

Lección 2: Intervalos de confianza

Módulo 3: Inferencia Estadística

Magdalena Cornejo

Precisión del estimador puntual

- Si la media muestral (\bar{X}) se usa para estimar a la media poblacional μ que no conocemos, ¿cuál es la probabilidad de acertarle?

Ejemplo

Si quiero conocer cuál es el nivel medio de consumo de cerveza anual per capita en cierto país y tomo una muestra aleatorio de 100 individuos que me arroja un $\bar{X} = 80,2$, ¿cuál es la probabilidad de que μ sea efectivamente 80,2?

Precisión del estimador puntual

- Si la media muestral (\bar{X}) se usa para estimar a la media poblacional μ que no conocemos, ¿cuál es la probabilidad de acertarle?

Ejemplo

Si quiero conocer cuál es el nivel medio de consumo de cerveza anual per capita en cierto país y tomo una muestra aleatorio de 100 individuos que me arroja un $\bar{X} = 80,2$, ¿cuál es la probabilidad de que μ sea efectivamente 80,2?

- Prácticamente ¡CERO!
- ¿Puedo saber “más o menos” en qué rango se encontraría el verdadero valor de la media?
- Sí: a través de la construcción de **intervalos de confianza**.

Motivación

- La **estimación por intervalo** es generalmente preferida a la estimación puntual ya que esta última no provee información respecto al error en la estimación.
- Veremos cómo construir intervalos de confianza para la media (μ) y para la proporción (p).

Definición

Un **intervalo de confianza** es un rango de valores que se determina en base a información muestral, en el cual es probable que el parámetro poblacional esté contenido.

Estructura de los intervalos de confianza

Estimador Puntual \pm Margen de Error

Los niveles de confianza más usados son 90%, 95% y 99%.

Veremos intervalos de confianza para:

- La media poblacional (μ):

$$\bar{X} \pm \text{margen de error}$$

- La proporción poblacional (p):

$$\hat{p} \pm \text{margen de error}$$

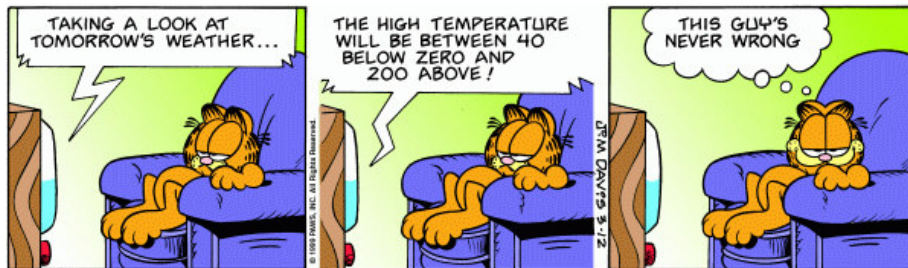
Intervalos de confianza para la media (μ)

$$P(l_1 < \mu < l_2) = 1 - \alpha$$

El intervalo me dice entre qué valores (l_1 y l_2) se encontraría el verdadero valor del parámetro μ con una probabilidad de $1 - \alpha$.

- $1 - \alpha$: **nivel de confianza** (usualmente 90%, 95% o 99%)
- α : **nivel de significación** (usualmente 10%, 5% o 1%)

Intervalos de confianza



Intervalos de confianza para la media (μ)

Hay dos casos posibles:

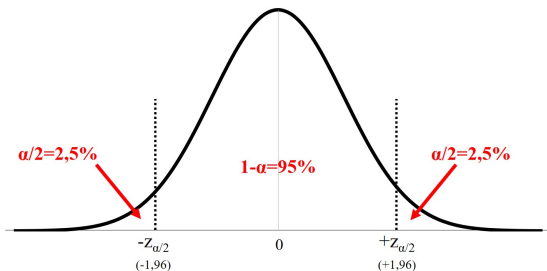
- **Caso 1.** La varianza poblacional (σ^2) es **conocida**.
- **Caso 2.** La varianza poblacional (σ^2) es **desconocida**.

Caso 1

Supone que σ^2 es **conocida**.

$$\bar{X} \pm z_{\alpha/2} \sigma / \sqrt{n}$$

Si queremos construir un intervalo de confianza al 95%:



Este caso raramente sucede, cuando desconocemos el valor de σ^2 es razonable reemplazando por la varianza muestral S^2 .

