Tema 3: Variable Aleatoria

(Nota acerca de la notación: log *se refiere al logaritmo neperiano*. Aunque no hay problema si preferís usar ln, log es la notación matemática internacional y es la más recomendada)

1. Una moneda se lanza tres veces. Si X es una variable aleatoria (VA) que indica el número de caras que aparecen, elabora una tabla que muestre la función de probabilidad de X y calcula su función de distribución. Usa la función *ecdf()* de *R* para dibujar la función de distribución.

Solution:

X	0	1	2	3
$P(x)$	1/8	3/8	3/8	1/8

$$F(x) = \begin{cases} 0 & x < 0 \\ 1/8 & 0 \le x < 1 \\ 4/8 & 1 \le x < 2 \\ 7/8 & 2 \le x < 3 \\ 1 & x \ge 3 \end{cases}$$

2. La siguiente ecuación muestra la función distribución de una variable X. Determina (a) su función de probabilidad, (b) $P(1 \le X \le 3)$, (c) $P(X \ge 2)$, (d) $P(X < 3)$, (e) $P(X > 1.4)$.

$$F(x) = \begin{cases} 0 & x < 1 \\ 1/8 & 1 \le x < 2 \\ 3/8 & 2 \le x < 3 \\ 6/8 & 3 \le x < 4 \\ 1 & x \ge 4 \end{cases}$$

Solution: (a)

X	1	2	3	4
$P(x)$	1/8	2/8	3/8	2/8

(b) 3/4, (c) 7/8, (d) 3/8, (e) 7/8

3. Una variable aleatoria (VA) X tiene función de densidad

$$f(x) = \begin{cases} ce^{-3x} & x \ge 0 \\ 0 & x < 0. \end{cases}$$

Calcula (a) la constante c, (b) $P(1 < X < 2)$, (c) $P(X \ge 3)$, (d) $P(X < 1)$, (e) La función de distribución de X, (f) Representa gráficamente las funciones de distribución y densidad de manera que describa la relación entre ellas.

Solution: (a) 3, (b) $e^{-3} - e^{-6}$, (c) e^{-9}, (d) $1 - e^{-3}$, (e) $F(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-3x} & x \ge 0 \end{cases}$

4. Una variable aleatoria (VA) X tiene función de densidad

$$f(x) = \begin{cases} cx^2 & 1 \leq x \leq 2 \\ cx & 2 \leq x \leq 3 \\ 0 & x < 0. \end{cases}$$

Calcula (a) la constante c, (b) $P(X > 2)$, (c) $P(1/2 < X < 3/2)$, (d) la función de distribución de X, (e) Representa gráficamente las funciones de distribución y densidad de manera que describa la relación entre ellas.

> **Solution:** (a) 6/29, (b), 15/29, (c) 19/116 (d)
>
> $$F(x) = \begin{cases} 0 & x < 1 \\ (2x^3 - 2)/29 & 1 \leq x \leq 2 \\ \frac{2}{29} + \frac{3}{29}x^2 & 2 \leq x \leq 3 \\ 1 & x \geq 3 \end{cases}$$

5. La demanda de combustible de una gasolinera, en m^3, tiene distribución

$$f(x) = \begin{cases} kx^2 & 0 \leq x \leq 27 \\ k27^4/x^2 & x \geq 27. \end{cases}$$

(a) Si la gasolinera tiene un depósito de 100 m^3, que se rellena al principio de cada semana, hallar la probabilidad de que se pueda atender toda la demanda semanal. (b) Cuál debería ser la capacidad del depósito para que haya probabilidad 0.95 de que no se agote el combustible en una semana.

> **Solution:** (a) 0.798, (b) $c = 405\ m^3$.

6. La función de distribución de una variable aleatoria X está dada por

$$F(x) = \begin{cases} 0 & x < 0 \\ cx^3 & 0 \leq x < 3 \\ 1 & x \geq 3 \end{cases}$$

Calcula (a) la constante c, (b) la función de densidad, (c) $P(X > 1)$, (d) $P(1 < X < 2)$.

> **Solution:** (a) 1/27, (b)
>
> $$f(x) = \begin{cases} x^2/9 & 0 \leq x < 3 \\ 0 & \text{en otro caso} \end{cases}$$
>
> (c) 26/77, (d) 7/27.

7. ¿Puede ser la siguiente función una función de distribución? ¿Por qué?

$$F(x) = \begin{cases} 0 & x < 0 \\ c(1 - x^2) & 0 \leq x \leq 1 \\ 1 & x > 1 \end{cases}$$

Solution: Dependiendo de los valores de c se incumplen diversas propiedades:

- $c = 0 \qquad F(x) = 0 \qquad \forall x \in (0, 1)$.
- $c > 0 \qquad F(x)$ es una función decreciente en $(0, 1)$
- $c < 0 \qquad F(x) < 0 \qquad \forall x \in (0, 1)$.

8. [Spiegel 2.2, 2.4]
   (a) Encuentra la función de probabilidad de niños y niñas en familas con 3 hijos, suponiendo igual probabilidad para niños y niñas.
   (b) Encuentra la función de distribución para la variable aleatoria del apartado a).
   (c) Dibuja la función de distribución.
   (d) Imagínate que el único dato del que partes es la función de distribución. ¿Cómo podrías hallar la función de probabilidad?

Solution:

X	0	1	2	3
$P(x)$	1/8	3/8	3/8	1/8

$$F(x) = \begin{cases} 0 & x < 0 \\ 1/8 & 0 \leq x < 1 \\ 4/8 & 1 \leq x < 2 \\ 7/8 & 2 \leq x < 3 \\ 1 & x \geq 3 \end{cases}$$

9. Se realizan 12 lanzamientos de una moneda equilibrada. Determinar la distribución (a) del número de caras obtenidas. (b) de la diferencia entre caras y cruces. (c) del número total de rachas. En cada caso, representar la función de probabilidad.

