

# Estadística II

## Tema 5. Estimación por intervalos

Facultad de Ciencias de la Economía y de la Empresa (FCEE)

Curso 2025–2026 · 4.5 ECTS · 2º cuatrimestre

**Francisco Rabadán Pérez · Raquel Ibar Alonso · Ester Muñoz Céspedes**

Departamento de Economía Aplicada I e Historia e Instituciones Económicas

# Tema 5 · Objetivo

## Al terminar este tema podrás:

- explicar **qué es un intervalo de confianza** y cuál es su interpretación correcta;
- distinguir entre intervalos **bilaterales y unilaterales**;
- construir intervalos de confianza en los **casos más habituales**, indicando los supuestos necesarios: para  $\mu$  en  $N(\mu, \sigma)$  con  $\sigma$  **conocida**; para  $\mu$  en  $N(\mu, \sigma)$  con  $\sigma$  **desconocida**; para  $\sigma^2$  (y  $\sigma$ ) en  $N(\mu, \sigma)$ ; para una **proporción  $\pi$**  en grandes muestras;
- **elegir el intervalo adecuado** según el parámetro, los supuestos y el tamaño muestral;
- **determinar el tamaño muestral necesario** para alcanzar un error máximo prefijado.

### Idea clave

Un intervalo de confianza aporta una estimación no puntual y, además, una medida de precisión.

## Cómo leer estas etiquetas

CORE

PLUS

ANEXO

**CORE:** ideas y procedimientos que sostienen el curso entero y que se practican de forma recurrente.

**PLUS:** material para profundizar o reforzar intuición; ayuda a consolidar.

**ANEXO:** material de referencia o consulta.

# Tema 5· Esquema

1. **Introducción e interpretación de los intervalos de confianza**
2. **Métodos de construcción de intervalos de confianza**
3. **Intervalos de confianza de longitud mínima**
4. **Intervalos de confianza en poblaciones normales**
5. **Intervalos de confianza en poblaciones no normales**
6. **Determinación del tamaño muestral**
7. **Síntesis y guía de aplicación (elección del intervalo)**

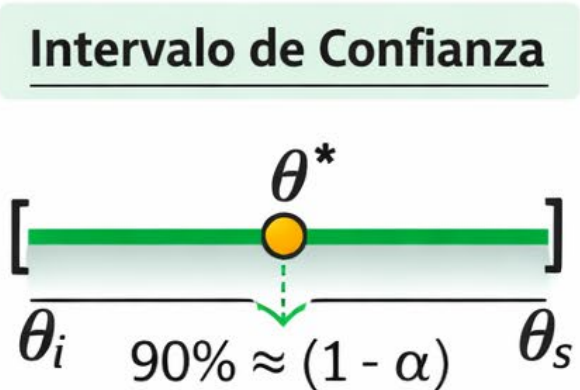
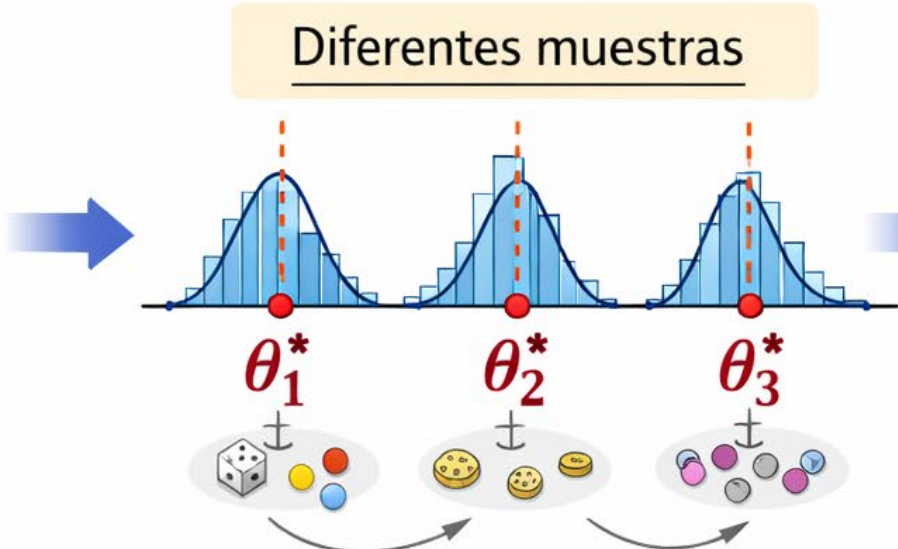
# Breviario

- **Nivel de confianza  $\gamma$** : probabilidad (antes de muestrear) de que el intervalo construido contenga al parámetro.
- **Relación**:  $\gamma = 1 - \alpha$ .
- **Significatividad ( $\alpha$ )**: probabilidad de que el intervalo no contenga al verdadero parámetro poblacional (quedarse “fuera” del I.C.).
- **Extremos del I.C.**:  $\theta_i(X)$  y  $\theta_s(X)$ . **Intervalo de confianza**:  $[\theta_i(X); \theta_s(X)]$ .
- **Amplitud (longitud) del I.C.**:  $D = \theta_s(X) - \theta_i(X)$ .
- En intervalos bilaterales simétricos, la **semi-amplitud (error máximo)** es  $e = D/2$ .
- **Intervalo bilateral vs unilateral**: bilateral acota por abajo y por arriba; unilateral es una sola cota:  $[\theta_i(X), +\infty)$  o  $(-\infty, \theta_s(X)]$ .
- **Colas iguales ( $\alpha/2$  y  $\alpha/2$ )**: **convenio estándar** en intervalos bilaterales.
- **Garantía mínima (Chebyshev)**: intervalos que contienen al parámetro poblacional con probabilidad al menos  $\gamma$ , sin suponer normalidad.

# ¿Por qué no basta con una estimación puntual?

## Idea:

- 1. Una **estimación puntual** resume la información de la muestra en un único número:  $\hat{\theta}^*$ .
- 2. Si **repetiésemos el muestreo**, obtendríamos otro valor:  $\hat{\theta}^*$  **cambia de una muestra a otra**.
- 3. **Necesitamos cuantificar la precisión**: construir un intervalo  $[\theta_i(X), \theta_s(X)]$  que contenga a  $\theta$  con una alta probabilidad  $(1 - \alpha) = \gamma$ .



✓ Un rango plausible para  $\theta$

? + Precisión

# 1. Introducción: estimación por intervalos

# 1. Introducción: estimación por intervalos

- Partimos de un **parámetro desconocido**  $\theta$  (p. ej.,  $\mu, \sigma^2, \pi$ ) y de una muestra aleatoria simple  $X$ .
- La estimación puntual  $\hat{\theta}$  es un **estadístico**: la estimación  $\hat{\theta}^*$  cambia de muestra a muestra.
- Queremos un procedimiento que, para cada muestra, produzca un **intervalo aleatorio**  $[\theta_i(X), \theta_s(X)]$  que contenga a  $\theta$  con alta probabilidad.
- Definimos:

$$\gamma = 1 - \alpha$$

Donde  $\alpha$  es la **probabilidad de que el intervalo no contenga a  $\theta$**  (es decir, “quedarse fuera del I.C.”).

- **Definición del I.C. :**

$$P(\theta_i(X) \leq \theta \leq \theta_s(X)) = \gamma$$

- **Lectura correcta:**

- **antes de muestrear**, el intervalo tiene **probabilidad  $\gamma$**  de contener a  $\theta$ ;
- **una vez observada** la muestra, obtenemos el **intervalo concreto para esa muestra**  $[\theta_i^*(X), \theta_s^*(X)]$ .

# 1.1. Interpretación correcta de un intervalo de confianza

- Un intervalo de confianza es un **intervalo aleatorio**  $[\theta_i(X), \theta_s(X)]$  construido a **partir de una muestra aleatoria simple**  $X$ .
- La afirmación :

$$P(\theta_i(X) \leq \theta \leq \theta_s(X)) = \gamma$$

**se interpreta antes de observar la muestra:** si repitiésemos el muestreo muchas veces, una fracción  $\gamma$  de los intervalos construidos contendría al verdadero  $\theta$ .

- **una vez observada** la muestra, el **intervalo numérico**  $[\theta_i^*(X), \theta_s^*(X)]$  queda fijado, entonces:
  - no tiene sentido hablar de la probabilidad  $\gamma$  de contener al verdadero valor del parámetro,
  - lo correcto es: “este método produce intervalos que contienen a  $\theta$  con probabilidad  $\gamma$ ”.

## Error típico a evitar

- **Incorrecto:** “la probabilidad de que  $\theta$  esté en  $[\theta_i(X), \theta_s(X)]$  es  $\gamma$ ”.

# 1.2 Intervalos bilaterales y unilaterales

$\alpha$  es la probabilidad de que el intervalo no contenga a  $\theta$ .

- Distinguimos:

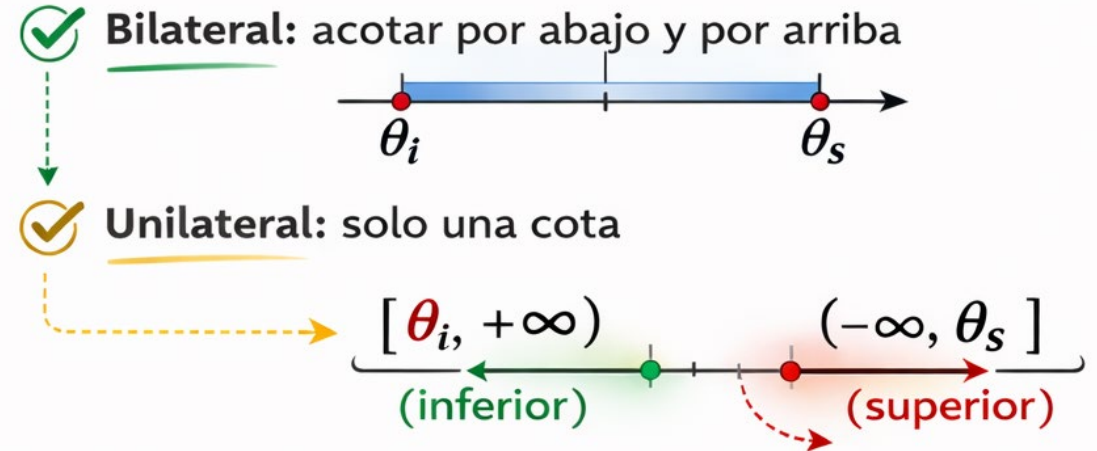
## 1) Intervalo bilateral (dos colas)

- Idea práctica*: queremos acotar  $\theta$  “por abajo y por arriba”.
- En el caso bilateral estándar (simétrico):  $\alpha/2$  en cada cola.

