

Curso 12

Correlación

El teorema central límite (CLT)

Para una muestra:

- suficientemente grande tomada al azar y
- con reemplazo tenemos que:

la distribución del promedio de las muestras es normal(se ve como una campana).

Distribución Gaussiana

- Tiene una ecuación.
- Es simétrica, mean está en la mitad = median.
- Proporciones = áreas

(Notebook)

Relación entre dos variables

- Asociación
- Tendencia
 - asociación positiva
 - asociación negativa
- Patrón
 - lineal
 - no-lineal

Visualizar, luego cuantificar

Correlación

- Mide la asociación lineal.
- Usa unidades estándar.
- está entre $[-1, 1]$
 - $r = 1$ scatter es una línea hacia arriba,
 - $r = -1$ scatter es una línea hacia abajo.
- $r = 0$, no hay correlación, no hay una asociación lineal.

(Notebook)

Fórmula

$$r = \text{mean}(x_sd * y_sd)$$

- Simétrica, qué pasa si intercambio x con y?
- Si $x > 0$ y $y > 0$, entonces r también.
- No tiene dimensiones.