



2025 ~ 40° Aniversario
de la Creación del Consejo
Interuniversitario Nacional



ESTADÍSTICA PROBABILIDAD Y ESTADÍSTICA

Notas de clase - *Estadística Inferencial*
Distribución en el muestreo - Intervalos de confianza

Ingenierías en: Recursos Hídricos, Informática, Ambiental
y Agrimensura

Año 2025

Prof. Juan Pablo Taulamet

consultas: taulamet@unl.edu.ar

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:

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

Planteo

Por último intentaremos confirmar que:

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

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

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

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:

$$\begin{aligned}\mu = \alpha_1 &= \sum_{\forall x_i} x_i^1 * f(x_i) = \\ &= \frac{22}{3} + \frac{24}{3} + \frac{26}{3} = 24\end{aligned}$$

Varianza de X

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

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

$$= \left[\frac{22^2}{3} + \frac{24^2}{3} + \frac{26^2}{3} \right] - 576 =$$
$$\frac{8}{3} \approx 2.67$$

Esperanza de la media muestral

(Caso con reposición)

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

M4: 22, 22 $\bar{X}_4 = 22$

M7: 24, 22 $\bar{X}_7 = 23$

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

M5: 24, 24 $\bar{X}_5 = 24$

M8: 26, 22 $\bar{X}_8 = 24$

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

M6: 26, 26 $\bar{X}_6 = 26$

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

$$\begin{aligned}\mu_{\bar{X}} &= \sum_{\forall \bar{x}_i} \bar{X}_i * f(\bar{X}_i) = \\ &= \frac{22*1}{9} + \frac{23*2}{9} + \frac{24*3}{9} + \frac{25*2}{9} + \frac{26*1}{9} = 24\end{aligned}$$

Varianza de media muestral

(Caso con reposición)

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

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

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

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

Esperanza de la media muestral

(Caso sin reposición)

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

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

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

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

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

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

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

$$\begin{aligned}\mu_{\bar{X}} &= \sum_{\forall \bar{x}_i} \bar{X}_i * f(\bar{X}_i) = \\ &= \frac{23*2}{6} + \frac{24*2}{6} + \frac{25*2}{6} = 24\end{aligned}$$

Varianza de media muestral

(Caso sin reposición)

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

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

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

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

Conclusiones

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

Caso con rep.

$$\begin{aligned}\mu_{\bar{X}} &= 24 \\ \sigma_{\bar{X}}^2 &= 4/3 \\ n &= 2\end{aligned}$$

Se verifica que:

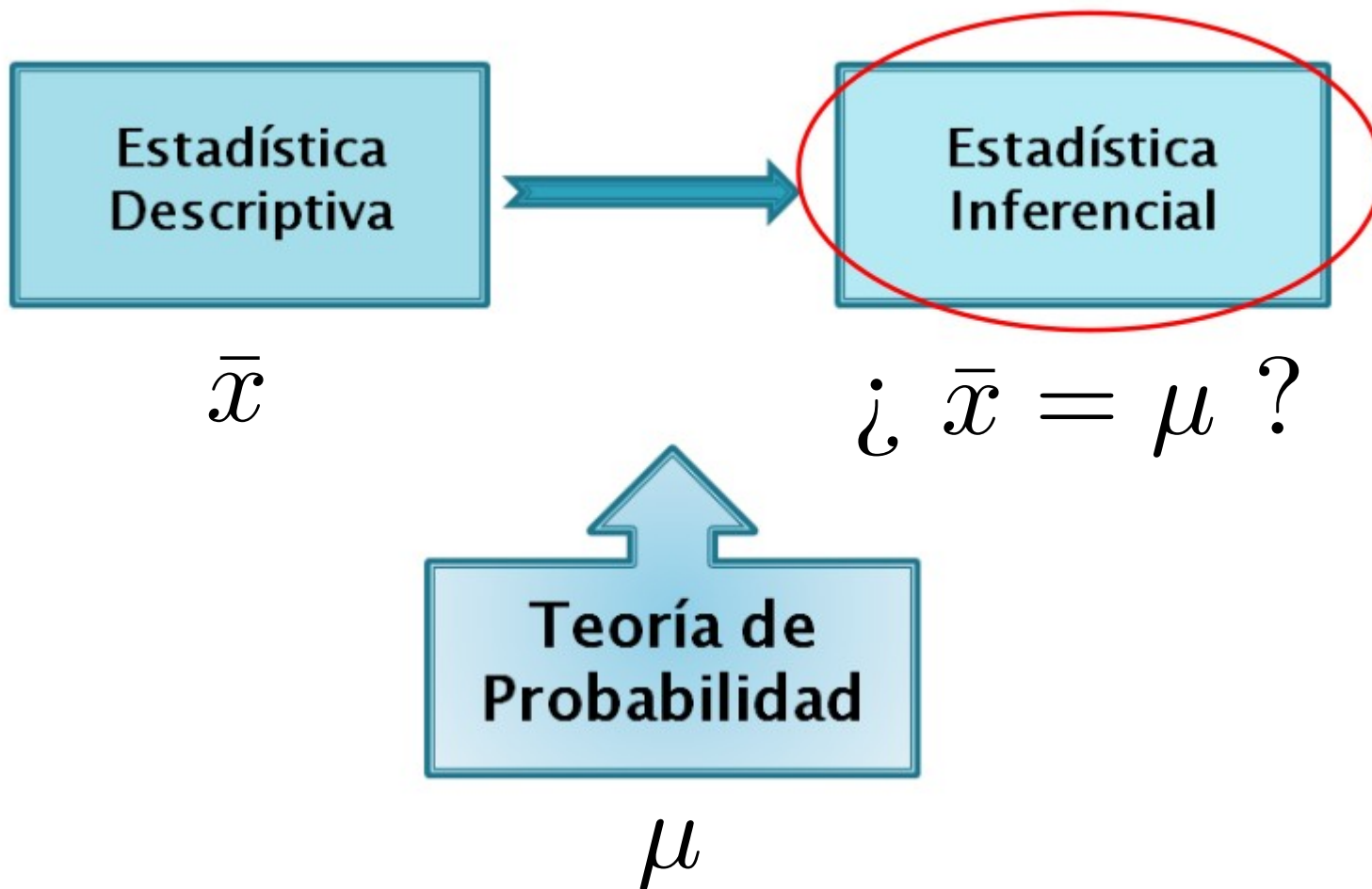
$$\begin{aligned}\mu_{\bar{X}} &= \mu \\ \sigma_{\bar{X}}^2 &= \frac{\sigma^2}{n}\end{aligned}$$

