40D
1983/2023
40 años de Democracia

UNIVERSIDAD NACIONAL DEL LITORAL
LUX INDEFICIENS · 1919

FICH

ESTADÍSTICA
ENCUENTRO DE TEORÍA

Unidad 6 – *Estadística Inferencial*

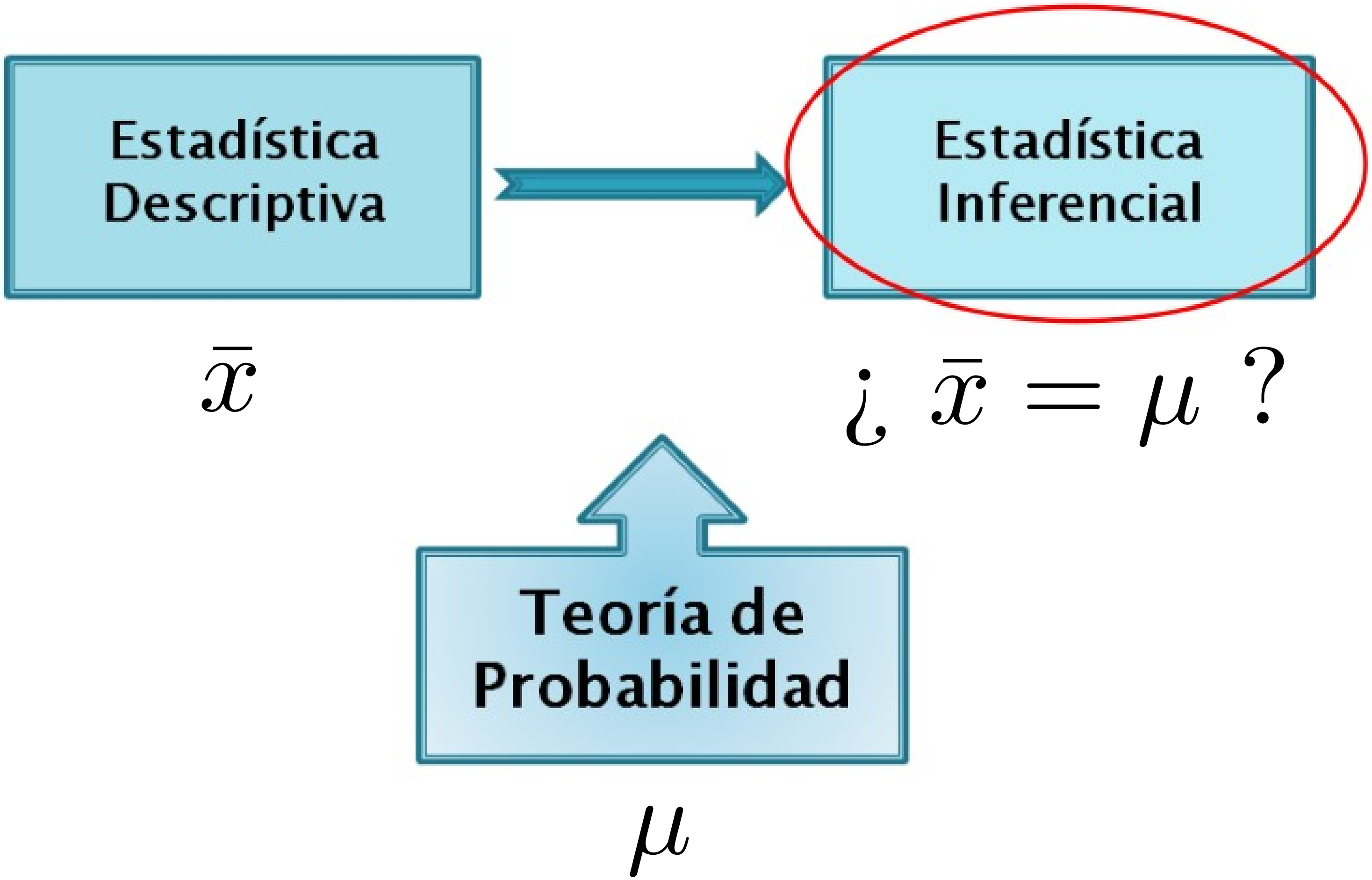
Teoría de Estimación

Ingeniería en Informática

Año 2023

Prof. Juan Pablo Taulamet

consultas: taulamet@unl.edu.ar


Estadística Descriptiva
Estadística Inferencial
$\bar{x}$
¿ $\bar{x} = \mu$?
Teoría de Probabilidad
μ

ESTIMACIÓN PUNTUAL

Consiste en utilizar el valor de una característica que depende de la muestra para estimar un el valor de un parámetro de la población.

Genéricamente:

$$\hat{\theta} = \theta$$

Por ejemplo:

$$\bar{x} = \mu$$

ESTIMACIÓN PUNTUAL

Genéricamente:

$$\hat{\theta} = \theta$$

En nuestro cursado:

$$\bar{x} = \mu \qquad \bar{x} - \bar{y} = \mu_x - \mu_y$$

$$S_x^2 = \sigma^2 \qquad \frac{S_x^2}{S_y^2} = \frac{\sigma_x^2}{\sigma_y^2}$$

$$p = \pi \qquad p_x - p_y = \pi_x - \pi_y$$

ESTIMADOR PUNTUAL

Es es una función de valores observados (muestra) que no depende de ningún parámetro desconocido.