Solution: (a) $P(X = k) = \binom{12}{k}(0.5)^k(0.5)^{12-k} = \binom{12}{k}(0.5)^{12}$.
(b) $P(X = k) = \binom{12}{6+k/2}0.5^{12} \qquad k = -12, -10, \cdots, 10, 12$.
(c) $P(X = k) = \binom{12-1}{k-1}0.5^{11} \qquad 1 \leq k \leq 12$.

10. La variable X toma valores naturales con probabilidad

$$p_n = \frac{1}{n(n+1)} \qquad n = 1, 2, 3 \cdots$$

Hallar la función de Y, definida como la parte entera de $X/2$

Solution: El suceso $Y = k$ equivale a $\{X = 2k\} \cup \{X = 2k+1\}$ con lo que $P(Y = k) = \frac{1}{2}P(X = k)$ para $k = 1, 2, 3, \ldots$.

11. Cinco bolas se colocan al azar en tres urnas. Hallar la distribución del número de urnas que contienen una bola.

Solution: El problema se soluciona al darse cuenta que, de los 3^5 casos posibles:

- Hay 3 que corresonden a que todas las bolas estén en la misma urna (¿Por qué?).
- Hay 30 casos con 4 bolas en una urna (¿Por qué?).
- Hay 60 casos con 3 bolas en una urna y dos en otra (¿Por qué?).
- Hay 60 casos con 4 bolas en una urna y una en cada una de las dos restantes (¿Por qué?).
- Hay 90 casos con dos bolas en dos urnas y una en la tercera.

Por tanto:

$$p(x) = \begin{cases} (3+60)/3^5 & x = 0 \\ (30+90)/3^5 & x = 1 \\ (60)/3^5 & x = 2 \end{cases}$$

12. Se lanzan dos monedas con probabilidades de cara p_1 y p_2. Estudiar el número de lanzamientos N que se realizan hasta que aparece cara en ambas monedas simultáneamente.

Solution:

$$P(N = n) = (1 - p_1 p_2)^{n-1} p_1 p_2 \qquad n = 1, 2, 3, \cdots$$

13. La función de probabilidad conjunta de dos variables aleatorias X e Y está dada por

$$f(x,y) = \begin{cases} cxy & x = 1,2,3 \text{ e } y = 1,2,3 \\ 0 & \text{en otro caso} \end{cases}$$

Calcula (a) la constante c, (b) $P(X = 2, Y = 3)$, (c) $P(X \leq 2, Y \leq 2)$, (d) $P(X \geq 2)$, (e) $P(Y \leq 2)$, (f) $P(X = 1)$, (g) $P(Y = 3)$

Solution: (a) 1/36, (b) 1/6, (c) 1/4, (d) 5/6, (e) 1/6, (f) 1/2, (g) 1/2

14. Encuentra las funciones de probabilidad marginal de (a) X e (b) Y, para las variables aleatorias del problema anterior. (c) Determina si X e Y son independientes.

Solution: (a)

$$f_x(x) = \begin{cases} x/6 & x = 1,2,3 \\ 0 & \text{en otro caso} \end{cases}$$

(b)

$$f_y(y) = \begin{cases} y/6 & y = 1,2,3 \\ 0 & \text{en otro caso} \end{cases}$$

15. La función de probabilidad conjunta de dos variables aleatorias X e Y está dada por

$$f(x,y) = \begin{cases} c(x^2+y^2) & 0 \le x \le 1 \text{ e} 0 \le y \le 1 \\ 0 & \text{en otro caso} \end{cases}$$

Calcula (a) la constante c, (b) $P(X < 1/2, Y > 1/2)$, (c) $P(1/4 < X < 3/4)$, (d) $P(Y < 1/2)$.

Solution: (a) 3/2, (b) 1/4, (c) 29/64, (d) 5/16.

16. Encuentra las funciones de distribución marginal de (a) X e (b) Y, para las variables aleatorias del problema anterior. (c) Determina si X e Y son independientes.

Solution: (a)

$$F_x(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{2}(x^3+x) & 0 \le x < 1 \\ 1 & x \ge 1 \end{cases}$$

(b)

$$F_y(y) = \begin{cases} 0 & y < 0 \\ \frac{1}{2}(y^3+y) & 0 \le y < 1 \\ 1 & y \ge 1 \end{cases}$$

17. Para la distribución del problema 13, encuentra la función de densidad condicional de (a) X dado Y, (b) Y dado X.

Solution: (a) $f(x \mid y) = f_x(x)$ para $x = 1, 2, 3$. (b) $f(y \mid x) = f_y(y)$ para $y = 1, 2, 3$.

18. Sea la función de densidad conjunta

$$f(x,y) = \begin{cases} x+y & 0 \le x \le 1; 0 \le y \le 1 \\ 0 & \text{en otro caso} \end{cases}$$

Encuentra la función de densidad condicional de (a) X dado Y, (b) Y dado X.

Solution: (a)

$$f(x \mid y) = \begin{cases} \frac{x+y}{y+1/2} & 0 \le x \le 1; 0 \le y \le 1 \\ 0 & x < 0 \text{ o } x > 1; 0 \le y \le 1 \end{cases}$$

(b)

$$f(y \mid x) = \begin{cases} \frac{x+y}{x+1/2} & 0 \le x \le 1; 0 \le y \le 1 \\ 0 & 0 \le x \le 1; y < 0 \text{ o } y > 1. \end{cases}$$

19. De una urna que contiene a bolas blancas y b negras se hacen extracciones sin reemplazamiento. Sea X_i el número de extracción en que aparece la i-ésima bola blanca. Calcula (a) La distribución conjunta de X_1 y X_2. (b) La distribución de X_1 condicionada por X_2.

