Probabilidad (3)

Jesús García de Jalón de la Fuente

IES Ramiro de Maeztu
Madrid

2020

Una **distribución de probabilidad** es la representación matemática de un experimento aleatorio.

Está compuesta de:

- La **variable aleatoria**. Puede ser discreta o continua y toma como valores los resultados del experimento. La variable aleatoria es **discreta** si toma un conjunto finito de valores. Es **continua** si toma como valores un intervalo de números reales.
- Una **función de probabilidad**. Se define de manera diferente según la variable aleatoria sea discreta o continua.

La variable aleatoria X toma un conjunto finito de valores $\{x_1, x_2, x_3, \ \ldots, \ x_n\}$.

La **función de probabilidad (pdf)** asigna una probabilidad a cada uno de estos valores:

$$f(x) = \mathrm{p}(X = x)$$

La probabilidad se puede asignar también mediante la **función de distribución (cdf)** que es la probabilidad de que la variable aleatoria tome un valor menor o igual que x:

$$F(x) = \mathrm{p}(X \leq x)$$

Supongamos que se lanza cuatro veces una moneda equilibrada y sea X la variable aleatoria que representa el número de caras obtenidas en los lanzamientos. La variable X puede tomar los valores 0, 1, 2, 3 y 4.

La función de probabilidad está dada en la siguiente tabla:

x_i	0	1	2	3	4
p_i	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$

Supongamos ahora que en una caja hay 3 bolas blancas y 5 bolas negras y se extraen al azar 3 bolas sin reemplazamiento. Sea Y la variable aleatoria que representa el número de bolas blancas extraídas. La variable Y puede tomar los valores 0, 1, 2 y 3.

La función de probabilidad está dada en la siguiente tabla:

y_i	0	1	2	3
p_i	$\frac{5}{28}$	$\frac{15}{28}$	$\frac{15}{56}$	$\frac{1}{56}$

Sea la variable aleatoria X cuya función de probabilidad está dada en la siguiente tabla:

x_i	x_1	x_2	x_3	x_4	$\dots$
p_i	p_1	p_2	p_3	p_4	$\dots$

La **media o valor esperado** y la **varianza** se definen por:

$$\mathrm{E}(X) = \mu = \sum p_i x_i \ ; \qquad \mathrm{Var}(X) = \sigma^2 = \sum p_i \left(x_i - \mathrm{E}(X)\right)^2 = \sum p_i x_i^2 - \mu^2$$

La **desviación típica** σ es la raíz cuadrada de la varianza.

En el ejemplo que hemos visto anteriormente

x_i	0	1	2	3	4
p_i	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$

La media sería:

$$\mathrm{E}(X) = \frac{1}{16}\cdot 0 + \frac{4}{16}\cdot 1 + \frac{6}{16}\cdot 2 + \frac{4}{16}\cdot 3 + \frac{1}{16}\cdot 4 = \frac{32}{16} = 2$$

y la varianza:

$$\mathrm{Var}(X) = \frac{1}{16}(0-2)^2 + \frac{4}{16}\cdot(1-2)^2 + \frac{6}{16}\cdot(2-2)^2 + \frac{4}{16}\cdot(3-2)^2 + \frac{1}{16}\cdot(4-2)^2 = 1$$

Mediante la otra fórmula:

$$\mathrm{Var}(X) = \frac{1}{16}\cdot 0^2 + \frac{4}{16}\cdot 1^2 + \frac{6}{16}\cdot 2^2 + \frac{4}{16}\cdot 3^2 + \frac{1}{16}\cdot 4^2 - 2^2 = \frac{80}{16} - 4 = 1$$

La desviación típica es la raíz de la varianza y también es igual a 1.

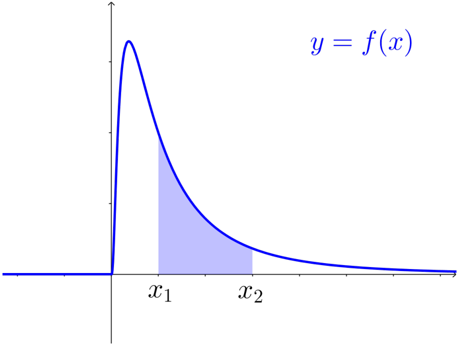
La variable aleatoria puede tomar las valores de un intervalo $[a, b]$ (a y b pueden ser infinitos).

