(1994–2024)
30 años de la Consagración Constitucional de la Autonomía y Autarquía Universitaria en Argentina.

UNIVERSIDAD NACIONAL DEL LITORAL
LUX INDEFICIENS · 1919

FICH

ESTADÍSTICA
PROBABILIDAD Y ESTADÍSTICA

Clase#13 - Unidad 6 - *Estadística Inferencial*
Distribución en el muestreo - Intervalos de confianza

Ingeniería en Informática

Año 2024

Prof. Juan Pablo Taulamet

consultas: taulamet@unl.edu.ar


Ud.
La promoción

Problema

Dada la siguiente población de elementos:

$$X = \{22, 24, 26\}$$

a) Enumerar todas las muestras posibles de tamaño dos, escogidas mediante muestreo aleatorio simple.

b) Obtener la esperanza y varianza de la población y de la media muestral.

Planteo

a) Diferenciaremos los dos casos de muestreo: con y sin reposición.

b) Se pide:

$$E(X), V(X), E(\bar{X}), V(\bar{X})$$

Es decir:

$$\mu, \sigma^2, \mu_{\bar{X}}, \sigma^2_{\bar{X}}$$

Planteo

Por último intentaremos confirmar que:

$$\mu_{\bar{X}} = \mu$$

$$\sigma^2_{\bar{X}} = \frac{\sigma^2}{n} \text{ (Con Rep.)}$$

$$\sigma^2_{\bar{X}} = \frac{\sigma^2}{n} * \frac{N-n}{N-1} \text{ (Sin Rep.)}$$

Esperanza de X

Los valores de la población tienen igual probabilidad de ocurrencia por tanto su función de cuantía es:

X	22	24	26
$f(X)$	$1/3$	$1/3$	$1/3$

Luego:

$$\mu = \alpha_1 = \sum_{\forall x_i} x_i^1 * f(x_i) =$$

$$= \frac{22}{3} + \frac{24}{3} + \frac{26}{3} = 24$$

Varianza de X

$$\sigma^2 = \alpha_2 - \alpha_1^2 =$$

$$[\sum_{\forall x_i} x_i^2 * f(x_i)] - 24^2 =$$

$$= [\frac{22^2}{3} + \frac{24^2}{3} + \frac{26^2}{3}] - 576 =$$

$$\frac{8}{3} \approx 2.67$$

Esperanza de la media muestral

(Caso con reposición)

M1: 22, 24 $\bar{X}_1 = 23$
M2: 22, 26 $\bar{X}_2 = 24$
M3: 24, 26 $\bar{X}_3 = 25$

M4: 22, 22 $\bar{X}_4 = 22$
M5: 24, 24 $\bar{X}_5 = 24$
M6: 26, 26 $\bar{X}_6 = 26$

M7: 24, 22 $\bar{X}_7 = 23$
M8: 26, 22 $\bar{X}_8 = 24$
M9: 26, 24 $\bar{X}_9 = 25$

X	22	23	24	25	26
$f(X)$	1/9	2/9	3/9	2/9	1/9

$$\mu_{\bar{X}} = \sum_{\forall \bar{x}_i} \bar{X}_i * f(\bar{X}_i) =$$

$$= \frac{22*1}{9} + \frac{23*2}{9} + \frac{24*3}{9} + \frac{25*2}{9} + \frac{26*1}{9} = 24$$

Varianza de media muestral

(Caso <u>con</u> reposición)

$$\sigma^2 = \alpha_2 - \alpha_1^2 =$$

$$[\sum_{\forall \bar{x}_i} x_i^2 * f(\bar{x}_i)] - 24^2 =$$

$$= [\frac{22^2 * 1}{9} + \frac{23^2 * 2}{9} + \frac{24^2 * 3}{9} +$$

$$\frac{25^2 * 2}{9} + \frac{26^2 * 1}{9}] - 576 = \frac{4}{3} \approx 1.34$$

consultas: taulamet@unl.edu.ar

Esperanza de la media muestral

(Caso sin reposición)

M1: 22, 24 $\overline{X}_1 = 23$

M2: 22, 26 $\overline{X}_2 = 24$

M3: 24, 26 $\overline{X}_3 = 25$

M4: 24, 22 $\overline{X}_4 = 23$

M5: 26, 22 $\overline{X}_5 = 24$

M6: 26, 24 $\overline{X}_6 = 25$

X	23	24	25
$f(X)$	$2/6$	$2/6$	$2/6$

$$\mu_{\bar{X}} = \sum_{\forall \bar{x}_i} \bar{X}_i * f(\bar{X}_i) =$$

$$= \frac{23*2}{6} + \frac{24*2}{6} + \frac{25*2}{6} = 24$$

Varianza de media muestral

(Caso sin reposición)

$$\sigma^2 = \alpha_2 - \alpha_1^2 =$$

$$[\sum_{\forall \bar{x}_i} x_i^2 * f(\bar{x}_i)] - 24^2 =$$

$$= [\frac{23^2 * 2}{6} + \frac{24^2 * 2}{6} + \frac{25^2 * 2}{6}] - 576 =$$

$$\frac{2}{3} \approx 0.67$$

Conclusiones

$$\mu = 24 \qquad \sigma^2 = \frac{8}{3}$$

Caso con rep.

