Intervalos de confianza para una población

IC para un promedio

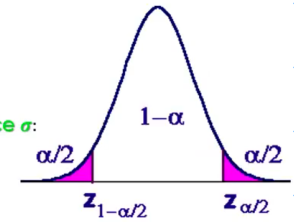
Caso1: El promedio muestral es normal y se conoce la varianza poblacional

Considere una población dada:

- **Variable aleatoria:** X ← n
- **Media poblacional:** μ
- **Varianza poblacional:** σ^2

$\sigma_1 \to \bar{x}_1$

$\sigma_2 \to \bar{x}_2$



Si $\bar{X}$ sigue una distribución normal para muestras de **tamaño n** y se **conoce σ**:

1. IC de $(1-\alpha)100\%$ para μ: $\bar{x} \pm z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$

   Centro de I.C. ← ↳ Radio
2. Tamaño de muestra para encontrar un IC para μ con un radio menor o igual a r: $n \geq \left(\frac{z_{\alpha/2} \cdot \sigma}{r}\right)^2$

→ Distribución muestral de medias $\bar{X}$ se comporta de forma normal

Si dicen que X es normal o si $n \geq 30$, se asume normalidad, sino t-student

Siempre se tiene]Lim inf, Lim sup[

]Li, LS[

Donde

$$Li = \bar{X} - \underbrace{Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}}_{\text{Radio}} \qquad Li = \bar{X} + \underbrace{Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}}_{\text{Radio}}$$

IC para un promedio: **observaciones**

Para que $\bar{X}$ siga una distribución normal, debe cumplirse una de las siguientes condiciones:

1. La **distribución original** dada por X sigue una distribución **normal** (Teorema: Promedio de Normales es normal)

2. **El tamaño de las muestras debe ser mayor o igual a 30** (Teorema del Límite Central) En estos casos que se cumpla la condición de $n \geq 30$, si NO SE CONOCE σ, se utiliza s en su lugar, considerando a este como una buena estimación de σ y se aplica el resultado explicado en la lámina anterior para el IC.

Promedio

$\bar{X}$ $\quad n \quad \sigma \quad Z_{\frac{\alpha}{2}}$ $\qquad \alpha = 1 - \text{Confianza}$

error = radio

Se tiene interés en estimar la vida útil de un producto nuevo. ¿Qué tamaño de muestra mínimo debe tomarse para estimar la media con un error de estimación máximo de $\frac{1}{10}$ desviaciones estándar y con una confiabilidad de 90%? → radio

$\alpha = 1 - \text{Confianza}$

$\alpha = 1 - 0{,}90$

$\alpha = 0{,}10$

$n \geq \left(\frac{Z_{\frac{\alpha}{2}} \cdot \sigma}{r} \right)^2$

$\frac{\alpha}{2} = \frac{0{,}10}{2} = 0{,}05$

$r = \frac{1}{10}\sigma \leftarrow 0{,}0$, se escribe así

$Z_{\frac{\alpha}{2}} = Z_{0{,}05} = -1{,}64485$

↘ Se mete en la aplicación en este caso con $P(X < x)$ o $P(X > x)$, no importa

Entonces

$n \geq \left(\frac{Z_{\frac{\alpha}{2}} \cdot \sigma}{r} \right)^2$

$n \geq \left(\frac{Z_{0{,}05} \cdot \sigma}{\frac{1}{10}\sigma} \right)^2$

$n \geq \left(\frac{-1{,}64485\sigma}{\frac{\sigma}{10}} \right)^2$

$n \geq \left(\frac{-1{,}64485 \cdot 10 \cancel{\sigma}}{\cancel{\sigma}} \right)^2$

$n \geq 270{,}55$ Se redondea hacia arriba

$\boxed{n \geq 271}$

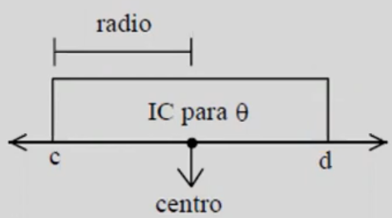
- En una muestra de tamaño 30 se obtuvo la estimación $\bar{x}$ =8cm con una desviación estándar s=1.2 cm
- Por el TLC: $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{30}\right)$
- así $\mathrm{Var}(\bar{X}) = \frac{\sigma^2}{30} \approx \frac{s^2}{30}$
- Requisitos
  - $\bar{X}$ es insesgado
  - Tiene una distrib. Simétrica respecto a μ
- Note que $A_{\alpha/2}$ coincide con $z_{\alpha/2}$
- IC del 95% para μ:
  - α=0.05 ⇒ $z_{\alpha/2} = z_{0,025} = -1.96$
  - $\bar{x} \pm z_{\alpha/2} \cdot \frac{s}{\sqrt{30}} = 8 \pm 1.96 \cdot \frac{1.2}{\sqrt{30}}$