Caso sin rep.

$$\begin{aligned}\mu_{\bar{X}} &= 24 \\ \sigma_{\bar{X}}^2 &= 2/3 \\ n &= 2 \\ \frac{N-n}{N-1} &= \frac{3-2}{3-1} = \frac{1}{2}\end{aligned}$$

Se verifica que:

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



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: ¡Aparecen nuevas VA!

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

$$S_x^2 = \sigma^2$$

$$p = \pi$$

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

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

$$p_x - p_y = \pi_x - \pi_y$$

ESTIMADOR PUNTUAL

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

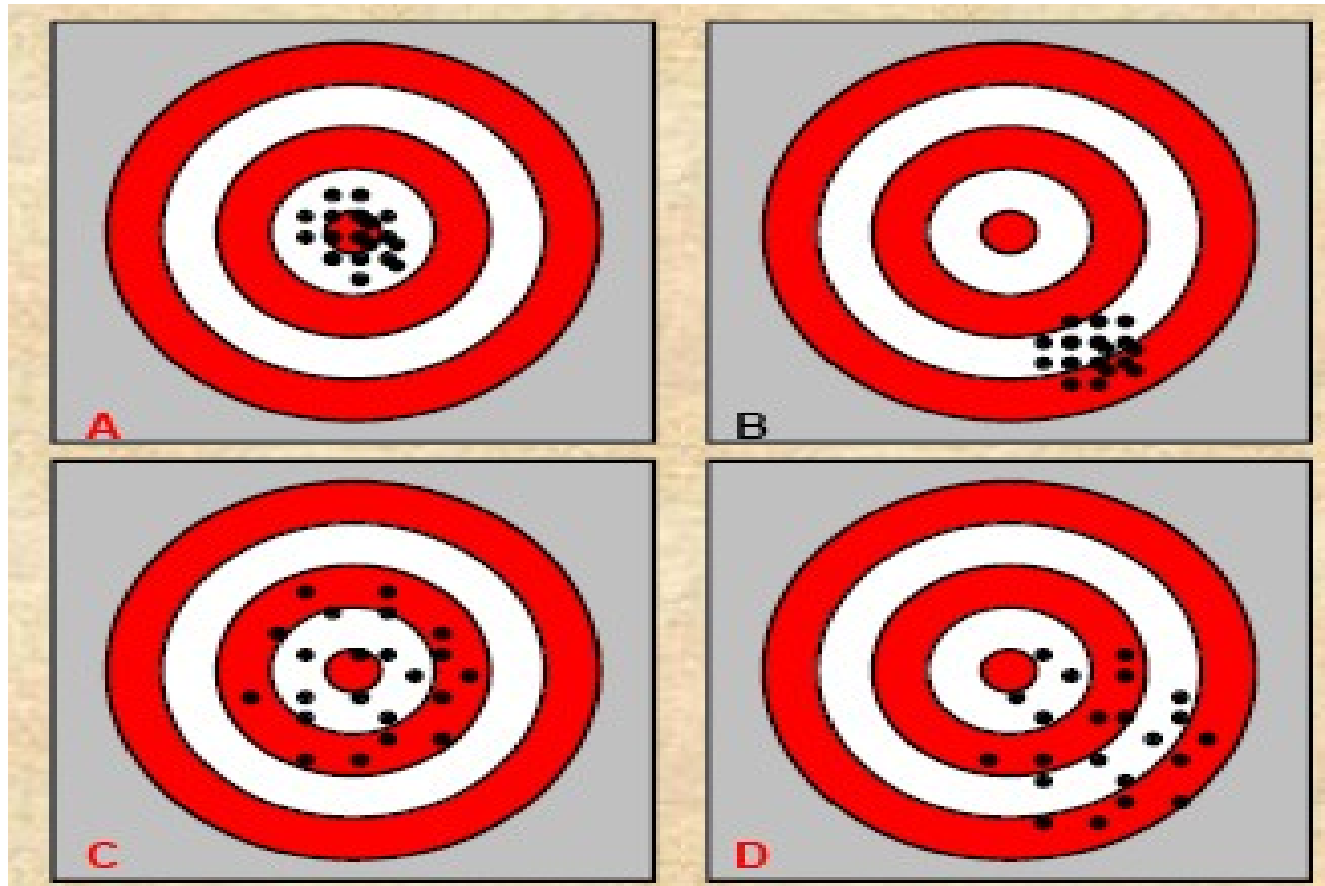
Suponiendo que $\hat{\theta}$ es un estimador de θ , será considerado un “buen estimador” si cumple con las siguientes **propiedades**:

Insesgado: La esperanza del estimador coincide con el parámetro $E(\hat{\theta}) = \theta$ a estimar.

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 la misma planteando la siguiente probabilidad:

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

INTERVALOS DE CONFIANZA

Si definimos la confianza como: $1 - \alpha$

Tendremos: $P(a \leq \theta \leq b) = 1 - \alpha$

Lo que podría expresarse como:

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

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

Siendo,

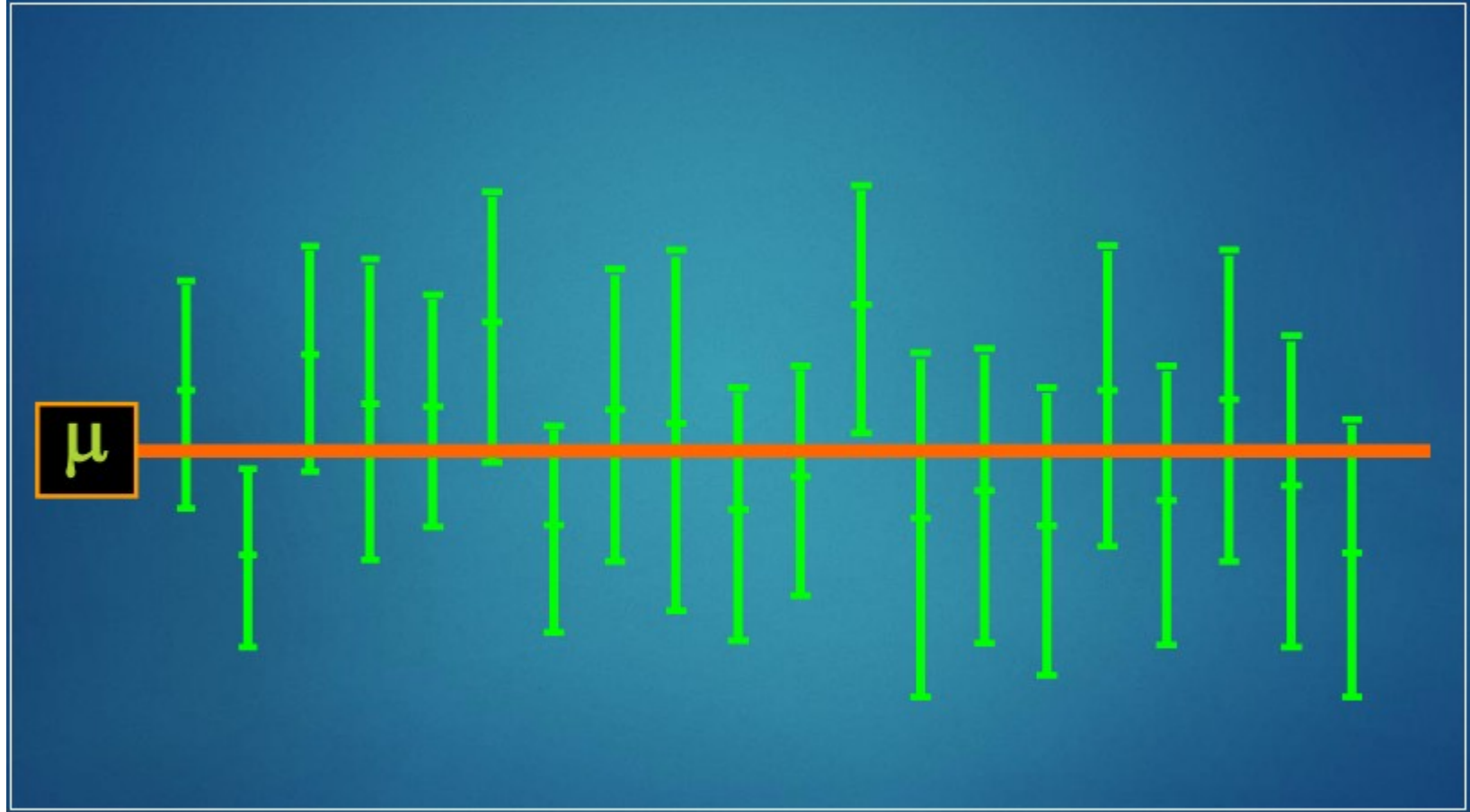
Error de estimación: $k\sigma_{\hat{\theta}}$

