

DISTRIBUCIONES MUESTRALES

Recordemos:

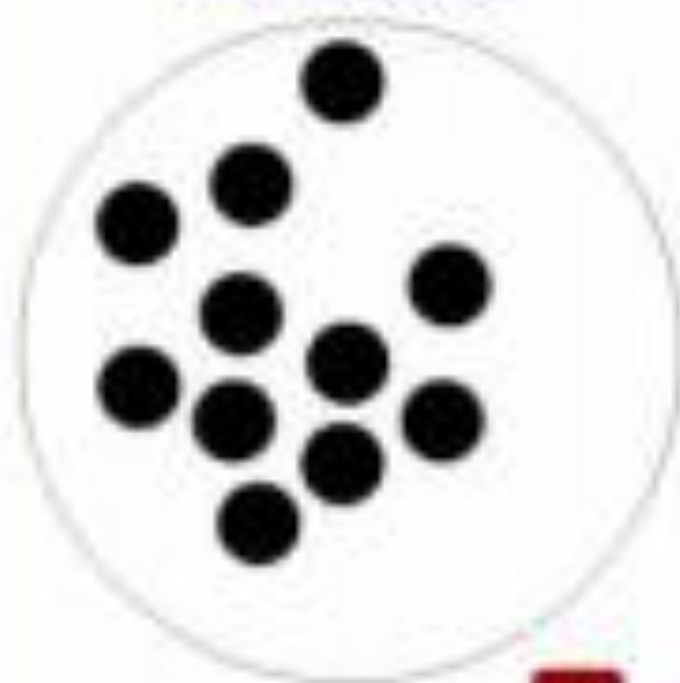
Parámetro es una medida de resumen numérica que se calcularía usando todas las unidades de la población.

Es un número fijo.

Generalmente no lo conocemos.

Estadística es una medida de resumen numérica que se calcula de las unidades de la muestra.

Población



Muestra



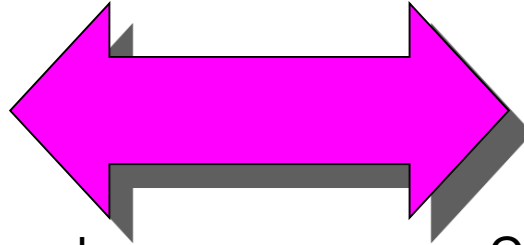
Inferencia



Tipos de Muestreos

PROBABILÍSTICOS

- Todas las unidades tienen igual probabilidad de participar en la muestra.
- La elección de cada unidad muestral es independiente de las demás
- Se puede calcular el error muestral



NO PROBABILÍSTICOS

- Cada unidad NO tiene igual probabilidad de participar en la muestra.
- No se puede calcular el error muestral
- Alto riesgo de invalidez producido por la introducción de sesgos

Usos de cada tipo de muestreo

Muestreo Probabilísticos

- Estimación de Parámetros.
- Comprobación de Hipótesis.

Muestreos No Probabilísticos

- Estudios Pilotos.
- Investigaciones en poblaciones de difícil registro o localización.

Inferencia estadística: es el proceso de sacar conclusiones de la población basados en la información de una muestra de esa población.

Objetivos de la inferencia:

- estimación de parámetros,
- intervalos de confianza y
- docimasia, test de hipótesis o pruebas de significación estadística.

Distribuciones muestrales

Una estadística muestral proveniente de una muestra aleatoria simple tiene un patrón de comportamiento (predecible) en repetidas muestras. Este patrón es llamado la distribución muestral de la estadística.

Si conocemos la distribución muestral podemos hacer inferencia.

Las distribuciones muestrales adoptan diferentes formas según las estadísticas investigadas y las características de la población estudiada.

Recuerden que al analizar una distribución nos interesa:

1. **Forma** (simétrica o sesgada)
2. **Posición central** - la media de una distribución muestral nos dice si el estadístico es un "buen" (insesgado) estimador del parámetro o es sesgado.
3. **Dispersión** - nos da una idea del error de muestreo.

Distribución muestral de la media muestral

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.

