

# Estadística inferencial vs descriptiva

Descriptiva: Parte de la estadística que arregla los datos de forma que puedan ser analizados e interpretados

Nos ayuda a determinar:

- La tendencia central de una variable
- La variabilidad de una variable
- La distribución de una variable

Inferencial: Parte de la estadística que busca predecir o deducir características o resultados esperados de una población, basados en los datos obtenidos de una muestra de esa población.

Nos ayuda a determinar:

- Muestreo
- Intervalos de confianza
- Validación de hipótesis
- Evitar sesgos

Inferencia estadística: nos ayuda a sacar conclusiones que se obtienen sobre los parámetros de la población de datos. Además lo que estamos haciendo es estudiar el grado de fiabilidad de los resultados extraídos del estudio.

## Uso en data science y machine learning

Tanto en un análisis como en un modelo predictivo, la estadística inferencial servirá para:

- Entender la distribución de los datos.
- Crear y validar hipótesis.
- Hacer experimentos.
- Elegir modelos predictivos adecuados según los datos.

## Estadísticos principales

### Experimento

Procedimiento que puede repetirse infinitamente y tiene un conjunto bien definido de resultados posibles, conocido como espacio muestral.

- Aleatorio: si tiene más de un resultado posible.
- Determinista: si solo tiene un resultado posible.

### Población y muestra

Muestra: Subconjunto de datos perteneciente a una población.

Condiciones:

- Número suficiente de registros para ser estadísticamente significativo.
- Representación no sesgada

### Evento

Cada uno de los posibles resultados de un experimento

### Variables

Es una característica que puede obtener diferentes valores.

Tipos:

- Cualitativas: atributos (no medibles).

- Cuantitativas: números (medibles)
  - Discretas: son números enteros
  - Continuas: son números que aceptan decimales (altura, peso...)

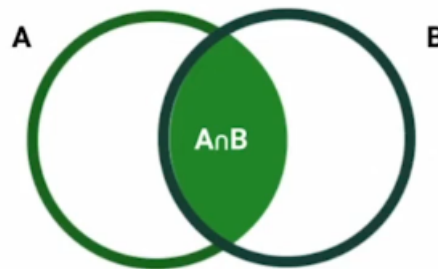
## Probabilidad

Mide qué tan posible es que ocurra un evento determinado.  
El análisis de los eventos probabilísticos se denomina estadística.

## Probabilidad condicionada

Posibilidad de que ocurra un evento como consecuencia de que otro evento haya sucedido.

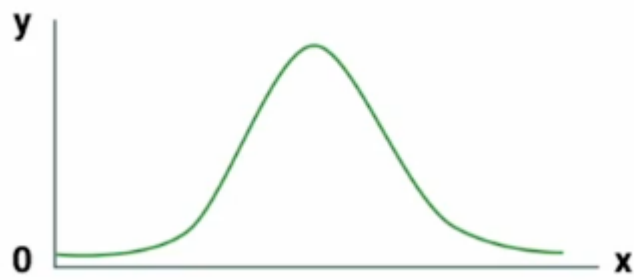
$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$



## Poblaciones normales

### Distribución normal

- Distribución normal = Distribución de Gauss
- Su moda = su media = su mediana
- Es asimétrica
- Tiene forma de campana



Ejemplos de población normal:

- Calorías ingeridas y peso
- Presión sanguínea
- Tamaño de los coches producidos por una máquina