Solution:

$$P(X_1 = j, X_2 = k) = \binom{a+b-k}{a-2} \Big/ \binom{a+b}{a}$$

$$P(X_1 = j, X_2 = k) = \frac{1}{k-1}$$

20. Para la distribución del problema 15, encuentra la función de densidad condicional de (a) X dado Y, (b) Y dado X.

Solution: (a)

$$f(x \mid y) = \begin{cases} \frac{x^2+y^2}{y^2+1/3} & 0 \leq x \leq 1; 0 \leq y \leq 1 \\ 0 & x < 0 \text{ o } x > 1; 0 \leq y \leq 1 \end{cases}$$

(b)

$$f(y \mid x) = \begin{cases} \frac{x^2+y^2}{x^2+1/3} & 0 \leq x \leq 1; 0 \leq y \leq 1 \\ 0 & 0 \leq x \leq 1; y < 0 \text{ o } y > 1. \end{cases}$$

21. Sea la función de densidad conjunta de X e Y:

$$f(x, y) = \begin{cases} e^{-(x+y)} & x \geq 0; y \geq 0.0 \qquad \text{en otro caso} \end{cases}$$

Encuentre la función de densidad condicional de (a) X dado Y, (b) Y dado X.

Solution: (a)

$$f(x \mid y) = \begin{cases} e^{-x} & x \geq 0; y \geq 0 \\ 0 & x < 0; y \geq 0 \end{cases}$$

(b)

$$f(y \mid x) = \begin{cases} e^{-y} & x \geq 0; y \geq 0 \\ 0 & x \geq 0; y < 0 \end{cases}$$

22. Sean 2 variables aleatorias independientes X e Y uniformemente distribuidas entre 0 y a. Calcular la densidad de $Z = X + Y$.

Solution:

$$f(z) = \begin{cases} \frac{z}{a^2} & 0 < z < a \\ \frac{2a-z}{a^2} & a < z < 2a \\ 0 & \text{en otro caso} \end{cases}$$

23. La cantidad que se produce de cierto material (en Kilogramos) es una variable aleatoria continua con función de densidad

$$f(x) = \frac{3x^2}{1000} \qquad 0 < x < 10$$

El precio del artículo es función de la cantidad producida $P = 40 - 2X$. Calcular la función de densidad del precio y del valor producido ($V = PX$).

Solution:

$$f_p(p) = \frac{3(40-p)^2}{8000} \qquad 20 < x < 40$$

$$f_v(v) = \frac{3}{16000}\frac{(40-\sqrt{1600-8v})^2}{\sqrt{1600-8v}} \qquad 0 < v < 200$$

24. [Spiegel 2.18] La función de probabilidad de una variable aleatoria X es

$$f(x) = \begin{cases} 2^{-x} & x = 1, 2, 3, \cdots \\ 0 & \text{en otro caso} \end{cases} .$$

Hallar la distribución de $U = x^4 + 1$.

Solution:

$$f(u) = \begin{cases} 2^{-(u-1)^{1/4}} & u = 2, 17, 82, \cdots \\ 0 & \text{en otro caso} \end{cases}$$

25. [Spiegel 2.19] Si X tiene distribución

$$f(x) = \begin{cases} x^2/81 & -3 < x < 6 \\ 0 & \text{en otro caso} \end{cases}$$

calcula la densidad de probabilidad de $U = \frac{1}{3}(12 - X)$.

Solution:

$$f(u) = \begin{cases} (12-3u)^2/27 & 2 < u < 5 \\ 0 & \text{en otro caso} \end{cases}$$

26. [Spiegel 2.21] Si X e Y tienen como distribución

$$f(x,y) = \begin{cases} xy/96 & 0 < x < 4, 1 < y < 5 \\ 0 & \text{en otro caso} \end{cases}$$

calcula la distribución de $U = X + 2Y$.

Solution:

$$f(u) = \begin{cases} (u-2)^2(u+4)/2304 & 2 < u < 6 \\ (3u-8)/144 & 6 < u < 10 \\ (348u - u^3 - 2128)/2304 & 10 < u < 14 \\ 0 \quad \text{en otro caso} \end{cases}$$

27. Sea

$$f(x) = \begin{cases} 2e^{-2x} & x \geq 0 \\ 0 & \text{en otro caso} \end{cases}$$

Calcula (a) $P(|X - \mu| > 1)$ de forma exacta y (b) usando la desigualdad de Chebyshev.

Solution: (a) 0.04979, (b) $P(|X - \mu| > 1) \leq 0.25$.

28. Sea X una variable discreta con función de probabilidad:

X	-2	-1	0	1	2
$P(x)$	1/5	1/6	1/5	1/15	11/30

Calcula las funciones de probabilidad para las siguientes variables aleatorias:

(a) $Y = X + 2$.

(b) $Z = X^2 + 3$.

Solution: (a)

Y	0	1	2	3	4
$P(y)$	1/5	1/6	1/5	1/15	11/30

(b)

Z	3	4	7
$P(z)$	1/5	7/30	17/30

29. Sea X una variable aleatoria (VA) con distribución

$$P(X = n) = p(1-p)^{n-1} \qquad n = 1, 2, 3, \ldots$$

Calcula las funciones de probabilidad para las siguientes VAs:

(a) $Y = (X - 5)^2$.

(b) $Z = X\text{mod}3$ (el módulo es el resto de dividir X entre 3).

Solution: (a) Para este apartado es fundamental hacer el gráfico de Y vs X para saber que posibles valores puede tomar Y. LLegamos a:

$$P(Y = m^2) = \begin{cases} p(1-p)^4 & m = 0 \\ p(1-p)^{4+m} + p(1-p)^{4-m} & 0 < m < 5 \\ p(1-p)^{4+m} & m \geq 5 \end{cases}$$

(b) Piensa en los posibles resultados del módulo y qué valores de X son compatibles con cada uno de ellos. Tendrás que hacer tres sumatorios para llegar a

$$P(Z = k) = \begin{cases} \frac{p(1-p)^2}{1-(1-p)^3} & k = 0 \\ \frac{p}{1-(1-p)^3} & k = 1 \\ \frac{p(1-p)}{1-(1-p)^3} & k = 2 \end{cases}$$

30. Si X es una variable aleatoria con distribución

$$P(X = x) = p(1-p)^x \qquad \text{para } x = 0, 1, 2, \ldots,$$

razona cuál es la distribución para la variable aleatoria

$$Y = (X-2)^2.$$

31. Se elige un punto P al azar en el cuadrado $[0,1]^2$. Si O es el origen, de coordenados. Determinar la distribución de (a) del área del rectángulo de lados paralelos a los ejes con vértices en O y P. (b) de la distancia de O a P. (c) del ángulo que forma la recta OP con el eje de abscisas.

