Estimación Puntual y por Intervalo para la Media.

Estimación: puntual y por intervalos

Como ya hemos visto, a partir de los estadísticos que hemos obtenido en la/s muestra/s queremos obtener una idea de los valores de los parámetros en la población.

Se trata de emplear los estadísticos para *estimar* los parámetros.

Veremos DOS tipos de estimadores:

1) Estimación puntual. Aquí obtendremos un punto, un valor, como estimación del parámetro.
2) Estimación por intervalos. Aquí obtendremos un intervalo dentro del cual estimamos (bajo cierta probabilidad) estará el parámetro.

Distribución Muestral de la Media Muestral

Recordando…

La distribución muestral de la media muestral es la distribución de los valores de las medias muestrales de todas las posibles muestras del mismo tamaño ***n*** tomadas de la misma población.

Si sacamos muestras aleatorias de tamaño ***n*** de una población con media **μ** y desviación estándar **σ**, entonces la distribución muestral de la media muestral tiene las siguientes propiedades:

Distribución Muestral de la Media Muestral

1. El promedio de todos los valores posibles de medias muestrales es igual al parámetro μ. En otras palabras, la media muestral $\bar{x}$ es un estimador insesgado de μ.

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

2. **Error estándar** de la media muestral: Es la desviación estándar de las posibles medias muestrales.

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

   El error estándar disminuye si el tamaño de la muestra aumenta.

3. Si la población original tiene distribución Normal, entonces para cualquier tamaño muestral n la distribución de la media muestral es también Normal

$$\text{Si } X \sim N(\mu, \sigma) \Rightarrow \bar{x} \sim N(\mu, \frac{\sigma}{\sqrt{n}})$$

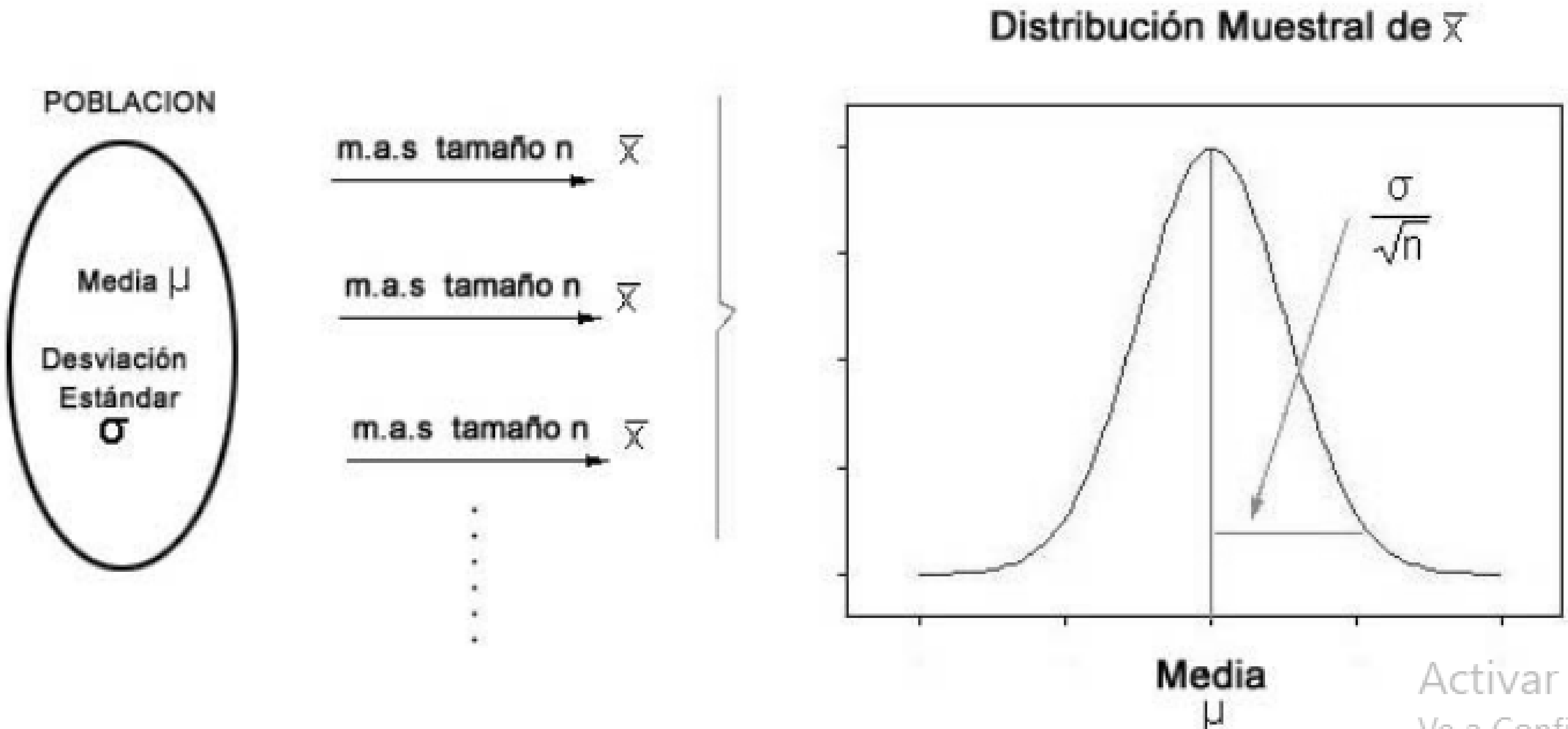
4. Si la población de origen no es Normal pero podemos calcular su media y desviación estándar y el tamaño muestral (n) es "suficientemente" grande la distribución de la media muestral es aproximadamente Normal

