

Intervalo de Confianza

📊 ¿Qué es un Intervalo de Confianza?

Un **intervalo de confianza (IC)** es un rango de valores que, con cierto nivel de confianza, tiene alta probabilidad de contener el verdadero valor de un parámetro poblacional. Es una herramienta fundamental en estadística inferencial para cuantificar la incertidumbre de las estimaciones.

🎯 Interpretación Correcta (MUY IMPORTANTE)

⚠️ Interpretación Correcta:

"Si repitiéramos el proceso de muestreo muchas veces y calculáramos un IC en cada muestra, aproximadamente el 95% de esos intervalos contendrían el verdadero valor del parámetro."

✗ Interpretación INCORRECTA:

"Hay un 95% de probabilidad de que el parámetro esté en este intervalo específico."

🔍 **Explicación:** El parámetro poblacional (μ , p , etc.) es una constante fija pero desconocida. El intervalo es lo que varía de muestra a muestra. El "95%" se refiere a la confianza en el procedimiento, no en un intervalo particular.

⚠️ Componentes de un Intervalo de Confianza

Nivel de Confianza (1- α)

Probabilidad de que el procedimiento genere un intervalo que contenga el parámetro verdadero.

Valores comunes: 90%, 95%, 99%

Margen de Error (ME)

Cantidad que se suma y resta al estimador puntual para formar el intervalo.

ME = valor crítico \times error estándar

Forma general de un IC:

$$IC = \text{Estimador} \pm \text{Margen de Error}$$

$$IC = \text{Estimador} \pm (\text{Valor Crítico}) \times (\text{Error Estándar})$$

Valores Críticos Comunes

Distribución Normal (z):

Nivel de Confianza	α	$\alpha/2$	$z_{\alpha/2}$
90%	0.10	0.05	1.645
95%	0.05	0.025	1.96
99%	0.01	0.005	2.576

Distribución t de Student (valores aproximados):

Grados de Libertad	$t_{0.05} \text{ (90\%)}$	$t_{0.025} \text{ (95\%)}$	$t_{0.005} \text{ (99\%)}$
5	2.015	2.571	4.032
10	1.812	2.228	3.169
20	1.725	2.086	2.845
30	1.697	2.042	2.750
∞	1.645	1.96	2.576

Nota: A medida que $gl \rightarrow \infty$, la distribución t converge a la normal estándar.

Intervalo de Confianza para la Media (μ)

Caso 1: σ conocida o n grande ($n \geq 30$)

Usar distribución Normal (Z)

$$IC = X \pm Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

$$ME = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

$$SE = \frac{\sigma}{\sqrt{n}}$$

Caso 2: σ desconocida y n pequeño ($n < 30$)

Usar distribución t de Student con $(n-1)$ grados de libertad

$$I.C = \bar{X} \pm t_{\alpha/2, n-1} \cdot \frac{s}{\sqrt{n}}$$

Donde:

s = desviación estándar muestral

gl = $n - 1$ (grados de libertad)

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}$$

Determinación del Tamaño de Muestra

Para obtener un margen de error específico (ME) con nivel de confianza dado:

Para la media:

$$n = \left(\frac{Z_{\alpha/2} \cdot \sigma}{M.E.} \right)^2$$

Para la proporción:

$$n = \left(\frac{Z_{\alpha/2}}{M.E.} \right)^2 \cdot p(1 - p)$$

Si no hay estimación previa de p , usar $p = 0.5$ (caso más conservador).

Ejemplo 1: IC para la media con σ conocida

Enunciado:

Una muestra de 100 estudiantes tiene una media de altura de 170 cm. La desviación estándar poblacional es $\sigma = 10$ cm. Calcular el IC al 95%.

Datos:

Media muestral = 170 cm

$\sigma = 10$ cm (conocida)

$n = 100$

Nivel de confianza = 95%

Valor crítico Z = 1.96

Cálculo:

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{10}{\sqrt{100}} = \frac{10}{10} = 1$$

$$ME = Z_{\alpha/2} \cdot SE = 1.96 \times 1 = 1.96$$

$$IC = 170 \pm 1.96 = (168.04, 171.96)$$

 **Interpretación:** Con 95% de confianza en el procedimiento, estimamos que la altura promedio de la población está entre 168.04 y 171.96 cm.

Ejemplo 2: IC para la media con σ desconocida (t de Student)

Enunciado:

Una muestra de 15 estudiantes tiene una media de 72 kg y una desviación estándar muestral de 8 kg. Calcular el IC al 95%.

