

Estadística Aplicada

Universidad Maimónides
2016

Clase 5 – Distribución de la Media Muestral

Pedro Elosegui

Métodos y Distribuciones de Muestreo

- En estadística nos gustaría contar con los parámetros de la población, pero esto no es siempre posible
 - Hay pruebas que son destructivas (controles de calidad)
 - La población es muy grande
 - Los costos de estudiar cada integrante de la población es muy alto
 - Llevaría mucho tiempo
- El análisis de la muestra es más que razonable si está bien diseñado el estudio

Métodos y Distribuciones de Muestreo

- Muestra probabilística
 - Cada integrante de la población tiene una misma probabilidad de ser incluido en la muestra
 - Muestreo aleatorio simple (usando tabla de números aleatorios)
 - Muestreo aleatorio sistemático (se seleccionan las k -ésimas observaciones)
 - Muestreo aleatorio estratificado (según subgrupos de la población o estratos)
- Error de Muestreo
 - Diferencia entre el valor del estadístico de muestra y su parámetro poblacional

Métodos y Distribuciones de Muestreo

- Error de Muestreo
 - Diferencia entre el valor del estadístico de muestra y su parámetro poblacional
 - Si existe, cómo se puede saber en base a una muestra si el verdadero valor del *parámetro* poblacional coincide con el *estadístico* muestral?

Distribuciones de Muestreo de Medias Muestrales

- Distribución de Muestreo de Medias
 - Es una distribución probabilística que consta de todas las medias muestrales posibles de un tamaño de muestra dado de una población y la probabilidad de ocurrencia asociada a cada media muestral
 - Tomamos muestras exhaustivas de una población, calculamos la media muestral en cada caso, la media de las medias muestrales coincide con la media poblacional
 - Ejemplo.
 - Población: 2, 3, 4, 5, 6
 - Muestras de a dos: (2,3); (2,4).....
 - Calcule media poblacional y media de muestras

Distribución de Media Muestral (Ej.)

- El despacho de abogados Hoya & Asociados tiene cinco socios.
- En su junta de socios semanal cada uno informa el número de horas que cobraron a los clientes por su servicios la semana anterior.

Socio	Horas
Dunn	22
Hardy	26
Kiers	30
Malinowski	26
Tillman	22

Distribución Muestral (Ej.)

- Calcule la media de horas de la población
- Calcule todas las muestras posibles tomadas de a dos abogados
- Calcule la media muestral para cada muestra y el error muestral
- Organice la distribución de muestreo (frecuencia). Grafique.
- Se afirma que la media muestral es igual a la media poblacional. Es correcta esta aseveración?

Teorema del Límite Central

- Aproximación a la distribución normal
 - Si se seleccionan todas las muestras posibles de una población de un tamaño determinado, la distribución de medias *muestrales* se acercará a una distribución normal
 - La aproximación aumenta cuando aumentamos el tamaño de la muestra
 - Cualquiera sea la distribución probabilística original de la población

Teorema del Límite Central

- Para una población con media μ y variancia σ^2 , la distribución de muestreo de las medias de todas las muestras posibles de tamaño n obtenidas de una población tendrá una distribución normal aproximada
- La media de la distribución de muestreo será igual a μ y la variancia igual a σ^2/n
- Si se supone que el tamaño de la muestra es suficientemente grande.

Estimaciones puntuales e intervalos de confianza

- Estimación puntual
 - Es el valor estadístico calculado a partir de la muestra para aproximar un parámetro poblacional
 - Por ej. la media muestral a la media poblacional, el error estándar y el desvío estándar

- Intervalo de confianza
 - Rango de valores obtenidos a partir de la muestra
 - El parámetro poblacional se encuentra en dicho rango con alguna probabilidad específica (nivel de confianza)

Intervalos de confianza

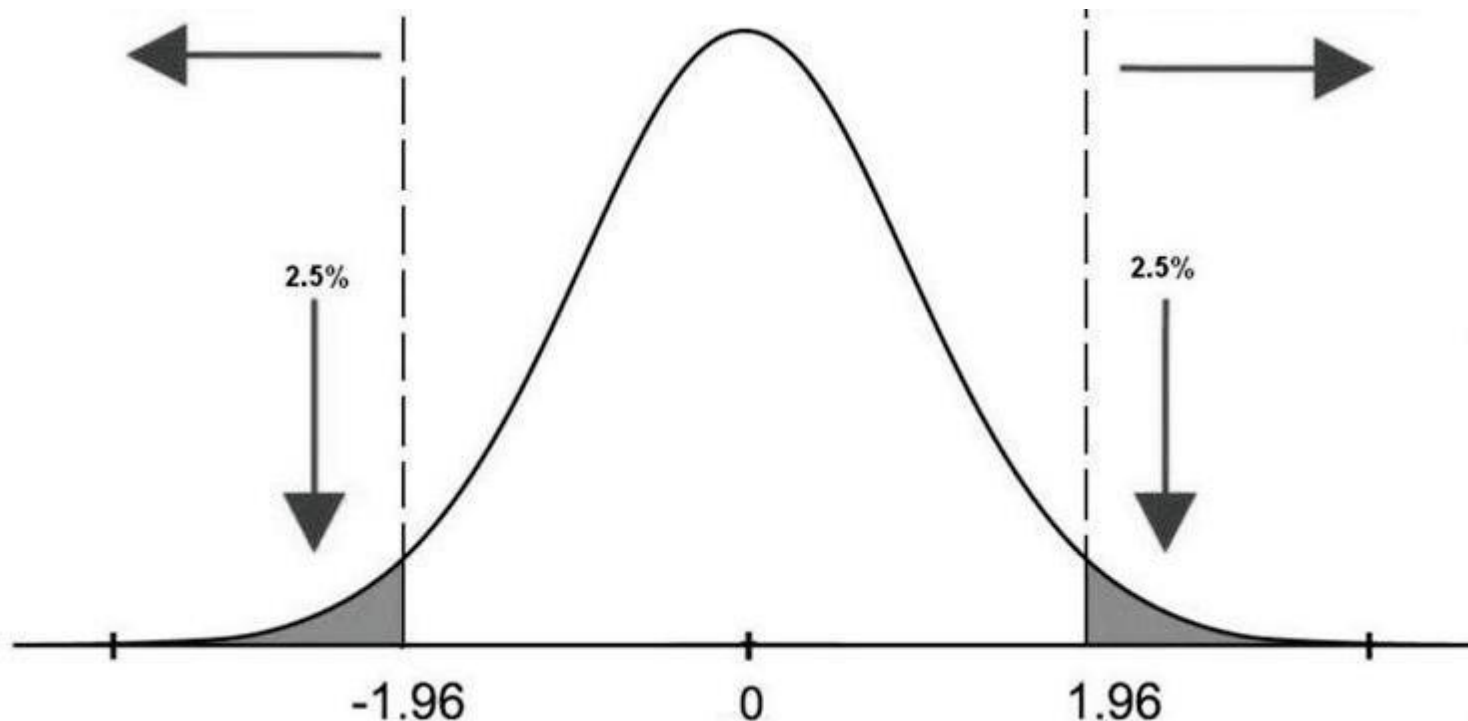
- Una **estimación de intervalo** establece la amplitud en la que es posible que se encuentre un parámetro poblacional.
- El intervalo dentro del cual se espera que esté un parámetro poblacional se llama **intervalo de confianza**.
- Los dos intervalos de confianza que más se usan son 95% y 99%.