Suponiendo que $\hat{\theta}$ es un estimador de θ, será considerado un “buen estimador” si cumple con las siguientes **propiedades**:

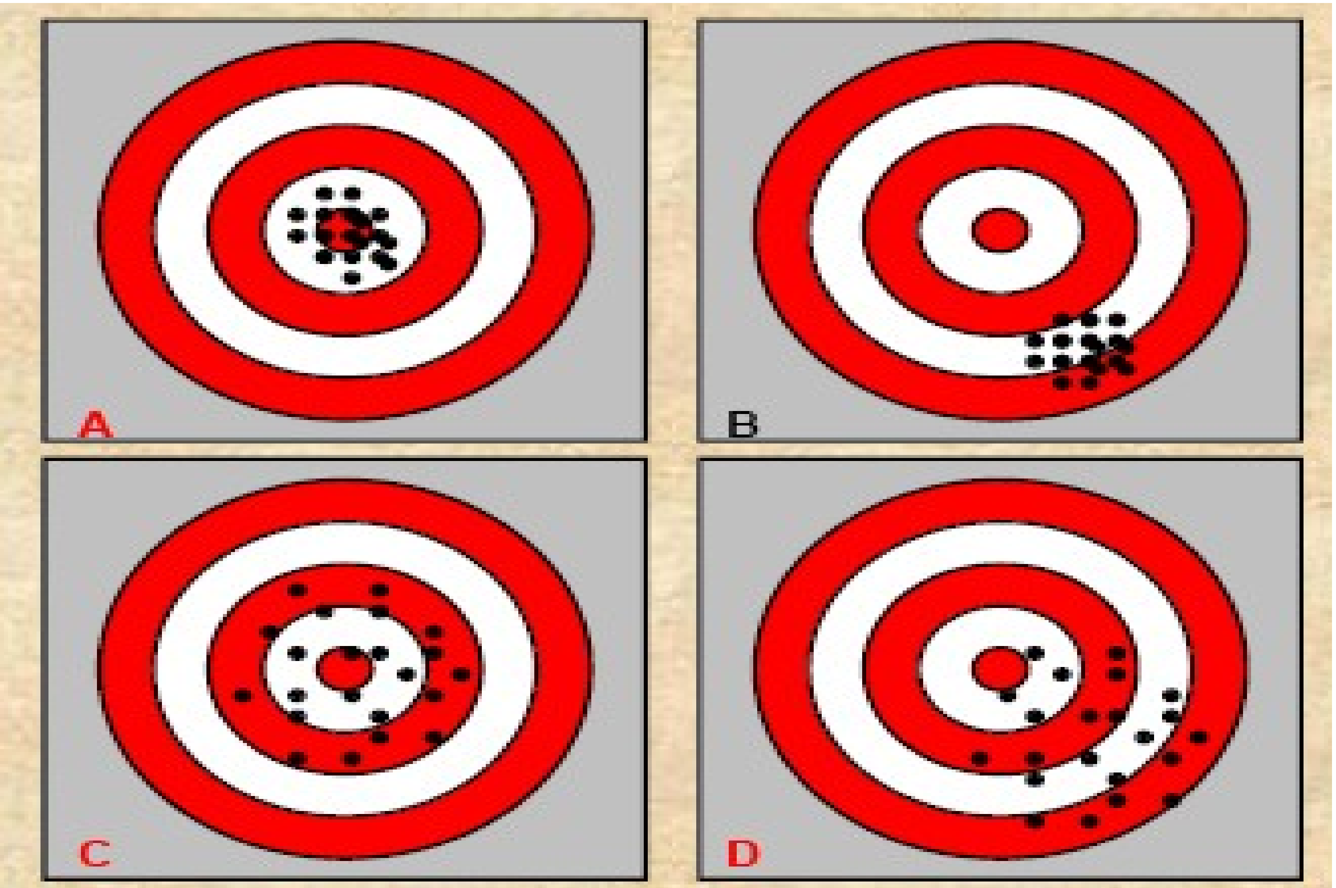
Insesgado: La esperanza del estimador coincide con el parámetro a estimar.

$$E(\hat{\theta}) = \theta$$

Eficiente: Su varianza es lo menor posible.

Consistente: Cuando el tamaño de la muestra crece arbitrariamente, se aproxima al valor del parámetro.

GRÁFICAMENTE



A: Estimador centrado y eficiente;
B: Estimador sesgado y eficiente
C: Estimador centrado e ineficiente;
D: Estimador sesgado e ineficiente

INTERVALOS DE CONFIANZA

El método de estimación por intervalos de confianza se basa en el valor de la estimación puntual pero considera además la distribución en el muestreo del estimador.