Solution: (a) $f(x) = -\log x \qquad 0 < x < 1.$
(b) $f(y) = \begin{cases} \pi y/2 & 0 \le y \le 1 \\ \pi y/2 - 2y \arccos(1/y) & 1 \le y \le \sqrt{2}. \end{cases}$
(c) $f(z) = \begin{cases} 1/(2\cos^2 z) & 0 \le z \le \pi/4 \\ 1/(2\sin^2 z) & \pi/4 \le z \le \pi/2 \end{cases}$

32. Un cañón forma con un plano horizontal un ángulo α que se elige al azar en $(0, \pi/2)$. Se dispara un proyectil con velocidad v. Determinar la distribución de la distancia al punto de impacto.

Solution: Para resolver el problema es necesario recordar que las coordenadas del proyectil (tal y como se ha visto en Física) son:

$$x = vt\cos\alpha \qquad y = vt\sin\alpha - gt^2/2.$$

A partir de esto es posible calcular la fórmula del punto de impacto en función de α. Usando la fórmula de cambio de variable se llega a:

$$f(x) = \frac{2g}{\pi\sqrt{v^4 - g^2x^2}} \qquad 0 < x < v^2/g.$$

33. [Spiegel 2.14] La función de densidad conjunta de X e Y es

$$f(x,y) = \begin{cases} \frac{xy}{96} & 0 < x < 4, \quad 1 < y < 5 \\ 0 & \text{en otro caso} \end{cases}$$

Calcula $P(X + Y < 3)$

Solution: La clave para resolver este ejercicio es dibujar la región donde se verifica la desigualdad $X + Y < 3$, para luego hacer la integral adecuada. Obtendremos: $P(X + Y < 3) = \frac{1}{48}$.

34. Si X e Y tienen función de densidad

$$f(x,y) = \lambda e^{-\lambda x}\mu e^{-\mu y} \qquad x > 0, y > 0$$

razona cuál es la función de densidad conjunta para las variables aleatorias

$$U = X + Y \qquad V = X - Y.$$

35. Una persona tiene dos prótesis de pierna. Sea X la variable aleatoria: "duración de la prótesis izquierda (en años)" e Y la variable aleatoria: "duración de la prótesis derecha (en años)". La distribución conjunta de ambas variables es

$$f(x,y) = \begin{cases} ce^{-2x}e^{-3y} & x \geq 0, y \geq 0 \\ 0 & \text{en otro caso} \end{cases}$$

   (a) Calcula c.
   (b) Calcula la función de <u>distribución</u> marginal de X.
   (c) Calcula $Pr(1 < X < 3)$.
   (d) Calcula la probabilidad de que la prótesis izquierda dure 3 años más que la prótesis derecha.

36. Calcula (a) $\mathbb{E}[X]$, (b) $\mathbb{E}[2X+5]$ y (c) $\mathbb{E}[X^2]$ de una variable aleatoria X definida por

$$\begin{cases} 1/3 & x = -2 \\ 1/2 & x = 3 \\ 1/6 & x = 1 \end{cases}$$

Solution: (a) 1, (b) 7, (c) 6.

37. Sea una variable aleatoria definida por

$$f(x) = \begin{cases} 3x^2 & 0 \leq x \leq 1 \\ 0 & \text{en otro caso} \end{cases}$$

Calcula (a) $\mathbb{E}[X]$, (b) $\mathbb{E}[3X-2]$ y (c) $\mathbb{E}[X^2]$.

Solution: (a) 3/4, (b) 1/4, (c) 3/5

38. Sea una variable aleatoria definida por

$$f(x) = \begin{cases} e^{-x} & x \geq 0 \\ 0 & \text{en otro caso} \end{cases}$$

Calcula (a) $\mathbb{E}[e^{2X/3}]$.

Solution: 3

39. Sean X e Y variables aleatorias con función de densidad:

$$f(x,y) = \begin{cases} \frac{3}{5}x(x+y) & 0 \leq x \leq 1; 0 \leq y \leq 2 \\ 0 & \text{en otro caso} \end{cases}$$

Calcula (a) $\mathbb{E}[X]$, (b) $\mathbb{E}[Y]$, (c) $\mathbb{E}[X+Y]$, (d) $\mathbb{E}[XY]$.

Solution: (a) 7/10, (b) 6/5, (c) 19/10, (d) 5/6.

40. En el problema 39, (a) $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$, (b) $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$

Solution: (a) Sí, es una propiedad general de la esperanza. (b) No, ya que X e Y no son independientes.

41. Sean X e Y variables aleatorias independientes con

$$p(x) = \begin{cases} 1/3 & x = 1, \\ 2/3 & x = 0 \end{cases}$$

$$p(y) = \begin{cases} 3/4 & y = 2, \\ 1/4 & y = -3 \end{cases}$$

Calcula (a) $\mathbb{E}[3X+2Y]$, (b) $\mathbb{E}[2X^2 - Y^2]$, (c) $\mathbb{E}[XY]$, (d) $\mathbb{E}[X^2Y]$.

Solution: (a) 5/2, (b) -55/12, (c) 1/4, (d) 1/4

42. Encuentra la media y la varianza de una distribución uniforme entre a y b

Solution: $\mu = (a+b)/2$, $\sigma^2 = (b-a)^2/12$.

43. Calcula la varianza de X si

$$f(x) = \begin{cases} e^{-x} & x \geq 0 \\ 0 & \text{en otro caso} \end{cases}$$

Solution: $\mu = 1$, $\sigma^2 = 1$.

44. Una variable aleatoria verifica que $\mathbb{E}[(X-1)^2] = 10$ y $\mathbb{E}[(X-2)^2] = 6$. Calcula (a) $\mathbb{E}[X]$, (b) $\text{Var}[X]$, (c) σ_x.

Solution: (a) 7/2, (b) 15/4, (c) $\sqrt{15}/2$.

