

# Análisis Predictivo y Gestión de Datos

## Sesión 0: Fundamentos de Estadística Básica

Oscar Leonardo Rincón León

8 de mayo de 2025

# Objetivo de la sesión

Revisar los fundamentos esenciales de estadística descriptiva e inferencial como base para el análisis predictivo aplicado a datos reales.

**Definición:** La media representa el valor promedio de una variable cuantitativa.

**Fórmula:**

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

donde  $x_i$  son los valores y  $n$  el número de observaciones.

**Mediana:** Valor que divide la distribución en dos partes iguales.

**Moda:** Valor que ocurre con mayor frecuencia en el conjunto de datos.

# Varianza y desviación estándar

**Varianza:** Mide la dispersión respecto a la media.

**Fórmulas:**

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad \text{y} \quad s = \sqrt{s^2}$$

donde  $s$  es la desviación estándar.

# Distribución normal

**Descripción:** Distribución continua simétrica en forma de campana.

**Fórmula:**

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

donde  $\mu$  es la media y  $\sigma$  la desviación estándar.

**Definición:** Proporción de veces que ocurre un evento en un conjunto de resultados posibles.

**Fórmula clásica de Laplace:**

$$P(A) = \frac{\text{número de casos favorables}}{\text{número total de casos posibles}}$$

# Covarianza y correlación

**Covarianza:** Mide si dos variables aumentan juntas.

**Correlación:** Estandariza la covarianza para compararla.

**Fórmulas:**

$$\text{Cov}(X, Y) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \quad , \quad r = \frac{\text{Cov}(X, Y)}{s_x s_y}$$



**Definición:** Rango dentro del cual se espera que se encuentre un parámetro poblacional con cierto nivel de confianza.

**Fórmula:**

$$\bar{x} \pm z \cdot \frac{s}{\sqrt{n}}$$

donde  $z$  depende del nivel de confianza (por ejemplo, 1.96 para 95 %).

## Pasos:

- 1 Plantear  $H_0$  y  $H_1$
- 2 Calcular estadístico de prueba
- 3 Determinar el valor p o comparar con valor crítico
- 4 Tomar decisión (rechazar o no  $H_0$ )

# Resumen de conceptos clave

- Media, mediana, moda
- Varianza y desviación estándar
- Distribución normal
- Probabilidad, correlación
- Intervalos de confianza, prueba de hipótesis