Estadística II

Juan Carlos Trejos Iglesias

2025-1

Índice

1. Distribuciones Muestrales

1.1. Distribución muestral de la media

Definición: La distribución muestral de la media es la distribución de las medias obtenidas de todas las posibles muestras de tamaño n tomadas de una población. Esta distribución tiende a ser normal si la población es normal o si n es suficientemente grande (por el Teorema Central del Límite).

$$Z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}}$$

- **Media de la distribución muestral:**
$$\mu_{\bar{x}} = \mu$$
La media de todas las medias muestrales es igual a la media poblacional.
- **Varianza de la distribución muestral:**
$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$$
La varianza de las medias muestrales disminuye con el tamaño de la muestra.
- **Desviación estándar de la media muestral (error estándar):**
$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$
Cuantifica cuánto varían las medias muestrales respecto a la media poblacional.

Ejemplo:

Supongamos que el tiempo promedio que una persona tarda en resolver un formulario es $\mu = 30$ minutos, con una desviación estándar de $\sigma = 10$ minutos. Si tomamos muestras de $n = 25$ personas:

- La media de la distribución muestral sigue siendo:
$$\mu_{\bar{x}} = \mu = 30$$
- La desviación estándar (error estándar) es:
$$\sigma_{\bar{x}} = \frac{10}{\sqrt{25}} = \frac{10}{5} = 2$$
- Si una muestra de 25 personas tiene una media muestral de $\bar{x} = 33$, el valor Z sería:
$$Z = \frac{33 - 30}{2} = \frac{3}{2} = 1{,}5$$

Esto indica que esa media muestral se encuentra 1.5 desviaciones estándar por encima de la media poblacional.

1.2. Distribución muestral de la proporción

Definición: La distribución muestral de la proporción describe el comportamiento de la proporción muestral $\hat{p}$, obtenida a partir de muestras aleatorias de tamaño n tomadas de una población con proporción poblacional p. Para muestras suficientemente grandes, esta distribución se aproxima a una distribución normal por el Teorema Central del Límite.

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$$

- **Media de la distribución muestral:**

$$\mu_{\hat{p}} = p$$

  La media de todas las proporciones muestrales es igual a la proporción poblacional.

- **Varianza de la distribución muestral:**

$$\sigma_{\hat{p}}^2 = \frac{p(1-p)}{n}$$

  La varianza disminuye al aumentar el tamaño de la muestra.

- **Desviación estándar (error estándar) de la proporción muestral:**

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$$

  Mide la variabilidad de la proporción muestral alrededor de la proporción poblacional.

Ejemplo:

En una ciudad, se sabe que el 60 % de la población está a favor de una nueva ley ($p = 0{,}6$). Se toma una muestra aleatoria de $n = 100$ personas.

- La media de la proporción muestral es:

$$\mu_{\hat{p}} = 0{,}6$$

- El error estándar es:

$$\sigma_{\hat{p}} = \sqrt{\frac{0{,}6 \cdot 0{,}4}{100}} = \sqrt{\frac{0{,}24}{100}} = \sqrt{0{,}0024} \approx 0{,}049$$

- Si en la muestra se encontró que $\hat{p} = 0{,}68$, el valor Z es:

$$Z = \frac{0{,}68 - 0{,}6}{0{,}049} \approx \frac{0{,}08}{0{,}049} \approx 1{,}63$$

Esto indica que la proporción muestral de 0.68 se encuentra aproximadamente 1.63 desviaciones estándar por encima de la proporción poblacional esperada.

1.3. Distribución Muestral de la Varianza

Supongamos que $X_1, X_2, \ldots, X_n$ son una muestra aleatoria de tamaño n proveniente de una población normal con media μ y varianza σ^2.

La varianza muestral se define como:

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2,$$

donde $\bar{X}$ es la media muestral.

Distribución

La variable aleatoria

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2}$$

sigue una distribución chi-cuadrado con $n-1$ grados de libertad:

$$\chi^2 \sim \chi^2_{n-1}.$$

Esto es fundamental para construir intervalos de confianza y realizar pruebas de hipótesis sobre la varianza poblacional cuando la población es normal.