## 2) Intervalo unilateral (una cola)

- Idea práctica*: solo nos interesa garantizar una cota (por ejemplo, “al menos” o “como mucho”).
  - Cota inferior**:  $[\theta_i(X), +\infty)$
  - Cota superior**:  $(-\infty, \theta_s(X)]$

 **Idea práctica**



# 1.2 ¿Cuándo usar un I.C. bilateral o unilateral?

1) **Intervalo bilateral** (dos colas) si queremos acotar  $\theta$  “por abajo y por arriba”.

Ejemplos típicos:

- estimar un coste medio  $\mu$  con margen de error,
- estimar una proporción  $\pi$  (tasa de paro, tasa de impago),
- estimar la variabilidad  $\sigma^2$ .

• **Intervalo unilateral** (una cola) si solo interesa una cota

- **Cota inferior:**  $[\theta_i(X), +\infty]$ : “garantizar que  $\theta$  es al menos...”
- **Cota superior:**  $[-\infty, \theta_s(X)]$ : “asegurar que  $\theta$  no supera...”

• **Regla práctica:**

- bilateral = acotación por ambos lados,
- unilateral = “cumplir un requisito” (mínimo o máximo).



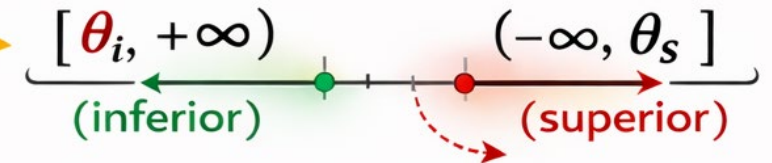
**Idea práctica**



**Bilateral:** acotar por abajo y por arriba



**Unilateral:** solo una cota



# Resumen de la Sección 1 (ideas clave)

1. **Verdadero / Falso:** “La probabilidad de que  $\theta$  esté dentro de  $[\theta_i(X), \theta_s(X)]$  es  $\gamma$ ”
2. **¿Qué representa  $\alpha$ ?**
  - a) La probabilidad total de que el intervalo no contenga a  $\theta$
  - b) La probabilidad de que  $\theta$  esté dentro del intervalo observado
  - c) La longitud media del intervalo
  - d) La varianza del estimador  $\hat{\theta}$
3. **¿Cuándo es más adecuado un intervalo unilateral (una cola)?**
  - a) Estimar un salario medio con “margen” por arriba y por abajo
  - b) Estimar la dispersión  $\sigma^2$  por arriba y por abajo
  - c) Garantizar que la tasa de defectos es como mucho un 2%
  - d) Dar un rango plausible para la tasa de paro
4. **En un intervalo bilateral estándar, ¿cómo se reparte la probabilidad fuera del intervalo?**
  - a)  $\alpha$  en una sola cola
  - b)  $\alpha/2$  en cada cola
  - c)  $\gamma/2$  en cada cola
  - d) Depende del tamaño muestral, no se puede fijar

# Resumen de la Sección 1 (ideas clave)

1. **Verdadero / Falso:** “La probabilidad de que  $\theta$  esté dentro de  $[\theta_i(X), \theta_s(X)]$  es  $\gamma$ ” ❌
2. **¿Qué representa  $\alpha$ ?**
  - a) La probabilidad total de que el intervalo no contenga a  $\theta$  ✅
  - b) La probabilidad de que  $\theta$  esté dentro del intervalo observado
  - c) La longitud media del intervalo
  - d) La varianza del estimador  $\hat{\theta}$
3. **¿Cuándo es más adecuado un intervalo unilateral (una cola)?**
  - a) Estimar un salario medio con “margen” por arriba y por abajo
  - b) Estimar la dispersión  $\sigma^2$  por arriba y por abajo
  - c) Garantizar que la tasa de defectos es como mucho un 2% ✅
  - d) Dar un rango plausible para la tasa de paro
4. **En un intervalo bilateral estándar, ¿cómo se reparte la probabilidad fuera del intervalo?**
  - a)  $\alpha$  en una sola cola
  - b)  $\alpha/2$  en cada cola ✅
  - c)  $\gamma/2$  en cada cola
  - d) Depende del tamaño muestral, no se puede fijar

## **2. Métodos de construcción de intervalos de confianza**

## 2.1 Métodos de construcción de intervalos de confianza

**Objetivo:** Construir un **intervalo aleatorio**  $[\theta_i(X), \theta_s(X)]$  con nivel  $\gamma = 1 - \alpha$  :

$$P(\theta_i(X) \leq \theta \leq \theta_s(X)) = \gamma$$

**Dos enfoques** (mismo objetivo, distinto punto de partida):

### 1. Método de la cantidad pivotal

- se construye un **estadístico**  $T(X; \theta)$  tal que su **distribución no depende de  $\theta$** ;
- Se eligen cuantiles  $K_1$  y  $K_2$  (tablas) y se plantea

$$P(K_1 \leq T(X; \theta) \leq K_2) = 1 - \alpha$$

- Se despeja  $\theta \rightarrow P(\theta_i(X) \leq \theta \leq \theta_s(X)) = 1 - \alpha \rightarrow [\theta_i(X), \theta_s(X)]$

### 2. Método general.

- se parte de la **distribución muestral del estimador**  $\hat{\theta} = \hat{\theta}(X)$  que está parametrizado por  $\theta$ .
- se buscan **límites**  $K_1(\alpha, \theta)$ ,  $K_2(\alpha, \theta)$  para  $\hat{\theta}$  con probabilidad  $\gamma$ .
- Se invierte el evento para acotar  $\theta \rightarrow [\theta_i(X), \theta_s(X)]$

- **Idea operativa común:** “estandarizar  $\rightarrow$  usar tablas ( $\alpha$  o  $\alpha/2$ )  $\rightarrow$  despejar/invertir”

## 2.2 Método de la cantidad pivotal

- Buscamos una cantidad pivotal  $T(X;\theta)$  cuya distribución no depende de  $\theta$ .
- Elegimos **valores críticos**  $K_1, K_2$  (a partir de  $\alpha$ ;  $\gamma$ , si aplica, de g.l.).

- **Planteamos:**

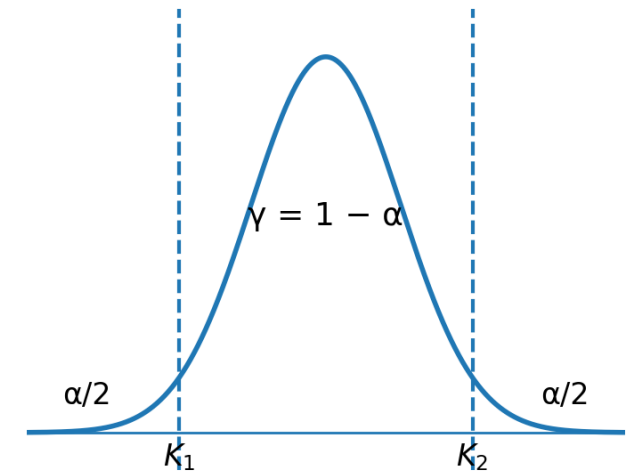
$$P(K_1 \leq T(X; \theta) \leq K_2) = \gamma \quad (\gamma = 1 - \alpha)$$

- Como  $T(X;\theta)$  depende de  $\theta$ , **se despeja  $\theta$**  en la doble desigualdad para obtener

$$P(\theta_i(X) \leq \theta \leq \theta_s(X)) = 1 - \alpha$$

- **Idea clave:** “pivote  $\rightarrow$  cuantiles tabulados  $\rightarrow$  despejar  $\theta$ ”.

Método de la cantidad pivotal (esquema)



$$P(K_1 \leq T(X; \theta) \leq K_2) = \gamma$$

1) Elegir una cantidad pivotal  $T(X; \theta)$  con distribución conocida

2) Fijar valores críticos con tablas ( $K_1, K_2$ ) a partir de  $\alpha$  (o  $\alpha/2$ )

3) Plantear la doble desigualdad  $K_1 \leq T(X; \theta) \leq K_2$

4) Despejar el parámetro  $\theta_i(X) \leq \theta \leq \theta_s(X)$

## 2.3 Método general de construcción

- Partimos del **estimador**  $\hat{\theta}$  y de su **distribución muestral**, que **suele depender** del parámetro  $\theta$ .
- Buscamos **dos funciones**  $K_1(\alpha; \theta)$ ,  $K_2(\alpha; \theta)$  tales que:

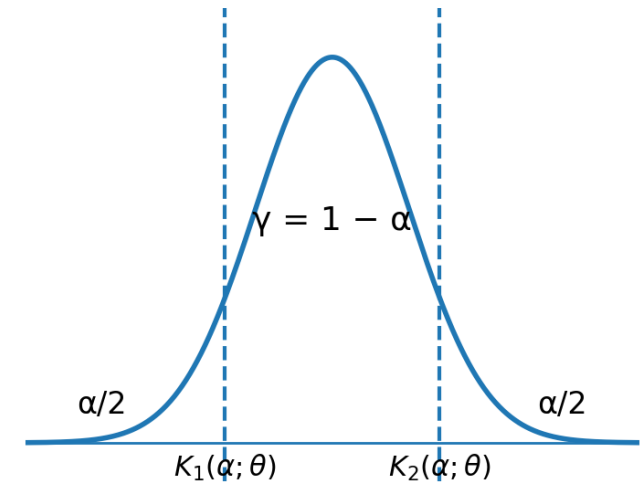
$$P[K_1(\alpha; \theta) \leq \hat{\theta} \leq K_2(\alpha; \theta)] = \gamma \quad (\gamma = 1 - \alpha)$$

- Como los límites dependen de  $\theta$ , **se invierte** el evento (se despeja  $\theta$ ) para obtener:

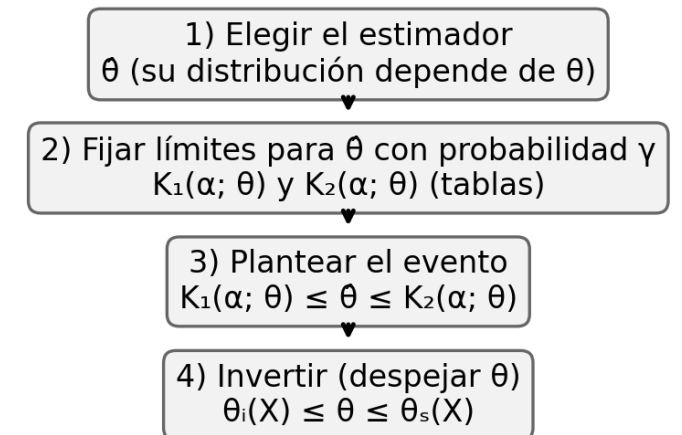
$$\theta_i(X) \leq \theta \leq \theta_s(X)$$

- **Idea clave:** “distribución de  $\hat{\theta} \rightarrow$  acotar  $\hat{\theta}$  con probabilidad  $\gamma \rightarrow$  invertir para acotar  $\theta$ ”.

