Distribuciones de Probabilidad

IES O Couto

curso 2019-2020

Variable Aleatoria Binomial

Experimento de Bernouilli

Un experimento aleatorio es un experimento de Bernouilli si cumple que:

I) Toda realización del experimento es independiente de realizaciones anteriores.

II) El experimento solo tiene dos posibles resultados (que se denominan éxito y fracaso).

III) Las probabilidades de éxito son las mismas en cualquier realización del experimento.

Definición de V.A. Binomial

Si repetimos n veces un experimento aleatorio de Bernouilli, y p es la probabilidad de éxito, la variable aleatoria $X =$ "Número de éxitos obtenidos" sigue una distribución Binomial de parámetros n y p. Abreviadamente $X \sim B(n, p)$.

Ejemplo 1

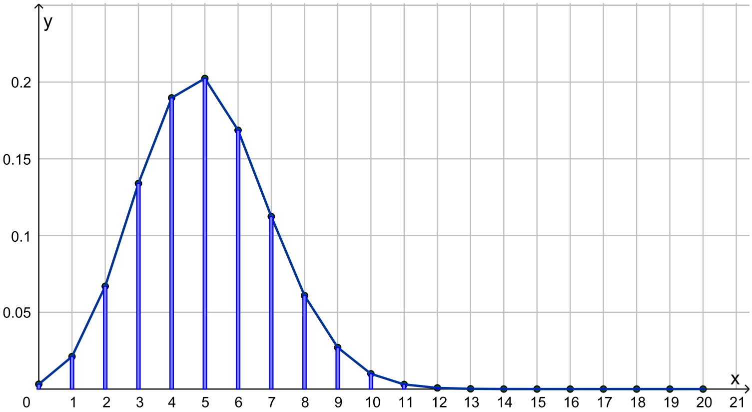
- Número de caras obtenidas al lanzar 20 veces una moneda equilibrada: $X \sim B(20, 0'5)$
- Número de preguntas acertadas al contestar al azar un examen tipo test de 30 preguntas, en el que cada una tiene cuatro posibles respuestas, pero solamente una correcta: $X \sim B(30, 0'25)$

Características de X $\sim$ B(n,p)

- La función masa de probabilidad es

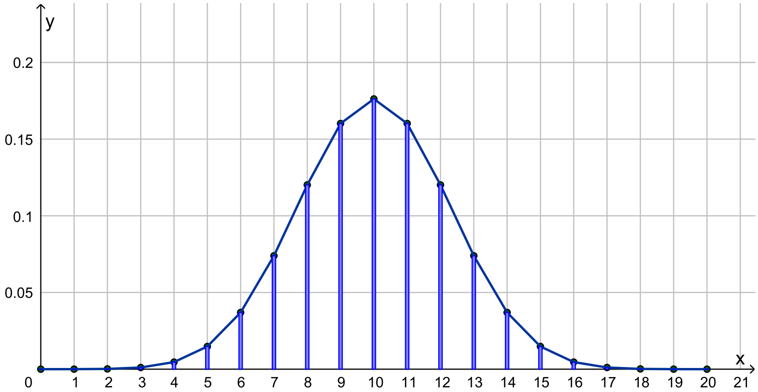
$$\mathcal{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \text{ para } k = 0,\ 1,\ 2, \ldots n$$

- $E(X) = np$
- $V(X) = npq$ con $q = 1 - p$ (q es la probabilidad de fracaso)



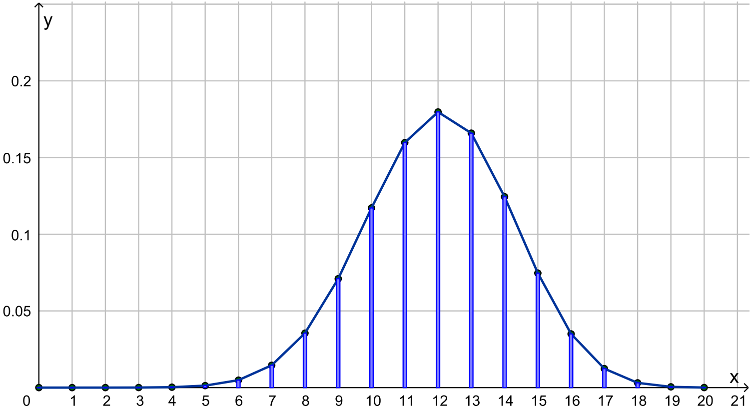
$X \sim B(20, 0.25)$

$P(X = k) = \binom{20}{k} 0.25^k \cdot (1 - 0.25)^{20-k}, \quad k = 0,\ 1,\ 2, \ldots 20$