Propiedades

- La distribución χ^2 es asimétrica y su forma depende de los grados de libertad.
- Para muestras grandes, la distribución se aproxima a una distribución normal.

Ejemplo

Suponga que se toma una muestra de tamaño $n = 10$ de una población normal y se obtiene una varianza muestral $S^2 = 25$. Si se conoce que la varianza poblacional es $\sigma^2 = 20$, la variable

$$\chi^2 = \frac{(10-1)\times 25}{20} = \frac{9\times 25}{20} = 11{,}25$$

sigue una distribución χ^2 con 9 grados de libertad.

Este valor puede usarse para pruebas de hipótesis o para construir intervalos de confianza sobre σ^2.

2. Intervalos de Confianza

2.1. Para una media con desviación estandar (σ) conocida

Cuando la desviación estándar poblacional σ es conocida, se utiliza la **distribución normal estándar** (Z) para construir el intervalo de confianza para la media poblacional μ.

Fórmula

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

- $\bar{X}$: media muestral.
- σ: desviación estándar poblacional (conocida).
- n: tamaño de la muestra.
- $Z_{\alpha/2}$: valor crítico de la distribución normal estándar.
- $\frac{\sigma}{\sqrt{n}}$: error estándar de la media.

Condiciones de uso

- La desviación estándar poblacional σ es conocida.
- La población es normal o la muestra es grande ($n \geq 30$), por el teorema central del límite.

Ejemplo

Supongamos que queremos estimar la media poblacional μ de un proceso, y sabemos que la desviación estándar poblacional es $\sigma = 12$. Se toma una muestra de tamaño $n = 36$ y se obtiene una media muestral $\bar{x} = 100$.

Queremos calcular un intervalo de confianza del 95 % para μ.

Datos:

- $\bar{x} = 100$
- $\sigma = 12$
- $n = 36$
- Nivel de confianza = 95 %

Paso 1: Determinar el valor crítico $Z_{\alpha/2}$

$$\alpha = 1 - 0{,}95 = 0{,}05, \quad \alpha/2 = 0{,}025$$

El valor crítico $Z_{\alpha/2}$ corresponde a la puntuación z tal que la probabilidad acumulada bajo la curva normal estándar hasta ese punto es:

$$P(Z < z_{\alpha/2}) = 1 - \alpha/2 = 0{,}975$$

Este valor se obtiene utilizando una **tabla Z de la distribución normal estándar** (acumulada desde la izquierda). Se busca el valor de z que deja 2.5 % en la cola derecha de la distribución, es decir, 97.5 % a la izquierda.

$$Z_{\alpha/2} = 1{,}96$$

Paso 2: Calcular el error estándar

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{12}{\sqrt{36}} = \frac{12}{6} = 2$$

Paso 3: Calcular el intervalo de confianza

$$IC = \bar{x} \pm Z_{\alpha/2} \cdot SE = 100 \pm 1{,}96 \cdot 2 = 100 \pm 3{,}92$$

$$\Rightarrow IC = (96{,}08,\ 103{,}92)$$

Por lo tanto, con un nivel de confianza del 95 %, se estima que la media poblacional μ se encuentra entre **96.08 y 103.92**.

$$IC = \bar{x} \pm Z_{\alpha/2} \cdot SE = 100 \pm 1{,}96 \times 2$$

$$IC = 100 \pm 3{,}92$$

$$IC = (96{,}08,\ 103{,}92)$$

2.2. Para la media con desviación estandar (σ) desconocida

Cuando la desviación estándar poblacional σ no es conocida, se utiliza la **distribución t de Student** en lugar de la distribución normal estándar (z).

Fórmula

$$IC = \bar{X} \pm t_{\alpha/2,\,df} \cdot \frac{s}{\sqrt{n}}$$

- $\bar{X}$: media muestral.
- s: desviación estándar muestral.
- n: tamaño de la muestra.
- $df = n - 1$: grados de libertad.
- $t_{\alpha/2,df}$: valor crítico t según el nivel de confianza y los grados de libertad.
- $\frac{s}{\sqrt{n}}$: error estándar de la media.