Método general (esquema)



$$P(K_1(\alpha; \theta) \leq \theta \leq K_2(\alpha; \theta)) = \gamma \rightarrow \text{invertir para acotar } \theta$$



## 2.4 Nota importante: valores críticos y colas

- En ambos métodos (pivotal y general) aparecen valores críticos  $K_1$  y  $K_2$  fijados por  $\alpha$ .
- En un intervalo de confianza,  $\alpha$  mide **la probabilidad de que el intervalo no contenga a  $\theta$** .
  
- **Caso bilateral (dos colas, estándar):**
  - $\alpha$  se reparte a ambos lados de los extremos del intervalo.
  - Si la distribución es simétrica  $\alpha/2$  a cada lado,
  
- **Caso unilateral:**
  - $\alpha$  se concentra en **una sola cola**,
  - se busca **un único valor crítico**.

## 2.5 Mini-test (Métodos de construcción)

**1. Una cantidad pivotal  $T(X;\theta)$  se caracteriza porque:**

- a) depende de  $\theta$  y su distribución también depende de  $\theta$
- b) depende de  $\theta$  pero su distribución no depende de  $\theta$
- c) no depende de  $\theta$  y su distribución depende de  $\theta$
- d) no depende de  $\theta$  y su distribución no depende de  $\theta$

**2. En el método general, el punto de partida es:**

- a) una tabla de cuantiles sin definir un estimador
- b) una función de verosimilitud  $L(\theta)$
- c) la distribución muestral del estimador  $\hat{\theta}$
- d) el reparto  $\alpha/2$  sin más información

**3. En un I.C. bilateral estándar, los valores críticos se eligen dejando**

- a)  $\alpha$  en cada cola
- b)  $\alpha/2$  en cada cola
- c)  $\gamma/2$  en cada cola
- d)  $\gamma$  en una cola

**4. ¿Qué significa “invertir” en el método general?**

- a) Cambiar  $K_1$  por  $K_2$
- b) Sustituir  $\theta$  por  $\hat{\theta}^*$
- c) Pasar de una desigualdad sobre  $\hat{\theta}$  a una desigualdad sobre  $\theta$
- d) Cambiar bilateral por unilateral

## 2.5 Mini-test (Métodos de construcción)

1. Una cantidad pivotal  $T(X;\theta)$  se caracteriza porque:

- a) depende de  $\theta$  y su distribución también depende de  $\theta$
- b) depende de  $\theta$  pero su distribución no depende de  $\theta$  ✓
- c) no depende de  $\theta$  y su distribución depende de  $\theta$
- d) no depende de  $\theta$  y su distribución no depende de  $\theta$

2. En el método general, el punto de partida es:

- a) una tabla de cuantiles sin definir un estimador
- b) una función de verosimilitud  $L(\theta)$
- c) la distribución muestral del estimador  $\hat{\theta}$  ✓
- d) el reparto  $\alpha/2$  sin más información

3. En un I.C. bilateral estándar, los valores críticos se eligen dejando

- a)  $\alpha$  en cada cola
- b)  $\alpha/2$  en cada cola ✓
- c)  $\gamma/2$  en cada cola
- d)  $\gamma$  en una cola

4. ¿Qué significa “invertir” en el método general?

- a) Cambiar  $K_1$  por  $K_2$
- b) Sustituir  $\theta$  por  $\hat{\theta}^*$
- c) Pasar de una desigualdad sobre  $\hat{\theta}$  a una desigualdad sobre  $\theta$  ✓
- d) Cambiar bilateral por unilateral

# **3. Intervalos de confianza de longitud mínima**

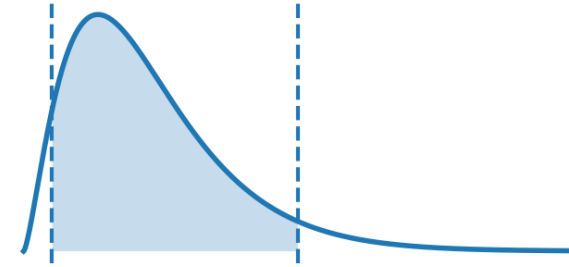
### 3. Intervalos de confianza de longitud mínima

- **Para un mismo nivel de confianza  $\gamma=1-\alpha$  puede haber más de un intervalo de confianza posible.**
  - En un intervalo **bilateral** (dos colas) suele haber una **familia de intervalos posibles**, según cómo se **reparta  $\alpha$**  en los extremos.
  - En un intervalo **unilateral** (una cola), una vez fijado el lado (inferior/superior), la construcción habitual queda mucho más determinada (**depende de un único cuantil**).
- **Entre todos los I.C. posibles** , elegimos **el más preciso**, es decir, el que **tenga menor amplitud**.
- Definimos la **amplitud** (longitud) **del intervalo**:
$$D = \theta_s(X) - \theta_i(X)$$
- **Objetivo** de esta sección:
  - **entender** por qué el intervalo no es único,
  - y justificar el **criterio de longitud mínima** (o el convenio bilateral estándar  $\alpha/2$  y  $\alpha/2$ ).

## 3.1 ¿Por qué no es único un I.C. bilateral?

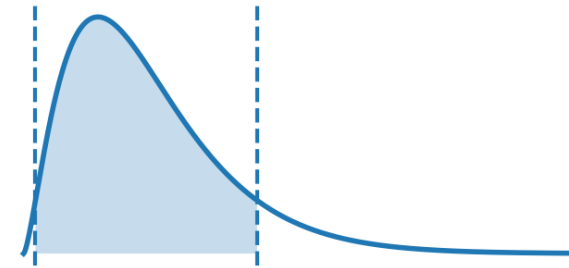
- En un **I.C. bilateral** dejamos fuera del intervalo una probabilidad  $\alpha$  (significatividad), **repartida entre ambos extremos**.
- Ese reparto **no tiene por qué ser simétrico**: podemos colocar más probabilidad en un lado y menos en el otro.
- Por eso **pueden existir muchos intervalos**  $[\theta_i(X), \theta_s(X)]$  con el mismo  $\gamma=1-\alpha$ , pero con distinta amplitud.
- Definimos la **amplitud (longitud) del intervalo**:
 
$$D = \theta_s(X) - \theta_i(X)$$
- **Objetivo** de esta sección:
  - **entender** por qué el intervalo no es único,
  - y justificar el **criterio de longitud mínima** o el **convenio bilateral estándar**  $\alpha/2$  y  $\alpha/2$ .

$\chi^2$ : mismos  $\gamma$ , distinta amplitud (D)  
IC colas iguales



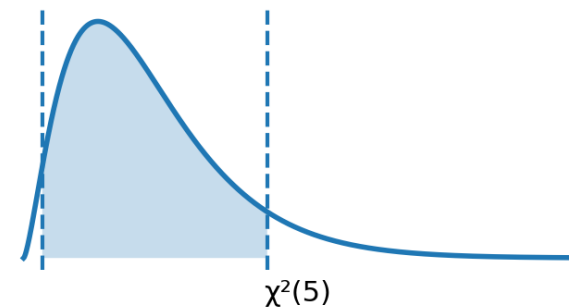
$P(a \leq X \leq b) = \gamma = 0.90 \mid a = 1.15, b = 11.07, D = 9.93$

IC longitud mínima



$P(a \leq X \leq b) = \gamma = 0.90 \mid a = 0.48, b = 9.43, D = 8.96$

IC asimétrico



$P(a \leq X \leq b) = \gamma = 0.90 \mid a = 0.75, b = 9.84, D = 9.08$

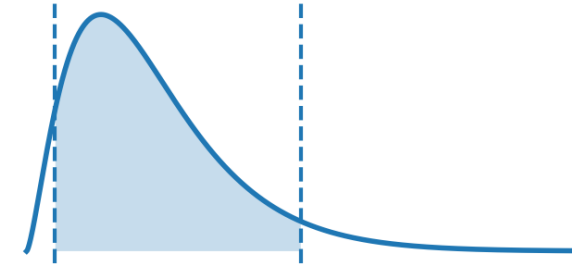
## 3.2 Criterio de longitud mínima

- Entre los intervalos bilaterales con el mismo nivel  $\gamma$ , elegimos el que tenga **menor amplitud**:

$$D = \theta_s(X) - \theta_i(X)$$

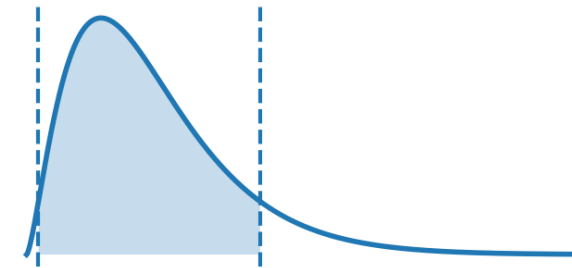
- **Idea:** el “mejor” intervalo es el que, manteniendo el mismo  $\alpha$ , tiene **menor amplitud**.
- **Nota práctica:**
  - En distribuciones **simétricas** (p. ej., Normal), el intervalo “**colas iguales**” suele **coincidir** con el de longitud mínima.
  - En distribuciones **asimétricas** (p. ej.,  $\chi^2$ ), **puede diferir**.