45. Sean X e Y variables aleatorias con función de densidad conjunta

$$f(x,y) = \begin{cases} x+y & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & \text{en otro caso} \end{cases}$$

Calcula (a) la varianza de X, (b) la varianza de Y, (c) la covarianza entre X e Y, (d) la correlación entre X e Y.

Solution: (a) 11/144, (b) 11/144, (c) -1/144, (f) -1/11.

46. Calcula (a) la covarianza y (b) el coeficiente de correlación de dos variables aleatorias X e Y si verifican $\mathbb{E}[X] = 2$, $\mathbb{E}[Y] = 3$, $\mathbb{E}[XY] = 10$, $\mathbb{E}[x^2] = 9$, $\mathbb{E}[Y^2] = 16$.

Solution: (a) 4, (b) $4/\sqrt{35}$.

47. Tres jugadores A, B, C lanzan, por turnos, dos dados y gana el primero que consiga obtener 9 como suma de ambas puntuaciones. En cada tirada, el que lanza debe poner un euro en la mesa y el ganador se lo lleva todo. ¿Cuál es el beneficio esperado de cada jugador?

Solution: La probabilidad de que el ganador se produzca con el lanzamiento n es $(8/9)^{n-1} \cdot (1/9)$ (¿Por qué?). Para el caso de A:

- Si A gana en la tirada $n = 3k + 1$ gana $2k$ euros (con $k = 0, 1, ...$).
- Si A pierde en la tirada $n = 3k + 2$ paga $k + 1$ euros (con $k = 0, 1, ...$).
- Si A pierde en la tirada $n = 3k$ paga $k1$ euros (¡ojo!, con $k = 1, 2, ...$).

Esto nos permite plantear un sumatorio infinito para calcular $\mathbb{E}[X_A]$. Repitiendo para B y C obtenemos: $\mathbb{E}[X_A] \approx -0.344$, $\mathbb{E}[X_B] \approx 0.026$, $\mathbb{E}[X_C] \approx 0.318$. Nótese que $\mathbb{E}[X_A] + \mathbb{E}[X_B] + \mathbb{E}[X_C] = 0$. A esto se le llama **juego de suma cero** (lo que gana un jugador lo pierden los otros) y tiene importantes aplicaciones en **teoría de juegos**.

48. De 800 familias, cada una con 5 hijos ¿Cuántas esperarías que tuvieran (a) 3 niños, (b) 5 niñas, (c) 2 o 3 niños? Supón que la probabilidad de niña y niño son iguales.

Solution: (a) 250, (b) 25, (c) 500.

49. Un vendedor de seguros vende pólizas a 5 hombres, todos con la misma edad e idénticas condiciones de buena salud. Según las tablas de seguros, la probabilidad de que un hombre de esta edad en particular esté vivo en 30 años es 2/3. Encuentre la probabilidad de que en 30 años estén vivos (a) los 5 hombres, (b) al menos 3 hombres, (c) solo 2 hombres, (d) al menos 1 hombre.

Solution: (a) 32/243, (b) 192/243, (c) 40/243, (d) 242/243.

50. Si el 3% de las bombillas eléctricas que fabrica una compañía son defectuosas, calcula la probabilidad de que en una muestra de 100 bombillas, (a) 0, (b) 1, (c) 2, (d) 4 sean defectuosas.

Solution: (a) 0.0476, (b) 0.1471, (c) 0.2251, (d) 0.1706.

51. Sea X una variable aleatoria normal estandarizada. Calcula la probabilidad de que (a) $P(X < -1.78)$, (b) $P(X < 0.56)$, (c) $P(X > -1.45)$, (d) $P(X \geq 2.16)$, (e) $P(-0.80 < X < 1.53)$.

Solution: (a) 0.0375, (b) 0.7123, (c) 0.9265, (d) 0.0154, (e) 0.7251.

52. Una máquina fabrica tornillos de los cuales el 10% son defectuosos. Encuentra la probabilidad de que en una muestra tomada al azar de 400 tornillos (a) como máximo 30, (b) entre 30 y 50, (c) 65 o más, de los tornillos sean defectuosos.

Solution: (a) 0.0524, (b) 0.9207, (d) $7.036756 \cdot 10^{-5}$

53. Demuestra, usando un cambio de variable, que si $X \sim \mathcal{N}(\mu, \sigma^2)$, entonces $Z = \frac{X-\mu}{\sigma}$ tiene distribución $Z \sim \mathcal{N}(0, 1)$.

54. El número de pacientes que llega a las urgencias de un hospital tiene distribución de Poisson de parámetro λ. Cada paciente tiene probabilidad p de requerir hospitalización. Calcula (a) La distribución del número de enfermos que ingresan en el hospital. (b) Si un día han ingresado r pacientes, hallar la probabilidad de que hayan llegado n personas al servicio de urgencia. Calcula el número medio de llegadas.

Solution: (a) Para resolver este apartado debemos emplear probabilidades condicionadas para calcular la distribución conjunta de el número de pacientes que llegan a urgencias y el número de pacientes ingresados. Marginalizando (con un sumatorio infinito), llegamos a $P(x = k) = e^{-\lambda p} \frac{(\lambda p)^k}{k!}$. (b) Tenemos que calcular una esperanza usando probabilidades condicionadas (¿Por qué?). Una vez calculada la probabilidad condicionada necesaria podemos calcular la esperanza con un sumatorio. Obtenemos $r + \lambda(1-p)$.

55. Un niño pequeño se pierde bajando el Monte do Gozo (Santiago de Compostela). El líder del equipo de búsqueda estima que hay una probabilidad de p de que el niño haya bajado por el lado Este y una probabilidad de $1-p$ de que haya bajado por el lado Oeste. En cada día de búsqueda, el líder tiene n personas en su equipo de búsqueda que buscarán al niño de forma independiente y, si el niño se encuentra en el lado del monte por el que están buscando, lo encontrarán con probabilidad u.

    (a) ¿Cómo debemos dividir a las n personas en dos grupos para tener la máxima probabilidad de encontrar al niño?

