

# Probabilidad

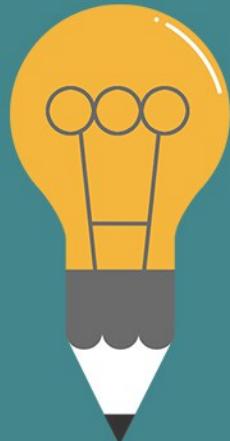
**Gerard Pascual Fontanilles  
EINES D'ANÀLISI I VISUALITZACIÓ DE DADES**

# Probabilidad vs estadística

## Diferencia entre probabilidad y estadística

- **Probabilidad:** Proporciona el marco teórico para entender la aleatoriedad y modelar la incertidumbre. Se centra en predecir resultados cuando se conocen las condiciones del experimento.
- **Estadística:** Utiliza los principios probabilísticos para analizar datos reales, inferir conclusiones y tomar decisiones cuando las condiciones o distribuciones son desconocidas.

La probabilidad va del modelo a los datos; la estadística, de los datos al modelo.



# Probabilidad

## Definición de probabilidad

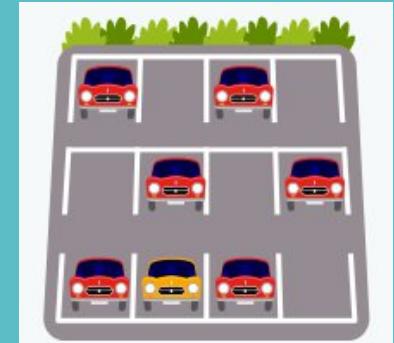
La probabilidad es el cálculo matemático que evalúa las posibilidades que existen de que un evento suceda cuando interviene el azar.

Con base en un valor entre 0 y 1, 0 menos seguridad y 1 toda la seguridad.

El cálculo de probabilidades se expresa en porcentaje y responde a la siguiente fórmula:

$$\text{Probabilidad} = \text{Casos favorables} / \text{casos posibles} \times 100.$$

Probabilidad salida coche rojo: 6/7  
Probabilidad salida coche amarillo: 1/7

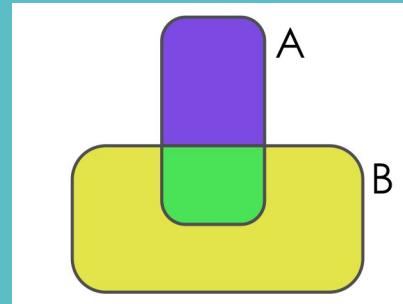


# Probabilidad condicionada

La probabilidad condicionada nos dice cuánto cambia la probabilidad de un suceso  $A$  cuando sabemos que otro suceso  $B$  ya ha ocurrido.

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

- $P(A|B)$ : Probabilidad de que ocurra A sabiendo que ocurrió B.
- $P(A \cap B)$ : Probabilidad de que ocurran ambos sucesos A y B al mismo tiempo (*intersección*).
- $P(B)$ : Probabilidad de que ocurra B (*evento condicionante*).



# Teorema de Bayes

El Teorema de Bayes permite actualizar la probabilidad de un suceso (*a priori*) cuando se dispone de nueva información o evidencia (*a posteriori*).

**P(A|B): Probabilidad posterior:** la probabilidad de A dada la evidencia B (*a posteriori*).

- **P(B|A): Probabilidad revisada:** probabilidad de observar B si A fuera cierto.
- **P(A): Probabilidad inicial:** conocimiento previo sobre A (*a priori*).
- **P(B): Probabilidad total** de que ocurra B.



$$P(A | B) = \frac{P(B | A)P(A)}{P(B)}$$

# Definición tipos y caracterización de variables aleatorias



**Valor esperado:** es el promedio teórico de una variable aleatoria, que representa el resultado medio que se obtendría si el experimento se repitiera infinitas veces.

**Media:** es el promedio aritmético de un conjunto de datos y mide su tendencia central.

**Desviación estándar:** mide cuánto se alejan los valores de los datos respecto a la media. Una desviación pequeña indica datos agrupados cerca de la media, y una grande, datos más dispersos.

$$E = \sum p_i \cdot x_i$$

$$\text{Media}(X) = \bar{x} = \frac{\sum_{i=1}^N X_i}{N}$$

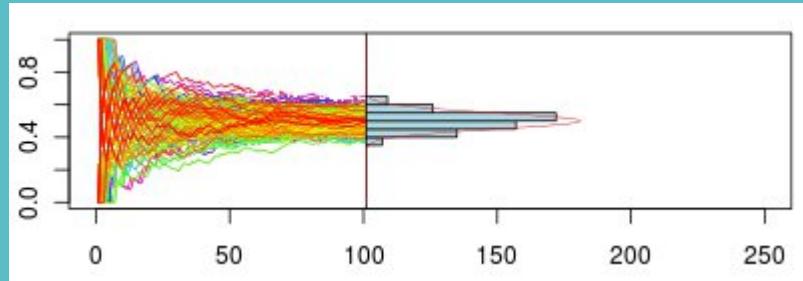
$$\sigma = \sqrt{\frac{\sum (x - \mu)^2}{N}}$$

# Ley de los grandes números

**La Ley de los Grandes Números** establece que, al aumentar el número de observaciones de un experimento aleatorio, la media de los resultados observados tiende a aproximarse al valor esperado teórico.

Cuanto mayor es el tamaño de la muestra, más se parece la media muestral a la media real (esperanza matemática) de la población.

Explica que el azar se “compensa” cuando se repite muchas veces un experimento.



# Teorema del límite central

**El Teorema del Límite Central** establece que cuando el tamaño de la muestra es suficientemente grande, la distribución de las medias muestrales tiende a seguir una distribución normal, independientemente de la forma de la distribución original de los datos.

No importa cómo sean los datos de partida, al tomar muchas muestras grandes y calcular sus medias, estas medias se distribuyen de forma aproximadamente normal.