$\mu_{\bar{X}} = 24$

$\sigma^2_{\bar{X}} = 4/3$

n = 2

Se verifica que:

$\mu_{\bar{X}} = \mu$

$\sigma^2_{\bar{X}} = \frac{\sigma^2}{n}$

Caso sin rep.

$\mu_{\bar{X}} = 24$

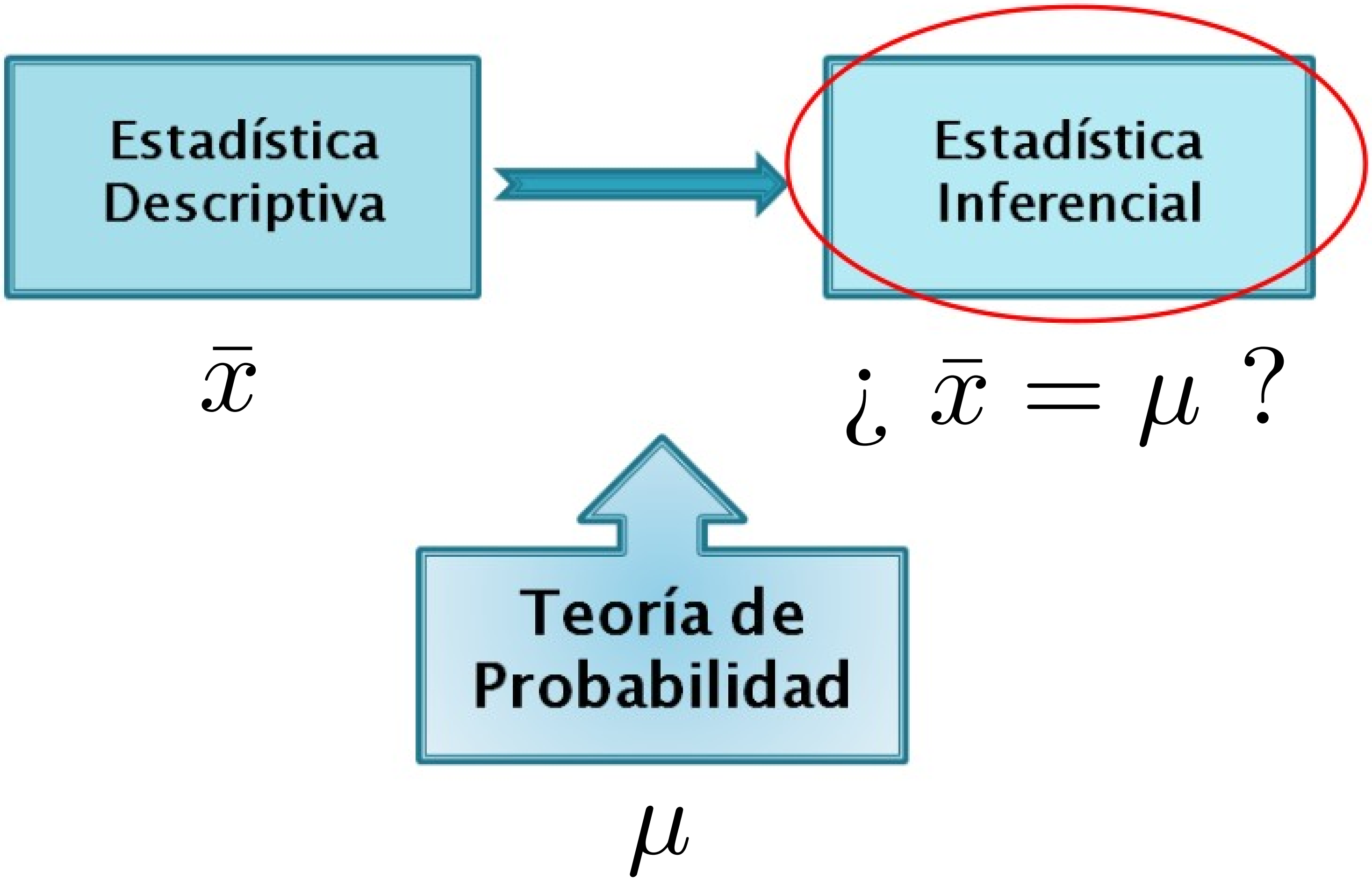
$\sigma^2_{\bar{X}} = 2/3$

n = 2

$\frac{N-n}{N-1} = \frac{3-2}{3-1} = \frac{1}{2}$

Se verifica que:

$\sigma^2_{\bar{X}} = \frac{\sigma^2}{n} * \frac{N-n}{N-1}$



consultas: taulamet@unl.edu.ar

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**:

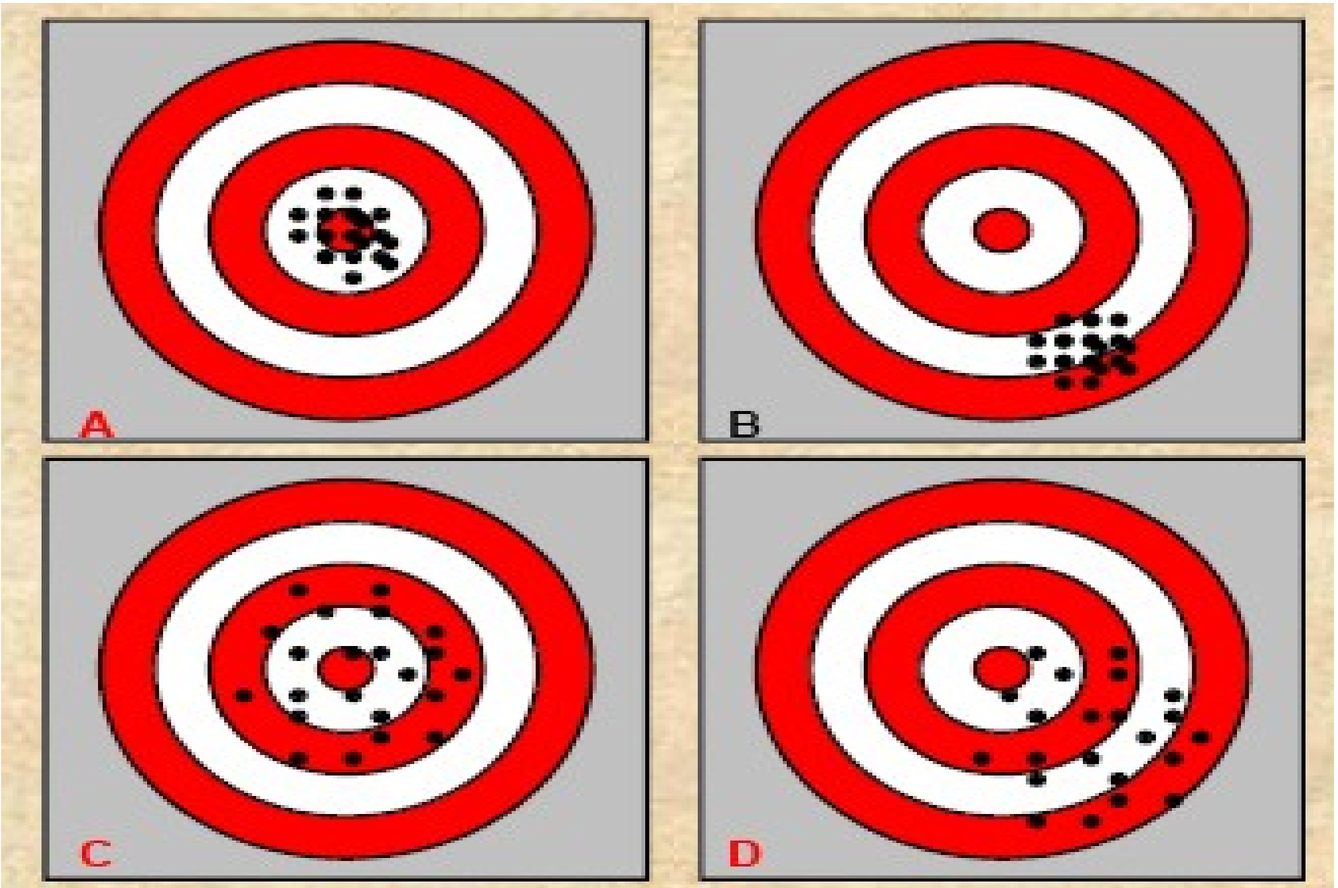
Insesgado: La esperanza del estimador coincide con el parámetro a estimar.