Solution: Esto es un problema de optimización, por lo que debes encontrar la expresión para $P(E)$ (siendo E="encontrar al niño"), derivarlo respecto a la variable que quieres optimizar, e igualar dicha derivada a cero. De aquí:

$$n_{este} = \frac{1}{2}\left(n + \frac{\log(1-p) - \log(p)}{\log(1-u)}\right).$$

En una aplicación real n_{este} debe ser un entero, pero la fórmula anterior seguramente sea un número real. En tal caso, todavía haría falta comparar cuál de los dos enteros más cercanos es mejor.

56. Cada vez que se tira de una manivela, una máquina produce cierta pieza redonda con un radio medio de $10\,cm$ y varianza $4\,cm^2$. La distribución de las longitudes es acampanada y simétrica.

(a) Calcula el número de piezas con un radio entre 7.5 y 11 cm que esperarías encontrarte entre un conjunto de 500 piezas.

(b) Un trabajador necesita 10 piezas con un radio mayor de 10.5 cm para construir una máquina. ¿Cuál es la probabilidad que necesite tirar más de 20 veces de la manivela?

(c) Si obtienes dos piezas de la máquina, calcula la probabilidad de que una sea mayor que la otra por más de 2 cm.

Resuelve el ejercicio analíticamente y mediante simulaciones.

57. Gosset cultiva T tomates en su huerto, teniendo una media de 20 tomates en él. Gosset verifica si hay defectos en sus tomates. La probabilidad de que un tomate tenga defectos es $p = 0.1$. Suponiendo que el hecho de que un tomate tenga defectos no influye en el resto de tomates, encuentra la cantidad esperada de tomates defectuosos mediante simulaciones.

Solution: $T \sim \mathcal{P}(20)$, ya que noy hay un límite superior para T. Sea D: nº de tomates defectuosos. D se distribuye como $D|T \sim \mathcal{B}(T, p)$. Simulando primero un vector de Ts y usando estos valores para muestrear $D \mid T$, podemos obtener muchas muestras de D. La media es $\mathbb{E}[D] = \lambda p = 2$.

58. Un agente de movilidad está monitorizando una intersección con el fin de detectar infracciones al volante (que ocurren con probabilidad 0.01). El agente quiere saber cuántos autos pasarán antes de que ocurra la primera de las infracciones. En particular,

(a) le gustaría encontrar la probabilidad de que la primera infracción ocurra después de que haya pasado el automóvil número 30.

(b) Después de que hayan pasado 20 coches sin una infracción, un segundo agente sustituye al primero. También le gustaría saber la probabilidad de que la primera violación en movimiento ocurra después del automóvil número 30 (contando desde su llegada). Dado que ya han pasado 20 coches sin ninguna infracción, esta probabilidad es igual a una probabilidad condicional.

Solution: La variable X: "nº de coches que han pasado antes de la primera infracción" tiene distribución $X \sim$ Geom(p=0.01). De donde a) $P(X \geq 30) = 1 - P(X < 30) = 1 - pgeom(29, 0.01) \approx 0.74$. b) Aquí se pide $P(X \geq 50 \mid X \geq 20) = \frac{1 - pgeom(49, 0.01)}{1 - pgeom(19, 0.01)} \approx 0.74$. Este ejercicio ilustra que la distribución geométrica *no tiene memoria* de eventos pasados (ver Wikipedia: *Memorylessness*).

59. El *Old Faithful* es un geyser en el parque de Yellowstone, en Estados Unidos. El *Old Faithful* tiene dos tipos de erupciones: erupciones largas, con probabilidad 0.6; y erupciones cortas, con probabilidad 0.4. Si la erupción es larga, la duración de la misma se puede aproximar por una distribución Normal de media 4 minutos y varianza 0.15. Si la erupción es corta, la duración de la misma se puede aproximar por una distribución Normal de media 2 minutos y varianza 0.2. Se pide:

(a) La distribución de la variable aleatoria "duración de una erupción". (Pista: haz uso de la regla de la probabilidad total y de la probabilidad condicionada).

(b) La duración media de una erupción.

(c) Calcula la probabilidad de que una erupción dure más de 3 minutos.

Solution: (a)

$$f(x) = P(\text{Erupción corta}) f(x \mid \text{Erupción corta}) + P(\text{Erupción larga}) f(x \mid \text{Erupción larga})$$

$$f(x) = 0.4 \cdot \mathcal{N}(x \mid 2, 0.2) + 0.6 \cdot \mathcal{N}(x \mid 4, 0.15),$$

donde $\mathcal{N}(x \mid \mu, \sigma^2)$ representa la función de densidad (que es una función de x) de una normal con media μ y varianza σ^2.
(b)

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x) dx = 0.4 \int_{-\infty}^{\infty} x \mathcal{N}(x \mid 2, 0.2) dx + 0.6 \int_{-\infty}^{\infty} x \mathcal{N}(x \mid 4, 0.15) dx = 0.4 \cdot 2 + 0.6 \cdot 4.$$

Nótese que no hace falta hacer las integrales, ya que son la definición de la esperanza de las normales.
(c) $P(X > 3) = \int_3^\infty f(x)dx = 0.4 \int_3^\infty \mathcal{N}(x \mid 2, 0.2)dx + 0.6 \int_3^\infty \mathcal{N}(x \mid 4, 0.15)dx$, de donde la probabilidad pedida es ≈ 0.6021.

60. Cierto test de Coeficiente Intelectual (CI) se divide en 2 partes. En cierta escuela, las puntuaciones de la parte de matemáticas, X, tienen media 25 y desviación típica 5. Las puntuaciones de la parte de lectura, Y, tienen distribución 55 y desviación típica 10 y, sorprendentemente, parecen ser independientes de los resultados en matemáticas. EL CI final se calcula como $CI = 2X + Y$. ¿Cuál es la probabilidad de que un alumno de la escela sea superdotado, $CI > 130$?

