17/02/2025

Herramientas estadísticas

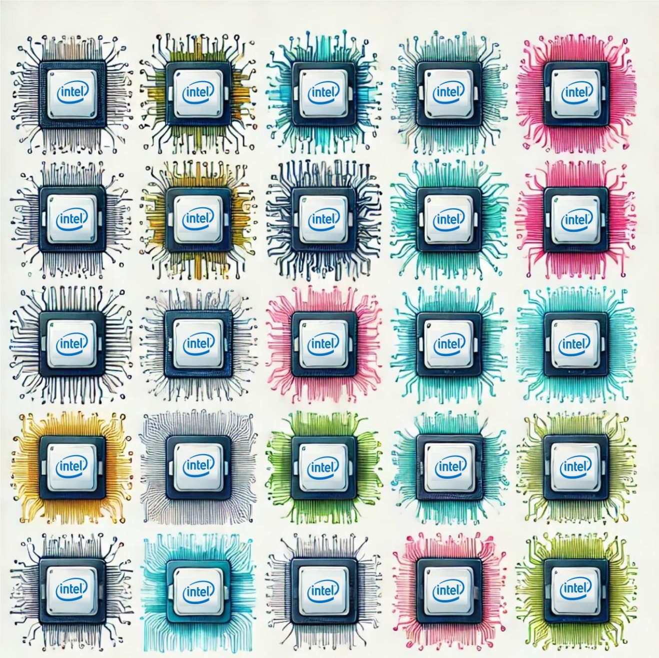
Aprea Mateo, Blaksley Felicitas, Maldonado Sofía

Contenidos

01

Población y muestra

Población



Muestra

- Fenómeno aleatorio ⟶ Tirar un dado
  - Un determinado resultado tiene probabilidad p ⟶ Obtener un 3
  - Un resultado alternativo tiene probabilidad q = 1 - p ⟶ Obtener otro número

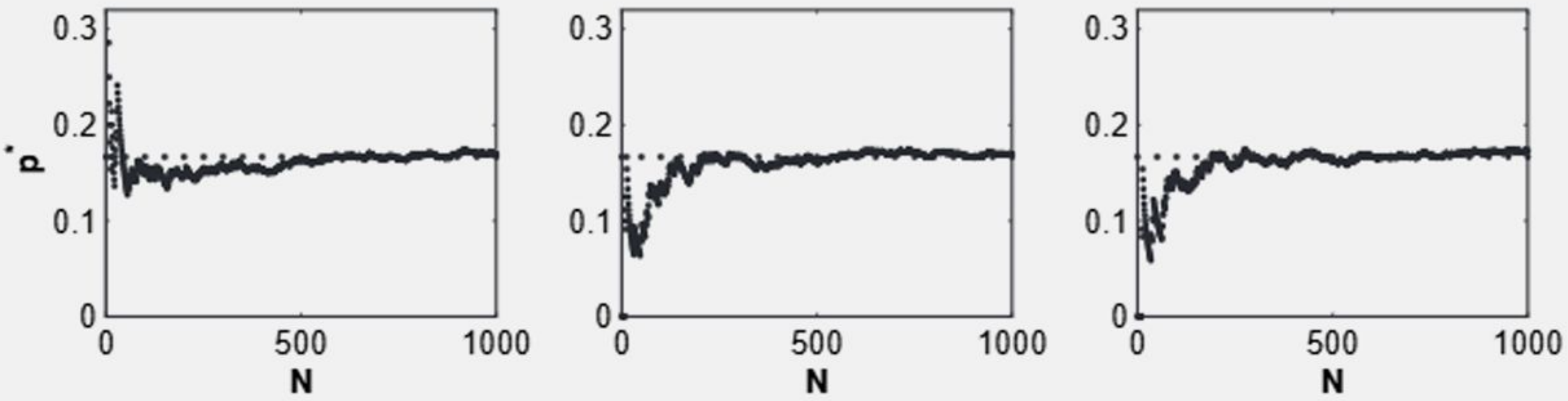
- Utilizamos una variable discreta:
  - K1 = 0 para los resultados negativos ⟶ Obtuve un número distinto a 3
  - K2 = 1 para los resultados positivos ⟶ Obtuve un 3