$$\text{Aún si } X \text{ no es } N(\mu, \sigma) \Rightarrow \bar{x} \dot{\sim} N(\mu, \frac{\sigma}{\sqrt{n}})$$

Distribución Muestral de la Media Muestral

Notas:

- Un tamaño muestral de 30 es considerado suficiente.
- El resultado en (4) se conoce como el **Teorema del Límite Central**.



Intervalo de Confianza para la media de una población

Como hemos mencionado la media muestral $\overline{x}$ basada en una muestra aleatoria, es un buen estimador puntual de la media poblacional μ. La pregunta es ¿qué tan buen estimador es? ¿Será el valor observado de la media muestral igual a la media poblacional? Probablemente no. ¿Será entonces cercano a μ? ¿Pero qué tan cercano?

Definiciones:

La media muestral $\overline{X}$ es un **estimador puntual** de la media de poblacional μ.

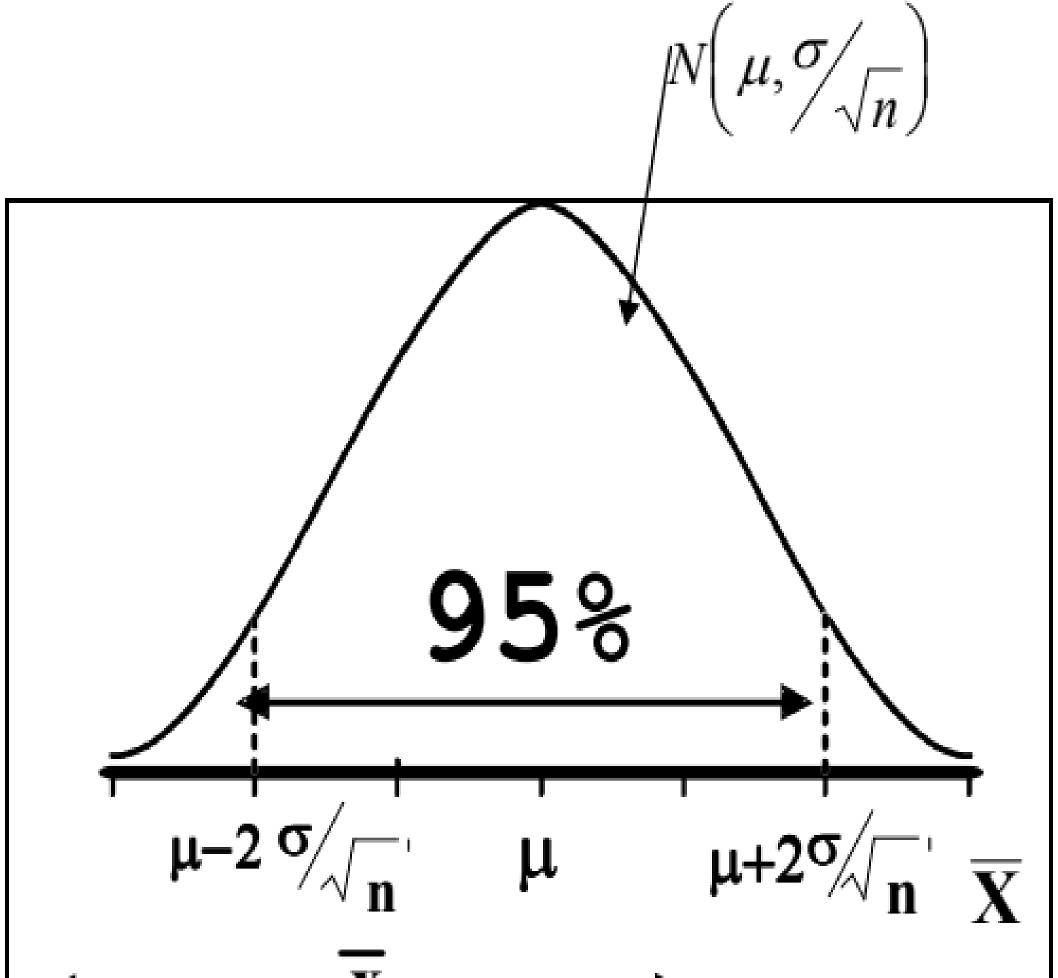
Un **estimador de intervalo de confianza** para la media poblacional μ es un intervalo de valores, calculados a partir de los datos de la muestra, entre los cuales podemos *confiar* que se encuentra la media poblacional μ.

