Inferencia Estadística

Solución: Intervalos de Confianza

DSLab

noviembre, 2024

Ejercicio 1: Conceptos Básicos

Pregunta: Explica brevemente qué es un intervalo de confianza y por qué es importante en la estadística inferencial. Da un ejemplo sencillo de un intervalo de confianza.

Solución:

¿Qué es un intervalo de confianza?

Un intervalo de confianza es un rango de valores, derivado de los datos muestrales, que se utiliza para estimar un parámetro desconocido de la población, como la media o la proporción. Este intervalo se construye de tal manera que, con un cierto nivel de confianza (por ejemplo, 95% o 99%), se espera que contenga el verdadero valor del parámetro poblacional.

Importancia en la estadística inferencial

1. **Estimación de Parámetros**: Proporciona una estimación del parámetro poblacional en lugar de un solo valor puntual, ofreciendo así más información sobre la precisión de la estimación.
2. **Cuantificación de la Incertidumbre**: Permite cuantificar la incertidumbre asociada a la estimación. Un intervalo de confianza más estrecho indica una estimación más precisa.

3. **Toma de Decisiones Informadas**: Ayuda en la toma de decisiones basadas en datos al proporcionar un rango de valores posibles para el parámetro poblacional. Esto es crucial en áreas como la investigación científica, la economía, la medicina y muchas otras.

Ejemplo Sencillo de un Intervalo de Confianza

Supongamos que estamos interesados en estimar la media de las alturas de los estudiantes en una universidad. Tomamos una muestra aleatoria de 30 estudiantes y encontramos que la media muestral ($\bar{x}$) es de 170 cm y la desviación estándar (σ) es de 10 cm.

Queremos calcular un intervalo de confianza del 95% para la media poblacional (μ).

1. **Determinar el nivel de confianza**: 95%
   - El nivel de significancia (α) es 0.05.
   - El valor crítico (z) para un nivel de confianza del 95% es aproximadamente 1.96.

2. **Calcular el error estándar**:

$$\text{Error estándar} = \frac{\sigma}{\sqrt{n}} = \frac{10}{\sqrt{30}} \approx 1.83$$

3. **Calcular el margen de error**:

$$\text{Margen de error} = z \times \text{Error estándar} = 1.96 \times 1.83 \approx 3.59$$

4. **Calcular el intervalo de confianza**:

$$\text{Intervalo de confianza} = \bar{x} \pm \text{Margen de error} = 170 \pm 3.59$$

$$\text{Intervalo de confianza} = (166.41, 173.59)$$

Interpretación

Con un 95% de confianza, podemos afirmar que la media de las alturas de todos los estudiantes en la universidad está entre 166.41 cm y 173.59 cm. Esto significa que si repitiéramos este proceso de muestreo muchas veces, aproximadamente el 95% de los intervalos calculados de esta manera contendrían la verdadera media poblacional.

Ejercicio 2: Intervalo de Confianza Básico

Pregunta: Dada una muestra aleatoria de tamaño 30 de una población normal con media muestral $\bar{x} = 50$ y desviación estándar $\sigma = 10$, calcula el intervalo de confianza del 95% para la media poblacional.

Solución

Datos proporcionados

- Tamaño de la muestra (n): 30
- Media muestral ($\bar{x}$): 50
- Desviación estándar (σ): 10
- Nivel de confianza: 95%

Paso 1: Determinar el valor crítico

Para un nivel de confianza del 95%, el valor crítico z (para una distribución normal estándar) es aproximadamente 1.96.

