VI. Estadística aplicada

#developed by **Roberto Angel Melendez-Armenta**

angelarmenta - Overview

angelarmenta has 6 repositories available. Follow their code on GitHub.

https://github.com/angelarmenta

6.1 Muestreo

El **muestreo** es el proceso de seleccionar un subconjunto representativo de una población con el objetivo de realizar inferencias sobre dicha población.

6.1.1 Tipos de muestreo

- **Muestreo aleatorio simple:** Cada elemento tiene la misma probabilidad de ser elegido.
- **Muestreo sistemático:** Se selecciona cada k-ésimo elemento de una lista ordenada.
- **Muestreo estratificado:** La población se divide en estratos homogéneos y se selecciona una muestra aleatoria de cada estrato.
- **Muestreo por conglomerados:** Se divide la población en grupos o conglomerados heterogéneos y se seleccionan algunos conglomerados completos.

6.1.2 Teorema del límite central

El teorema del límite central establece que, cuando el tamaño de la muestra es suficientemente grande ($n \geq 30$), la distribución de las medias muestrales de cualquier población con media μ y desviación estándar σ se aproxima a una distribución normal, independientemente de la forma de la distribución original.

Esto permite utilizar la distribución normal para realizar inferencias sobre medias, aun cuando la población no sea normal.

6.1.3 Distribución muestral de la media

La distribución muestral de la media es la distribución de todas las posibles medias que se podrían obtener al tomar muestras del mismo tamaño de una población.

Propiedades:

- Media de la distribución: $\mu_{\bar{x}} = \mu$
- Error estándar de la media: $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

Si la población tiene distribución normal o el tamaño de la muestra es grande ($n \geq 30$), entonces la media muestral $\bar{x}$ sigue una distribución aproximadamente normal.

6.1.4 Distribución muestral de una proporción

Es la distribución de las proporciones muestrales $\hat{p}$ calculadas a partir de todas las posibles muestras de tamaño n extraídas de una población con proporción p.

Propiedades:

- Media: $\mu_{\hat{p}} = p$
- Error estándar: $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$

La distribución de $\hat{p}$ se aproxima a una distribución normal cuando se cumplen las condiciones: $np > 5$ y $n(1-p) > 5$.

6.2 Estimación

La estimación consiste en inferir valores de parámetros poblacionales a partir de estadísticos muestrales.

6.2.1 Estimación puntual

Es un solo valor que se utiliza como aproximación del parámetro poblacional.

- Ejemplo: $\bar{x}$ estima a μ

6.2.2 Estimación por intervalo

Es un rango de valores que tiene una cierta probabilidad (nivel de confianza) de contener el verdadero parámetro poblacional.

6.2.3 Intervalo de confianza para una media

- Si σ conocida: $\bar{x} \pm Z_{\alpha/2} \cdot \left(\frac{\sigma}{\sqrt{n}}\right)$
- Si desconocida: $\bar{x} \pm t_{\alpha/2,n-1} \cdot \left(\frac{s}{\sqrt{n}}\right)$

6.2.4 Intervalo de confianza para una proporción

Cuando se desea estimar la proporción de éxitos en una población a partir de una muestra, se construye un intervalo de confianza que proporciona un rango dentro del cual probablemente se encuentre la proporción verdadera, p, con un cierto nivel de confianza (por ejemplo, 95%).

La proporción muestral se denota como $\hat{p} = \frac{x}{n}$, donde:

- x: número de éxitos de la muestra.
- n: tamaño de la muestra.

El **intervalo de confianza para una proporción** está dado por la siguiente fórmula:

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

Donde:

- $\hat{p}$: proporción muestral.
- $Z_{\alpha/2}$: valor crítico de la distribución normal estándar para el nivel de confianza deseado (por ejemplo, 1.96 para 95%).
- n: tamaño de la muestra.

Condiciones de aplicabilidad: Para que la aproximación normal sea válida, deben cumplirse las condiciones: $np \geq 5$ y $n(1-p) \geq 5$. Esto asegura que la distribución muestral de se aproxima bien a una distribución normal.

6.3 Prueba de hipótesis

La prueba de hipótesis es una herramienta estadística para tomar decisiones o extraer conclusiones sobre una población, basándose en información muestral.

6.3.1 Errores tipo I y II

- Error tipo I (**α**): Rechazar la hipótesis nula H_0 cuando es verdadera.
- Error tipo II (**β**): No rechazar H_0 cuando en realidad es falsa.
- Potencia de la prueba: $1 - \beta$, es la probabilidad de detectar correctamente un efecto.

6.3.2 Pasos para realizar una prueba de hipótesis

1. Formular H_0 y H_a.
2. Elegir el nivel de significancia de α.
3. Seleccionar el estadístico de la prueba (Z, t, etc.)
4. Establecer la región crítica o calcular el valor p.
5. Recolectar datos y calcular el estadístico.
6. Tomar una decisión:
    a. Si valor de $p < \alpha$ entonces rechazar H_0.
    b. Si valor de $p \geq \alpha$ entonces no rechazar H_0.

6.3.3 Prueba de hipótesis para una media

Si la desviación estándar es desconocida entonces utilizar t de Student:

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

Donde:

- $\bar{x}$: media muestral.
- μ_0: media hipotética.
- s: desviación estándar muestral.
- n: tamaño de la muestra.

6.3.4 Prueba de hipótesis para una proporción

Se usa el estadístico Z cuando se cumplen las condiciones de normalidad:

$$Z = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Donde:

- $\hat{p}$: proporción muestral.
- p_0: proporción poblacional hipotética.
- n: tamaño de la muestra.