El **nivel de confianza** es la probabilidad de que el método de estimación nos dé un intervalo de confianza que contiene al parámetro (μ en este caso).

Intervalo de Confianza para la media de una población con desviación estándar conocida

Para construir un intervalo de confianza para μ usamos la distribución muestral de la media muestral $\bar{x} \sim N(\mu, \frac{\sigma}{\sqrt{n}})$,

entonces aproximadamente 95% de los valores de estarán a 1,96 desviaciones estándar de μ.



Intervalo de Confianza para la media de una población con desviación estándar conocida

Un intervalo de confianza $(1-\alpha)$*100% para μ está dado por:

$$\overline{X} \pm z_{1-\frac{\alpha}{2}}\left(\frac{\sigma}{\sqrt{n}}\right)$$

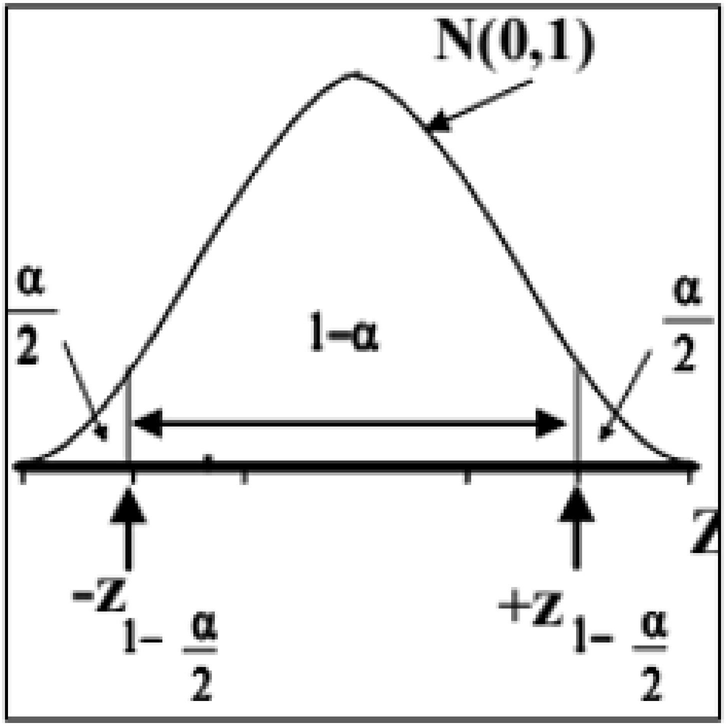
donde $z_{1-\frac{\alpha}{2}}$ es un percentil de la distribución N(0,1)

Este intervalo será exacto si los datos provienen de una **muestra aleatoria** de una población Normal. Será aproximado para cualquier otra población y tamaños de muestra $n > 30$ (Teorema de Límite Central).

