

1 INFERENCIA ESTADÍSTICA

Objetivo: Obtener conclusiones válidas para toda la población a partir del estudio de una muestra.

¿Cómo?: Mediante los métodos de estimación puntual y de intervalos de confianza

2 ESTIMACIÓN PUNTUAL

Se recurre a cálculos con los datos de la muestra para inferir algunos aspectos de la población. Al cálculo realizado se le llama estadístico.

Definición: Dado λ , un parámetro de la población. Decimos que el estadístico $\hat{\lambda}$ es un estimador suyo si es un parámetro obtenido a partir de una muestra y cuyo es asignar un valor aproximado a λ .

La estimación puntual sirve de poco mientras desconozcamos cuál es el grado de aproximación del estimador al parámetro real. Por ese motivo se procede a la estimación mediante un intervalo.

3 ESTIMACIÓN POR INTERVALOS DE CONFIANZA

A partir de una muestra aleatoria de tamaño n , podemos estimar el valor de un parámetro de la población del siguiente modo:

- Dando un intervalo dentro del cual confiamos que esté el parámetro. Se llama **intervalo de confianza**.
- Hallando la probabilidad de que tal cosa ocurra. A dicha probabilidad se la llama **nivel de confianza**.

3.1. Estimación de la media por intervalo de confianza

Se desea estimar la media, μ , de una población cuya desviación típica, σ , es conocida. Para ello, se recurre a una muestra de tamaño n de la cual se obtiene la media, \bar{x} (media muestral). Si la población de partida es normal, o si el tamaño de la muestra es $n \geq 30$, el intervalo de confianza de μ , con un nivel de confianza de $(1 - \alpha) \cdot 100\%$, es:

$$\left(\bar{x} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \right)$$

siendo:

- \bar{x} : La media de los datos de la muestra
- $z_{\alpha/2}$: El valor de la distribución normal $Z \rightsquigarrow N((0, 1))$
- σ : La desviación típica de la distribución de la población (o si no se conoce de un estimador sesgado de la misma)
- n : El tamaño de la muestra

Error máximo cometido:

$$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

Es el radio del entorno dado por el intervalo de confianza. **NOTA:** Disminuye al aumentar el tamaño de la muestra y por tanto si se quiere garantizar un error determinado para un nivel de confianza habrá que tomar muestras de al menos un **tamaño determinado de la muestra**.

Ejemplo 2: Se desea realizar una investigación para estimar el peso medio de los recién nacidos de madres fumadoras. Se admite un error máximo de 50 gramos, con una confianza del 95 %. Si por estudios anteriores se sabe que la desviación típica del peso medio de tales recién nacidos es de 400 gramos, ¿qué tamaño mínimo de muestra se necesita en la investigación?

Solución: $E = 50$, Confianza=95 y $\sigma = 400$
 $\alpha = 1 - 0,95 = 0,05 \rightarrow \frac{\alpha}{2} = \frac{0,05}{2} = 0,025$.

Por tanto, el valor crítico será:

$$P(Z \leq z_{\alpha/2}) = 0,95 + 0,025 = 0,975 \rightarrow z_{\alpha/2} = 1,96$$

A partir de la definición de error máximo admitido:

$$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \rightarrow n = \left(\frac{z_{\alpha/2} \cdot \sigma}{E} \right)^2$$

Luego:

$$n = \left(\frac{1,96 \cdot 400}{50} \right)^2 \approx 245,8534 \rightarrow n = 246$$

3.2. Estimación de la proporción por intervalo de confianza

Se desea estimar la proporción, p , que una determinada característica se cumple en una población. Para ello, se recurre a una muestra de tamaño n de la cual se calcula la \bar{p} (proporción muestral). Si la población de partida es normal, o si el tamaño de la muestra es $n \geq 30$, el intervalo de confianza de p , con un nivel de confianza de $(1 - \alpha) \cdot 100\%$, es:

$$\left(\bar{p} - z_{\alpha/2} \cdot \sqrt{\frac{\bar{p} \cdot (1 - \bar{p})}{n}}, \bar{p} + z_{\alpha/2} \cdot \sqrt{\frac{\bar{p} \cdot (1 - \bar{p})}{n}} \right)$$

En este caso el error máximo admitido es:

$$E = z_{\alpha/2} \cdot \sqrt{\frac{\bar{p} \cdot (1 - \bar{p})}{n}}$$

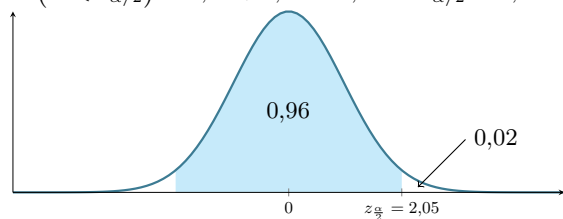
Ejemplo: Para estimar la proporción de personas con sobrepeso en una población se ha tomado una muestra aleatoria simple de tamaño 100 personas, de las cuales 21 tienen sobrepeso. Calcular el intervalo de confianza al 96 % para la proporción de personas con sobrepeso en la población.

Solución: $\bar{p} = 0,21$, Confianza=96

$$\alpha = 1 - 0,96 = 0,04 \rightarrow \frac{\alpha}{2} = \frac{0,04}{2} = 0,02 .$$

Por tanto, el valor crítico será:

$$P(Z \leq z_{\alpha/2}) = 0,96 + 0,02 = 0,98 \rightarrow z_{\alpha/2} \approx 2,05$$



A partir de la definición de error máximo admitido:

$$E = z_{\alpha/2} \cdot \sqrt{\frac{\bar{p} \cdot (1 - \bar{p})}{n}} \approx 2,05 \cdot \sqrt{\frac{0,21 \cdot 0,79}{100}} \approx 0,08$$

Luego el intervalo es:

$$(0,21 - 0,08, 0,21 + 0,08) = (0,13, 0,29)$$

Por tanto con una alta probabilidad, en concreto 0.96, el porcentaje de individuos con sobrepeso en la población se encuentra entre el 13 % y el 29 %