$X \sim B(20, 0.5)$

$P(X = k) = \binom{20}{k} 0.5^k \cdot (1 - 0.5)^{20-k}, \quad k = 0,\, 1,\, 2, \ldots 20$



$X \sim B(20, 0.6)$

$P(X = k) = \binom{20}{k} 0.6^k \cdot (1 - 0.6)^{20-k}, \quad k = 0,\ 1,\ 2, \ldots 20$

Variables Aleatorias Absolutamente Continuas

Una variable aleatoria es absolutamente continua si su función de distribución F es continua en todo $\mathbb{R}$, y derivable con derivada continua en todo $\mathbb{R}$ salvo en a lo sumo un conjunto numerable de puntos.

Función densidad

Esto equivale a decir que existe una función $f : \mathbb{R} \longrightarrow \mathbb{R}$ no negativa e integrable tal que

$$F(x) = \int_{-\infty}^{x} f(t)dt$$

- La función f se llama función densidad de la v.a. X.
- Como consecuencia, toda probabilidad será un área.

Teorema de caracterización de función densidad

Toda función $f : \mathbb{R} \longrightarrow \mathbb{R}$ que satisface

I) $f(x) \geq 0, \forall x \in \mathbb{R}$

II) f es Riemann integrable en todo $\mathbb{R}$

III) $\displaystyle\int_{-\infty}^{+\infty} f(t)dt = 1$

es función densidad de una variable aleatoria absolutamente continua X, cuya función de distribución es

$$F(x) = \int_{-\infty}^{x} f(t)dt$$

Propiedades de una función densidad

Toda función de densidad cumple como consecuencia de la definición, y de las propiedades de la integral de Riemann, las siguientes propiedades:

- $f(x) \geq 0 \quad \forall x \in \mathbb{R}$
- $\displaystyle\int_{-\infty}^{+\infty} f(x)dx = 1$
- $\displaystyle\mathcal{P}(a,b] = F(b) - F(a) = \int_{-\infty}^{b} f(t)dt - \int_{-\infty}^{a} f(t)dt = \int_{a}^{b} f(t)dt$

Es decir, la probabilidad de que los valores de la variable estén comprendidos entre a y b es el área bajo la función densidad entre $x = a$ y $x = b$.

- $\mathcal{P}(X = a) = 0$. En efecto, por ser F una función continua:

$$\mathcal{P}(X = a) = \lim_{h \to 0^+} \mathcal{P}(a - h, a] = \lim_{h \to 0^+} (F(a) - F(a - h)) =$$

$$F(a) - F(a) = 0$$

- Si f es continua en x_o, F es derivable en x_o, y además $F'(x_o) = f(x_o)$ (como consecuencia del Teorema Fundamental del Cálculo Integral).

Variable Normal

Definición

Una variable aleatoria X absolutamente continua se dice que sigue una distribución normal de media μ y desviación típica σ ($X \sim N(\mu, \sigma)$) si su función de densidad es

$$f(x) = \frac{1}{\sqrt{2\,\pi\,\sigma^2}}\, e^{-\frac{(x-\mu)^2}{2\,\sigma^2}} \quad -\infty < x < +\infty$$

Características de $X \sim N(\mu, \sigma)$

- $E(X) = \mu$
- $Var(X) = \sigma^2$

Representación gráfica de la densidad de $N(\mu, \sigma)$

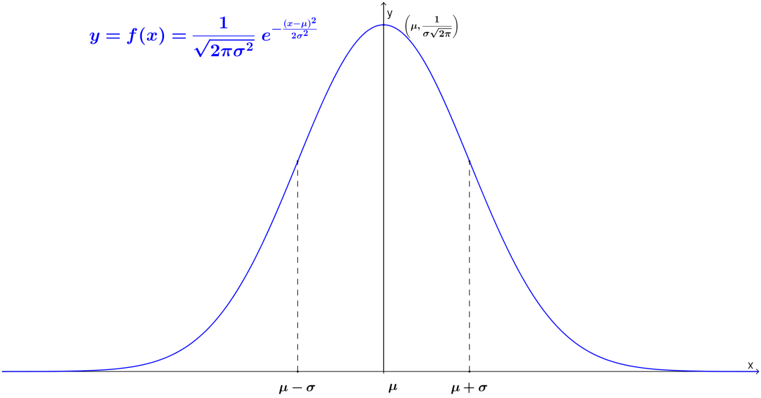


Imagen: Densidad de $X \sim N(\mu, \sigma)$

Tipificación de una variable $N(\mu, \sigma)$

Sea $X \sim N(\mu, \sigma)$, y sea Z la variable aleatoria transformada mediante la expresión $Z = \dfrac{X - \mu}{\sigma}$.