Solution: Las puntuaciones de CI son uno de los ejemplos clásicos de normalidad, por lo que asumiremos normalidad para X e Y. Por otra parte, el enunciado afirma que X y Y son independientes, de donde se sigue que $CI \sim \mathcal{N}(2\mu_x + \mu_y, 4\sigma_x^2 + \sigma_y^2) = \mathcal{N}(105, 200)$. Por tanto,

$$P(CI > 130) \approx 0.039.$$

61. Un ingeniero informático desea estudiar la longitud de las consultas en una base de datos (midiendo la longitud en "palabras"). Supongamos que las consultas contienen, en promedio, aproximadamente 5 palabras. Halla la función de probabilidad de X: número de palabras en una consulta.

Solution: En principio, la distribución más adecuada parece la Poisson con $\lambda = 5$. Sin embargo, no es posible que exista una consulta de cero palabras. Por eso, modelaremos X como una distribución de Poisson *restringida* que no toma valores de 0. Es decir, que

$$P(X = k) = P(Y = k \mid Y > 0)$$

donde $Y \sim \mathcal{P}(5)$. De aquí

$$P(X = k) = \frac{5^k e^{-5}}{k!(1 - e^{-5})}.$$

Nótese que la media de X no será exactamente 5, pero el modelo es suficientemente bueno ya que el enunciado nos dice que la media de 5 es *aproximada*. Si quisiésemos que la media de X fuese exactamente 5 deberíamos calcular la esperanza de

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!(1 - e^{-\lambda})}$$

y despejar la λ adecuada. Nótese que esta aproximación es mucho más compleja que nuestra estrategia.

62. Una bolsa tiene N bombones, siendo $N-1$ de chocolate negro y uno de ellos de riquísimo licor. Antonio está obsesionado con los bombones de licor. Razona detalladamente cuál es la distribución de la variable aleatoria X: número de extracciones hasta que se obtiene el bombón de licor...

    (a) ... si cada vez que Antonio saca un bombón de chocolate negro, lo devuelve a la bolsa.

    (b) ... si cada vez que Antonio saca un bombón de chocolate negro, lo deja fuera de la bolsa para descartarlo.

> **Solution:** (a) $P(X=k) = \left(1-\frac{1}{N}\right)^{k-1}\frac{1}{N}$. (b)$P(X=k) = 1/N$

63. **La dama degustadora de té**. Este es un famoso problema enunciado por el gran Sir Ronald A. Fisher: Una dama afirma que al probar una taza de té con leche es capaz de saber si la leche se echó antes o después que la infusión de té. Consideremos el problema de diseñar un experimento mediante el cual esta afirmación pueda ser probada. Nuestro experimento consiste en mezclar ocho tazas de té, cuatro de una manera y cuatro en la otra, y presentarlas a la dama para que las pruebe en un orden aleatorio. La dama debe dividir las 8 tazas en dos conjuntos de 4: uno para té+leche y otro para leche+té.

    (a) Sea X: número de tazas de leche+té identificadas correctamente. Identifica la distribución de X asumiendo que la dama hace la selección al azar (la dama no es capaz de distinguir los dos tipos de infusiones, solo estaba presumiendo).

    (b) En base al punto anterior. Halla $P(X=k)$.

> **Solution:** (a) X es hipergeométrica: piensa que tenemos una urna con 4 leche+té y 4 té+leche; tras probarlos, y dado que la dama no es capaz de distinguir entre los dos grupos, simplemente saca de la urna 4 infusiones al azar.
> (b) Se sigue que $P(X=k) = \binom{4}{k}\binom{4}{4-k}/\binom{8}{4}$ para $k=0,1,2,3,4$.

64. Una solicitud para un puesto de trabajo se acepta con probabilidad 0.15. Las evaluaciones se realizan de una en una, en orden de llegada, hasta completar el número de puestos que se ofrecen (10 puestos de trabajo). Encuentra la probabilidad de que no se necesiten evaluar más de 100 solicitantes para asignar todos los puestos.

> **Solution:** Se trata de una binomial negativa de parámetros 10 y 0.15. La probabilidad pedida es 0.9449.

65. Gertrude Cox va todos los días al banco. En el banco hay dos colas, una cola estándar y una cola prioritaria. Independientemente del número de personas en el banco, la cola estándar tiene un tiempo medio de espera de 10 minutos, mientras que la prioritaria tiene un tiempo medio de espera de 5 minutos. Desgraciadamente, el banco está intentando ahorrar costes y solo abre ocasionalmente la cola prioritaria. Cada día se decide si se abre la cola prioritaria con independencia de lo que haya sucedido en días anteriores. La probabilidad de abrir la cola prioritaria es 0.7. Gertrude es cliente VIP, así que tiene derecho a usar la cola prioritaria. Si Gertrude va al banco y encuentra la cola prioritaria abierta se pone en ella, en otro caso usa la cola estándar.

    (a) Sabiendo que los tiempos de espera se pueden modelar con una distribución exponencial, obtén la distribución del tiempo de espera (diario) de Gertrude.

    (b) Calcula la probabilidad de que Gertrude tenga que esperar más de 4 minutos en una visita al banco. Calcula el valor sin usar R, haciendo uso de las funciones de densidad/distribución de los chuletarios.

(c) Calcula la probabilidad de que, en 7 visitas al banco, Gertrude sea atendida en menos de 4 minutos en tres ocasiones.

   i. Calcula el valor sin usar R, haciendo uso de las funciones de densidad/distribución de los chuletarios.

   ii. Indica el (único) comando de R que tendrías que usar para calcular esta misma probabilidad.

Solution:

(a) $f(x) = \frac{0.7}{5}\exp(-x/5) + \frac{0.3}{10}\exp(-x/10)$

(b) 0.5156.

(c) i. 0.28097; ii. *dbinom(3, 7, 0.4843)*.

66. De cara a unas elecciones, se estima que los partidos A, B, C y D recibirán el 20, 25, 30 y 25% de los votos, respectivamente. En una muestra de 10 votantes, encuentra la probabilidad de que haya dos votantes del partidoA, 2 votantes de B, tres votantes de C y tres votantes de D.

Solution: Si X_A, X_B, X_C, X_D denota el número de votantes en la muestra, entonces

$$(X_A, X_B, X_C, X_D) \sim \text{Multi}(10, 0.20, 0.25, 0.30, 0.25).$$

Por tanto $P(X_A = 2, X_B = 2, X_C = 3, X_D = 3) \approx 0.0266$.