$\chi^2$ : mismos  $\gamma$ , distinta amplitud (D)  
IC colas iguales



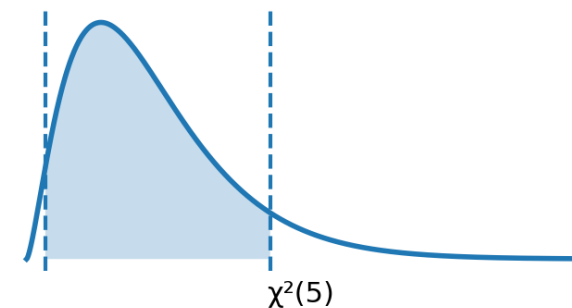
$P(a \leq X \leq b) = \gamma = 0.90 \mid a = 1.15, b = 11.07, D = 9.93$

IC longitud mínima



$P(a \leq X \leq b) = \gamma = 0.90 \mid a = 0.48, b = 9.43, D = 8.96$

IC asimétrico

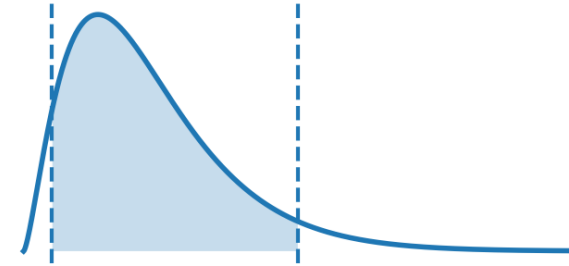


$P(a \leq X \leq b) = \gamma = 0.90 \mid a = 0.75, b = 9.84, D = 9.08$

## 3.3 Convenio: intervalo bilateral estándar

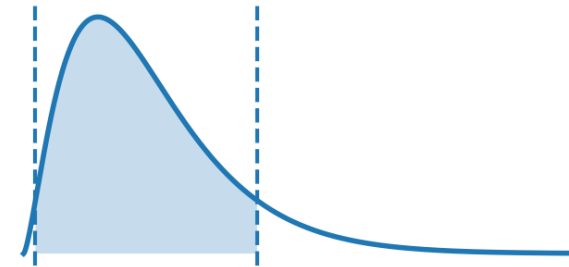
- En la **práctica**, cuando se pide un intervalo bilateral y no se especifica nada más, **se usa el intervalo de colas iguales**.
- **Significa** que la significatividad  $\alpha$  se reparte de forma **simétrica**:  $\alpha/2$  a cada lado.
- **Ventajas**:
  - es un **criterio simple y único** (evita arbitrariedad),
  - es **el más usado** en tablas y manuales,
  - en distribuciones **simétricas** coincide con el **intervalo central** (colas iguales).
- **Importante**: en distribuciones asimétricas (p. ej.,  $\chi^2$ ), este convenio no tiene por qué dar el intervalo de menor amplitud.

$\chi^2$ : mismos  $\gamma$ , distinta amplitud (D)  
IC colas iguales



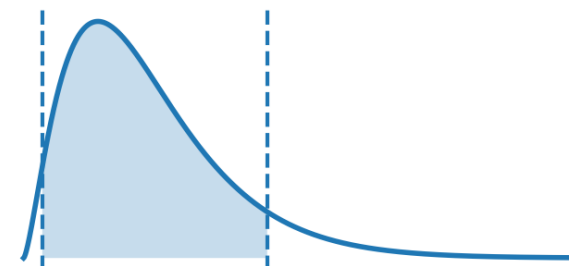
$P(a \leq X \leq b) = \gamma = 0.90 \mid a = 1.15, b = 11.07, D = 9.93$

IC longitud mínima



$P(a \leq X \leq b) = \gamma = 0.90 \mid a = 0.48, b = 9.43, D = 8.96$

IC asimétrico



$\chi^2(5)$

$P(a \leq X \leq b) = \gamma = 0.90 \mid a = 0.75, b = 9.84, D = 9.08$

# 3.T1 Mini-test (longitud mínima)

1. En un I.C. bilateral, ¿por qué puede no ser único el intervalo para un mismo  $\gamma$ ?

- a) Porque  $\gamma$  cambia con la muestra
- b) Porque el reparto de la probabilidad  $\alpha$  entre los extremos puede variar
- c) Porque  $\theta$  es aleatorio
- d) Porque siempre hay que usar  $\alpha/2$

2. La amplitud del intervalo se define como:

- a)  $D = \theta_i(X) + \theta_s(X)$
- b)  $D = \theta_s(X) - \theta_i(X)$
- c)  $D = \gamma - \alpha$
- d)  $D = \hat{\theta}^* - \theta$

3. En una distribución simétrica, el intervalo de colas iguales coincide con :

- a) el más amplio
- b) el intervalo central, y a menudo el menos amplio
- c) siempre asimétrico
- d) solo unilateral

4. En una distribución asimétrica (p. ej.,  $\chi^2$ ), el intervalo de colas iguales:

- a) siempre es el de longitud mínima
- b) nunca es válido
- c) puede no coincidir con el de longitud mínima
- d) obliga a usar un intervalo unilateral

# 3.T1 Mini-test (longitud mínima)

1. En un I.C. bilateral, ¿por qué puede no ser único el intervalo para un mismo  $\gamma$ ?

- a) Porque  $\gamma$  cambia con la muestra
- b) Porque el reparto de la probabilidad  $\alpha$  entre los extremos puede variar
- c) Porque  $\theta$  es aleatorio
- d) Porque siempre hay que usar  $\alpha/2$

2. La amplitud del intervalo se define como:

- a)  $D = \theta_i(X) + \theta_s(X)$
- b)  $D = \theta_s(X) - \theta_i(X)$
- c)  $D = \gamma - \alpha$
- d)  $D = \hat{\theta}^* - \theta$

3. En una distribución simétrica, el intervalo de colas iguales coincide con :

- a) el más amplio
- b) el intervalo central, y a menudo el menos amplio
- c) siempre asimétrico
- d) solo unilateral

4. En una distribución asimétrica (p. ej.,  $\chi^2$ ), el intervalo de colas iguales:

- a) siempre es el de longitud mínima
- b) nunca es válido
- c) puede no coincidir con el de longitud mínima
- d) obliga a usar un intervalo unilateral

# **4. Intervalos de confianza en poblaciones normales**

## 4.1 Caso $N(\mu, \sigma)$ : $\sigma^2$ conocida (I.C. para $\mu$ )

- **Supuesto:**  $x_1, \dots, x_n$  es m.a.s. de una población  $N(\mu, \sigma)$  y  $\sigma^2$  conocida.
- **Estimador:**  $\hat{\mu} = \bar{x}$

### 1. Pivote:

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

### 2. Intervalo teórico (colas iguales)

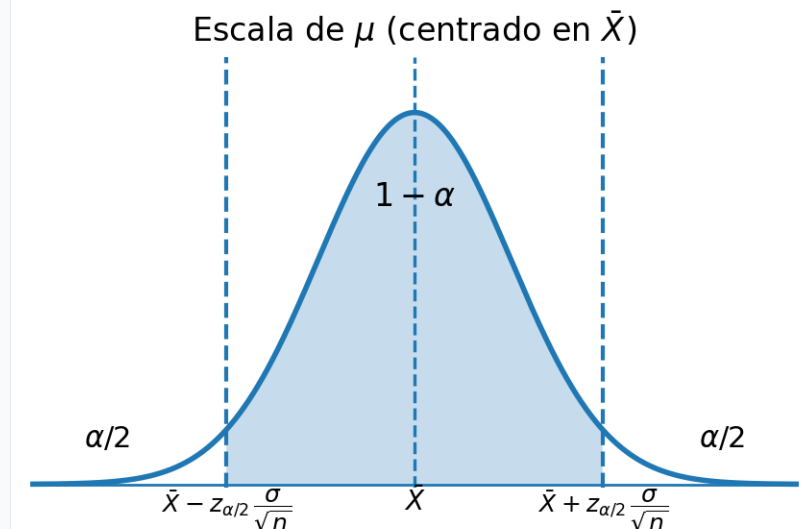
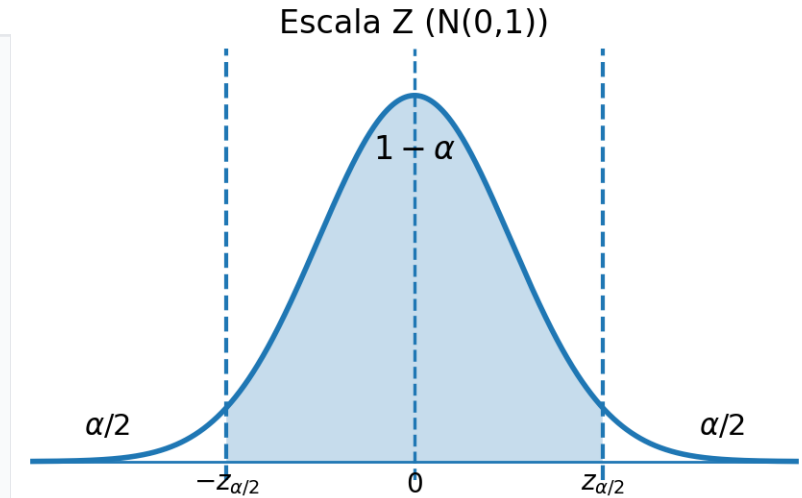
$$P\left(-z_{\alpha/2} \leq \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \leq z_{\alpha/2}\right) = 1 - \alpha$$

$$P\left(\bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

### 3. Resultado $IC_{1-\alpha}(\mu)$

$$\mu \in \left(\bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right); \gamma = 1 - \alpha$$

**Interpretación:**  $\bar{x}$  es el **centro** y el  $z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$  es la **semi amplitud** del I.C.



## 4.1.1 Ejercicio: I.C. para $\mu$ con $\sigma$ conocida en $N(\mu, \sigma)$

**Enunciado:** Una consultora analiza el gasto mensual medio (en €) en transporte de los estudiantes de un campus. Estudios previos indican que el gasto puede modelizarse como  $N(\mu, \sigma)$  y que la desviación típica poblacional es conocida,  $\sigma = 18$  €. Se toma una m.a.s. de tamaño  $n=36$  y se obtiene  $\bar{x} = 72$  €.