Si $\widehat{\Theta}$ es insesgado $\left(E\left(\widehat{\Theta}\right) = \theta\right)$ y su función de distribución es simétrica con respecto a θ entonces un intervalo de confianza de $100\,(1-\alpha)\,\%$ para θ tiene extremos

$$\widehat{\theta} \pm A_{\alpha/2} \cdot \overline{\sigma}$$

IC del 95%:
$]7.57059, 8.42941[$
cm

Definición 20. *Dado un intervalo de confianza $I =]c, d[$ se define el centro del IC por $\frac{c+d}{2}$ y el radio o error de estimación por $\frac{d-c}{2}$.*



Características de los IC

- Caso particular $\widehat{\theta} \pm A_{\alpha/2} \cdot \overline{\sigma}$
- Elementos:
  - Confianza: a mayor confianza mayor radio del IC
  - Variancia del estimador: Si varía poco entonces el IC tiene un radio pequeño (más preciso)
  - Tamaño de la muestra: A mayor tamaño de la muestra se obtiene un radio más pequeño.
- PROBLEMA: Confianza vs Precisión
- SOLUCIÓN: Manejar un tamaño de muestra adecuado

Todas son $X \sim N(u, \sigma^2)$

Ejemplo 51. Los pesos de las bolsas de frijoles marca Sabores siguen una normal de media desconocida y variancia 0.0004 kg^2. Un inspector de la Oficina del Consumidor tomó una muestra de 20 bolsas y obtuvo un peso promedio de 1.9 kg.

1. Determine un IC del 90 % para el peso promedio de las bolsas de frijoles Sabores.

Si piden promedio, entonces

$$X \sim N(u, \sigma^2) \to \bar{X} \sim N\left(u, \frac{\sigma^2}{n}\right)$$

$u = ?$ $\quad \sigma^2 = 0{,}0004$ $\quad n = 20$ $\quad \bar{x} = 1{,}9$

$$X \sim N(u, 0{,}0004) \to \bar{X} \sim N\left(u, \frac{0{,}0004}{20}\right)$$

$\alpha = 1 - 0{,}90 = 0{,}10$

1. IC de $(1-\alpha)100\%$ para μ: $\bar{x} \pm z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$

$\frac{\alpha}{2} = \frac{0{,}10}{2} = 0{,}05$

$z_{0{,}05}$ en la app con $P(X<x) \wedge P(X>x)$
$\approx -1{,}645 \wedge 1{,}645$

Entonces $]a, b[$

Desviacion estandart

$$a = 1{,}9 - 1{,}645 \cdot \frac{\sqrt{0{,}0004}}{\sqrt{20}} = 1{,}89264$$

$$b = 1{,}9 + 1{,}645 \cdot \frac{\sqrt{0{,}0004}}{\sqrt{20}} = 1{,}9073$$

$\boxed{\therefore \text{ IC para } u \text{ es } I =]1.89264, 1.9073[}$

Ejemplo 52. Se tiene interés en estimar la vida útil promedio de los bombillos marca Ilumina. ¿Qué tamaño de muestra mínimo debe tomarse para estimar dicho promedio con un error de estimación máximo de un noveno de desviación estándar y una confiabilidad del 90 %?

radio = error

$r = \frac{1}{9}\sigma \qquad \alpha = 1 - 0{,}90$

$\alpha = 0{,}10 \qquad n \geq \left(\frac{Z_{\frac{\alpha}{2}} \cdot \sigma}{r}\right)^2$

$\frac{\alpha}{2} = \frac{0{,}10}{2} = 0{,}05 = Z_{0{,}05} = \underset{R}{\pm}\, 1{,}645$

no importa en IC

$n \geq \left(\frac{1{,}645 \cdot \cancel{\sigma}}{\frac{\cancel{\sigma}}{9}}\right)^2$

$n \geq (9 \cdot 1{,}645)^2$

$n \geq 219{,}188$

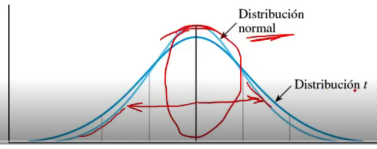
$\boxed{n \geq 220}$

Distribución $t - student$

Considere una población dada por la variable aleatoria $\bar{X}$ que sigue una distribución normal con media poblacional μ. Se tiene que:

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

Tiene una distribución $t - student$ con $v = n - 1$ $grados\ de\ libertad\ (gl)$, la cuál es simétrica con respecto a su media que es cero.