67. Un globo se infla hasta alcanzar un radio aleatorio R que toma valores de forma uniforme en $(0, 2)$, sea V el volumen del globo (recuerda que $V = \frac{4}{3}\pi R^3$). Calcula el volumen esperado analíticamente y mediante simulaciones.

68. Jacob Bernoulli propuso el siguiente juego de dados. El jugador paga un euro y lanza un dado. Luego lanza un conjunto de n dados, donde n es el número que muestra el primer dado. El número total de puntos que muestran los n dados se utiliza para determinar el resultado del juego. Si el número es menor que doce pierde el euro apostado, mientras que si el número es igual o mayor que doce recibe dos euros. Calcula, usando simulaciones en R, la cantidad de dinero que esperarías ganar al jugar a este juego.

69. Los coches pasan por cierta intersección a un ritmo de 16 automóviles por minuto. ¿Cuál es la probabilidad de que al menos 1000 coches crucen la intersección en la próxima hora?

Solution: Sea X: nº de coches en una hora. Elegimos una distribución de Poisson para modelar X, ya que no hay un límite superior claro en los valores que puede tomar y además, nos dan un valor medio que permite calcular λ. Para λ, el número medio de coches en una hora será 960 de donde $X \sim \mathcal{P}(960)$. La probabilidad por la que se pregunta es

$$P(X \geq 1000) \approx 0.1018.$$

70. Se eligen dos números al azar, X e Y, en el intervalo $[0, 1]$. Sea A la variable aleatoria "Área del rectángulo de lados paralelos a los ejes con vértices en el origen y (X, Y)" (ver Figura 1). Razona cuál es la función de distribución de la variable aleatoria A.

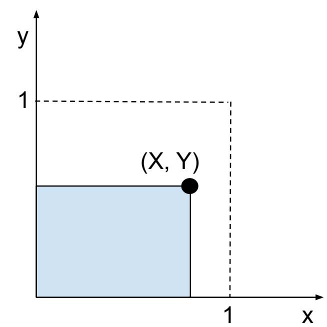

Figure 1: Área del rectángulo.

71. El tiempo medio que un estudiante tarda en completar un examen es 45 minutos. En una clase de 10 estudiantes, ¿cuál es la probabilidad de que al menos un estudiante acabe en menos de 20 minutos?

Solution: 0.9883 (asumiendo independencia)

72. Se tira una moneda (sin trucar) hasta que aparecen cuatro caras. Calcula, analíticamente y mediante simulaciones,

    (a) La probabilidad de que se necesiten 10 lanzamientos.

    (b) La probabilidad de que se necesiten al menos 10 lanzamientos.

    - Halla la función de probabilidad de Y: "número de cruces obtenidas" para el caso general en el que la probabilidad de cara sea p.

Solution: (a) 0.082, (b) 0.254, (c) $P(Y = y) = \binom{y+4-1}{3} p^4 (1-p)^y$.

73. ¿Cuántos lanzamientos, en promedio, debemos lanzar un dado hasta obtener el primer 6?

Solution: 6

74. La distribución exponencial se dice que no tiene memoria. ¿Es por tanto cierto que $\{X > s + t\}$ y $\{X > t\}$ son eventos independientes?

Solution: No es cierto.

75. Un método muy empleado para estimar el número de individuos de poblaciones esquivas es el de *captura-recaptura*. Para estimar la población total N, se capturan r individuos que se etiquetan y se devuelven a la población total. Más tarde, los investigadores recapturan una segunda muestra de tamaño n y se cuenta el número de individuos K que están etiquetados. Se aplica este método para contar el número de peces en un lago. En la "captura" se cuentan y etiquetan 600 peces. En la recaptura se pescan 800 peces, de los cuáles 100 están etiquetados. ¿Cuál es la población total de peces?¿En qué distribución te has basado para tus cálculos?

Solution: 4800 (hipergeométrica).

76. Von Neumann usa su router de alta velocidad para jugar a juegos en línea. El router transmite ceros y unos mediante el envío de señales -1 y +1, respectivamente. Asumimos que cualquier bit tiene probabilidad p de ser cero. La línea telefónica introduce ruído Gaussiano de media 0 y varianza σ^2 (por tanto, el receptor en el otro extremo recibe una señal que es la suma de la señal transmitida y el ruido del canal). El receptor decide que la señal es -1 si la señal recibida es menor que a. ¿Cuál es la probabilidad de que se produzca un error en la transmisión de la información? Usa R para hallar el valor que minimiza la probabilidad de error si $p = 2/5$ y $\sigma^2 = 1/4$.

Solution: $P(error) = p\left(1 - F_z\left(\frac{a+1}{\sigma}\right)\right) + (1-p)F_z\left(\frac{a-1}{\sigma}\right)$ siendo $F_z(x)$ la función de distribución de una normal estandarizada.

77. Las personas con tipo sanguíneo 0-negativo son donantes universales. En cierto país, el 7.2% de la población es de tipo 0-negativo. Un médico quiere encontrar a 10 personas 0-negativas, para lo que hace un cribado entre sus pacientes. ¿Cuál es la probabilidad de encontrar a estas 10 personas entre las primeras 100 personas cribadas?

Solution: 0.183

78. Si $X \sim Exp(\lambda)$, encuentra la función de densidad y la función de distribución de $Y = cX$ $(c > 0)$.

Solution: $Y \sim Exp(\lambda/c)$, por lo que

$$f(y) = \frac{\lambda}{c}\exp(-\frac{\lambda}{c}y)$$

$$F(y) = 1 - \exp(-\frac{\lambda}{c}y)$$

79. Cierta franquicia de comida rápida regala una figura de tu película favorita con cada comida. Hay 10 personajes en total. ¿Cuántas comidas tienes que comprar hasta conseguir los 10 personajes? Resuelve el problema analíticamente y mediante simulaciones.

Solution: 29.3 comidas

Fuentes

- [Spiegel]: Spiegel et al. Probabilidad y estadística, Schaum, segunda edición, 2003.