Dada una variable aleatoria con distribución normal, podremos estimar ***con una cierta confianza*** un parámetro desconocido de su planteando la siguiente probabilidad:

$$\mathrm{P}\ (\ \mathrm{a} \leq \theta \leq b) = confianza$$

INTERVALOS DE CONFIANZA

Si expresamos la confianza simbólicamente como:

$$1 - \alpha$$

Tendremos:

$$\mathrm{P}\ (\ \mathrm{a} \leq \theta \leq b) = 1 - \alpha$$

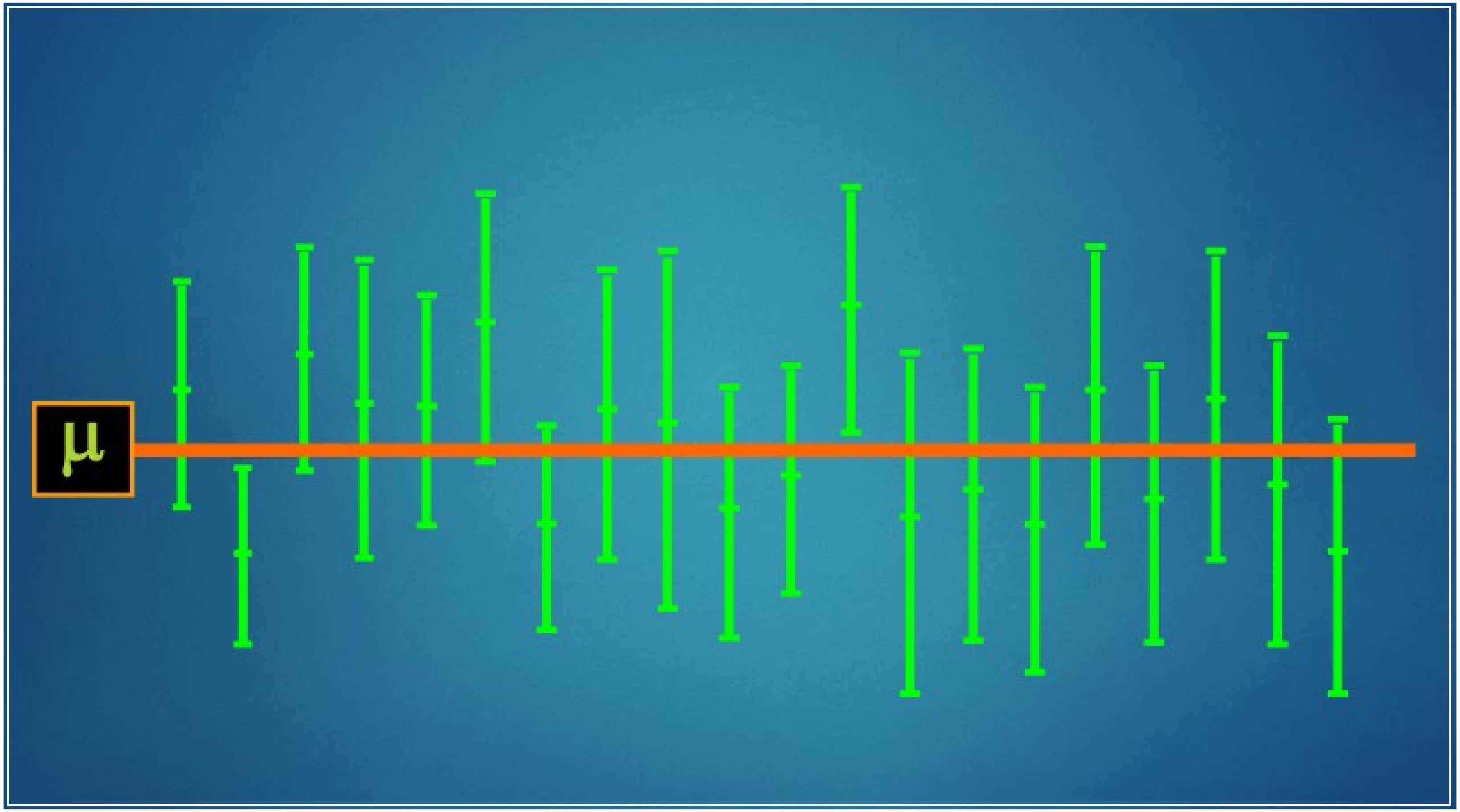
Lo que podría expresarse como:

$$P(|\theta - \hat{\theta}| \leq k\sigma_{\hat{\theta}}) = 1 - \alpha$$

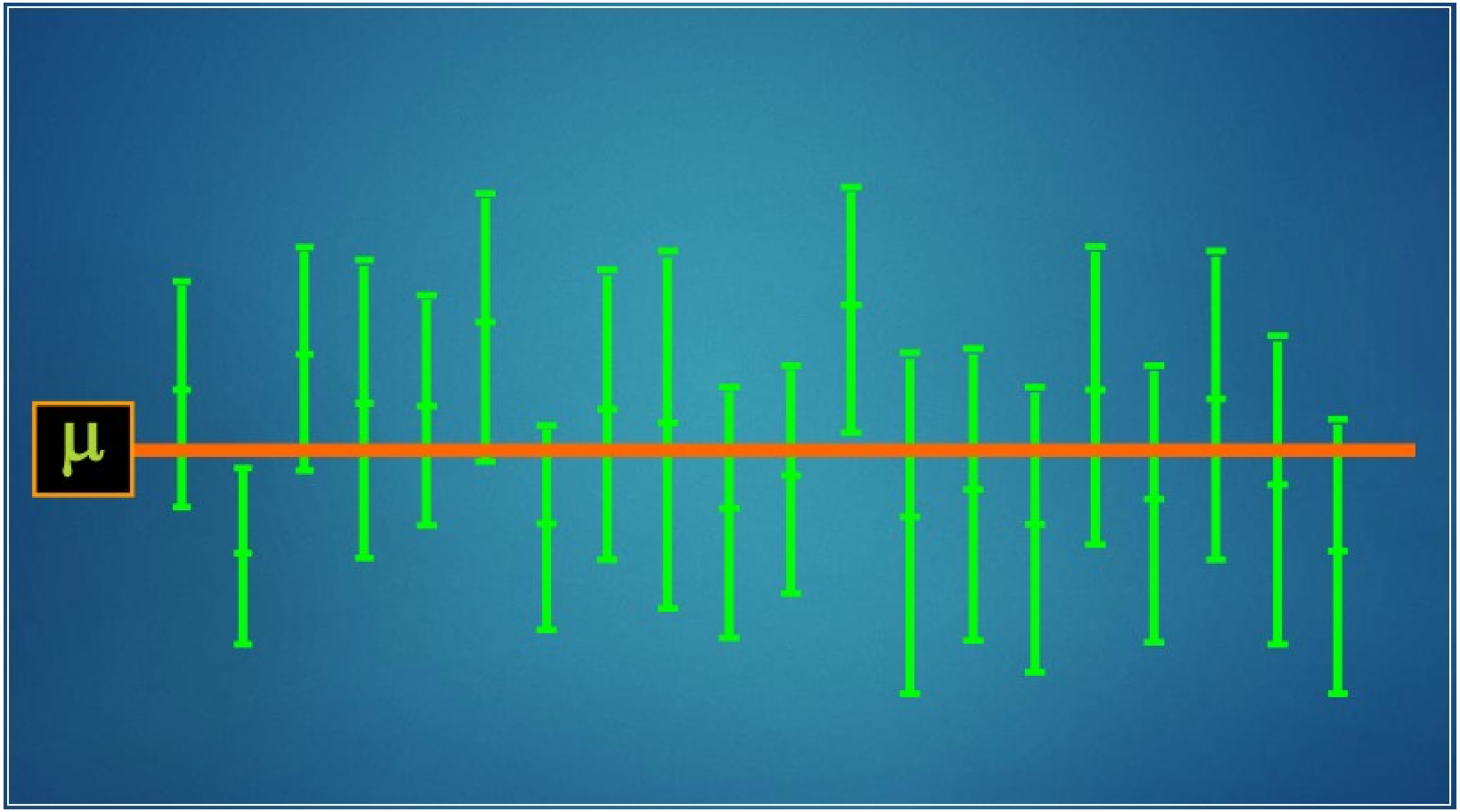
$$P(-k\sigma_{\hat{\theta}} \leq \theta - \hat{\theta} \leq k\sigma_{\hat{\theta}}) = 1 - \alpha$$

