Distribuciones de muestreo de estadísticas

Sergio Cendales

Abril 6, 2020

Índice

1. Definiciones preliminares.
2. Distribución de muestreo.
3. Definición de $\overline{X}$.
4. Distribución de muestreo de $\overline{X}$.
5. Teorema del Límite Central.
6. Ejemplo 1
7. Ejemplo 2

Definiciones preliminares

- **Objetivo de la estadística** Realizar inferencia de parámetros desconocidos con buenos estimadores.
- **Muestreo aleatorio simple:** Cada muestra de tamaño n tiene la misma probabilidad de ser seleccionada.
- **Variable aleatoria:** $X : \Omega \to R$
- **Parámetro θ:** Característica numérica de la distribución de la población de manera que describe parcial o completamente a la función de masa $p(X;\theta)$ o densidad de probabilidad $f(X;\theta)$.
- **Estadística (T):** Función de la variable aleatoria. Al ser función de variable aleatoria, ¡también es variable aleatoria! $T = u(X)$.
- **Estimador:** Estadístico empleado para estimar parámetros desconocidos.

Distribución de muestreo

Es la distribución de probabilidad de la estadística T que puede obtenerse como resultado de realizar un número infinito de muestreos aleatorios independientes de tamaño n.

Definición de $\overline{X}$

Sea $X_1, X_2, ..., X_n$ una muestra aleatoria que consiste en n variables aleatorias independientes e identicamente distribuídas (IID) tal que $E(X_i) = \mu$ y $Var(X_i) = \sigma^2$ $\forall i$ en $1 : n$
Entonces la estadística

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} \tag{1}$$

se define como la media muestral de las n variables aleatorias IID.
Por lo tanto, $E(\overline{X}) = \mu$ y $Var(\overline{X}) = \frac{\sigma^2}{n}$, donde μ y σ^2 son la media y la varianza de la distribución de la población a partir de la cual fue tomada la muestra (entre más grande sea la muestra, la precisión de la media muestral para la estimar la media poblacional aumenta).

Distribución de muestreo de $\overline{X}$

Teorema

Sea $X_1, X_2, ..., X_n$ una muestra aleatoria que consiste en n variables aleatorias independientes y normalmente distribuídas (IID) tal que $E(X_i) = \mu$ y $Var(X_i) = \sigma^2$ conocida $\forall i$ en $1 : n$, entonces

$$\overline{X} \sim N(\mu, \frac{\sigma^2}{n}) \tag{2}$$

Ver demostración con función generadora de momentos

¿Cual es la distribución de muestreo de $\overline{X}$ cuando no puede especificarse la distribución de probabilidad de a población a partir de la cual se obtiene la muestra?

Teorema del Límite Central

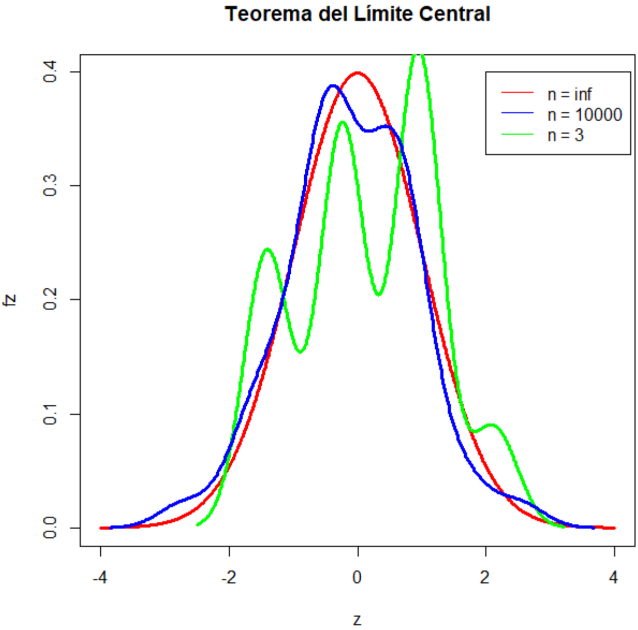
Teorema

Sea $X_1, X_2, ..., X_n$ una muestra aleatoria que consiste en n variables aleatorias independientes identicamente distribuídas con una distribución de probabilidad no especificada y que tiene una media μ y una varianza σ^2. Conforme el tamaño de la muestra tiende a ∞

$$\overline{X} \sim N(\mu, \frac{\sigma^2}{n}) \tag{3}$$

Por lo tanto, conforme el tamaño de la muestra tiende a ∞, la estadística

$$Z = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1) \tag{4}$$



Teorema del Límite Central
n = inf
n = 10000
n = 3
fz
z
0.0
0.1
0.2
0.3
0.4
-4
-2
0
2
4

Ejemplo 1

Un inspector general de pesos y medias visita una planta de empacado para verificar que el peso neto de las cajas sea el indicado en éstas. El gerente de la planta asegura al inspector que el peso promedio de cada caja es de 750 gr con una desviación estándar de 5 gr. El inspector selecciona, al azar, 100 cajas y encuentra que el peso promedio es de 748 gr. Bajo estas condiciones, ¿qué tan probable es tener un peso de 748 o menos? ¿qué actitud debe tomar el inspector?

$$P(\overline{X} \leq 748) = P(Z \leq -4) = \int_{-\infty}^{-4} \frac{1}{\sqrt{2\pi}} e^{-(x)^2/2} dx = 3.167124e{-}05 \quad (5)$$

Ejemplo 2

Durante los 12 meses pasados el volumen promedio diario de ventas de un restaurante fue de 2.000 pesos. El gerente piensa que los próximo 25 días serán típicos con respecto al volumen de ventas normal. Al finalizar los 25 días, el volumen promedio de ventas y su desviación estándar fueron de 1800 y 200 pesos respectivamente. Supóngase que el volumen de ventas diario es una variable aleatoria normal. Si usted fuese el gerente ¿tendría alguna razón para creer, con base en este resultado, que hubo una disminución en el volumen promedio de ventas diario?

Fin