alpha_1,\lambda)$ e $Y \sim \Gamma(\alpha_2,\lambda)$. Probar que $X+Y$ y $\frac{X}{X+Y}$ son independientes con distribución $\Gamma(\alpha_1+\alpha_2,\lambda)$ y $\beta(\alpha_1,\alpha_2)$, respectivamente[2].

c) Teniendo en cuenta el ejercicio 7 de la Práctica 4, deducir que si $Z_1,\dots,Z_n$ son variables aleatorias independientes con distribución normal estándar entonces $Z_1^2+\cdots+Z_n^2$ tiene distribución $\Gamma(\frac{n}{2},\frac{1}{2})$[3].

Ejercicio 12. Sean X e Y variables aleatorias independientes con distribución $\mathrm{N}(0,\sigma^2)$. Sea (ρ,θ) la expresión de (X,Y) en coordenadas polares, es decir $(X,Y)=(\rho\cos(\theta),\rho\,\mathrm{sen}(\theta))$, con $\rho\geq 0$ y $0\leq\theta<2\pi$.

a) Probar que ρ y θ son variables aleatorias independientes y hallar su distribución.

b) Hallar la probabilidad de que el par (X,Y) caiga en el círculo de centro en el origen y radio σ.

Ejercicio 13. Sean X_1 y X_2 variables aleatorias independientes con distribución $\mathcal{U}[-1,1]$. Sea $U=\frac{X_1+X_2}{2}$. Verificar que

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

Ejercicio 14.

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

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

c) Deducir que si X e Y son variables aleatorias con distribución $\mathrm{N}(\mu,\sigma^2)$ entonces $X+Y$ y $X-Y$ son independientes[4].

d) Deducir[5] que si $X\sim\mathrm{N}(\mu_1,\sigma_1^2)$ e $Y\sim\mathrm{N}(\mu_1,\sigma_1^2)$ entonces $aX+bY+c\sim\mathrm{N}(a\mu_1+b\mu_2+c,a^2\sigma_1^2+b^2\sigma_2^2)$ para todo $a,b,c\in\mathbb{R}$.

Ejercicio 15. Sean $X_1,\dots,X_n$ variables aleatorias absolutamente continuas, independientes e idénticamente distribuidas con función de densidad f y consideremos el vector aleatorio $\overline{X}=(X^{(1)},\dots,X^{(n)})$ conformado por sus estadísticos de orden. Mostrar que $\overline{X}$ es absolutamente continuo y que su función de densidad viene dada por

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

[2]Esto muestra que suma de variables aleatorias Gamma independientes tiene distribución Gamma.

[3]Esta distribución particular se conoce como la distribución χ^2 con n grados de libertad y se la nota $\chi^2(n)$.

[4]De hecho, mostraremos más adelante que vale la equivalencia: si X e Y son variables aleatorias independientes con una misma distribución F entonces $X+Y$ y $X-Y$ son independientes si y sólo si F es una distribución normal.

[5]En particular, esto muestra que suma de variables aleatorias independientes con distribución normal tiene distribución normal.