Pasos para hallar $t_{\alpha/2,\,df}$

1. Calcular $\alpha = 1 -$ nivel de confianza.
2. Dividir α entre 2: $\alpha/2$.
3. Calcular los grados de libertad: $df = n - 1$.
4. Buscar en la tabla t-Student el valor de t correspondiente a la probabilidad acumulada $1 - \alpha/2$ y grados de libertad df.

Condiciones de uso

- Se desconoce la desviación estándar poblacional σ.
- Generalmente se aplica cuando $n < 30$, aunque también se puede usar para muestras grandes si σ no está disponible.

Ejemplo

Suponga que se tiene:

- $\bar{X} = 50$
- $s = 12$
- $n = 16$
- Nivel de confianza: 95 %

Entonces:

$$df = n - 1 = 15, \quad t_{0{,}025,15} \approx 2{,}131$$

$$IC = 50 \pm 2{,}131 \cdot \frac{12}{\sqrt{16}} = 50 \pm 2{,}131 \cdot 3 = 50 \pm 6{,}393$$

Intervalo de confianza: (43,61, 56,39)

Nota

El valor $t_{\alpha/2,df}$ se obtiene de una tabla t de Student, o utilizando software como Excel, Python, R, o calculadoras estadísticas.

2.3. Para la diferencia de medias ($\mu_1 - \mu_2$) con (σ conocidas)

Cuando queremos estimar la diferencia entre las medias de dos poblaciones normales y conocemos las desviaciones estándar poblacionales (σ_1 y σ_2), usamos la distribución normal estándar para construir el intervalo de confianza.

Supuestos:

- Las dos poblaciones son normales o las muestras son grandes ($n_1 \geq 30$ y $n_2 \geq 30$).
- Las muestras son independientes.
- Se conocen las desviaciones estándar poblacionales σ_1 y σ_2.

Fórmula:

$$\left(\bar{X}_1 - \bar{X}_2\right) \pm Z_{\alpha/2} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Donde:

- $\bar{X}_1, \bar{X}_2$: medias muestrales de los grupos 1 y 2.
- σ_1, σ_2: desviaciones estándar poblacionales.
- n_1, n_2: tamaños de muestra de cada grupo.
- $Z_{\alpha/2}$: valor crítico de la distribución normal estándar correspondiente al nivel de confianza deseado.

Este intervalo nos da un rango en el cual es razonable suponer que está la verdadera diferencia entre las medias poblacionales $\mu_1 - \mu_2$, con el nivel de confianza establecido (por ejemplo, 95 %).

Ejemplo

Supongamos que se desea comparar el rendimiento promedio en una prueba matemática entre estudiantes de dos regiones. Se toman muestras independientes de cada región:

- Muestra 1 (EE.UU.): $n_1 = 50$, $\bar{X}_1 = 570$, $\sigma_1 = 102$
- Muestra 2 (Europa): $n_2 = 50$, $\bar{X}_2 = 540$, $\sigma_2 = 115$
- Nivel de confianza: $95\,\% \Rightarrow Z_{\alpha/2} = 1{,}96$

Paso 1: Calcular el error estándar (EE)

$$EE = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{102^2}{50} + \frac{115^2}{50}} = \sqrt{208{,}08 + 264{,}5} = \sqrt{472{,}58} \approx 21{,}74$$

Paso 2: Calcular el intervalo de confianza

$$(\bar{X}_1 - \bar{X}_2) \pm Z_{\alpha/2} \cdot EE = (570 - 540) \pm 1{,}96 \cdot 21{,}74 = 30 \pm 42{,}61$$

Resultado:

$$IC_{95\,\%} = (-12{,}61, 72{,}61)$$

Interpretación: Con un 95 % de confianza, la diferencia en medias poblacionales podría estar entre $-12{,}61$ y $72{,}61$. Como el intervalo incluye el 0, no hay evidencia suficiente para afirmar que hay una diferencia significativa entre las dos medias poblacionales.