Error del estimador: $\sigma_{\hat{\theta}}$

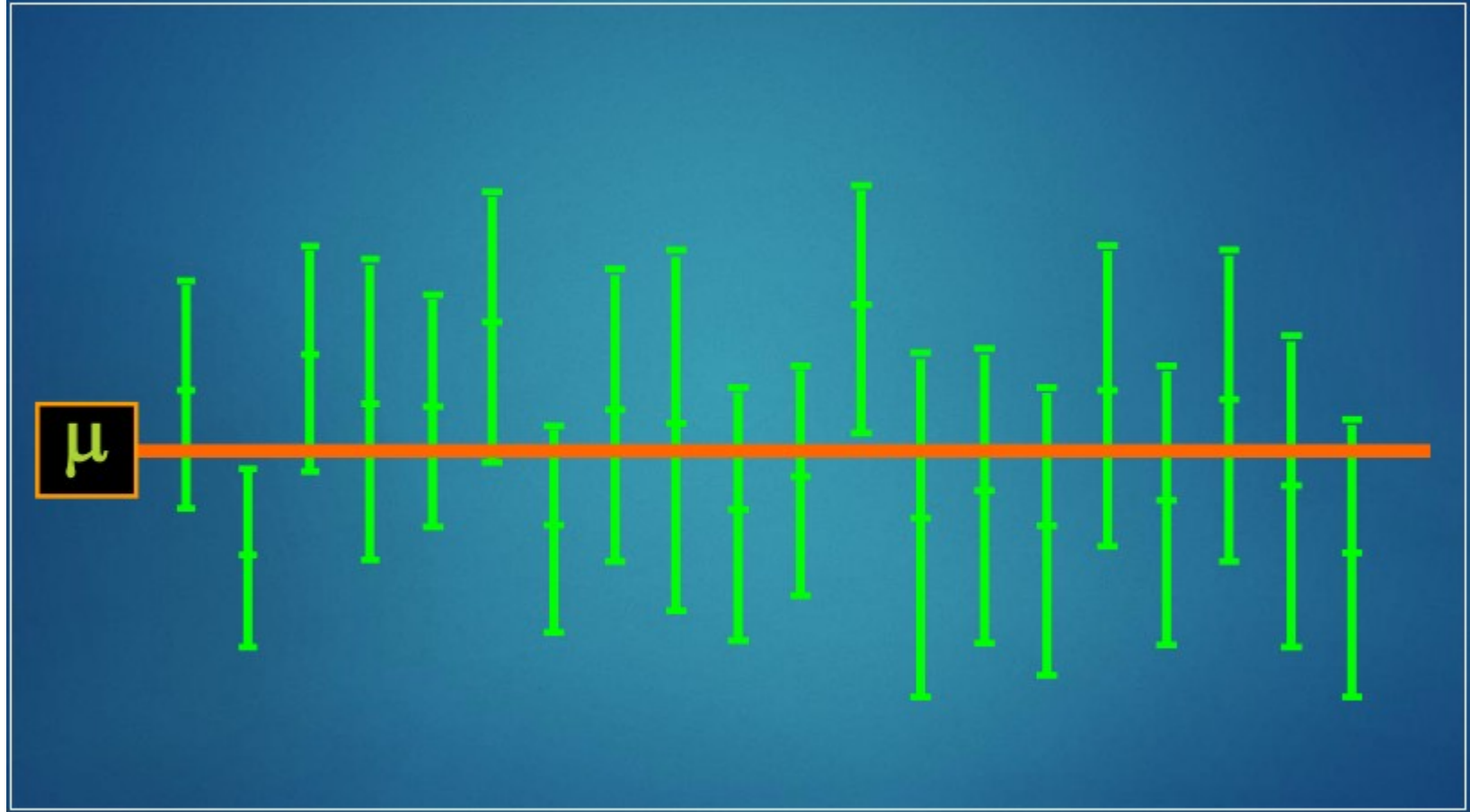
Porcentil de la confianza: k

Amplitud del intervalo: $2k\sigma_{\hat{\theta}}$

LA NOCIÓN DE CONFIANZA

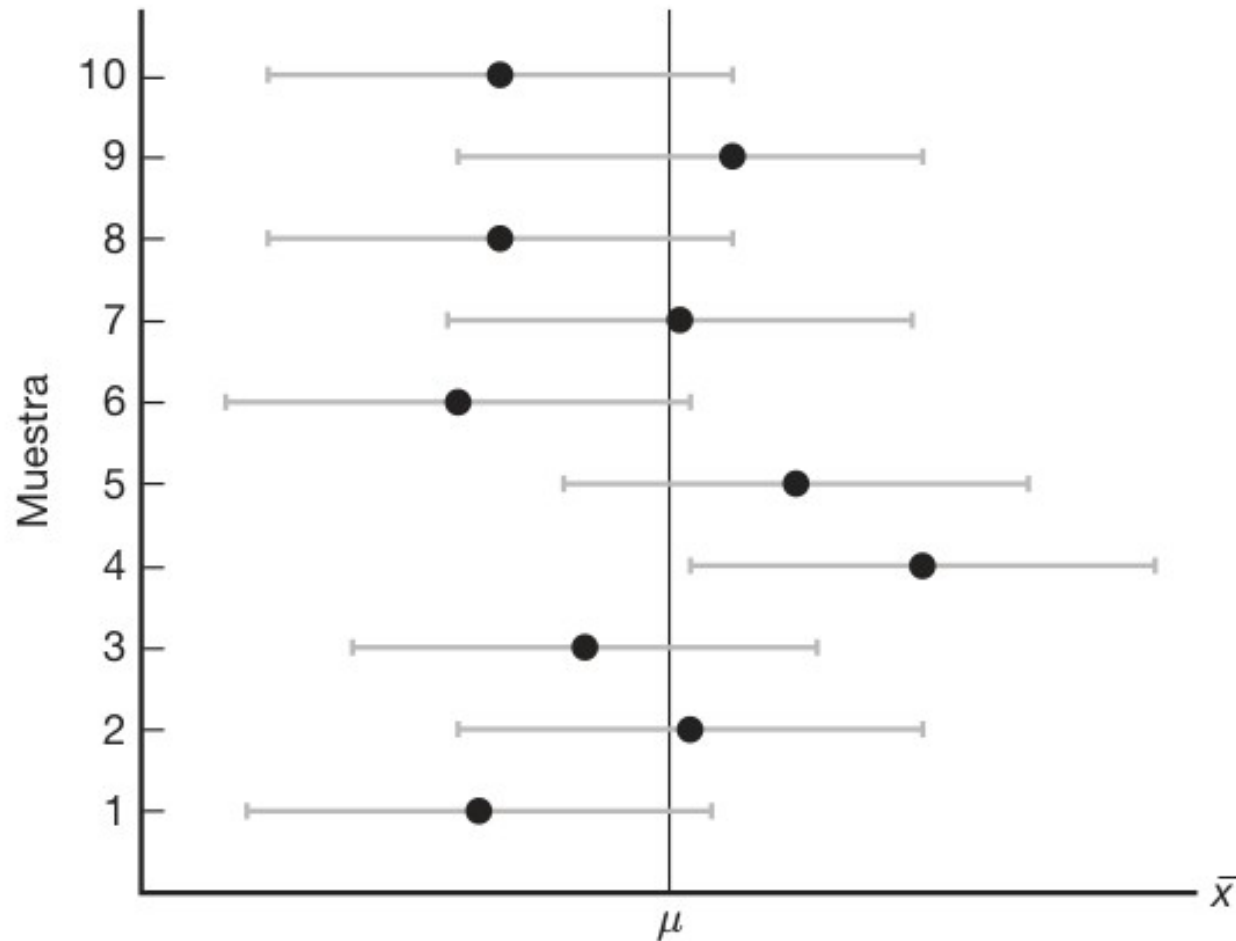


LA NOCIÓN DE CONFIANZA

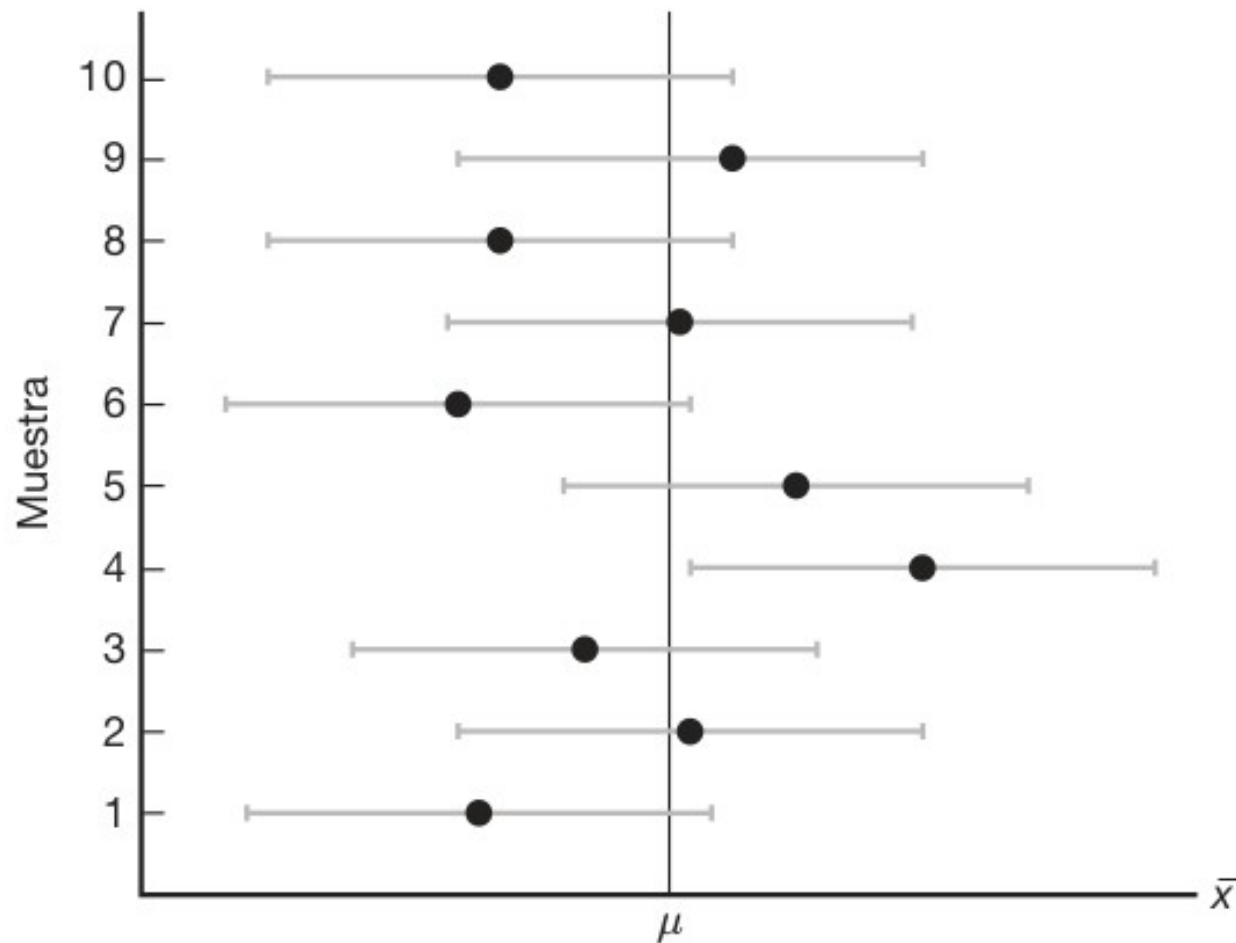