Entonces:

I) $Z \sim N(0, 1)$

II) $\mathcal{P}(X \leq x) = \mathcal{P}\left(Z \leq \frac{x-\mu}{\sigma}\right)$

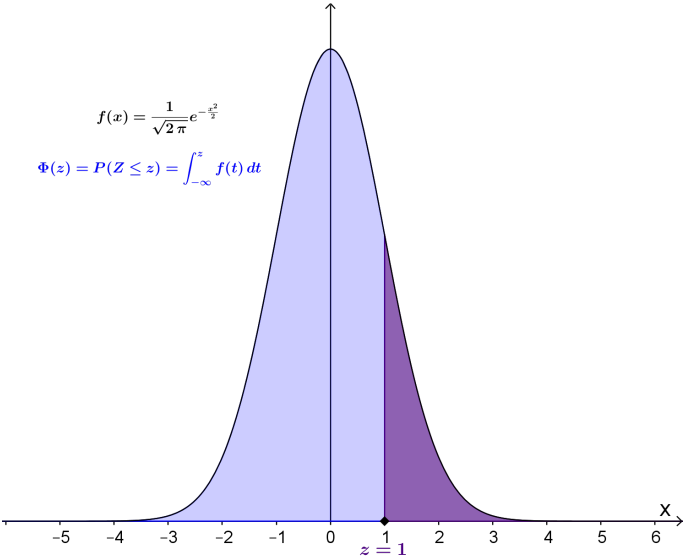
Cálculo de probabilidades de una Z $\sim$ N(0,1)

Dado que no se conoce una función primitiva de la densidad de la N(0,1), los valores de su distribución $\Phi(z)$ se encuentran tabulados.
Z es simétrica respecto $z = 0$, por eso en las tablas, solo aparecen los valores de $\mathcal{P}(Z \leq z)$ para $z \geq 0$. Por tanto, para cualquier $z \geq 0$:

- $\mathcal{P}(Z \geq z) = 1 - \mathcal{P}(Z \leq z)$
- $\mathcal{P}(Z \leq -z) = \mathcal{P}(Z \geq z) = 1 - \mathcal{P}(Z \leq z)$
- $\mathcal{P}(Z \geq -z) = \mathcal{P}(Z \leq z)$
- $\mathcal{P}(|Z| \leq z) = \mathcal{P}(Z \leq z) - \mathcal{P}(Z \leq -z) = \mathcal{P}(Z \leq z) - \mathcal{P}(Z \geq z) = \mathcal{P}(Z \leq z) - (1 - \mathcal{P}(Z \leq z)) = 2\mathcal{P}(Z \leq z) - 1$

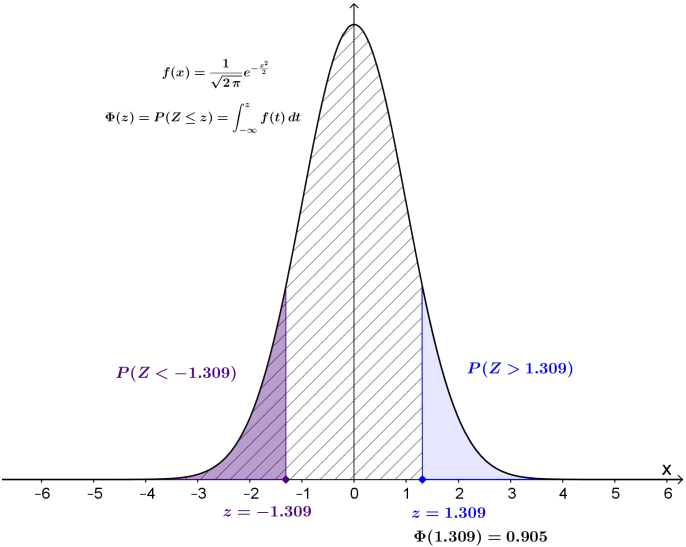
$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