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

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:

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

consultas: taulamet@unl.edu.ar

INTERVALOS DE CONFIANZA

Si expresamos la confianza simbólicamente como:

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

Tendremos:

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

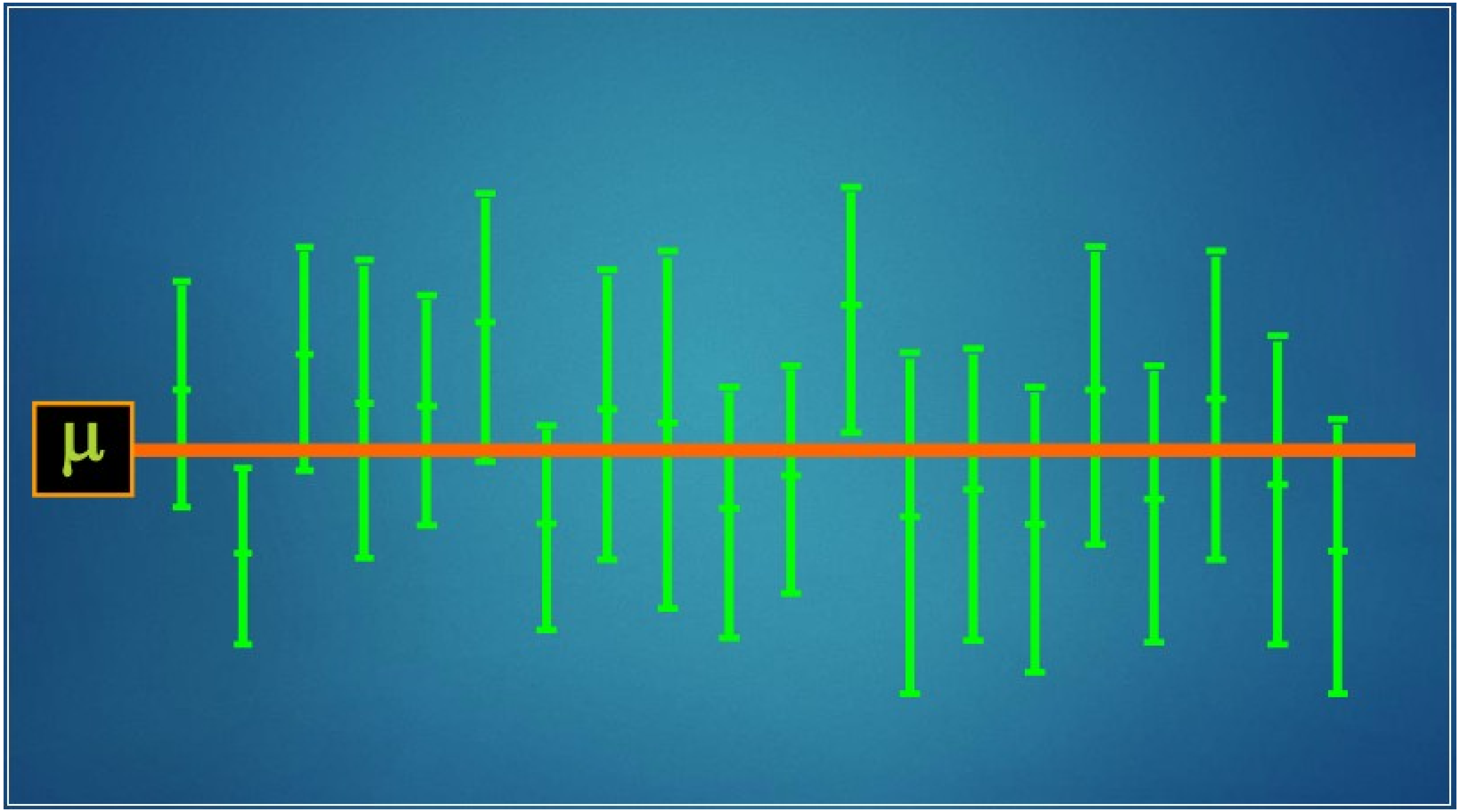
Lo que podría expresarse como:

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

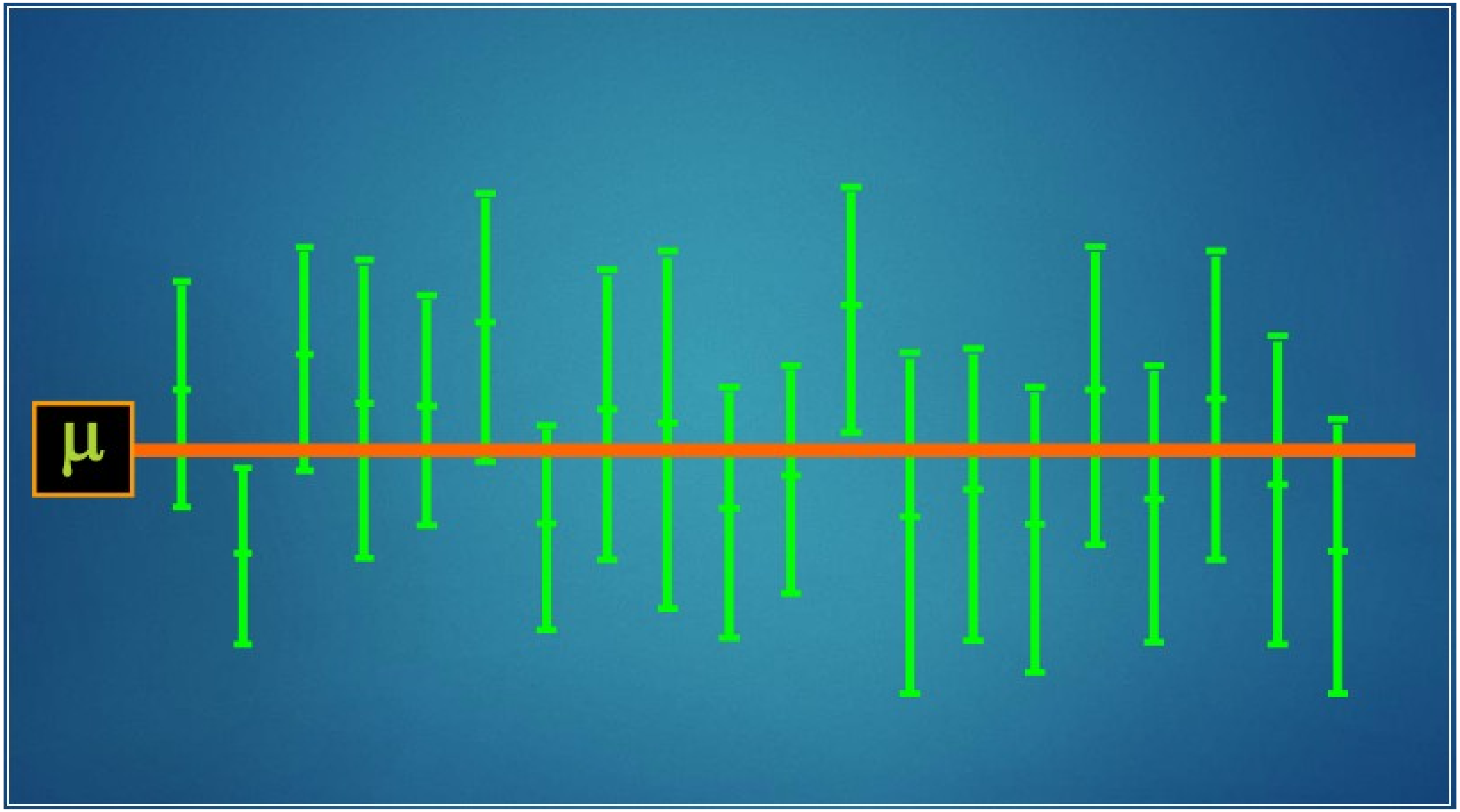
$$P(-k\sigma_{\hat{\theta}} \leq \theta - \hat{\theta} \leq k\sigma_{\hat{\theta}}) = 1 - \alpha$$