Caso 2

- Es un caso más realista (no conocemos σ^2).
- **No** usamos una distribución normal estándar, sino

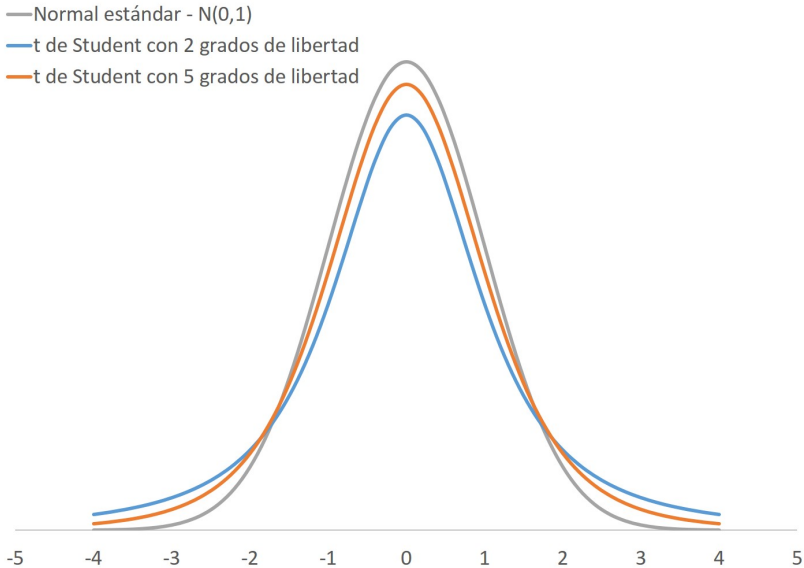
$$t = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

otra distribución que se denomina **t de Student con (n-1) grados de libertad**.

- Noten que \bar{X} está estandarizada usando S en lugar de σ .
- El intervalo de confianza correspondiente será:

$$\bar{X} \pm t_{n-1, \alpha/2} S/\sqrt{n}$$

Distribución t de Student



Caso 1 (Z): con varianza poblacional conocida

Ejemplo

Se conoce, de estudios anteriores, que el costo variable de construcción de determinado tipo de vivienda prefabricada, por metro cuadrado, se distribuye normalmente con un desvío estándar de \$135. Se tomó una muestra aleatoria de 12 viviendas con las que se calculó un costo variable promedio de \$1440.

¿Entre qué valores estará el costo variable medio de la construcción de dicho tipo de vivienda si se lo estima con una confianza del 95%?

Caso 1 (Z): con varianza poblacional conocida

Ejemplo

Se conoce, de estudios anteriores, que el costo variable de construcción de determinado tipo de vivienda prefabricada, por metro cuadrado, se distribuye normalmente con un desvío estándar de \$135. Se tomó una muestra aleatoria de 12 viviendas con las que se calculó un costo variable promedio de \$1440.

¿Entre qué valores estará el costo variable medio de la construcción de dicho tipo de vivienda si se lo estima con una confianza del 95%?

Solución:

Sabemos que $\sigma = 135$, $n = 12$, $\bar{X} = 1440$ y $1 - \alpha = 0.95$. Entonces,

$$\begin{aligned} \bar{X} \pm z_{\alpha/2} \sigma / \sqrt{n} \\ 1440 \pm 1.96 \times 135 / \sqrt{12} \end{aligned}$$

Por lo que el intervalo de confianza al 95% resulta: (\$1363,6; \$1516,4).

Caso 2 (T): con varianza poblacional desconocida

Aplicación en Excel:

- Vuelva a trabajar sobre la base de datos de Excel que utilizó en el Módulo 1 en la cual calculó el retorno diario de IBM entre el 4 de enero de 2016 y 22 de julio de 2016.
- Utilice nuevamente el complemento de Excel “Herramientas para análisis”, pero esta vez calcule el intervalo de confianza al 95% del retorno medio diario de IBM.
- ¿Y cuál es el intervalo de confianza al 99%?

Intervalos de confianza para la proporción (p)

Si ahora estamos interesados en construir un intervalo de confianza para la proporción poblacional (p), el intervalo de confianza asociado será:

$$\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}$$

Ejemplo

Una encuesta realizada a cabo en el mes pasado en una ciudad, se entrevistó a 871 adultos. Con respecto a una pregunta dada, se determinó que 53% de los entrevistados tiene una imagen positiva del gobierno de turno. Calcular el intervalo de confianza al 95% para la proporción de todos los adultos de la ciudad que tienen una imagen positiva del gobierno actual.

Solución

Del ejemplo anterior sabemos que $n = 871$, $\hat{p} = 0.53$ y $1 - \alpha = 0.95$.
Reemplazando en el intervalo de confianza de la proporción:

$$\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}$$
$$0.53 \pm 1.96 \times \sqrt{\frac{0.53 \times 0.47}{871}}$$

Entonces, el intervalo de confianza al 95% resulta: (52,94%; 53,06%).

Para pensar...

¿Qué pasará con el **margen de error** (aumentará o disminuirá) si...

- aumento el tamaño de la muestra (n)?

Para pensar...

¿Qué pasará con el **margen de error** (aumentará o disminuirá) si...

- aumento el tamaño de la muestra (n)?
- tomo otra muestra de igual tamaño, pero con mayor dispersión (S)?

Para pensar...

¿Qué pasará con el **margen de error** (aumentará o disminuirá) si...

- aumento el tamaño de la muestra (n)?
- tomo otra muestra de igual tamaño, pero con mayor dispersión (S)?
- trabajo con un mayor nivel de confianza ($1 - \alpha$)?