2.4. Para la diferencia de medias $(\mu_1 - \mu_2)$ con (σ desconocidas y no iguales

Cuando se comparan dos medias poblacionales, μ_1 y μ_2, a partir de muestras independientes donde las varianzas poblacionales σ_1^2 y σ_2^2 son desconocidas y no se asume que sean iguales, se utiliza la distribución t de Student con grados de libertad aproximados (método de Welch).

Fórmula del intervalo de confianza al nivel $1 - \alpha$:

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2,\nu} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Donde:

- $\bar{x}_1, \bar{x}_2$: medias muestrales
- s_1^2, s_2^2: varianzas muestrales

- n_1, n_2: tamaños muestrales
- $t_{\alpha/2,\nu}$: valor crítico de la t-student para ν grados de libertad y nivel α
- ν: grados de libertad aproximados, calculados con la fórmula de Welch:

$$\nu \approx \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1-1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2-1}}$$

Ejemplo: Supón que se tienen dos muestras independientes con los siguientes datos:

- Muestra 1: $\bar{x}_1 = 570$, $s_1 = 102$, $n_1 = 50$
- Muestra 2: $\bar{x}_2 = 540$, $s_2 = 115$, $n_2 = 50$
- Nivel de confianza: $95\,\% \Rightarrow \alpha = 0{,}05$

Paso 1: Calcular el error estándar (EE):

$$EE = \sqrt{\frac{102^2}{50} + \frac{115^2}{50}} = \sqrt{\frac{10404}{50} + \frac{13225}{50}} = \sqrt{208{,}08 + 264{,}5} = \sqrt{472{,}58} \approx 21{,}74$$

Paso 2: Calcular los grados de libertad:

$$\nu \approx \frac{(208{,}08 + 264{,}5)^2}{\frac{(208{,}08)^2}{49} + \frac{(264{,}5)^2}{49}} \approx \frac{472{,}58^2}{883{,}7 + 1427{,}9} \approx \frac{223334}{2311{,}6} \approx 96{,}6$$

Paso 3: Determinar $t_{0{,}025,96} \approx 1{,}985$ (usando tabla t-student)

Paso 4: Construir el intervalo:

$$IC = (570 - 540) \pm 1{,}985 \cdot 21{,}74 = 30 \pm 43{,}14 \Rightarrow (-13{,}14,\ 73{,}14)$$

Interpretación: Con un 95 % de confianza, la diferencia entre las medias poblacionales está entre -13.14 y 73.14 puntos. Como el intervalo incluye 0, no hay evidencia estadísticamente significativa de una diferencia entre las medias poblacionales.

2.5. Margen de error para estimar la media poblacional

El margen de error (ME) representa el rango de variación permitido alrededor de la media muestral $\bar{x}$, dentro del cual se espera que se encuentre la media poblacional μ con un determinado nivel de confianza.

1. Cuando la desviación estándar poblacional σ es conocida

Si la desviación estándar poblacional σ es conocida y la muestra proviene de una población normal o el tamaño de la muestra es suficientemente grande, el margen de error se calcula como:

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

donde:

- $z_{\alpha/2}$ es el valor crítico de la distribución normal estándar correspondiente al nivel de confianza deseado.
- σ es la desviación estándar poblacional conocida.
- n es el tamaño de la muestra.

2. Cuando la desviación estándar poblacional σ es desconocida

Si σ es desconocida y se utiliza la desviación estándar muestral s, y la muestra proviene de una población normal, el margen de error se calcula con la distribución t de Student:

$$ME = t_{\alpha/2,\, n-1} \cdot \frac{s}{\sqrt{n}}$$

donde:

- $t_{\alpha/2,\, n-1}$ es el valor crítico de la distribución t de Student con $n-1$ grados de libertad, correspondiente al nivel de confianza deseado.
- s es la desviación estándar muestral.
- n es el tamaño de la muestra.