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

LA NOCIÓN DE CONFIANZA

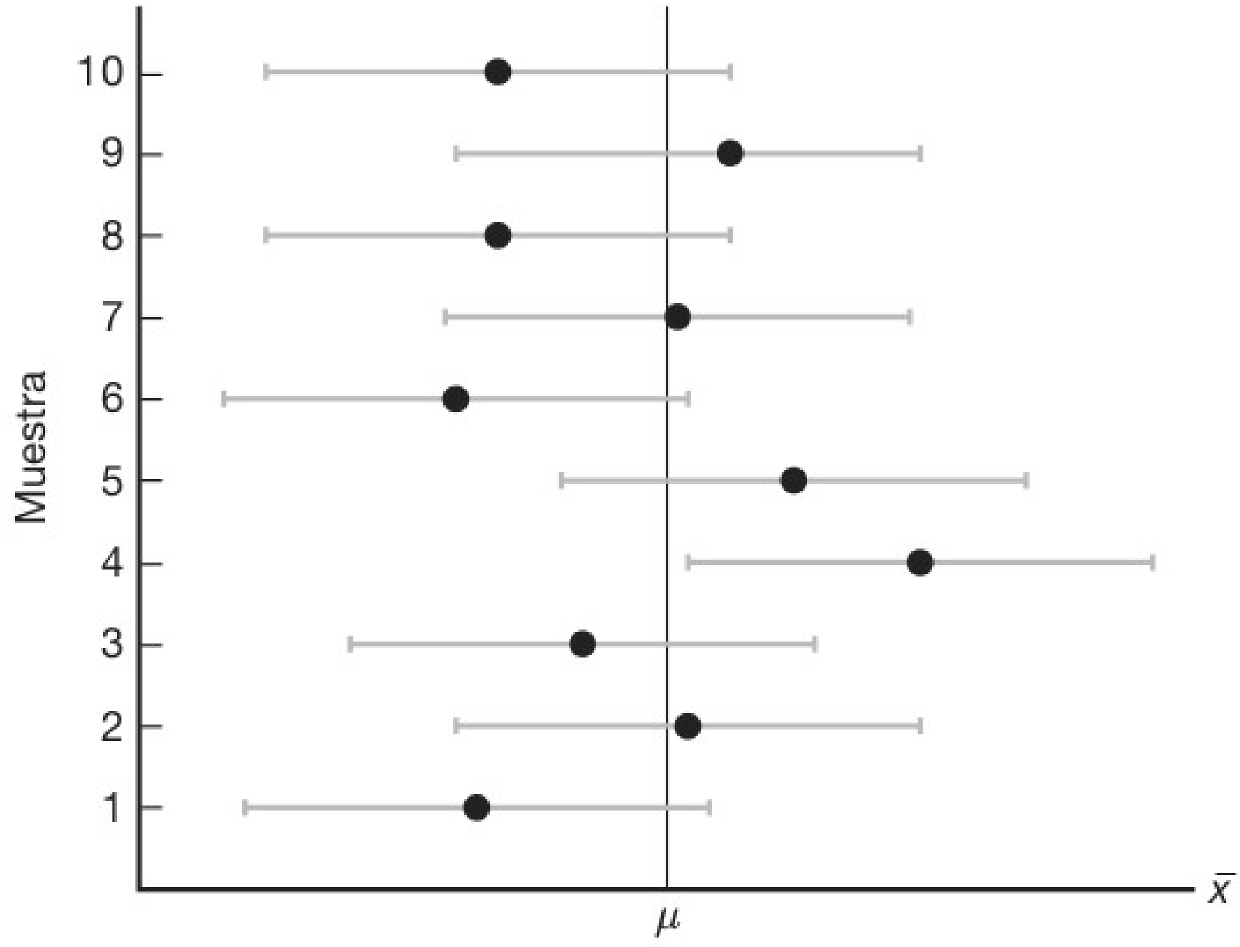


LA NOCIÓN DE CONFIANZA



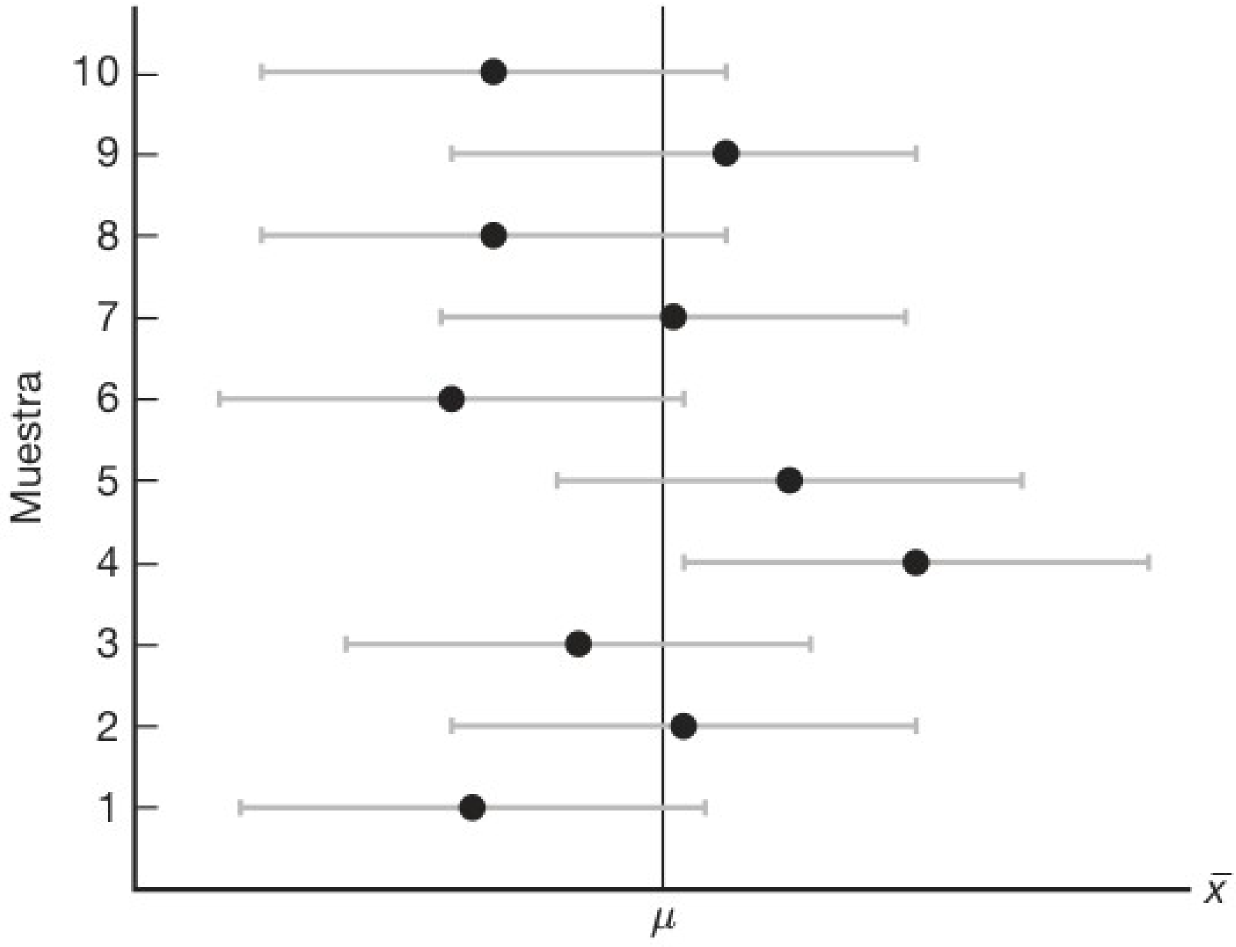
$$1 - \alpha = \frac{18}{20}$$

LA NOCIÓN DE CONFIANZA



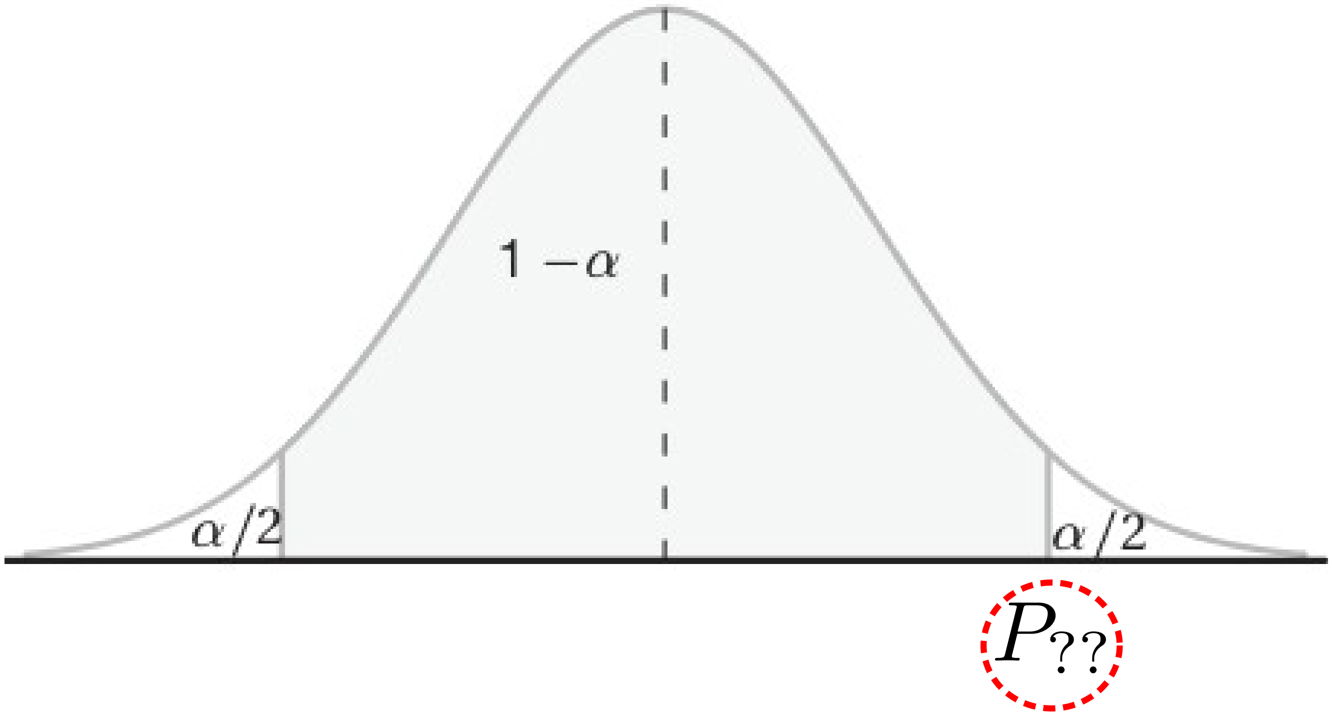
consultas: taulamet@unl.edu.ar

LA NOCIÓN DE CONFIANZA

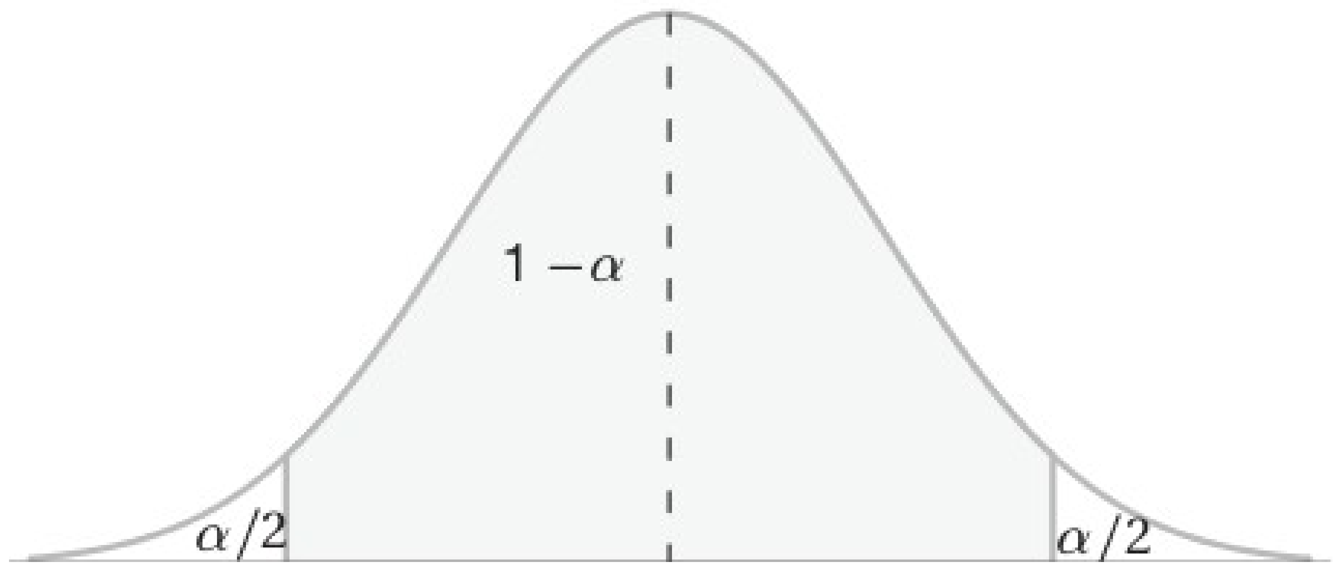


$$1 - \alpha = \frac{9}{10}$$

LA NOCIÓN DE CONFIANZA



INTERVALOS DE CONFIANZA



θ $\quad k = P_{1-\frac{\alpha}{2}}$

Genéricamente:

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