1. Construye el intervalo de confianza bilateral del 95% para  $\mu$ .

$$\alpha = 0,05, \quad z_{\alpha/2} = z_{0,025} \approx 1,96, \quad \text{semiamplitud} = 1,96 \cdot \frac{18}{\sqrt{36}} = 1,96 * 3 = 5,88$$

(desarrollar el cálculo del intervalo)

$$IC_{95\%}(\mu) = [72 - 5,88 ; 72 + 5,88] = [66,12 ; 77,88] \text{€}$$

2. Interpreta el resultado en una frase (en términos del método, no del parámetro aleatorio).

“Si repitiésemos el muestreo y construyésemos el IC del 95% del mismo modo, aproximadamente el 95% de los intervalos contendrían a  $\mu$ .”

3. (Mini) ¿Qué ocurriría con la semi-amplitud si el tamaño muestral se duplicara?

$$\text{semiamplitud nueva} = \frac{5,88}{\sqrt{2}} \approx 4,16\text{€}$$

## 4.2 Caso $N(\mu, \sigma)$ : $\sigma^2$ desconocida (I.C. para $\mu$ )

- **Supuesto:**  $x_1, \dots, x_n$  es m.a.s. de una población  $N(\mu, \sigma)$  y  $\sigma^2$  desconocida.
- **Estimador:**  $\hat{\mu} = \bar{x}$
- **Estadístico muestral:**  $S_1$  (cuasidesviación típica)

### 1. Pivote:

$$T = \frac{\bar{x} - \mu}{S_1/\sqrt{n}} \sim t_{n-1}$$

### 2. Intervalo teórico (colas iguales)

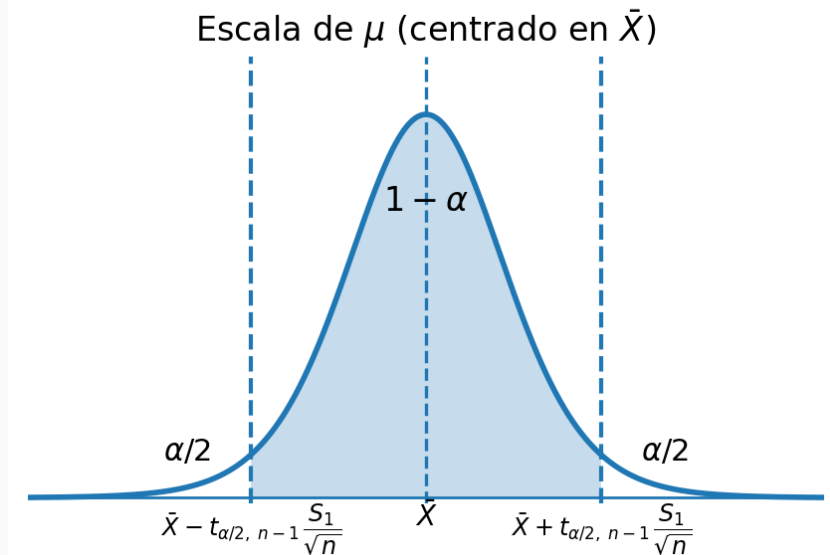
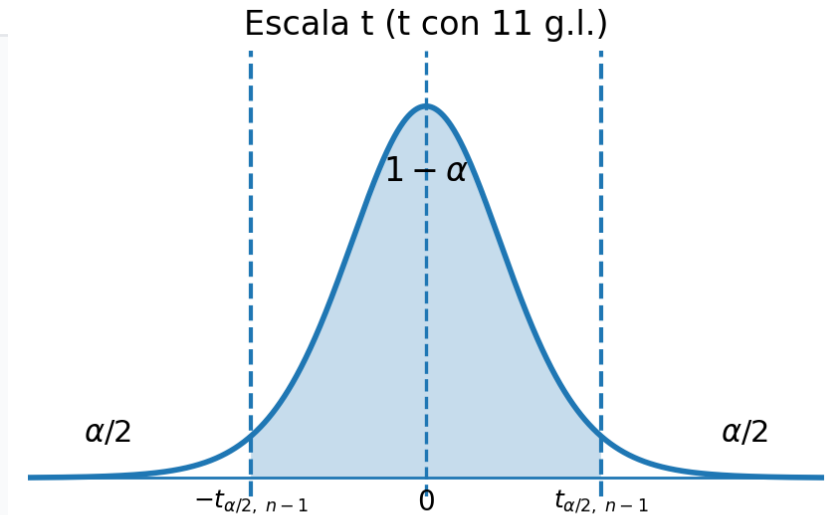
$$P\left(-t_{\alpha/2, n-1} \leq \frac{\bar{x} - \mu}{S_1/\sqrt{n}} \leq t_{\alpha/2, n-1}\right) = 1 - \alpha$$

$$P\left(\bar{x} - t_{\alpha/2, n-1} \frac{S_1}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{\alpha/2, n-1} \frac{S_1}{\sqrt{n}}\right) = 1 - \alpha$$

### 3. Resultado $IC_{1-\alpha}(\mu)$

$$\mu \in \left(\bar{x} - t_{\alpha/2, n-1} \frac{S_1}{\sqrt{n}}, \bar{x} + t_{\alpha/2, n-1} \frac{S_1}{\sqrt{n}}\right); \gamma = 1 - \alpha$$

**Interpretación:**  $\bar{x}$  es el **centro** y  $t_{\alpha/2, n-1} \frac{S_1}{\sqrt{n}}$  es **semi amplitud** del I.C.



## 4.2.1 Ejercicio: I.C. para $\mu$ con $\sigma$ desconocida en $N(\mu, \sigma)$

**Enunciado:** Una empresa estudia el tiempo medio (en minutos) que sus clientes esperan en caja. Se asume  $N(\mu, \sigma)$  con  $\sigma$  desconocida.

**Muestra:**  $n = 12$ ,  $\bar{x} = 8.4$ ,  $S_1 = 2,5$ .

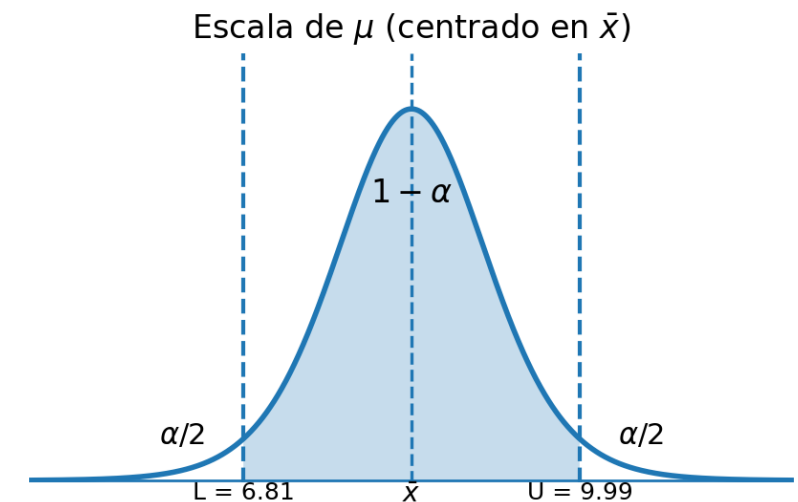
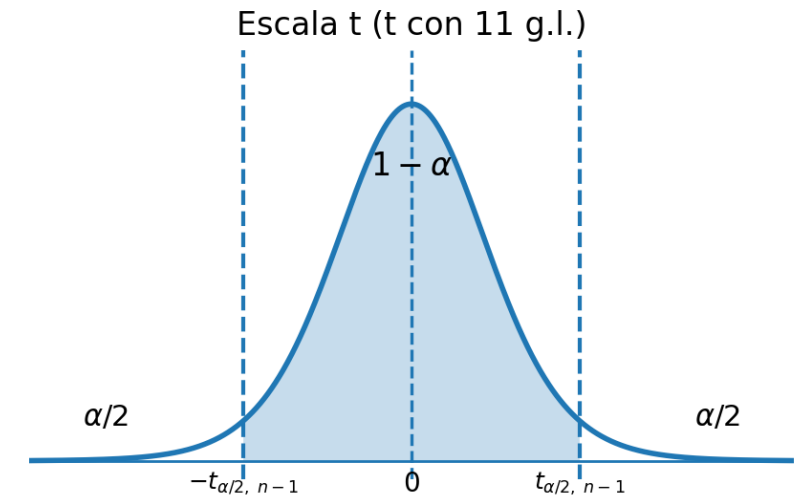
1. Construye el intervalo de confianza bilateral del 95% para  $\mu$ .

$$\alpha = 0,05, \quad g.l. = n - 1 = 11 \quad t_{\alpha/2, n-1} = t_{0,025; 11} \approx 2,201,$$

$$\text{semiamplitud} = 2,201 \cdot \frac{2,5}{\sqrt{12}} = 2,201 * 0,7217 \approx 1,59$$

(desarrollar el cálculo del intervalo)

$$IC_{95\%}(\mu) = [8,4 - 1,59 ; 8,4 + 1,59] = [6,81 ; 9,99] \text{ minutos}$$



# Test rápido (Z vs t)

- 1. ¿Cuándo usamos  $z_{\alpha/2}$  en un I.C. para  $\mu$ ?**
  - a) Siempre que  $n \geq 30$
  - b) Cuando  $\sigma$  es conocida y  $\xi \sim N(\mu, \sigma)$
  - c) Cuando  $\sigma$  es desconocida
  - d) Solo en intervalos unilaterales
- 2. Si  $\sigma$  es desconocida y asumimos  $N(\mu, \sigma)$ , el pivote usa:**
  - a) Normal estándar
  - b)  $\chi^2$
  - c) t de Student con  $n-1$  g.l.
  - d) F de Snedecor
- 3. En el caso t, el “n-1” aparece porque:**
  - a) el estimador  $\bar{x}$  tiene  $n-1$  valores
  - b) se estima  $\sigma$  con  $S_1$ , que usa  $n-1$  g.l.
  - c) la normal tiene  $n-1$  parámetros
  - d) es una convención sin motivo
- 4. Al aumentar n, manteniendo el resto, el I.C. para  $\mu$ :**
  - a) tiene mayor amplitud
  - b) no cambia
  - c) tiene menor amplitud
  - d) se vuelve unilateral