Intervalo de Confianza para la media de una población con desviación estándar conocida

Revisemos los percentiles de la tabla N(0,1):

Nivel de Confianza $1-\alpha$	$z_{1-\frac{\alpha}{2}}$
0,80	
0,90	
0,95	
0,98	
0,99	
0,999	



Intervalo de Confianza para la media de una población con desviación estándar conocida

EJEMPLO

Suponga que una máquina de bebidas esta calibrada de tal manera que la cantidad de líquido entregado es aproximadamente normal con desviación estándar 0,15 decilitros (nota: 1 decilitro es 0,1 litro). De acuerdo a esta información:

a) Calcule un intervalo de 95% de confianza para la media de la cantidad de líquido entregado basado en una muestra aleatoria de 36 vasos con promedio de 2,25 decilitros.
b) ¿Un intervalo de 90% de confianza será más ancho o más angosto que el calculado en la parte (a)?
c) ¿Qué tamaño de muestra necesitaría si quisiera obtener un margen de error para un intervalo de 95% de confianza sea 0,02?

Intervalo de Confianza para la media de una población con desviación estándar desconocida

Una vez más nos encontramos con el problema de que en general no conocemos la desviación estándar de la población (σ), por lo tanto estimamos σ con la desviación estándar de la muestra seleccionada (s), y tenemos que usar la distribución t de Student en vez de la Normal.

Un intervalo de confianza $(1-\alpha)$*100% para μ está dado por:

$$\bar{x} \pm t_{1-\alpha/2}^{(n-1)}\left(s/\sqrt{n}\right)$$

donde $t_{1-\alpha/2}^{(n-1)}$ es el percentil apropiado de la distribución t con (n-1) grado de libertad.

Este intervalo nos da valores posibles para la media de la población μ basados en la media muestral $\bar{x}$. Se basa en el supuesto de que los datos provienen de una muestra aleatoria de una población con distribución normal con desviación estándar σ *desconocida*. Si el tamaño de la muestra es grande, el supuesto de normalidad no es crucial, sin embargo debemos preocuparnos si la distribución de los datos es sesgada o tiene valores extremos.

El **margen de error** de la media poblacional es: $E = t_{1-\alpha/2}^{(n-1)}\left(s/\sqrt{n}\right)$.

Intervalo de Confianza para la media de una población

Corrección para población finita (cpf) (Opcional)

La fórmula para calcular el tamaño muestral supone que el tamaño de la población es grande o es un muestreo con reemplazo. Cuando queremos sacar una muestra de una población "pequeña" en conveniente usar la cpf

$$n_c = \frac{n}{1 + n/N}$$

Resumen Intervalos de Confianza

Tabla resumen:

Situación	Parámetro	Test Estadístico	Intervalo de Confianza
Muestra aleatoria de una población normal con σ conocida	Media μ	$z = \dfrac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$	$\bar{x} \pm z_{1-\alpha/2}\left(\sigma / \sqrt{n}\right)$
Muestra aleatoria de una población normal con σ desconocida	Media μ	$t = \dfrac{\bar{x} - \mu_0}{s / \sqrt{n}}$	$\bar{x} \pm t_{1-\alpha/2}^{(n-1)}\left(S / \sqrt{n}\right)$

EJERCICIOS

1. Suponga que el peso de las cajas de cierta clase de cereal está normalmente distribuido con desviación estándar de 8,22 gramos. Se toma una muestra aleatoria de 25 cajas y se obtiene una media de 278,39 gramos.

a) Calcule un intervalo de 95% de confianza.
b) Suponga que la media de 278,39 gramos resultó de una muestra de tamaño n=36 en vez de n=25 cajas. ¿Qué efecto tiene el mayor tamaño muestral en el intervalo de confianza?
c) ¿Cuál es el margen de error del intervalo de confianza basado en n=25 observaciones?
d) ¿Qué nos dice el intervalo de 95% de confianza acerca de la media del peso?

EJERCICIOS

2. En una muestra aleatoria de 35 funcionarios de una institución, su sueldo medio era de $950.000 mensuales. Suponiendo que la desviación típica de los sueldos de todos los funcionarios es de $450.000, hallar un intervalo de 95% de confianza para el sueldo medio en la institución.

3. En una muestra aleatoria de 45 estudiantes universitarios, la altura media era de 170cm. Suponiendo que la desviación típica de las alturas de todos los universitarios es de 5cm, hallar un intervalo de 99% de confianza para la altura media.

EJERCICIOS

4. Una muestra aleatoria de 9 tarros de duraznos proporciona los siguientes pesos en gramos:

88 90 90 86 87 88 91 92 89

Encuentre un intervalo de confianza del 95% para la media del peso en gramos de la población.

5. Se toma una muestra aleatoria simple de 10 baterías y se obtienen las siguientes duraciones (en meses):

33, 34, 26, 37, 30, 39, 26, 31, 36, 19

Encuentre un intervalo de confianza del 90% para la duración media de ese modelo de batería.