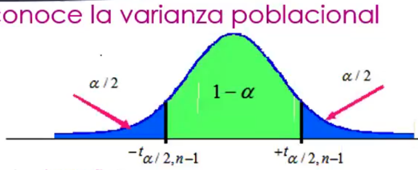


IC para un promedio

Caso2: La población es normal y se desconoce la varianza poblacional

Considere una población dada:

- **Variable aleatoria**: X
- **Media poblacional**: μ
- **Varianza poblacional**: σ^2



Si $\bar{X}$ sigue una distribución normal para muestras de **tamaño n** y se **desconoce σ**:

$s \leftarrow \sigma$ de la muestra

$$\text{IC de } (1-\alpha)100\% \text{ para } \mu: \underbrace{\bar{x}}_{\text{centro}} \pm \underbrace{t_{\frac{\alpha}{2}, n-1} \frac{s}{\sqrt{n}}}_{\text{Radio}}$$

Donde $\bar{x}$ es el valor que toma $\bar{X}$ para una tamaño de muestra n, $t_{\frac{\alpha}{2}, n-1}$ es el valor de la distribución $t - Student$ con $n - 1$ gl.

En normal (Z) $\rightarrow$ r: $Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$

En t-student (T) $\rightarrow$ r: $t_{\frac{\alpha}{2}, n-1} \cdot \frac{S}{\sqrt{n}}$

Una campaña sobre consumo de una bebida, afirma en su publicidad por televisión que su consumo diario durante un mes, produce una pérdida promedio de al menos cinco libras de peso. Para analizar esta afirmación, se toma un grupo control de ocho personas y se les suministra el producto diariamente por un mes obteniendo los siguientes datos:

Peso Inicial (lb)	165	195	188	170	185	163	155	177
Peso Final (lb)	164	190	187	163	185	159	148	174

8 personas

1. Suponiendo que la pérdida de peso sigue una distribución normal, encuentre un intervalo de confianza de 95% para la pérdida de peso promedio.. R/]1.22143, 5.77857[

Peso Inicial (lb)	165	195	188	170	185	163	155	177
Peso Final (lb)	164	190	187	163	185	159	148	174
Dif. peso (lb)	1	5	1	7	0	4	7	3

$n = 8$, $n < 30 \to$ t- student : $\bar{x} \pm t_{\frac{\alpha}{2}, n-1} \frac{s}{\sqrt{n}}$

Requisitos para t-student;

X: normal ✓

σ: Desconocida ✓

Entonces

Desviacion estandart de la muestra

$\alpha = 1 - \text{confianza}$

$\alpha = 1 - 0{,}95$

$\alpha = 0{,}05$

$\bar{x} = ?$ $n = 8$

$s = ?$

$gl = n - 1 = 8 - 1 = 7$

$\bar{x}$ es promedio, entonces seria

$$\frac{1+5+1+7+0+4+7+3}{8} = 3{,}5$$

S es desviacion estandart de la muestra

menu -> 6 -> 2 -> meter datos -> optn -> 3

2,72554

Ahora

$\alpha = 0{,}05$ $\quad \bar{x} = 3{,}5$ $\quad n = 8$

$gl = 7$ $\quad s = 2{,}72554$

$$\frac{\alpha}{2} = \frac{0{,}05}{2} = 0{,}025$$

$$t_{\frac{\alpha}{2}} = t_{0{,}025}$$

$t_{\frac{\alpha}{2}, n-1} \frac{s}{\sqrt{n}}$

← Primero se mete solo $t_{\frac{\alpha}{2}}, n-1$ y luego se multiplica

$$t_{0{,}025,7} = \pm 2{,}36462$$

gl, v en la app

Nota: No importa $P(X > x)$ v $P(X < x)$, solo que si se escoge el equivocado, el intervalo queda alrevez y nada mas se le da la vuelta

IC de $(1-\alpha)100\%$ para μ: $\bar{x} \pm t_{\frac{\alpha}{2}, n-1} \frac{s}{\sqrt{n}}$

$$L_i = 3{,}5 - 2{,}36462 \cdot \frac{2{,}72554}{\sqrt{8}} \approx 1{,}2214$$

$$L_s = 3{,}5 + 2{,}36462 \cdot \frac{2{,}72554}{\sqrt{8}} \approx 5{,}7786$$

El IC para 95% corresponde a

$$]1{,}2214, 5{,}7786[$$

2. ¿Se oponen los datos a la información dada en la publicidad?.