Ejemplo

Suponga que una muestra de tamaño $n = 64$ proviene de una población normal con desviación estándar poblacional conocida $\sigma = 12$. Se desea calcular el margen de error para estimar la media poblacional con un nivel de confianza del 98 %.

- Nivel de confianza: 98 %, entonces $\alpha = 1 - 0{,}98 = 0{,}02$.
- $\alpha/2 = 0{,}01$.
- El valor crítico $z_{\alpha/2}$ para 98 % se encuentra en la tabla Z y es aproximadamente 2,33.
- Tamaño de la muestra: $n = 64$.

Cálculo del margen de error:

$$ME = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = 2{,}33 \cdot \frac{12}{\sqrt{64}} = 2{,}33 \cdot \frac{12}{8} = 2{,}33 \cdot 1{,}5 = 3{,}495.$$

Por lo tanto, el margen de error es aproximadamente 3,495.

Esto significa que el intervalo de confianza para la media poblacional μ será $\bar{x} \pm 3{,}495$.

2.6. Para una proporción

Dado un experimento aleatorio con una proporción muestral $\hat{p}$ obtenida a partir de una muestra aleatoria simple de tamaño n, el intervalo de confianza para la proporción poblacional p se puede estimar bajo ciertas condiciones.

Para usar la distribución normal

La aproximación normal es válida si se cumple que:

$$\begin{aligned} np &\geq 5 \\ n(1-p) &\geq 5 \end{aligned}$$

En la práctica, dado que p es desconocido, se suele verificar usando $\hat{p}$:

$$\begin{aligned} n\hat{p} &\geq 5 \\ n(1-\hat{p}) &\geq 5 \end{aligned}$$

Fórmula del intervalo de confianza (con aproximación normal)

Si se cumplen las condiciones anteriores, se puede construir un intervalo de confianza de $(1-\alpha) \cdot 100\,\%$ para p usando la distribución normal estándar:

$$\hat{p} \pm Z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \tag{1}$$

Donde:

- $\hat{p}$: proporción muestral
- n: tamaño de la muestra
- $Z_{\alpha/2}$: valor crítico de la distribución normal estándar correspondiente al nivel de confianza deseado (por ejemplo, $Z_{0{,}025} = 1{,}96$ para un 95 % de confianza)

Ejemplo

En una encuesta realizada a 200 personas, 130 afirmaron estar satisfechas con un servicio. Se desea construir un intervalo de confianza del 95 % para la proporción real de personas satisfechas.

- Tamaño de la muestra: $n = 200$
- Proporción muestral: $\hat{p} = \frac{130}{200} = 0{,}65$
- Nivel de confianza: 95 % $\Rightarrow \alpha = 0{,}05$, entonces $Z_{\alpha/2} = 1{,}96$

Verificamos las condiciones para usar la aproximación normal:

$$n\hat{p} = 200 \cdot 0{,}65 = 130 \geq 5$$
$$n(1-\hat{p}) = 200 \cdot 0{,}35 = 70 \geq 5$$

Calculamos el error estándar:

$$\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0{,}65 \cdot 0{,}35}{200}} = \sqrt{\frac{0{,}2275}{200}} \approx \sqrt{0{,}0011375} \approx 0{,}0337$$

Construimos el intervalo:

$$0{,}65 \pm 1{,}96 \cdot 0{,}0337 \Rightarrow 0{,}65 \pm 0{,}0661$$

$$\text{Intervalo de confianza: } (0{,}5839,\ 0{,}7161)$$

Interpretación: Con un 95 % de confianza, se estima que la proporción real de personas satisfechas con el servicio está entre 58.39 % y 71.61 %.

2.7. Para la diferencia de proporciones $(\hat{p}_1 - \hat{p}_2)$

Cuando se desea comparar dos proporciones poblacionales, p_1 y p_2, a partir de muestras independientes y grandes, se puede construir un intervalo de confianza para la diferencia $p_1 - p_2$ utilizando la distribución normal estándar.