Estimaciones puntuales e intervalos de confianza

- Intervalo de confianza
 - Un 95% de las medias muestrales se ubican en un entorno de 1.96 desvíos estándar de la media poblacional
 - Un 99 % de las medias muestrales se ubican en un entorno de 2.58 desvíos estándar de la media poblacional

Estimaciones puntuales e intervalos de confianza

- Intervalo de confianza
 - Un 95% de las medias muestrales se ubican en un entorno de 1.96 desvíos estándar de la media poblacional



Intervalo de Confianza

- Error estándar de la media muestral
 - Desviación estándar de la distribución de muestreo de las medias muestrales (variabilidad)
 - Se calcula mediante

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$
 - $\sigma_{\bar{x}}$ es el símbolo del error estándar de las medias muestrales.
 - σ es la desviación estándar de la población.
 - n es el tamaño de la muestra.

Intervalo de Confianza

- Error estándar de la media muestral
 - Si σ no se conoce y $n \geq 30$, la desviación estándar de la muestra, denotada por s , se usa para aproximar la desviación estándar poblacional.
 - La fórmula para el error estándar se convierte en:

$$s_{\bar{x}} = \frac{s}{\sqrt{n}}$$

Intervalo de Confianza

- Intervalo de confianza de un 95% para una media

$$\bar{X} \pm 1.96 \frac{s}{\sqrt{n}}$$

- Intervalo de confianza de un 99% para la media

$$\bar{X} \pm 2.58 \frac{s}{\sqrt{n}}$$

- Intervalo de confianza, fórmula general

$$\bar{X} \pm Z \frac{s}{\sqrt{n}}$$

Intervalo de Confianza

- Si se toma una muestra de 256 salarios de asistentes contables y se obtiene una media muestral de \$ 12000 y una varianza muestral de \$1560
 - Cuál es el salario medio estimado de todos los asistentes contables (estimación puntual de la media poblacional)?
 - Cuál sería un intervalo de confianza del 95% para la media poblacional?
 - Interprete los resultados

Intervalo de Confianza

- El director de la escuela de administración desea estimar el número medio de horas por semana que estudian los alumnos. Una muestra de 49 estudiantes dio una media de 24 h con desviación estándar de 4 h.
 - Cuál es la estimación puntual de horas de estudio?
 - Cuál es el intervalo de confianza de 95% para el número promedio de horas por semana que estudian los alumnos?
 - Cuál es el límite inferior y superior?

Intervalo de Confianza

- Si se usa un IC de 95% para la media poblacional, se tiene

$$22.88 < \mu < 25.12$$

- Los puntos terminales del intervalo de confianza son los límites de confianza. El límite inferior es 22.88 y el límite superior es 25.12

Factor de corrección de población finita

- Una población con una cota superior (fija) se define como *finita*.
- Para una población finita, donde el número total de objetos es N y el tamaño de la muestra es n , se hace el siguiente ajuste a los errores estándar de las medias muestrales.
- Error estándar de las medias muestrales:

$$s_x = \frac{c}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

Selección del tamaño de muestra

- Existen tres factores que determinan el tamaño de una muestra, ninguno de ellos tiene una relación directa con el tamaño de la población.
- El grado de confianza elegido.
 - El error máximo permitido.
 - La variación o dispersión en la población.

Variación en la población

- Tamaño de la muestra para la media: una fórmula computacional conveniente para determinar n es:

$$n = \left(\frac{Z \cdot S}{E} \right)^2$$

- donde: E es el error permitido, Z es el valor normal estándar asociado con el grado de confianza seleccionado y S es la desviación estándar estimada del estudio piloto.

Ejercicio de Tamaño de la Muestra

- Un grupo de consumidores desea estimar la media mensual en los recibos de luz para una casa unifamiliar.
- Según estudios similares la desviación estándar se estima en \$20.00.
- Se desea un nivel de confianza de 99%, con una precisión de \$5.00.
- Qué tamaño de muestra se requiere?

Ejercicio de Tamaño de la Muestra