# Test rápido (Z vs t)

- 1. ¿Cuándo usamos  $z_{\alpha/2}$  en un I.C. para  $\mu$ ?**
  - a) Siempre que  $n \geq 30$
  - b) Cuando  $\sigma$  es conocida y  $\xi \sim N(\mu, \sigma)$
  - c) Cuando  $\sigma$  es desconocida
  - d) Solo en intervalos unilaterales
- 2. Si  $\sigma$  es desconocida y asumimos  $N(\mu, \sigma)$ , el pivote usa:**
  - a) Normal estándar
  - b)  $\chi^2$
  - c) t de Student con  $n-1$  g.l.
  - d) F de Snedecor
- 3. En el caso t, el “n-1” aparece porque:**
  - a) el estimador  $\bar{x}$  tiene  $n-1$  valores
  - b) se estima  $\sigma$  con  $S_1$ , que usa  $n-1$  g.l.
  - c) la normal tiene  $n-1$  parámetros
  - d) es una convención sin motivo
- 4. Al aumentar n, manteniendo el resto, el I.C. para  $\mu$ :**
  - a) tiene mayor amplitud
  - b) no cambia
  - c) tiene menor amplitud
  - d) se vuelve unilateral

## 4.3 Caso $N(\mu, \sigma)$ : I.C. para $\sigma^2$ (y para $\sigma$ )

- **Supuesto:**  $x_1, \dots, x_n$  es m.a.s. de una población  $N(\mu, \sigma)$ .
- **Estimador:**  $\widehat{\sigma^2} = S_1^2$

### 1. Pivote (Fisher–Cochran):

$$\frac{(n-1)S_1^2}{\sigma^2} \sim \chi_{n-1}^2$$

### 2. Intervalo teórico (colas iguales)

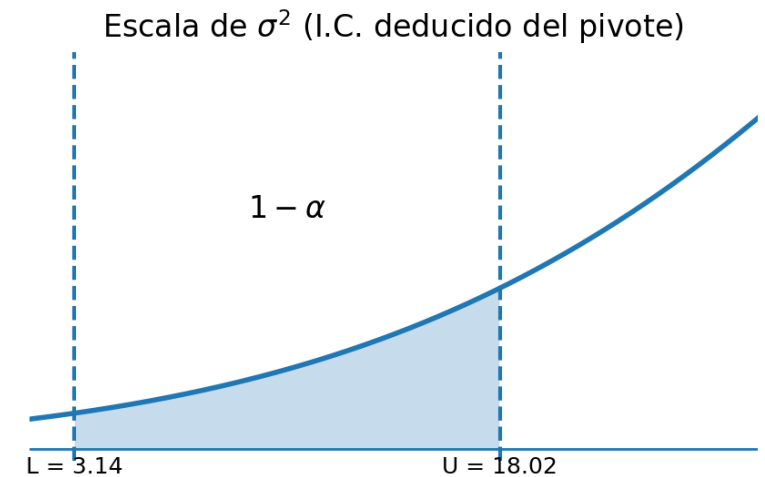
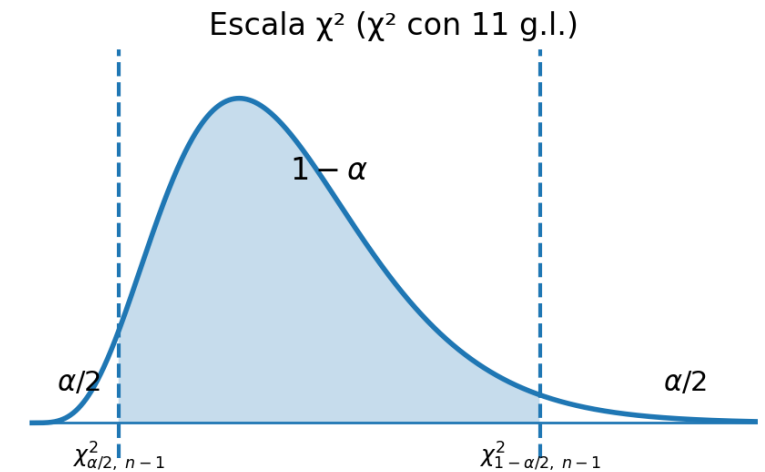
$$P\left(\chi_{(1-\alpha)/2, n-1}^2 \leq \frac{(n-1)S_1^2}{\sigma^2} \leq \chi_{\alpha/2, n-1}^2\right) = 1 - \alpha$$

$$P\left(\frac{(n-1)S_1^2}{\chi_{\alpha/2, n-1}^2} \leq \sigma^2 \leq \frac{(n-1)S_1^2}{\chi_{(1-\alpha)/2, n-1}^2}\right) = 1 - \alpha$$

### 3. Resultado $IC_{1-\alpha}(\sigma^2)$

$$\sigma^2 \in \left(\frac{(n-1)S_1^2}{\chi_{\alpha/2, n-1}^2}, \frac{(n-1)S_1^2}{\chi_{(1-\alpha)/2, n-1}^2}\right); \gamma = 1 - \alpha$$

### 4. I.C. para $\sigma$ : tomar raíz en ambos extremos.



$$P(\chi_{\alpha/2}^2 \leq (n-1)S_1^2/\sigma^2 \leq \chi_{1-\alpha/2}^2) = 0.95 \quad (n=12, S_1^2=6.25)$$

## 4.3 Caso $N(\mu, \sigma)$ : I.C. para $\sigma^2$ (y para $\sigma$ )

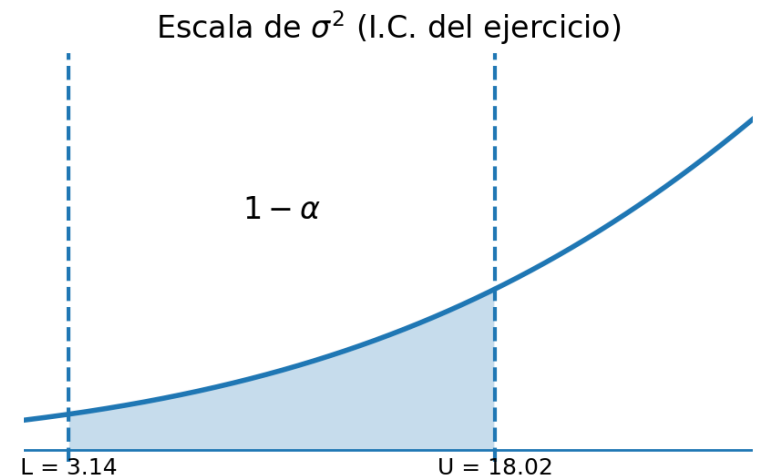
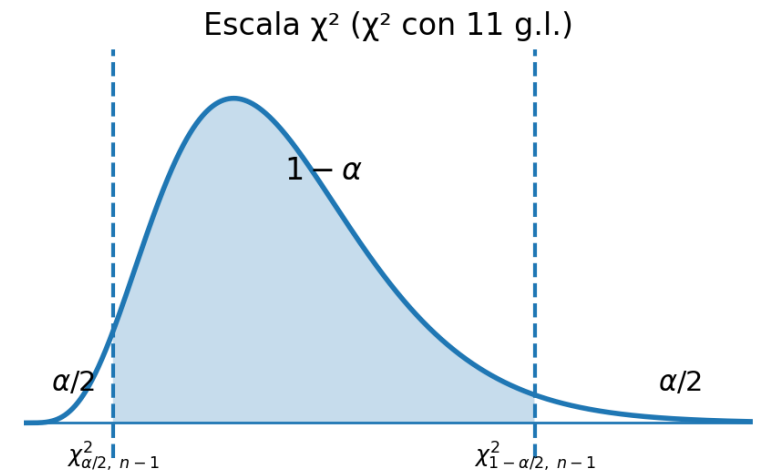
Una empresa de logística quiere estimar la variabilidad del tiempo de entrega (en horas) para un servicio estándar. Se asume que los tiempos siguen  $N(\mu, \sigma)$ . En una m.a.s. de  $n=12$  entregas se obtiene  $S_1^2 = 6,25$ . Construye el I.C. bilateral del 95% para  $\sigma^2$  y para  $\sigma$ .

### 1. Resultado (I.C. para $\sigma^2$ )

$$\begin{aligned} IC_{95\%}(\sigma^2) &= \left( \frac{(n-1)S_1^2}{\chi_{0,975}^2; 11}, \frac{(n-1)S_1^2}{\chi_{0,025}^2; 11} \right) \\ &= \left( \frac{11 \cdot 6,25}{21,920}, \frac{11 \cdot 6,25}{3,816} \right) = [3,14; 18,02] \text{ horas}^2 \end{aligned}$$

### 2. Resultado $IC_{1-\alpha}(\sigma^2)$

$$IC_{95\%}(\sigma) = \left[ \sqrt{3,14}; \sqrt{18,02} \right] \approx [1,77; 4,24] \text{ horas}$$



IC 95% para  $\sigma^2$ : [3.14, 18.02] ( $n=12, S_1^2=6.25$ )

## 4.T2 Test rápido ( $\chi^2$ y variabilidad)

1. El pivote para construir un I.C. de  $\sigma^2$  en  $N(\mu, \sigma)$  es:
  - a)  $(\bar{x} - \mu) / (\sigma/\sqrt{n})$
  - b)  $(n-1)S_1^2 / \sigma^2$
  - c)  $\sigma^2 / ((n-1)S_1^2)$
  - d)  $S_1 / \sigma$
2. En ese pivote, los grados de libertad son:
  - a)  $n$
  - b)  $n+1$
  - c)  $n-1$
  - d) Depende de  $\mu$
3. El I.C. para  $\sigma^2$  en  $\chi^2$  suele ser:
  - a) simétrico en longitud (centrado)
  - b) asimétrico en longitud (no centrado)
  - c) siempre de longitud mínima
  - d) unilateral por defecto
4. Para obtener un I.C. de  $\sigma$  a partir del de  $\sigma^2$ :
  - a) se divide por 2
  - b) se toma logaritmo
  - c) se toma raíz en ambos extremos
  - d) se resta  $\bar{x}$