Fórmula del intervalo de confianza al nivel $1-\alpha$:

$$(\hat{p}_1 - \hat{p}_2) \pm Z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

Condiciones de aplicabilidad:

- Las muestras deben ser independientes.

- Cada muestra debe ser suficientemente grande:

$$n_i \cdot \hat{p}_i > 5 \quad \text{y} \quad n_i \cdot (1 - \hat{p}_i) > 5 \quad \text{para } i = 1, 2$$

Ejemplo: Una encuesta encuentra que el 52 % de 400 personas en la ciudad A están a favor de una nueva ley, mientras que en la ciudad B, el 47 % de 350 personas encuestadas están a favor. Construya un intervalo de confianza del 95 % para la diferencia en proporciones.

Datos:

$$\hat{p}_1 = 0{,}52, \quad n_1 = 400, \quad \hat{p}_2 = 0{,}47, \quad n_2 = 350, \quad Z_{0{,}025} = 1{,}96$$

Paso 1: Calcular el error estándar (EE):

$$\begin{aligned} EE &= \sqrt{\frac{0{,}52(1-0{,}52)}{400} + \frac{0{,}47(1-0{,}47)}{350}} \\ &= \sqrt{\frac{0{,}2496}{400} + \frac{0{,}2491}{350}} \\ &= \sqrt{0{,}000624 + 0{,}0007117} \\ &= \sqrt{0{,}0013357} \\ &\approx 0{,}03654 \end{aligned}$$

Paso 2: Calcular el intervalo de confianza:

$$(0{,}52 - 0{,}47) \pm 1{,}96 \cdot 0{,}03654 = 0{,}05 \pm 0{,}0716$$

$$IC_{95\,\%} = (-0{,}0216,\ 0{,}1216)$$

Interpretación: Con un 95 % de confianza, la diferencia real en proporciones entre las dos ciudades está entre -2.16 % y 12.16 %. Como el intervalo incluye el 0, no hay evidencia estadísticamente significativa de que una ciudad tenga una mayor proporción de apoyo que la otra.

3. Pruebas de hipótesis

4. Anexos

4.1. 1. Medidas de tendencia central

- **Media (promedio):**

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

- **Mediana:** Valor central cuando los datos están ordenados. Si n es impar: Mediana = $x_{(n+1)/2}$ Si n es par: Mediana = $\frac{x_{n/2}+x_{(n/2)+1}}{2}$
- **Moda:** Valor que más se repite en el conjunto de datos.

4.2. 2. Medidas de dispersión

- **Rango:**
$$R = x_{\text{máx}} - x_{\text{mín}}$$
- **Varianza (poblacional):**
$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2$$
- **Varianza (muestral):**
$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$
- **Desviación estándar:**
$$\sigma = \sqrt{\sigma^2} \quad \text{o} \quad s = \sqrt{s^2}$$
- **Coeficiente de variación:**
$$CV = \frac{s}{\bar{x}} \times 100\,\%$$

4.3. 3. Medidas de posición

- **Cuartil Q_k:**
$$Q_k = \text{Valor que deja } \frac{k}{4} \text{ de los datos por debajo,} \quad k = 1, 2, 3$$
- **Percentil P_k:**
$$P_k = \text{Valor que deja } \frac{k}{100} \text{ de los datos por debajo}$$
- **Decil D_k:**
$$D_k = \text{Valor que deja } \frac{k}{10} \text{ de los datos por debajo}$$

4.4. 4. Medidas de forma

- **Asimetría (Skewness):**

$$\text{Asimetría} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{s}\right)^3$$

- **Curtosis:**

$$\text{Curtosis} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{s}\right)^4$$

4.5. 5. Indicadores de inferencia estadística

- **Error estándar de la media:**

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

- **Estadístico Z:**

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

- **Estadístico t (cuando no se conoce σ):**

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}, \quad \text{con } n-1 \text{ grados de libertad}$$

- **Intervalo de confianza para la media (con σ conocida):**

$$IC = \bar{x} \pm z \cdot \frac{\sigma}{\sqrt{n}}$$