- Población: Infinitas posibles repeticiones del experimento
- Muestra: N repeticiones idénticas

Distribución de la muestra

- Resultados negativos en n_1^* repeticiones ⟶ Obtuve un número distinto a 3
  - $q^* = n_1^*/N$
- Resultados positivos en n_2^* repeticiones ⟶ Obtuve un 3
  - $p^* = n_2^*/N$

$$\mathcal{P}^*(k_2) = p^*$$



02

Media y varianza muestral

Media y varianza poblacional

Media (m)

$$\sum_j x_j p_j$$

$$\int_{-\infty}^{+\infty} x\, f(x)\, \mathrm{d}x$$

Varianza (D)

$$\sum_j (x_j - m)^2 p_j$$

$$\int_{-\infty}^{+\infty} (x - m)^2 f(x)\, \mathrm{d}x$$

Media y varianza muestral

Media (m*)

$$\frac{1}{N}\sum_{i=1}^{N} x_i$$

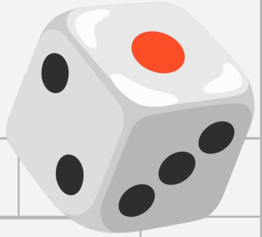
$$\sum_{j=1}^{\mathcal{N}} x_j p_j^*$$

Varianza (D*)

$$\frac{1}{N}\sum_{i=1}^{N} (x_i - m^*)^2$$

$$\sum_{j=1}^{\mathcal{N}} (x_j - m^*)^2 \, p_j^*$$

Ejemplo: tiradas de dados



Población

$$p_j = \frac{1}{6} \quad \forall\, j$$

$$m = \sum_{j=1}^{6} \underbrace{x_j}_{=j} \frac{1}{6} = 3.5$$

Muestra

$$\{\underset{x_1}{\underset{\uparrow}{2}}, 5, 3, 2, 6, 1, 4, 5, 3, \underset{x_{10}}{\underset{\uparrow}{1}}\}$$

$$m^* = \frac{1}{10} \sum_{j=1}^{10} x_j = 3.2$$

La media muestral (m*) y la varianza muestral (D*) son **variables aleatorias**

Propiedades

$$m[aX] = a\,m[X]$$

$$m[X+Y] = m[X] + m[Y]$$

$$D[aX] = a^2\,D[X]$$

$$D[X+Y] = D[X] + D[Y] \;*$$

* si X e Y son **independientes**

Media muestral (m*)

$$
\begin{aligned}
\mathbf{m}[m^*] &= \mathbf{m}\left[\frac{X_1 + X_2 + \cdots + X_N}{N}\right] \\
&= \frac{1}{N}\left\{\mathbf{m}[X_1] + \mathbf{m}[X_2] + \cdots + \mathbf{m}[X_N]\right\} \\
&= \frac{1}{N}\{N\, m\} = m\,.
\end{aligned}
$$

Media muestral (m*)

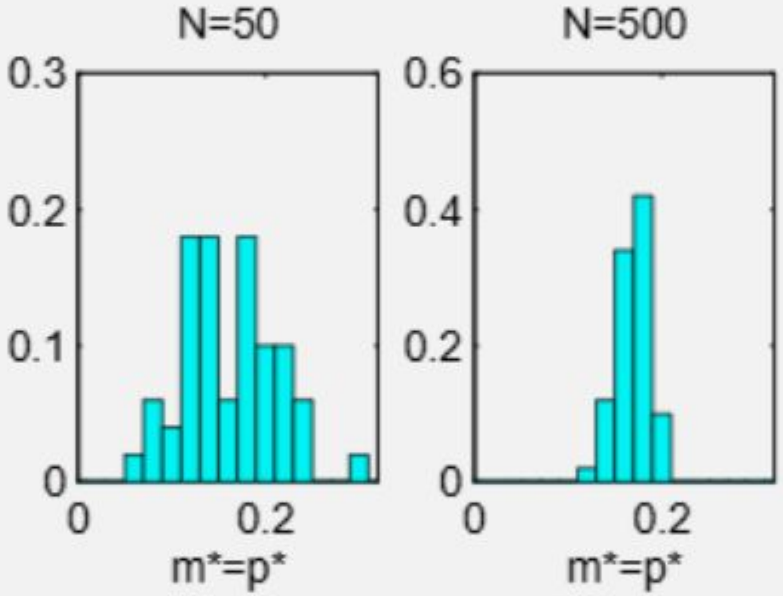
$$\begin{aligned}
\mathbf{D}[m^*] &= \mathbf{D}\left[\frac{X_1 + X_2 + \cdots + X_N}{N}\right] \\
&= \frac{1}{N^2}\left\{\mathbf{D}[X_1] + \mathbf{D}[X_2] + \cdots + \mathbf{D}[X_N]\right\} \\
&= \frac{1}{N^2}\{N\,D\} = \frac{D}{N}\,.
\end{aligned}$$