Error de estimación: $k\sigma_{\hat{\theta}}$

LA NOCIÓN DE CONFIANZA

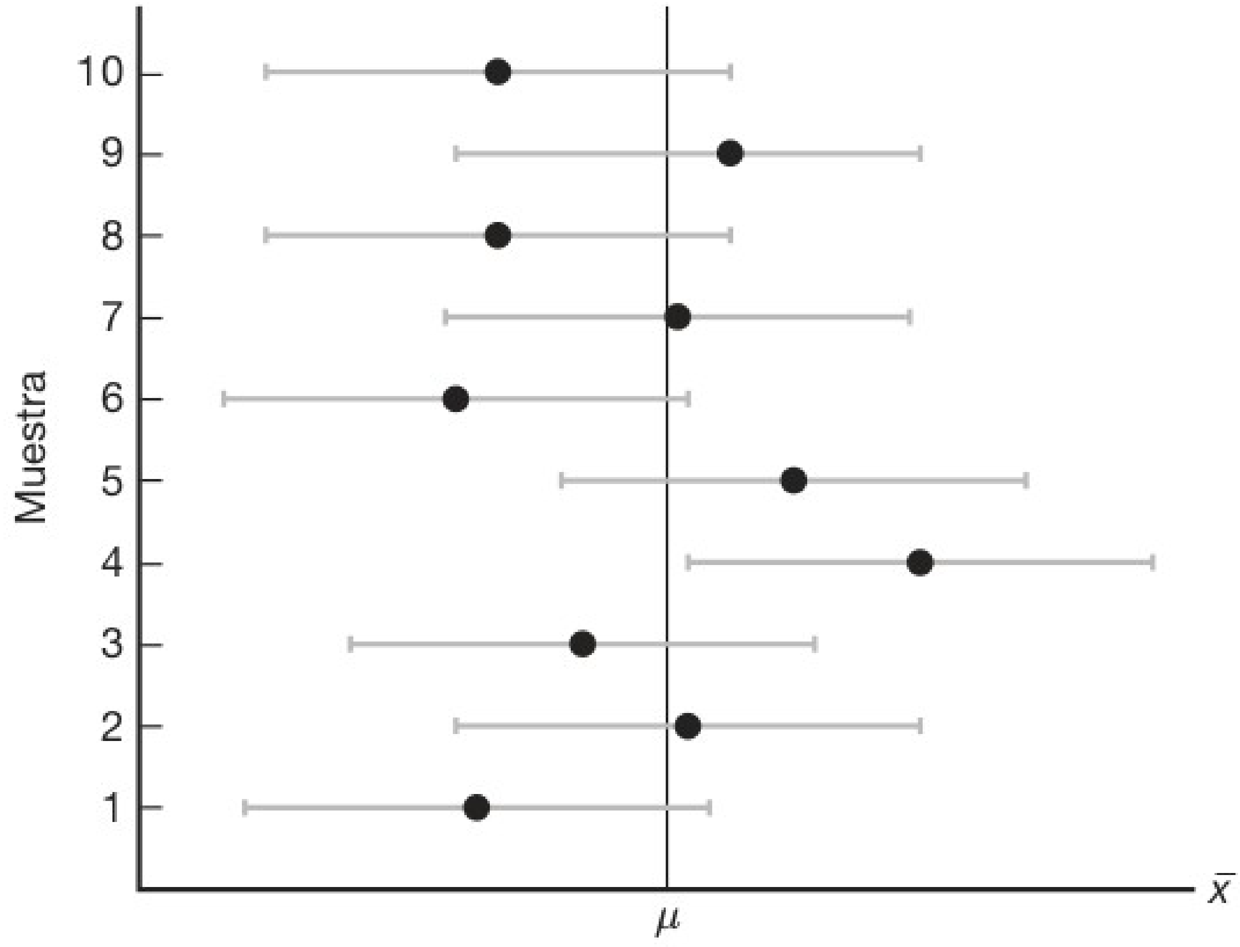


LA NOCIÓN DE CONFIANZA

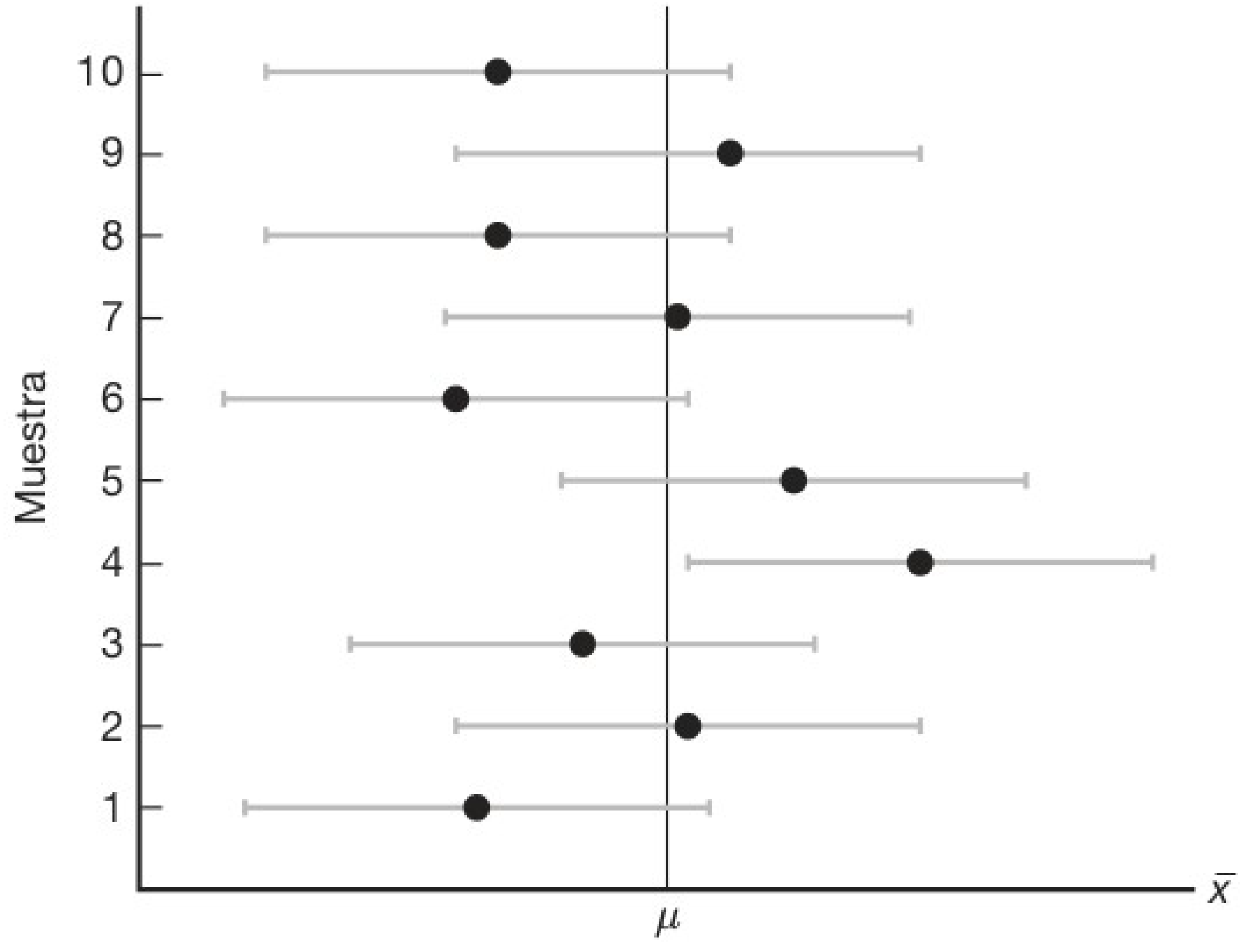


$$1 - \alpha = \frac{18}{20}$$

LA NOCIÓN DE CONFIANZA

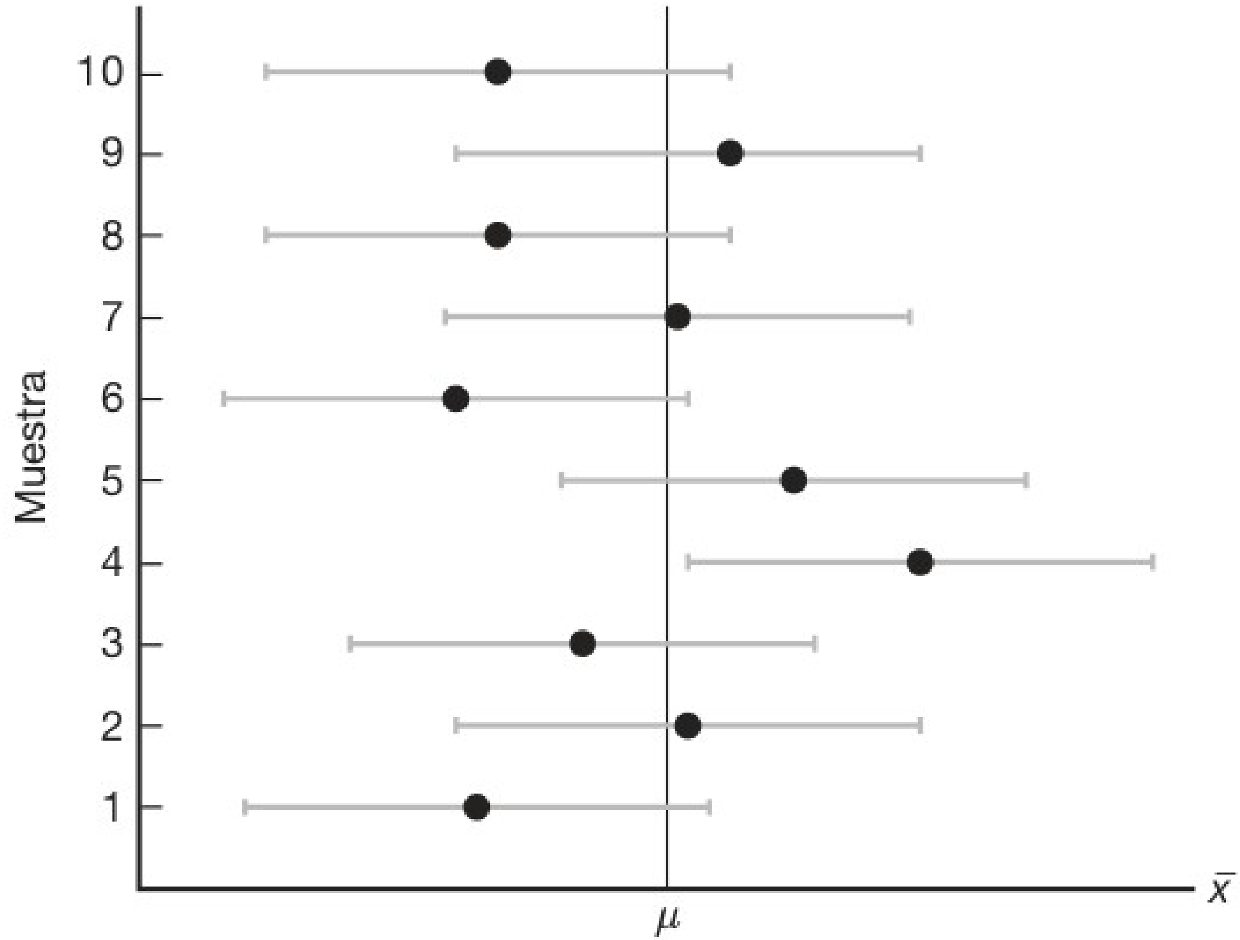


LA NOCIÓN DE CONFIANZA



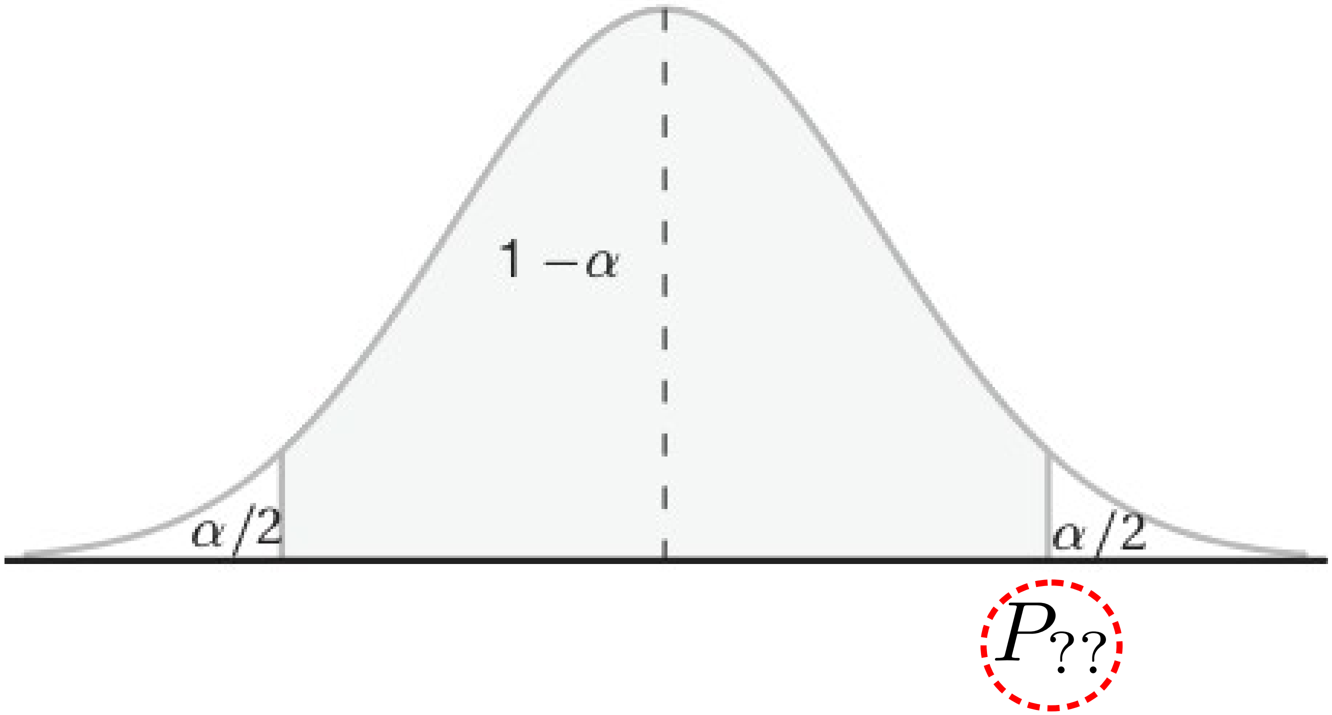