consultas: taulamet@unl.edu.ar

CASOS CONCRETOS

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

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

CASOS CONCRETOS

Parámetro	Supuesto	Estimador
μ	$\sigma \; desconocido, n > 30$	$\bar{x}$

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

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

CASOS CONCRETOS

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

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

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

CASOS CONCRETOS

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

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

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

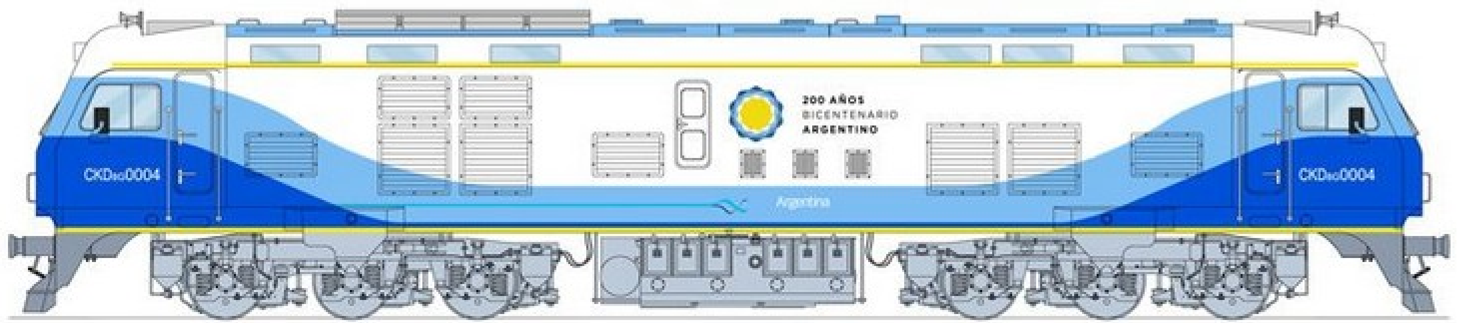
CASOS CONCRETOS

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

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

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

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



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:

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

consultas: taulamet@unl.edu.ar