Datos:

Media muestral = 72 kg

s = 8 kg (σ desconocida)

n = 15, entonces gl = 14

95% confianza

Valor crítico t = 2.145

Cálculo:

$$SE = \frac{s}{\sqrt{n}} = \frac{8}{\sqrt{15}} = \frac{8}{3.873} \approx 2.066$$

$$ME = t_{\alpha/2, gl} \cdot SE = 2.145 \times 2.066 \approx 4.43$$

$$IC = 72 \pm 4.43 = (67.57, 76.43)$$

Interpretación: Con 95% de confianza, el peso promedio poblacional está entre 67.57 y 76.43 kg. Notar que el intervalo es más amplio que si usáramos Z debido al tamaño muestral pequeño.

Ejemplo 3: Determinación del tamaño de muestra

Enunciado:

Queremos estimar la altura promedio con un margen de error de 2 cm y 95% de confianza. Si $\sigma = 10$ cm, ¿qué tamaño de muestra necesitamos?

Cálculo:

$$n = \left(\frac{Z_{\alpha/2} \cdot \sigma}{ME} \right)^2 = \left(\frac{1.96 \times 10}{2} \right)^2$$

$$n = \left(\frac{19.6}{2} \right)^2 = (9.8)^2 = 96.04$$

Por lo tanto: $n \approx 97$ (redondeamos hacia arriba)

Resultado: Se necesitan al menos 97 estudiantes para lograr un margen de error de 2 cm con 95% de confianza.

Intervalo de Confianza para la Proporción (p)

Se utiliza cuando el parámetro de interés es una proporción poblacional (p). El estimador es la proporción muestral $p = x/n$.

$$IC = p \pm Z_{\alpha/2} \sqrt{\frac{p(1-p)}{n}}$$

Condiciones de validez:

$$np \geq 10 \text{ y } n(1-p) \geq 10$$

Muestra aleatoriedad simple

Observaciones independientes

$$ME = Z_{\alpha/2} \sqrt{\frac{p(1-p)}{n}}$$

Ejemplo 4: IC para una proporción

Enunciado:

En una encuesta de 200 personas, 60 prefieren el producto A. Calcular el IC al 95% para la proporción poblacional.

Verificación de condiciones:

$$p = 60/200 = 0.3$$

$$np = 200(0.3) = 60 \geq 10 \checkmark$$

$$n(1-p) = 200(0.7) = 140 \geq 10 \checkmark$$

Cálculo:

$$SE = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.3(0.7)}{200}} = \sqrt{\frac{0.21}{200}} = \sqrt{0.00105} \approx 0.0324$$

$$ME = 1.96 \times 0.0324 \approx 0.0635$$

$$IC = 0.3 \pm 0.0635 = (0.2365, 0.3635)$$

 Interpretación: Con 95% de confianza, la proporción poblacional que prefiere el producto A está entre 23.65% y 36.35%.

Intervalo de Confianza para Diferencia de Medias

Muestras Independientes (varianzas desconocidas pero iguales)

$$IC = (\bar{X}_1 - \bar{X}_2) \pm t_{\alpha/2, gl} \cdot s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

Donde:

Desviación estándar combinada:

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

Grados de libertad: $gl = n_1 + n_2 - 2$

Muestras Pareadas (datos dependientes)

Cuando las observaciones están emparejadas (antes/después, gemelos, etc.)

$$IC = \bar{d} \pm t_{\alpha/2, n-1} \cdot \frac{s_d}{\sqrt{n}}$$

Donde:

\bar{d} = media de las diferencias

s_d = desviación estándar de las diferencias

Ejemplo 5: IC para diferencia de medias (muestras independientes)

Enunciado:

Grupo A ($n_1=20$): media = 85, $s_1=10$. Grupo B ($n_2=25$): media = 78, $s_2=12$. Calcular IC 95% para $\mu_1 - \mu_2$.

Cálculo:

$$s_p = \sqrt{\frac{(20-1)(10)^2 + (25-1)(12)^2}{20+25-2}} = \sqrt{\frac{1900 + 3456}{43}}$$

$$s_p = \sqrt{124.56} \approx 11.16$$

$$SE = s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = 11.16 \sqrt{\frac{1}{20} + \frac{1}{25}}$$

$$SE = 11.16 \sqrt{0.09} \approx 3.35$$

Valor crítico: t ($gl=43, 95\%$) ≈ 2.017

$$IC = (85 - 78) \pm 2.017(3.35) = 7 \pm 6.76 = (0.24, 13.76)$$

Interpretación: Con 95% de confianza, la diferencia entre las medias poblacionales está entre 0.24 y 13.76. Como el intervalo no contiene 0, hay evidencia de que las medias son diferentes.

Intervalo de Confianza para Diferencia de Proporciones

$$IC = (p_1 - p_2) \pm Z_{\alpha/2} \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$

Condiciones: Ambas muestras deben cumplir $np \geq 10$ y $n(1-p) \geq 10$ para cada proporción.