Media muestral (m*): ejemplo

$$K = \begin{cases} 1 & \text{si sale un } 5 \\ 0 & \text{si no} \end{cases}$$

$$p = \frac{1}{6}$$

$$m = p = 0.167$$



Varianza muestral (D*)

$$D^* = \frac{1}{N}\sum_{i=1}^{N}(x_i - m^*)^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - m)^2 - (m^* - m)^2$$

$$\mathbf{m}[D^*] = \mathbf{m}\left[\frac{1}{N}\sum_{i=1}^{N}(x_i - m)^2\right] - \mathbf{m}\left[(m^* - m)^2\right] =$$

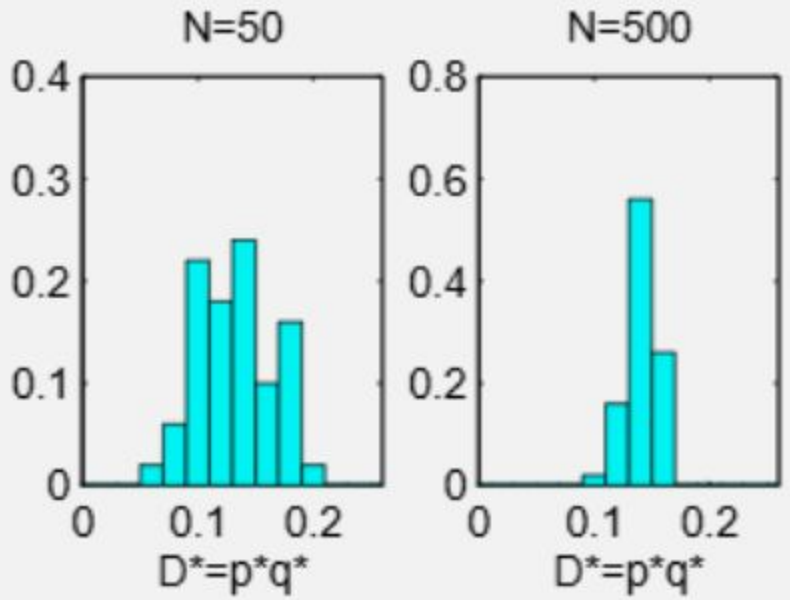
$$= D - \frac{1}{N}D = \boxed{\frac{N-1}{N}D}$$

Varianza muestral (D*): ejemplo

$$K = \begin{cases} 1 & \text{si sale un 5} \\ 0 & \text{si no} \end{cases}$$

$$p = \frac{1}{6} \qquad q = \frac{5}{6}$$

$$D = pq = 0.139$$



Muestra $\longrightarrow$ Población

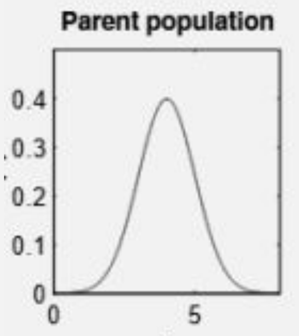
$N \to \infty$

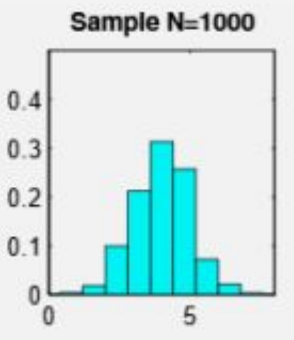
03

Estimación de parámetros

¿Cómo puedo conseguir la mayor cantidad de información sobre la población, solo estudiando la muestra?