$$1 - \alpha = \frac{9}{10}$$

LA NOCIÓN DE CONFIANZA

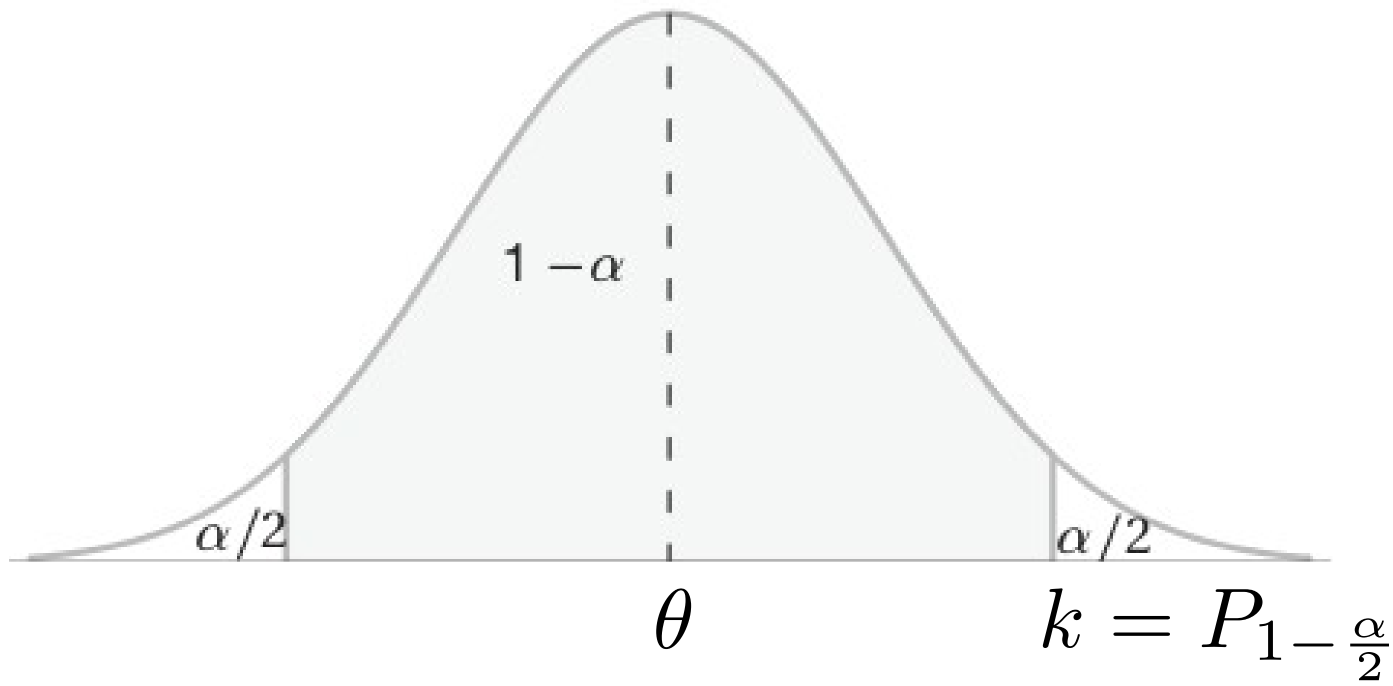


$$1 - \alpha = \frac{9}{10}$$

LA NOCIÓN DE CONFIANZA



INTERVALOS DE CONFIANZA



Genéricamente:

$$(\hat{\theta} - k\sigma_{\hat{\theta}}; \hat{\theta} + k\sigma_{\hat{\theta}})$$

CASOS CONCRETOS

Parámetro	Supuesto	Estimador
μ	σ *conocido*	$\overline{x}$

Intervalo de confianza
$\left(\overline{x} \pm \lvert Z_{\left(1-\frac{\alpha}{2}\right)} \rvert \dfrac{\sigma}{\sqrt{n}} \right)$

CASOS CONCRETOS

Parámetro	Supuesto	Estimador
μ	σ *desconocido*, $n > 30$	$\bar{x}$

Intervalo de confianza
$\left(\bar{x} \pm \left\lvert Z_{\left(1-\frac{\alpha}{2}\right)} \right\rvert \frac{S}{\sqrt{n}}\right)$