$$\Phi(z) = P(Z \leq z) = \int_{-\infty}^{z} f(t)\, dt$$



$$\Phi(1) = P(Z \leq 1) = 0.841$$

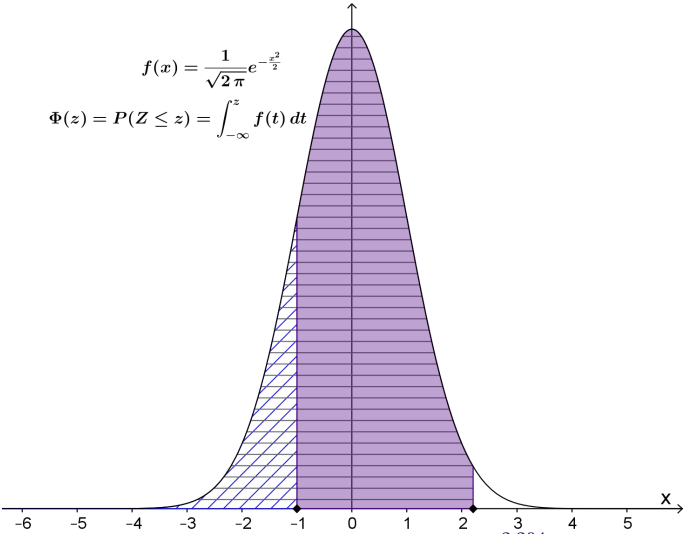
$$P(Z > 1) = 1 - P(Z \leq 1) = 1 - 0.841 = 0.159$$



$$P(Z > 1.309) = 1 - P(Z \leq 1.309) = 1 - 0.905 = 0.095$$

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

$$\Phi(z) = P(Z \leq z) = \int_{-\infty}^{z} f(t)\, dt$$

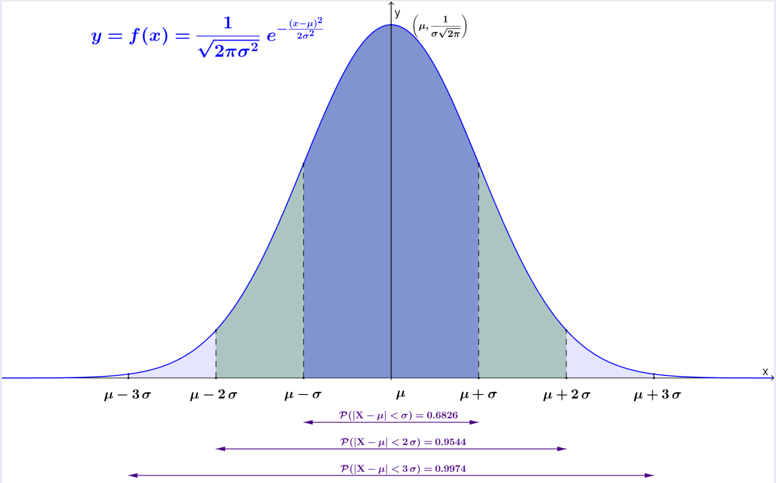


$z_1 = -1$ $\quad$ $z_2 = 2.204$

$\Phi(-1) = 0.159$ $\quad$ $\Phi(2.204) = 0.986$

$$P(-1 \leq Z \leq 2.204) = \Phi(2.204) - \Phi(-1) = 0.828$$

Algunas áreas importantes de N(μ,σ)



Aproximación de una B(n,p) a una N($np, \sqrt{npq}$)

Para n suficientemente grande, una distribución binomial $B(n, p)$ puede aproximarse por una normal $N(np, \sqrt{n\,p\,q})$ (con $q = 1 - p$).

Condiciones para la aproximación

- Puede usarse la aproximación cuando se cumpla $np > 5$ y $nq > 5$
- Dado que la normal es una distribución de tipo continuo, el uso directo de la aproximación anterior asignaría probabilidad cero a puntos aislados y a los extremos de intervalos cerrados. Se usa la corrección por continuidad o corrección de Yates.

Corrección de Yates

$$\mathcal{P}(X = k) = \mathcal{P}\left(k - 0'5 \leq X \leq k + 0'5\right) =$$

$$\mathcal{P}\left(\frac{k - 0'5 - n\,p}{\sqrt{n\,p\,q}} \leq Z \leq \frac{k + 0'5 - n\,p}{\sqrt{n\,p\,q}}\right)$$

con $k = 0,\ 1,\ 2,\ \ldots,\ n$

- Las probabilidades de tipo $\mathcal{P}(X \leq a)$ se aproximarán antes de tipificar por $\mathcal{P}(X \leq a + 0'5)$.
- Las probabilidades de tipo $\mathcal{P}(X < a)$ se aproximarán antes de tipificar por $\mathcal{P}(X \leq a - 0'5)$.
- Las probabilidades de tipo $\mathcal{P}(X > a)$ se aproximarán antes de tipificar por $\mathcal{P}(X \geq a + 0'5)$
- Las probabilidades de tipo $\mathcal{P}(X \geq a)$ se aproximarán antes de tipificar por $\mathcal{P}(X \geq a - 0'5)$