## 4.T2 Test rápido ( $\chi^2$ y variabilidad)

1. El pivote para construir un I.C. de  $\sigma^2$  en  $N(\mu, \sigma)$  es:
- $(\bar{x} - \mu) / (\sigma/\sqrt{n})$
  - $(n-1)S_1^2 / \sigma^2$  ✓
  - $\sigma^2 / ((n-1)S_1^2)$
  - $S_1 / \sigma$
2. En ese pivote, los grados de libertad son:
- $n$
  - $n+1$
  - $n-1$  ✓
  - Depende de  $\mu$
3. El I.C. para  $\sigma^2$  en  $\chi^2$  suele ser:
- simétrico en longitud (centrado)
  - asimétrico en longitud (no centrado) ✓
  - siempre de longitud mínima
  - unilateral por defecto
- Nota: es simétrico en probabilidad** (colas iguales)
4. Para obtener un I.C. de  $\sigma$  a partir del de  $\sigma^2$ :
- se divide por 2
  - se toma logaritmo
  - se toma raíz en ambos extremos ✓
  - se resta  $\bar{x}$

# 5. Intervalos de confianza en poblaciones no normales

# 5.2 Chebyshev: I.C. para $\mu$ con garantía mínima

- Supuesto mínimo: existe media  $\mu$  y desviación típica  $\sigma$  (sin asumir forma de la distribución).
- Desigualdad de Chebyshev:

$$P(|\xi - \mu| < k\sigma) \geq 1 - \frac{1}{k^2}$$

- Aplicada a la media muestral  $\bar{x}$  (si  $\sigma$  es conocida):

$$P\left(|\bar{x} - \mu| < k \frac{\sigma}{\sqrt{n}}\right) \geq 1 - \frac{1}{k^2}$$

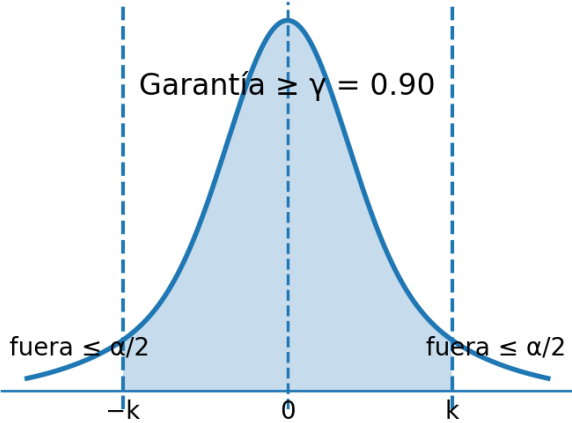
- Para garantizar un nivel  $\gamma$ , elegimos  $k$  tal que  $1 - \frac{1}{k^2} = \gamma$ , es decir

$$k = \sqrt{\frac{1}{1 - \gamma}}$$

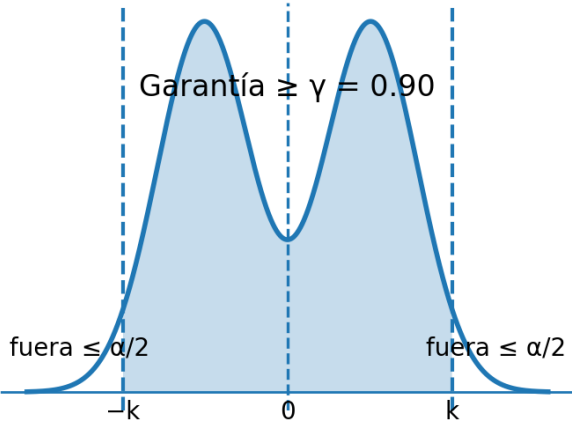
- Resultado (garantía al menos  $\gamma$ ):

$$\mu \in \left[ \bar{x} - k \frac{\sigma}{\sqrt{n}}; \bar{x} + k \frac{\sigma}{\sqrt{n}} \right] \text{ con probabilidad } \geq \gamma$$

Chebyshev: cualquier distribución (unimodal)



Chebyshev: cualquier distribución (bimodal)



## 5.2.1 Mini-ejemplo (Chebyshev)

- Una empresa estima el coste medio  $\mu$  de un proceso. No se asume forma de la distribución, pero se conoce  $\sigma = 30$  €. Muestra:  $n = 36$ ,  $\bar{x} = 120$  €. Queremos construir un intervalo para  $\mu$  que contenga al parámetro con probabilidad al menos  $\gamma = 0.90$ .

1. Calcular  $k$

$$k = \sqrt{\frac{1}{1-\gamma}} = \sqrt{\frac{1}{0,10}} = \sqrt{10} \approx 3,162$$

2. Semi-amplitud:

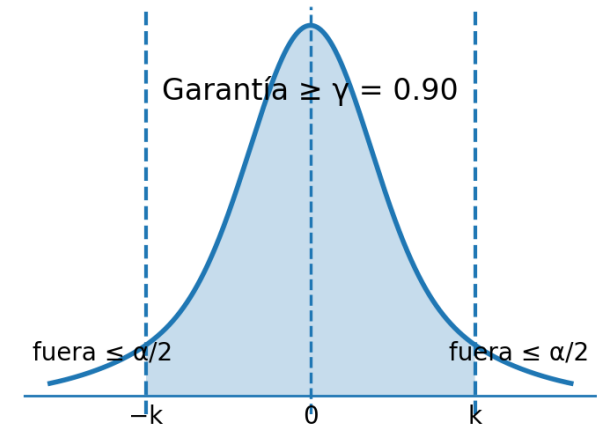
$$k \frac{\sigma}{\sqrt{n}} = 3,162 \frac{30}{6} = 3,162 \cdot 5 \approx 15,81$$

3. Intervalo (cota  $\gamma \geq 0,90$ ):

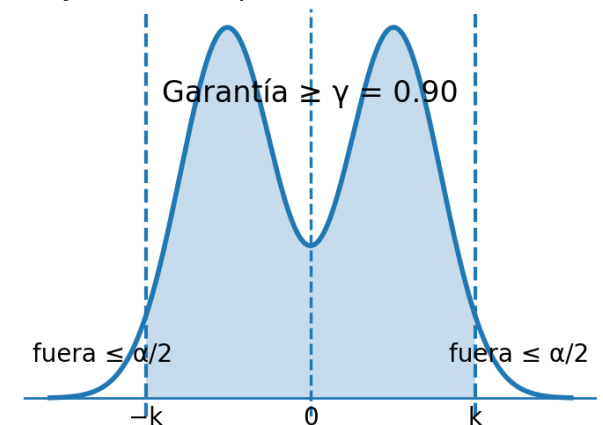
$$\mu \in [120 - 15,81; 120 + 15,81] = [104,19; 135,81]; \quad \gamma \geq 0,90$$

**Idea clave:** la garantía es " $\gamma \geq cota$ " y el intervalo suele ser más ancho que en el caso normal.

Chebyshev: cualquier distribución (unimodal)



Chebyshev: cualquier distribución (bimodal)



Chebyshev:  $P(|X-\mu| < k\sigma) \geq 1 - 1/k^2$ . Para  $\gamma=0.90 \Rightarrow k = \sqrt{1/(1-\gamma)} = 3.16$

## 5.3 Grandes muestras (MV): I.C. aproximado para $\theta$

- En condiciones habituales, el estimador de máxima verosimilitud  $\hat{\theta}_{MV}$  es asintóticamente normal :

$$\hat{\theta}_{MV} \approx N\left(\theta, \sqrt{V(\hat{\theta}_{MV})}\right)$$

- Equivalentemente, el pivote aproximado:

$$T(X; \theta) = \frac{\hat{\theta}_{MV} - \theta}{\sqrt{V(\hat{\theta}_{MV})}} \approx N(0,1)$$

- En un I.C. bilateral estándar (colas iguales):

$$P(-z_{\alpha/2} \leq T(X; \theta) \leq z_{\alpha/2}) \approx 1 - \alpha$$

- Despejando  $\theta$  :  $IC_{1-\alpha}(\theta)$

$$\theta \in \left[ \hat{\theta}_{MV} - z_{\alpha/2} \sqrt{V(\hat{\theta}_{MV})} ; \hat{\theta}_{MV} + z_{\alpha/2} \sqrt{V(\hat{\theta}_{MV})} \right]$$

## 5.3.1 Ejercicio: I.C. aproximado por grandes muestras

- Una cadena de supermercados quiere estimar el gasto medio  $\theta = \mu$  (en €) por cliente en compras online. Se toma una muestra grande de  $n=400$  tickets y se obtiene:
  - $\hat{\theta} = \bar{x} = 52,3$  €
  - $S_1 = 18$  € (cuasidesviación típica muestral).
  - Se pide: Construir un I.C. bilateral aproximado del 95% para  $\theta$ .
- Solución (aprox. normal):

$$\hat{V}(\hat{\theta}) = \frac{S_1^2}{400} = \frac{18^2}{400} = \frac{324}{400} = 0,81 \rightarrow \sqrt{\hat{V}(\hat{\theta})} = \sqrt{0,81} = 0,9$$

$$z_{0,025} \approx 1,96, \quad \text{semi - amplitud} = 1,96 \cdot 0,9 = 1,764$$

$$IC_{95\%}(\mu) = [52,3 \pm 1,764] = [50,54 ; 54,06]$$

- Interpretación: intervalo aproximado (mejora con  $n$  grande)

## 5.4 Proporción $\pi$ : I.C. aproximado (grandes muestras)

- **Contexto:** variable dicotómica (0/1). Parámetro:  $\pi$
- **Estimador:**  $\hat{\pi} = \bar{x} = p$  (proporción muestral)
- **Aproximación (n grande):**

$$\hat{\pi} \approx N \left( \pi, \sqrt{\frac{\pi(1-\pi)}{n}} \right)$$

- **Sustituyendo  $\pi$  por  $\hat{\pi}$ :**

$$IC_{1-\alpha}(\pi) \in \left[ \hat{\pi} - z_{\alpha/2} \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}; \hat{\pi} + z_{\alpha/2} \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}} \right]$$

- **Condición práctica** (para que la aproximación sea razonable):  $n\hat{\pi}$  y  $n(1-\hat{\pi})$  “suficientemente grandes”.
  - **En ejercicios:**  $n\hat{\pi} \geq 10$  y  $n(1-\hat{\pi}) \geq 10$

## 5.4.1 Ejercicio: I.C. para una proporción $\pi$

Un banco quiere estimar la proporción de clientes que entran en mora en el primer año de un préstamo. En una muestra de  $n=500$  préstamos,  $x = 60$  entran en mora.