Intervalo de Confianza para la Varianza (σ^2)

Se utiliza la distribución chi-cuadrado (χ^2) con $(n-1)$ grados de libertad.

$$IC = \left(\frac{(n-1)s^2}{\chi^2_{\alpha/2, n-1}}, \frac{(n-1)s^2}{\chi^2_{1-\alpha/2, n-1}} \right)$$

Nota: La distribución χ^2 es asimétrica, por eso el intervalo no es simétrico respecto a s^2 .

Relación con Pruebas de Hipótesis

Los intervalos de confianza y las pruebas de hipótesis están estrechamente relacionados:

Si un valor hipotético ($H_0: \mu = \mu_0$) **está fuera del IC al $(1-\alpha) \times 100\%$** , entonces rechazaríamos H_0 con nivel de significancia α .

Si el valor **está dentro del IC**, no rechazaríamos H_0 .

Para diferencia de medias/proportiones: si el IC contiene 0, no hay evidencia significativa de diferencia.

Ejemplo: $IC_{95\%} = (68, 72)$. Si probamos $H_0: \mu = 70$, no rechazamos porque $70 \in IC$. Si probamos $H_0: \mu = 75$, rechazamos porque $75 \notin IC$.

Supuestos y Condiciones de Validez

Para IC de la Media:

Muestra aleatoria simple: Cada individuo tiene la misma probabilidad de ser seleccionado.

Independencia: Las observaciones son independientes entre sí.

Normalidad:

Si $n \geq 30$: El Teorema del Límite Central justifica el uso de la normal

Si $n < 30$: La población debe ser aproximadamente normal (verificar con gráficos)

Para IC de Proporción:

Muestra aleatoria simple

Independencia de observaciones

Tamaño muestral suficiente: $np \geq 10$ y $n(1-p) \geq 10$

Para IC de Varianza:

La población debe ser **estrictamente normal** (más sensible que para la media)

Muestra aleatoria simple

Factores que Afectan la Amplitud del IC

El intervalo es MÁS AMPLIO cuando:

↑ **Mayor nivel de confianza** (95% → 99%): Necesitamos mayor amplitud para tener más confianza

↑ **Mayor variabilidad** (σ o s más grande): Más dispersión requiere más margen

↓ **Menor tamaño de muestra** (n pequeño): Menos información = más incertidumbre

El intervalo es MÁS ESTRECHO cuando:

↓ Menor nivel de confianza (99% → 90%)

↓ Menor variabilidad de los datos

↑ Mayor tamaño de muestra

Nota importante: Para reducir el ME a la mitad, necesitas **cuadruplicar** el tamaño de muestra (porque n aparece en \sqrt{n}).

Ejemplo 6: Comparación de niveles de confianza

Enunciado:

Con media = 50, $\sigma = 10$, $n = 25$, calcular IC al 90%, 95% y 99%.

Cálculo:

Error estándar: $SE = 10/\sqrt{25} = 2$

Confianza	z	ME	Intervalo	Amplitud
90%	1.645	3.29	(46.71, 53.29)	6.58
95%	1.96	3.92	(46.08, 53.92)	7.84
99%	2.576	5.15	(44.85, 55.15)	10.30

Conclusión: A mayor confianza, mayor amplitud del intervalo. El compromiso es entre precisión (intervalo estrecho) y confianza.

Errores Comunes

Interpretación probabilística incorrecta: Decir "hay 95% de probabilidad de que μ esté en este intervalo". El parámetro es fijo, la confianza se refiere al procedimiento.

Confundir nivel de confianza ($1-\alpha$) con nivel de significancia (α): Son complementarios pero conceptualmente diferentes.

Usar Z cuando debería usar t: Con σ desconocida y $n < 30$, siempre usar distribución t.

Ignorar supuestos: No verificar normalidad en muestras pequeñas o condiciones de tamaño para proporciones.

Confundir IC para media con IC para proporción: Usan fórmulas y condiciones diferentes.

Olvidar que $n-1$ son los grados de libertad: No usar n directamente.

Pensar que un IC más amplio es "peor": Depende del contexto; más confianza requiere más amplitud.

No redondear n hacia arriba: En cálculos de tamaño muestral, siempre redondear al entero superior.

Guía de Decisión Rápida

Parámetro	Condiciones	Distribución	Fórmula Base
Media (μ)	σ conocida O $n \geq 30$	Normal (z)	$\text{Media} \pm z \times \sigma/\sqrt{n}$
Media (μ)	σ desconocida Y $n < 30$	t de Student	$\text{Media} \pm t \times s/\sqrt{n}$
Proporción (p)	$np \geq 10, n(1-p) \geq 10$	Normal (z)	$p \pm z \times \sqrt{[p(1-p)/n]}$
Varianza (σ^2)	Población normal	Chi-cuadrado	$[(n-1)s^2/\chi^2_{\text{sup}}, (n-1)s^2/\chi^2_{\text{inf}}]$