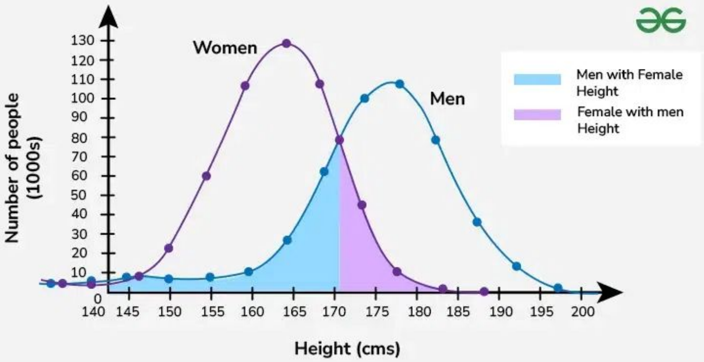
Ejemplos





$\lambda_1 = m$
$\lambda_2 = D$

$\lambda^*_1 = m^*$
$\lambda^*_2 = D^*$



$\lambda_1 = m_1$ $\lambda_x = \ldots$
$\lambda_2 = D_1$

$\lambda^*_x = ???$

Estimadores

¿Qué son?

Una función que se aplica a una muestra de datos para aproximar un parámetro desconocido de una población.

Propiedades

01

Consistente

$\lim_{N\to\infty} \tilde{\lambda} = \lambda$

02

Insesgado

$\mathbf{m}[\tilde{\lambda}] = \lambda$

03

Efectivo

$\mathbf{D}[\tilde{\lambda}]$ es mínima

Media muestral (m*)

$$\tilde{m} \;=\; m^* \;=\; \sum_i x_i/N\ .$$

$$\begin{aligned}
\mathbf{m}[m^*] \;&=\; \mathbf{m}\left[\frac{X_1 + X_2 + \cdots + X_N}{N}\right] \\
&= \frac{1}{N}\,\{\mathbf{m}[X_1] + \mathbf{m}[X_2] + \cdots + \mathbf{m}[X_N]\} \\
&= \;\frac{1}{N}\,\{N\,m\} \;=\; m\ .
\end{aligned}$$

Varianza muestral (D*)

$$D^* = \frac{1}{N}\sum_{i=1}^{N}(x_i - m^*)^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - m)^2 - (m^* - m)^2$$

$$\mathbf{m}[D^*] = \mathbf{m}\left[\frac{1}{N}\sum_{i=1}^{N}(x_i - m)^2\right] - \mathbf{m}\left[(m^* - m)^2\right] =$$

$$= D - \frac{1}{N}D = \boxed{\frac{N-1}{N}D}$$

Criterio de máxima verosimilitud

$$g(x_1, x_2, \ldots, x_N) \;=\; f(x_1)\, f(x_2) \cdots f(x_N)\,.$$

Función de verosimilitud

$$\mathbf{max}\;\; [g(x_1, x_2, \ldots, x_N; \lambda_1, \lambda_2, \ldots, \lambda_n)] \;\Rightarrow\; \tilde{\lambda}_1, \tilde{\lambda}_2, \ldots, \tilde{\lambda}_n\,.$$

Ejemplo

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-m)^2}{2\sigma^2}\right]$$

$$g(x_1, x_2, \ldots, x_N; m, \sigma) = \frac{1}{\sigma^N(\sqrt{2\pi})^N} \exp\left[-\sum_{i=1}^{N} \frac{(x_i-m)^2}{2\sigma^2}\right]$$

$$\tilde{m} = \frac{1}{N}\sum\nolimits_i x_i = m^*$$

$$\tilde{\sigma} = \sqrt{\frac{1}{N}\sum\nolimits_i (x_i - m)^2}$$

Gracias

Espacio de preguntas