$$1 - \alpha = \frac{18}{20}$$

LA NOCIÓN DE CONFIANZA

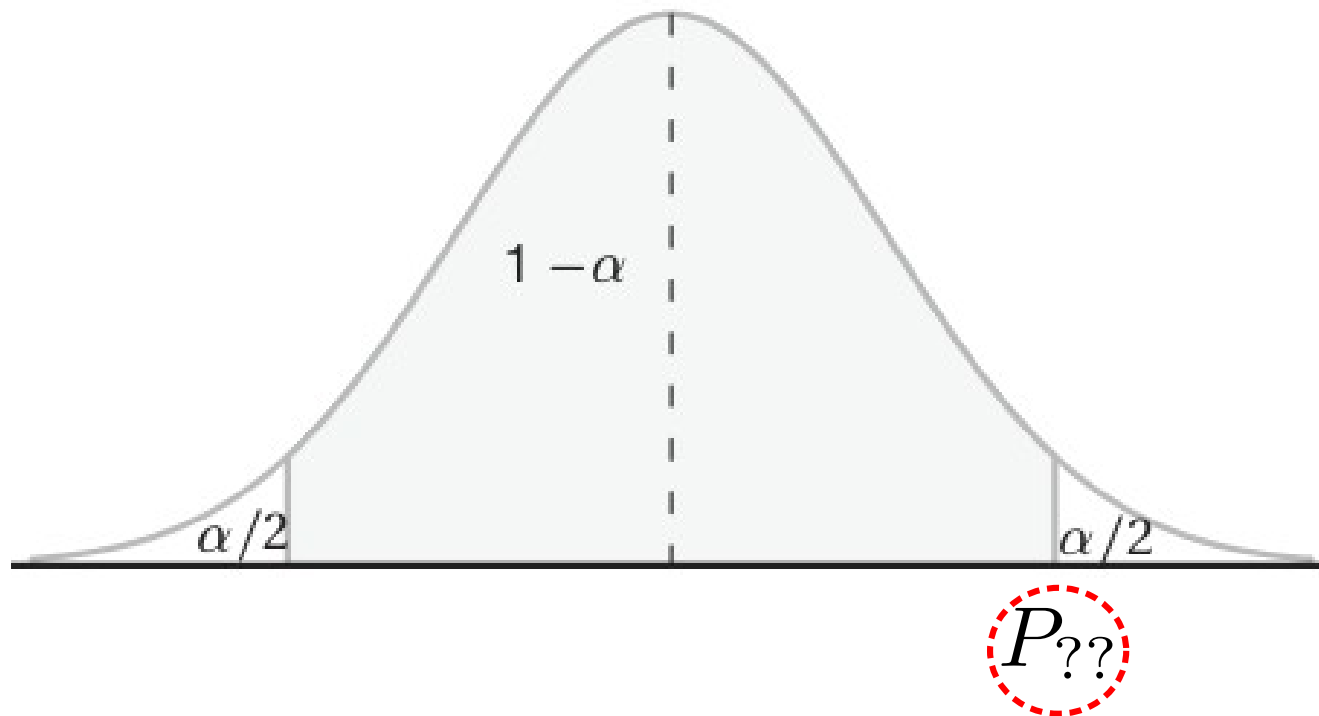


LA NOCIÓN DE CONFIANZA