Como $5 \in IC$, los datos no se oponen a lo afirmado por la compañia

Ejemplo 54. Se desea estimar la estatura promedio μ de los estudiantes de quinto año de la ciudad C. En una muestra de 32 estudiantes se observó una estatura promedio de 1.4 m con una desviación estándar de 0.6 m. ¿De qué tamaño debe ser la muestra para hallar un intervalo de confianza del 80 % para μ con un radio menor o igual a 15 cm?

$n = 32 \geq 30$, entonces se usa la normal

$\alpha = 1 - 0{,}8 = 0{,}2 \rightarrow \frac{\alpha}{2} = 0{,}1 \rightarrow z_{0{,}1} = \pm 1{,}28$

$u = 1{,}4 \quad \sigma = 0{,}6 \quad r = 0{,}15$

$$\boxed{n \geq \left(\frac{z_{\frac{\alpha}{2}} \cdot \sigma}{r} \right)^2}$$

$$n \geq \left(\frac{1{,}28 \cdot 0{,}6}{0{,}15} \right)^2$$

$$\boxed{n \geq 26{,}2144 \quad y \quad n \geq 30}$$

Ejercicio 13. Las duraciones de ocho baterías de computadora marca Dutec cargadas son 151, 153, 175, 134, 170, 172, 156 y 114 minutos. Suponga que las duraciones se distribuyen normalmente.

1. Encuentre un intervalo de confianza del 90 % para la duración promedio de las baterías. $R/\quad]139.19, 167.06[$

$n = 8 < 30 \rightarrow$ t-student

IC de $(1-\alpha)100\%$ para μ: $\bar{x} \pm t_{\frac{\alpha}{2}, n-1} \frac{s}{\sqrt{n}}$

$$\bar{x} = \frac{151+153+175+134+170+172+156+114}{8} = 153,125$$

$s = 20,81$

$gl = n-1 = 8-1 = 7$

$\alpha = 1 - 0,90 = 0,10 \rightarrow \frac{\alpha}{2} = 0,05$

$t_{0,05,7} = \pm 1,89458$

IC de $(1-\alpha)100\%$ para μ: $\bar{x} \pm t_{\frac{\alpha}{2}, n-1} \frac{s}{\sqrt{n}}$

$$153,125 - 1,89458 \cdot \frac{20,81}{\sqrt{8}} = 139,19$$

$$153,125 + 1,89458 \cdot \frac{20,81}{\sqrt{8}} = 167,06$$

$\boxed{\text{R/ El IC para 90\% corresponde a }]139,19, 167.06[}$

Como si dan, se usa $n \geq \left(\frac{z_{\frac{\alpha}{2}} \cdot \sigma}{r} \right)^2$

2. Si se supone que la desviación estándar de las duraciones es 20 min, ¿de qué tamaño debió ser una muestra para que el intervalo de confianza del 90 % tuviera radio menor que 7.5 min? *R/* 20

$\sigma = 20 \quad r = 7{,}5$

$\alpha = 1 - 0{,}90 = 0{,}10 \rightarrow \frac{\alpha}{2} = 0{,}05 \rightarrow z_{0{,}05} = 1{,}64485$

$$n \geq \left(\frac{1{,}64485 \cdot 20}{7{,}5} \right)^2$$

$n \geq 19{,}239$

$\boxed{n \geq 20}$

3. En el cálculo del tamaño de muestra anterior se utiliza $z_{\alpha/2}$. ¿Por qué el n no tiene que ser mayor a 30 para justificar el uso de la normal estándar en el cálculo?

Por que se supone normalidad y la desviacion estandar es conocida

4. Con base en el IC obtenido, ¿considera aceptable o rechaza las siguientes posibles afirmaciones que se pueden indicar en el manual de uso de la batería?

(a) Las baterías Dutec cargadas duran en promedio 3 horas.
R/ Se rechaza

(b) Las baterías Dutec cargadas duran en promedio 2.5 horas. minutos
R/ Es aceptable, pues 150 $\in$ *IC*

(c) Las baterías Dutec cargadas duran en promedio más de 2.5 horas.
R/ Se rechaza, pues en el IC hay valores menores o iguales a 150

(d) ¿La duración promedio de las baterías Dutec cargadas puede ser 3 horas? por 160 minutos *R/* Sí, pero es poco factible