1. Estimar  $\pi$  y construir el I.C. bilateral del 95% (aprox. grandes muestras).
2. Comprobar la condición:  $n\hat{\pi} \geq 10$  y  $n(1 - \hat{\pi}) \geq 10$

• Solución:

$$\hat{\pi} = \frac{x}{n} = \frac{60}{500} = 0,12$$

• Condición:

$$n\hat{\pi} = 500 \cdot 0,12 = 60 \geq 10 ; n(1 - \hat{\pi}) = 500 \cdot 0,88 = 440 \geq 10$$

• Desviación típica estimada:

$$\sqrt{\hat{V}(\hat{\pi})} = \sqrt{\frac{\hat{\pi}(1 - \hat{\pi})}{n}} = \sqrt{\frac{0,12 \cdot 0,88}{500}} \approx \sqrt{0,0002112} \approx 0,01453$$

• Con  $z_{0,025} = 1,96$ :

$$\text{semi amplitud} = 1,96 \cdot 0,01453 \approx 0,0285$$

$$IC_{95\%}(\pi) \in [0,12 - 0,0285 ; 0,12 + 0,0285] = [0,0915 ; 0,1485]$$

(En porcentaje: [9,15% ; 14,85%] )

## 5.T1 Test rápido (exacto / aproximado / garantía mínima)

1. Si asumimos  $N(\mu, \sigma)$  y  $\sigma$  es conocida, el I.C. para  $\mu$  usando  $Z$  es:
  - a) garantía mínima
  - b) aproximado
  - c) exacto
  - d) unilateral por defecto
2. El I.C. de  $\pi$  basado en la aproximación normal es:
  - a) exacto siempre
  - b) aproximado (mejora con  $n$  grande)
  - c) garantía mínima
  - d) independiente de  $n$
3. El intervalo obtenido con Chebyshev para  $\mu$  tiene interpretación:
  - a) exacto con nivel  $\gamma$
  - b) aproximado con nivel  $\gamma$
  - c) contiene a  $\mu$  con probabilidad al menos  $\gamma$
  - d) solo sirve si la distribución es normal
4. La condición  $n \cdot p \geq 10$  y  $n(1-p) \geq 10$  se usa para:
  - a) elegir entre  $t$  y  $\chi^2$
  - b) justificar la aproximación normal en el I.C. de  $\pi$
  - c) calcular  $\bar{x}$
  - d) decidir si  $\sigma$  es conocida

## 5.T1 Test rápido (exacto / aproximado / garantía mínima)

1. Si asumimos  $N(\mu, \sigma)$  y  $\sigma$  es conocida, el I.C. para  $\mu$  usando  $Z$  es:
  - a) garantía mínima
  - b) aproximado
  - c) exacto
  - d) unilateral por defecto
2. El I.C. de  $\pi$  basado en la aproximación normal es:
  - a) exacto siempre
  - b) aproximado (mejora con  $n$  grande)
  - c) garantía mínima
  - d) independiente de  $n$
3. El intervalo obtenido con Chebyshev para  $\mu$  tiene interpretación:
  - a) exacto con nivel  $\gamma$
  - b) aproximado con nivel  $\gamma$
  - c) contiene a  $\mu$  con probabilidad al menos  $\gamma$
  - d) solo sirve si la distribución es normal
4. La condición  $n \cdot \hat{p} \geq 10$  y  $n(1 - \hat{p}) \geq 10$  se usa para:
  - a) elegir entre  $t$  y  $\chi^2$
  - b) justificar la aproximación normal en el I.C. de  $\pi$
  - c) calcular  $\bar{x}$
  - d) decidir si  $\sigma$  es conocida

# 6. Determinación del tamaño muestral

## 6.1 Tamaño muestral para estimar $\mu$ con $\sigma^2$ conocida en $N(\mu, \sigma)$

- **Objetivo:** elegir  $n$  para que el I.C. **bilateral** de  $\mu$  con un nivel de **confianza**  $\gamma$ , tenga semi-amplitud (**error máximo**)  $e$ .
- En  $N(\mu, \sigma)$  con  $\sigma$  conocida:

$$IC_{1-\alpha}(\mu): \mu \in \left[ \bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right]; \gamma = 1 - \alpha$$

- **Por tanto, la semi-amplitud es:**

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

- Despejando  $n$ :

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

(Se redondea hacia arriba al entero siguiente.)

## 6.1.1 Ejercicio: tamaño muestral para $\mu$

- Una empresa quiere estimar el coste medio  $\mu$  (en €) de una operación. Se asume que la población se distribuye  $N(\mu, \sigma)$  y se conoce  $\sigma = 30$  €. Se desea un I.C. bilateral del 95% con error máximo (semi-amplitud)  $e = 3$  €.
- Calcular el tamaño muestral mínimo  $n$ .
- **Solución :**

$$\alpha = 0,05 ; \quad z_{0,025} = 1,96; \quad n = \left( z_{\alpha/2} \frac{\sigma}{e} \right)^2 = \left( 1,96 \frac{30}{3} \right)^2 = 19,6^2 = 384,16$$

- Redondeo hacia arriba :

$$n = 385$$

## 6.2 Tamaño muestral para estimar una proporción $\pi$

- **Objetivo:** elegir  $n$  para que el I.C. bilateral  $\pi$  de con un nivel de confianza  $\gamma$ , tenga semi-amplitud (**error máximo**)  $e$ .
- **Aproximación (grandes muestras):**

$$IC_{1-\alpha}(\pi): \pi \in \left[ \hat{\pi} - z_{\alpha/2} \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}; \hat{\pi} + z_{\alpha/2} \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}} \right]$$

- La semi-amplitud se aproxima por:

$$e = z_{\alpha/2} \sqrt{\frac{\pi(1-\pi)}{n}}$$

- Despejando  $n$

$$n = \frac{z_{\alpha/2}^2 \pi(1-\pi)}{e^2}$$

- Si  $\pi$  es desconocida y no hay información previa, se usa el caso conservador:

$$\pi(1-\pi) \leq 0,25 \rightarrow n = \frac{z_{\alpha/2}^2 \cdot 0,25}{e^2}$$

(Se redondea hacia arriba.)

## 6.2.1 Ejercicio: tamaño muestral para $\pi$

- Una empresa de telecomunicaciones quiere estimar la proporción de clientes que se dan de baja en un trimestre. No hay información previa sobre  $\pi$ . Se desea un I.C. bilateral del 95% con error máximo  $e=0,03$  (3 puntos porcentuales).
- Calcular el tamaño muestral mínimo  $n$ .

- **Solución :**

$$\alpha = 0,05 ; \quad z_{0,025} = 1,96; \quad n = \frac{z_{\alpha/2}^2 \cdot 0,25}{e^2} = \frac{1,96^2 \cdot 0,25}{0,03^2} = \frac{3,8116 \cdot 0,25}{0,0009} = 1067,11$$

- Redondeo hacia arriba :

$$n = 1068$$

# **7. Guía rápida: elección del intervalo de confianza**

# 7.1 Mapa de decisión (qué intervalo usar)

## 1) ¿Qué parámetro quieres estimar?

- $\mu$  (media)
- $\sigma^2 / \sigma$  (variabilidad)
- $\pi$  (proporción)

## 2) ¿Qué supuestos puedes hacer?

- Población  $N(\mu, \sigma)$
- No normal / desconocida
- Tamaño muestral grande (aproximación)

## 3) Elegir el intervalo:

- $\mu$  en  $N(\mu, \sigma)$ 
  - $\sigma$  conocida  $\rightarrow Z$
  - $\sigma$  desconocida  $\rightarrow t_{n-1}$
- $\sigma^2$  en  $N(\mu, \sigma) \rightarrow \chi_{n-1}^2$
- $\pi \rightarrow$  grandes muestras (aprox. normal) + comprobar  $n \cdot \hat{\pi}$  y  $n(1 - \hat{\pi})$
- **Distribución desconocida** (sin modelo)  $\rightarrow$  Chebyshev (garantía mínima)

## 7.2 Checklist final (construcción e interpretación)

1. Identifica el **parámetro poblacional desconocido**  $\theta$  ( $\mu, \sigma^2, \pi, \dots$ ).
2. Fija la **distribución/supuesto** (Normal, grandes muestras, Chebyshev).
3. Elige el **pivote** y la **tabla** correspondiente ( $Z, t, \chi^2$ ).
4. **Determina  $\alpha$**  y, si es **bilateral estándar**, usa  $\alpha/2$  en cada cola.
5. Construye el **intervalo**  $[\theta_i(X), \theta_s(X)]$  y presenta el **intervalo numérico**.
6. **Interpreta correctamente**: “el método produce intervalos que contienen a  $\theta$  con nivel de confianza  $\gamma = 1 - \alpha$ ”.

## 7.3 Cierre del tema

- **Idea clave:** un intervalo de confianza es “estimación puntual + precisión”.
- $\gamma = 1 - \alpha$ :
  - $\gamma$ : nivel de confianza
  - $\alpha$ : mide la probabilidad de que el intervalo no contenga a  $\theta$ .
- **Casos principales:**
  - $\mu$ : Z ( $\sigma$  conocida) / t ( $\sigma$  desconocida)
  - $\sigma^2$ :  $\chi^2$
  - $\pi$ : grandes muestras
  - distribución desconocida: Chebyshev (garantía mínima)

# Recursos

## Bibliografía:

- Ruiz-Maya, L., Martín-Pliego López, F. J. (3.ª ed.). Fundamentos de Inferencia Estadística. Thompson–Paraninfo.
- **StatsCalculators — Statistical Tables:**
  - Normal estándar, t de Student, Chi-cuadrado y F de Fisher–Snedecor
  - [https://www.statscalculators.com/resources/statistical-tables?utm\\_source=chatgpt.com](https://www.statscalculators.com/resources/statistical-tables?utm_source=chatgpt.com)  
(recurso externo; enlace informativo)

## Atribuciones y transparencia:

- **Gráficas** de este anexo: generadas con Python (SciPy/Matplotlib) con apoyo de ChatGPT.
- **Imágenes externas:** cuando se incluyen, se usan bajo licencia CC-BY / CC-BY-SA, citadas individualmente.