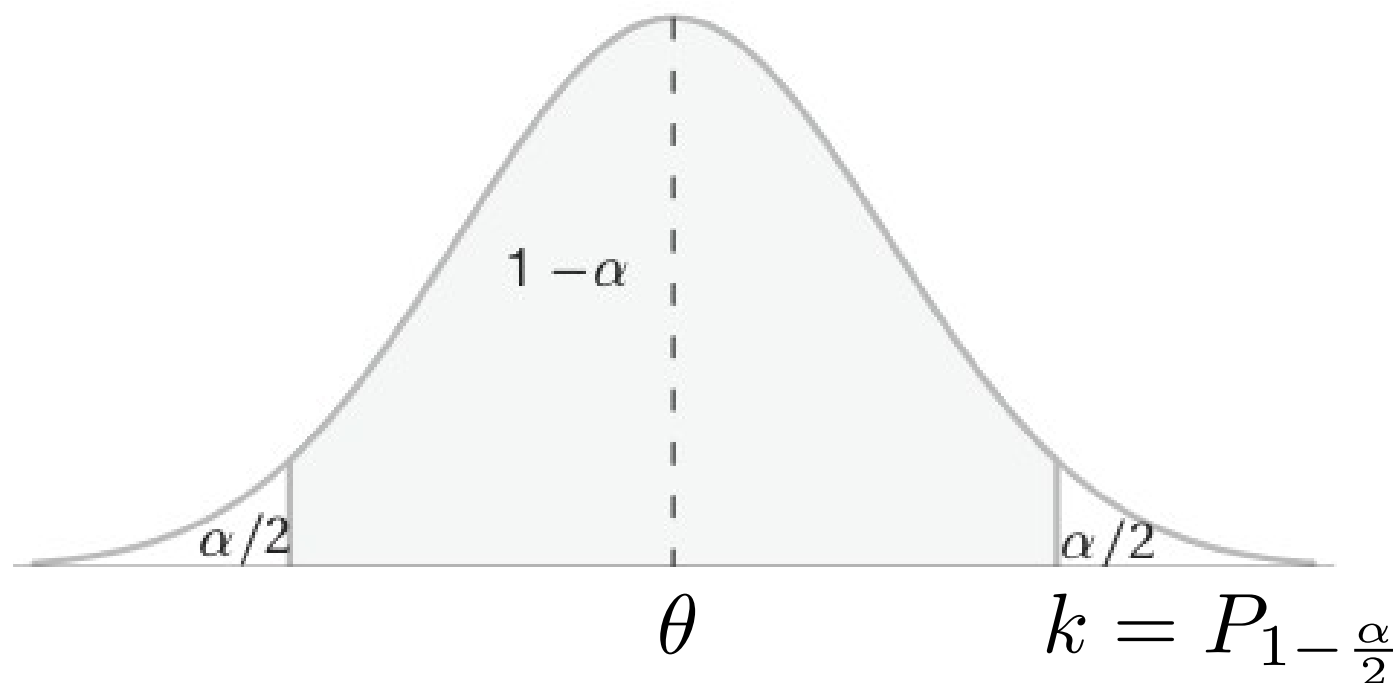


$$1 - \alpha = \frac{9}{10}$$

LA NOCIÓN DE CONFIANZA



INTERVALOS DE CONFIANZA



Genéricamente:

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

CASOS CONCRETOS

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

Intervalo de confianza

$$\left(\bar{x} \pm |Z_{\left(1-\frac{\alpha}{2}\right)}| \frac{\sigma}{\sqrt{n}} \right)$$

CASOS CONCRETOS

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

Intervalo de confianza

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

CASOS CONCRETOS

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

Intervalo de confianza

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

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

$$\left(p \pm |Z_{\left(1-\frac{\alpha}{2}\right)}| \sqrt{\frac{\pi(1-\pi)}{n}} \right)$$

CASOS CONCRETOS

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

Intervalo de confianza

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

CASOS CONCRETOS

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

Intervalo de confianza

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

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

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

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

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

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

CASOS CONCRETOS

Parámetro	Supuesto	Estimador
$\pi_1 - \pi_2$	<i>Poblaciones normales</i>	$p_1 - p_2$

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

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

CASOS CONCRETOS

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

Intervalo de confianza

$$\left(\frac{S_x^2}{S_y^2}(I) \leq \frac{\sigma_x^2}{\sigma_y^2} \leq \frac{S_x^2}{S_y^2}(II) \right)$$

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

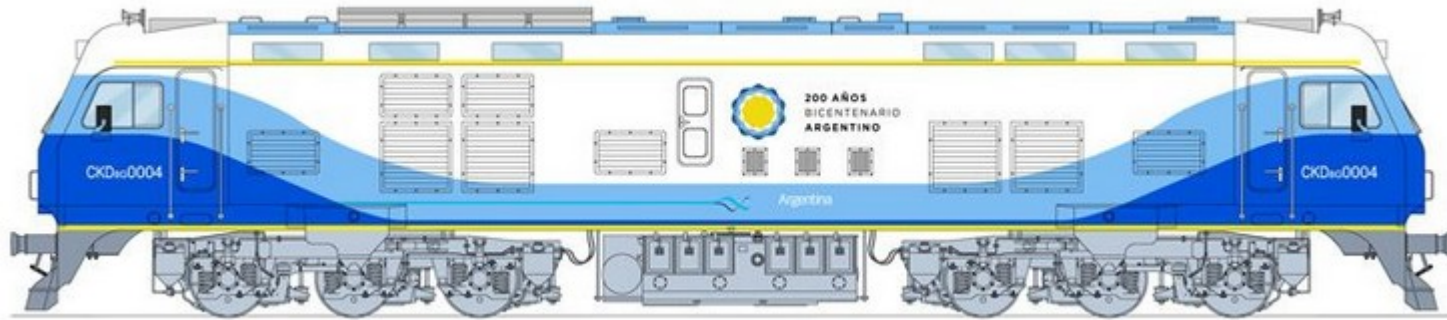
$$(II) = F_{\left(\frac{\alpha}{2}; n_x-1; n_y-1\right)}$$

SESGO

n	n-1	n/n-1	n-1/n
3	2	1,500	0,667
5	4	1,250	0,800
10	9	1,111	0,900
15	14	1,071	0,933
18	17	1,059	0,944
20	19	1,053	0,950
25	24	1,042	0,960
28	27	1,037	0,964
30	29	1,034	0,967
35	34	1,029	0,971
50	49	1,020	0,980
80	79	1,013	0,988
100	99	1,010	0,990
1000	999	1,001	0,999

SESGO

n	n-1	n/n-1	n-1/n				
3	2	1,500	0,667				
5	x	f	h	x*h	x ² *h		
10		-4	1	0,1	-0,4	1,6	
15		-3	1	0,1	-0,3	0,9	
18		-1	1	0,1	-0,1	0,1	
20		1	1	0,1	0,1	0,1	
20		2	1	0,1	0,2	0,4	
25		4	1	0,1	0,4	1,6	
28		4	1	0,1	0,4	1,6	
30		6	1	0,1	0,6	3,6	
35		10	1	0,1	1	10	
35		124	1	0,1	12,4	1538	
50			media	14,3	1558	media ²	204,49
80			media	14,3		S ²	1353,01
100						S' ²	1503,34
1000			var	1503		Sesgo	1,11
						1/Sesgo	0,90



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.

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



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

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 encontrará entre -8,2 y 34,7, en símbolos:

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