La probabilidad queda determinada por la **función de densidad (pdf)** $f(x)$ de tal forma que:

$$\mathrm{p}(x_1 \leq X \leq x_2) = \int_{x_1}^{x_2} f(x)\,\mathrm{d}x$$

Puesto que la suma de todas las probabilidades debe ser igual a 1, se cumple que:

$$\int_a^b f(x)\,\mathrm{d}x = 1$$

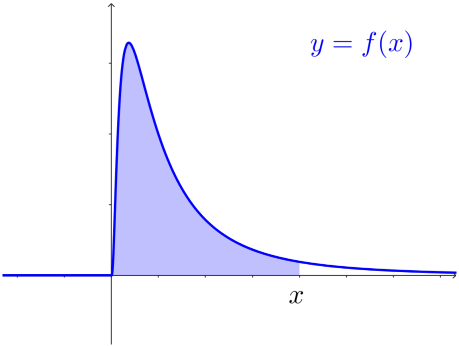


La **función de distribución (cdf)** representa la probabilidad acumulada:

$$F(x) = \mathrm{p}(X \leq x) = \int_a^x f(x)\, \mathrm{d}x$$

también:

$$\mathrm{p}(x_1 \leq X \leq x_2) = F(x_2) - F(x_1)$$



La función de densidad es la derivada de la función de distribución:

$$f(x) = F'(x)$$

Las fórmulas de la media y la varianza son iguales que las de la variable discreta sustituyendo la suma por una integral:

$$\mathrm{E}(X) = \int_a^b x f(x)\,\mathrm{d}x$$

$$\mathrm{Var}(X) = \int_a^b (x - \mathrm{E}(X))^2 f(x)\,\mathrm{d}x = \int_a^b x^2 f(x)\,\mathrm{d}x - \mathrm{E}(X)^2$$

La desviación típica es la raíz cuadrada de la varianza.

Ejemplo

Una variable aleatoria continua tiene una función de densidad de probabilidad dada por:

$$\mathrm{p}(x) = \begin{cases} k(2x - x^2) & \text{si } 0 \leq x \leq 2 \\ 0 & \text{en el resto} \end{cases}$$

(a) Hallar el valor de k.

(b) Calcular $\mathrm{p}\left(\frac{1}{4} < X < \frac{1}{2}\right)$.

Solución:

(a) Puesto que la suma de las probabilidades debe ser igual a 1:

$$k\int_0^2 (2x - x^2)\,\mathrm{d}x = k\left[\frac{2x^2}{2} - \frac{x^3}{3}\right]_0^2 = \frac{4}{3}\cdot k = 1\ ; \qquad k = \frac{3}{4}$$

(b)

$$\mathrm{p}\left(\frac{1}{4} < X < \frac{1}{2}\right) = \frac{3}{4}\int_{\frac{1}{4}}^{\frac{1}{2}} (2x - x^2)\,\mathrm{d}x = \frac{3}{4}\left[\frac{2x^2}{2} - \frac{x^3}{3}\right]_{\frac{1}{4}}^{\frac{1}{2}} = \frac{3}{4}\left(\frac{1}{4} - \frac{\frac{1}{8}}{3} - \frac{1}{16} - \frac{\frac{1}{64}}{3}\right) = \frac{29}{256}$$

Ejemplo

La función densidad de probabilidad de una variable aleatoria X viene dada por:

$$f(x) = \begin{cases} ax\cos x & 0 \leq x \leq \frac{\pi}{2} \\ 0 & \text{resto de valores} \end{cases}$$

(a) Calcular a (b) Calcular la función de distribución (c) Hallar $\mathrm{p}\left(X < \dfrac{\pi}{4}\right)$.

Solución:

(a) Puesto que la suma de todas las probabilidades debe ser igual a 1:

$$a\int_0^{\frac{\pi}{2}} x\cos x \, \mathrm{d}x = 1$$

Integrando por partes:

$$a\Big[x \operatorname{sen} x + \cos x\Big]_0^{\frac{\pi}{2}} = a\left(\frac{\pi}{2} - 1\right) = 1$$

y de aquí:

$$a = \frac{1}{\frac{\pi}{2} - 1} = \frac{2}{\pi - 2}$$

(b) La función de distribución es entonces:

$$\begin{aligned} \mathrm{p}(X < x) &= \frac{2}{\pi - 2}\Big[x \operatorname{sen} x + \cos x\Big]_0^x \\ &= \frac{2}{\pi - 2}\left(x \operatorname{sen} x + \cos x - 1\right) \end{aligned}$$

(c) Con esa función de distribución:

$$\mathrm{p}\left(X < \frac{\pi}{4}\right) = \frac{2}{\pi - 2}\left(\frac{\pi}{4}\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2} - 1\right) \simeq 0{,}460$$

Agradecimiento

Gracias por vuestra atención