Paso 2: Calcular el error estándar de la media

El error estándar de la media se calcula como:

$$\text{Error estándar} = \frac{\sigma}{\sqrt{n}} = \frac{10}{\sqrt{30}}$$

Paso 3: Calcular el margen de error

El margen de error se calcula como:

$$\text{Margen de error} = z \times \text{Error estándar} = 1.96 \times \frac{10}{\sqrt{30}}$$

Paso 4: Calcular el intervalo de confianza

El intervalo de confianza se calcula como:

$$\text{Intervalo de confianza} = \bar{x} \pm \text{Margen de error}$$

Ahora, realicemos los cálculos:

Cálculo del error estándar de la media

$$\text{Error estándar} = \frac{10}{\sqrt{30}} \approx 1.8257$$

Cálculo del margen de error

$$\text{Margen de error} = 1.96 \times 1.8257 \approx 3.5774$$

Cálculo del intervalo de confianza

$$\text{Intervalo de confianza} = 50 \pm 3.5774 = (46.4226, 53.5774)$$

Resultado

El intervalo de confianza del 95% para la media poblacional es aproximadamente $(46.42, 53.58)$. Esto significa que con un 95% de confianza, podemos afirmar que la verdadera media poblacional se encuentra dentro de este rango.

Ejercicio 3: Nivel de Confianza Diferente

Pregunta: Utilizando la misma muestra del ejercicio anterior, calcula el intervalo de confianza del 99% para la media poblacional. Compara ambos intervalos y explica las diferencias.

Solución: Para calcular el intervalo de confianza del 99% para la media poblacional utilizando la misma muestra, seguimos los mismos pasos que antes pero con el valor crítico correspondiente al 99% de confianza.

Datos proporcionados

- Tamaño de la muestra (n): 30
- Media muestral ($\bar{x}$): 50
- Desviación estándar (σ): 10
- Nivel de confianza: 99%

Paso 1: Determinar el valor crítico

Para un nivel de confianza del 99%, el valor crítico z (para una distribución normal estándar) es aproximadamente 2.576.

Paso 2: Calcular el error estándar de la media

El error estándar de la media es el mismo que antes:

$$\text{Error estándar} = \frac{10}{\sqrt{30}} \approx 1.8257$$

Paso 3: Calcular el margen de error

El margen de error se calcula como:

$$\text{Margen de error} = z \times \text{Error estándar} = 2.576 \times 1.8257 \approx 4.7015$$

Paso 4: Calcular el intervalo de confianza

El intervalo de confianza se calcula como:

$$\text{Intervalo de confianza} = \bar{x} \pm \text{Margen de error}$$

Cálculo del intervalo de confianza

$$\text{Intervalo de confianza} = 50 \pm 4.7015 = (45.2985, 54.7015)$$

Comparación de los intervalos

- **Intervalo de confianza del 95%**: $(46.42, 53.58)$
- **Intervalo de confianza del 99%**: $(45.30, 54.70)$

Explicación de las diferencias

1. **Ancho del intervalo**:
    - El intervalo de confianza del 99% es más amplio que el del 95%. Esto se debe a que al aumentar el nivel de confianza, necesitamos un intervalo más amplio para asegurarnos de que contiene el parámetro poblacional con mayor certeza.
2. **Precisión vs. Confianza**:
    - Un intervalo de confianza más amplio (como el del 99%) proporciona mayor seguridad de que contiene la verdadera media poblacional, pero a costa de ser menos preciso.
    - Un intervalo más estrecho (como el del 95%) es más preciso pero conlleva un menor nivel de certeza de que contiene la verdadera media poblacional.
3. **Valor crítico z**:

- El valor crítico para el 99% de confianza (2.576) es mayor que el valor crítico para el 95% de confianza (1.96), lo que contribuye al aumento del margen de error y, por ende, al ancho del intervalo.

En resumen, al aumentar el nivel de confianza del 95% al 99%, el intervalo de confianza se vuelve más amplio, lo que refleja un mayor grado de certeza de que el intervalo incluye la verdadera media poblacional.

Ejercicio 4: Interpretación Práctica

Pregunta: Un investigador encuentra que el intervalo de confianza del 95% para la media de una variable es (45, 55). Explica qué significa este intervalo en términos prácticos y cómo debería interpretarlo el investigador.