CASOS CONCRETOS

Parámetro	**Supuesto**	**Estimador**
μ	σ *desconocido*, $n < 30$	$\bar{x}$

Intervalo de confianza
$\left(\bar{x} \pm \lvert t_{\left(1-\frac{\alpha}{2};n-1\right)} \rvert \dfrac{S'}{\sqrt{n}}\right)$

consultas: taulamet@unl.edu.ar

CASOS CONCRETOS

Parámetro	Supuesto	Estimador
σ^2	Población normal	S^2

Intervalo de confianza
$\left(\frac{n.S^2}{\chi^2_{\left(1-\frac{\alpha}{2};n-1\right)}} \leq \sigma^2 \leq \frac{n.S^2}{\chi^2_{\left(\frac{\alpha}{2};n-1\right)}} \right)$

CASOS CONCRETOS

Parámetro	**Supuesto**	**Estimador**
π	Población normal	p

Intervalo de confianza
$\left(p \;\pm\; \lvert Z_{\left(1-\frac{\alpha}{2}\right)} \rvert \; \sqrt{\frac{\pi(1-\pi)}{n}} \right)$

CASOS CONCRETOS

Parámetro	**Supuesto**	**Estimador**
$\mu_x - \mu_y$	σ_x, σ_y *conocidas*	$\bar{x} - \bar{y}$

Intervalo de confianza
$\left(\left(\bar{x} - \bar{y}\right) \pm \left\| Z_{\left(1-\frac{\alpha}{2}\right)} \right\| \sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}} \right)$

CASOS CONCRETOS

Parámetro	Supuesto	Estimador
$\mu_x - \mu_y$	σ_x, σ_y *desconocidas* $n > 30$	$\bar{x} - \bar{y}$

Intervalo de confianza
$\left((\bar{x} - \bar{y}) \pm \left\| Z_{\left(1-\frac{\alpha}{2}\right)} \right\| \sqrt{\frac{S_x^2}{n_x} + \frac{S_y^2}{n_y}} \right)$

CASOS CONCRETOS

Parámetro	Supuesto	Estimador
$\mu_x - \mu_y$	σ_x, σ_y *desconocidas* *pero iguales* $n < 30$	$\bar{x} - \bar{y}$

Intervalo de confianza
$\left((\bar{x} - \bar{y}) \pm \lvert t_{\left(1-\frac{\alpha}{2}\right)} \rvert \cdot S_w \cdot \sqrt{\frac{1}{n_x} + \frac{1}{n_y}} \right)$

$$Sw = \sqrt{\frac{n_x \cdot S_x^2 + n_y \cdot S_y^2}{n_x + n_y - 2}}$$

CASOS CONCRETOS

Parámetro	Supuesto	Estimador
$\mu_x - \mu_y$	σ_x, σ_y *desconocidas y distintas* $n < 30$	$\bar{x} - \bar{y}$

Intervalo de confianza
$\left((\bar{x} - \bar{y}) \pm \left| t_{\left(1-\frac{\alpha}{2}\right),\upsilon} \right| \sqrt{\frac{S_x^2}{n_x} + \frac{S_y^2}{n_y}} \right)$

consultas: taulamet@unl.edu.ar

CASOS CONCRETOS

Parámetro	**Supuesto**	**Estimador**
$\pi_1 - \pi_2$	*Poblaciones normales*	$p_1 - p_2$

Intervalo de confianza
$\left(\Delta p \pm \left\lvert Z_{\left(1-\frac{\alpha}{2}\right)} \right\rvert \; (I) \right)$

$$(I) = \sqrt{\frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2}}$$

CASOS CONCRETOS

Parámetro	Supuesto	Estimador
$\dfrac{\sigma_x^2}{\sigma_y^2}$	*Poblaciones normales*	$\dfrac{S_x^2}{S_y^2}$

Intervalo de confianza
$\left(\dfrac{S_x^2}{S_y^2}(I) \leq \dfrac{\sigma_x^2}{\sigma_y^2} \leq \dfrac{S_x^2}{S_y^2}(II)\right)$

$$(I) = \frac{1}{F_{\left(\frac{\alpha}{2};n_x-1;n_y-1\right)}}$$

$$(II) = F_{\left(\frac{\alpha}{2};n_x-1;n_y-1\right)}$$

consultas: taulamet@unl.edu.ar

Durante los últimos 10 días de junio un tren llegó tarde a su destino en los siguientes tiempos (en minutos; un número negativo significa que el tren llegó temprano ese número de minutos):

-3	6	4	10	-4	124	2	-1	4	1

Obtener un intervalo de confianza para estimar la esperanza poblacional de los tiempos de de llegada tarde a destino del tren en minutos, con un 90% de confianza.

$$\left(\bar{x} \pm \left|t_{\left(1-\frac{\alpha}{2};n-1\right)}\right| \frac{S'}{\sqrt{n}}\right)$$

consultas: taulamet@unl.edu.ar

$$\left(\bar{x} \pm \left| t_{\left(1-\frac{\alpha}{2};n-1\right)} \right| \frac{S'}{\sqrt{n}} \right)$$

Media	14,3
Error estándar	12,2611

CONF	90,00%
T95	1,83
Error	22,48
LI	-8,2
LS	34,7

Respuesta: Con un 90% de confianza, el promedio poblacional de los tiempos de demora en llegar a destino del tren se econtrará entre -8,2 y 34,7, en símbolos:

$$\text{IC p/ } \mu, 1-\alpha = 90\% : (-8,2; 34,7)$$