Distribución muestral de la media muestral

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:

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:

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

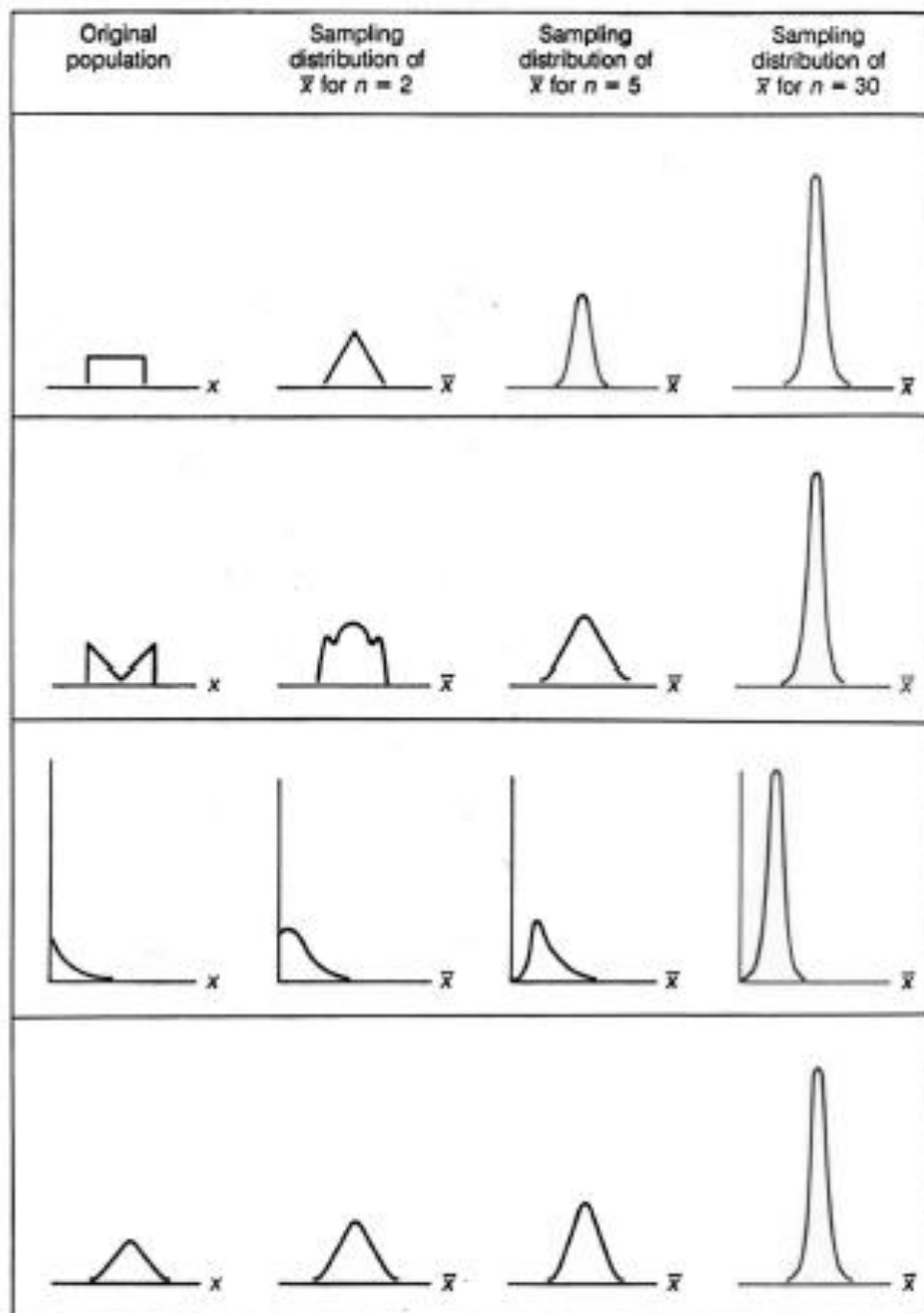
4. Si la población de origen **no** es Normal, pero n es “suficientemente” grande la distribución de la media muestral es aproximadamente Normal:

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

Nota:

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

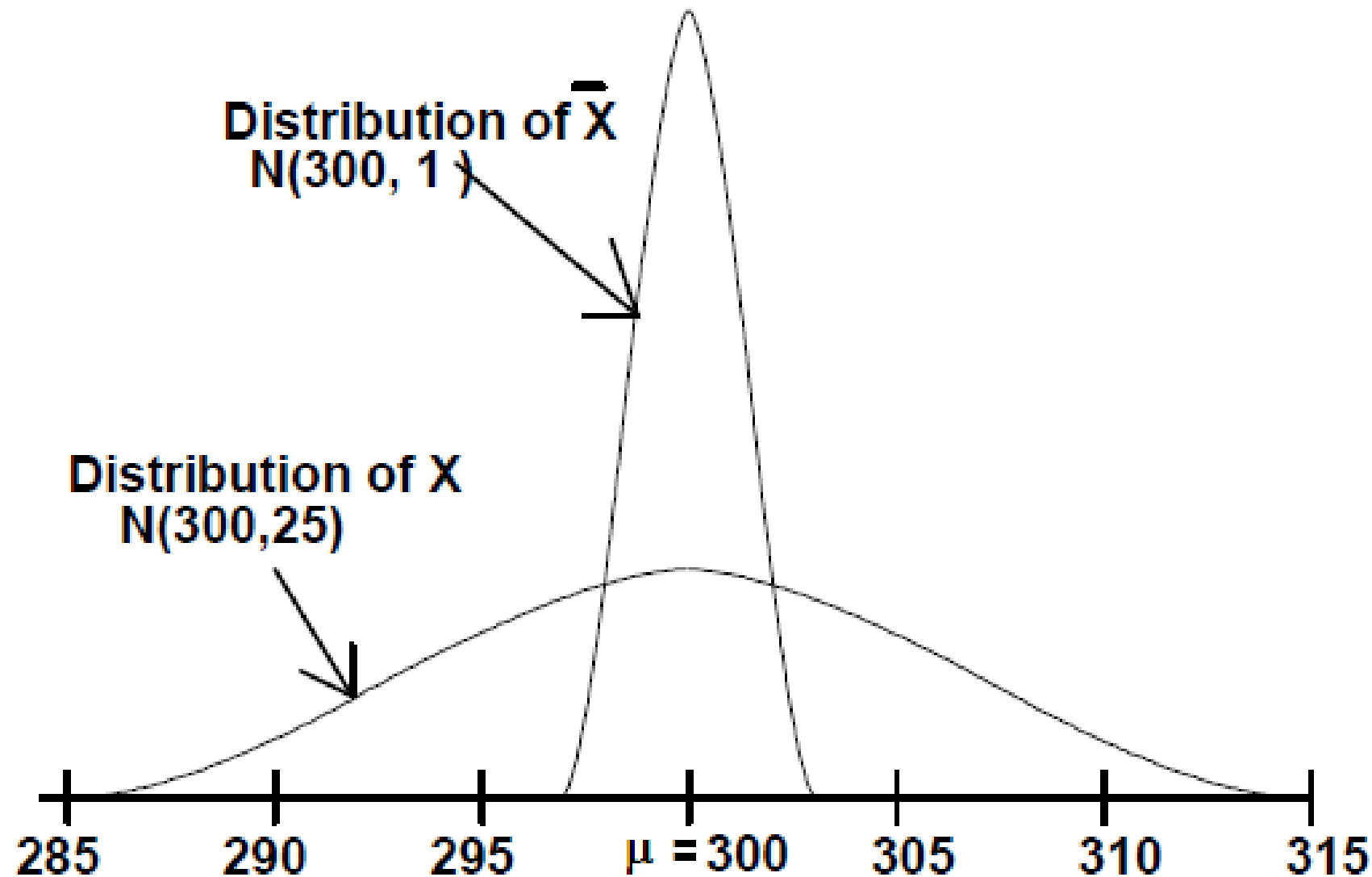
FIGURE 6.10
Sampling Distributions of \bar{x} for Different Populations and Different Sample Sizes



✓ Suponga que X = peso de carga de camionetas en kilos, tiene distribución normal con media = 300 k y varianza = 25.

Se toma una muestra aleatoria de 25 camionetas cargadas y se calcula la media muestral.

Esquema de las distribuciones de la variable aleatoria X y de la media muestral:



✓ Suponga que X = la edad de las madres en los nacimientos en Chile tiene distribución normal con media = 26,5 años y desviación estándar 6,3 años.

a) Describa la distribución de la edad de la madre.

- b) ¿Cuál es la probabilidad de que una madre elegida al azar tenga más de 30 años?
- c) Suponga que tomamos una muestra aleatoria de $n=25$ madres ¿cuál es la probabilidad de que la media muestral sea mayor a 30?
- d) ¿porqué las respuestas en (b) y (c) son distintas?