Solución:

Interpretación del Intervalo de Confianza

Cuando un investigador encuentra que el intervalo de confianza del 95% para la media de una variable es (45, 55), esto significa lo siguiente:

1. **Rango de Valores Probables**:
    - El intervalo (45, 55) representa un rango de valores dentro del cual se estima que se encuentra la verdadera media poblacional de la variable con un nivel de confianza del 95%.
2. **Nivel de Confianza**:
    - El nivel de confianza del 95% indica que si se repitiera el proceso de muestreo muchas veces y se calcularan intervalos de confianza del 95% para cada muestra, aproximadamente el 95% de esos intervalos incluirían la verdadera media poblacional.
3. **Incertidumbre**:
    - El intervalo refleja la incertidumbre inherente a la estimación. Debido a la variabilidad de la muestra, no podemos estar 100% seguros de que la media poblacional esté exactamente en un punto específico, pero podemos afirmar con un 95% de confianza que está dentro del intervalo calculado.

Cómo Interpretarlo el Investigador

El investigador debería interpretar el intervalo de confianza (45, 55) de la siguiente manera:

1. **Estimación de la Media Poblacional**:

  - La mejor estimación de la media poblacional basada en la muestra es que se encuentra entre 45 y 55.
2. **Toma de Decisiones**:
   - Si el investigador necesita tomar decisiones basadas en la media de la población, puede hacerlo con la confianza de que la verdadera media está muy probablemente dentro de este rango. Por ejemplo, si está considerando un límite superior o inferior para una variable crítica, el intervalo le ayuda a entender dónde podría estar la verdadera media.
3. **Informe de Resultados**:
   - Al comunicar los resultados a otros, el investigador debe mencionar que el intervalo de confianza del 95% para la media es (45, 55). Esto proporciona una visión clara de la precisión y fiabilidad de la estimación basada en los datos de la muestra.

Ejemplo Práctico

Supongamos que el investigador está estudiando la media de una variable como la puntuación de satisfacción del cliente para un producto. Si el intervalo de confianza del 95% para la media de satisfacción es (45, 55), puede interpretar que:

- **Satisfacción General**: La satisfacción promedio de los clientes está en algún lugar entre 45 y 55 en la escala utilizada, con un 95% de confianza.
- **Comparación de Productos**: Si se está comparando este producto con otro cuya satisfacción promedio conocida es 60, el investigador puede argumentar que es muy probable (con un 95% de confianza) que la media de satisfacción del producto en estudio sea significativamente menor.
- **Intervenciones Futuras**: Si la satisfacción mínima aceptable se define en 50, el investigador puede señalar que hay una posibilidad considerable de que la media real esté por debajo de este umbral, sugiriendo una necesidad de mejorar el producto.

Conclusión

En resumen, el intervalo de confianza proporciona una estimación del rango en el que se espera encontrar la verdadera media poblacional con un cierto nivel de confianza. Es una herramienta clave en la estadística inferencial que ayuda a los investigadores a comunicar y tomar decisiones informadas basadas en sus datos muestrales.

Ejercicio 5: Varianza Conocida

Pregunta: Se sabe que la desviación estándar de las alturas de una población es de 5 cm. Se toma una muestra aleatoria de 40 individuos con una media muestral de 170 cm.

Calcula el intervalo de confianza del 95% para la media poblacional.

Solución:

Para calcular el intervalo de confianza del 95% para la media poblacional, utilizamos los datos proporcionados:

Datos Proporcionados

- Desviación estándar poblacional (σ): 5 cm
- Tamaño de la muestra (n): 40
- Media muestral ($\bar{x}$): 170 cm
- Nivel de confianza: 95%

Paso 1: Determinar el valor crítico

Para un nivel de confianza del 95%, el valor crítico z (para una distribución normal estándar) es aproximadamente 1.96.

Paso 2: Calcular el error estándar de la media

El error estándar de la media se calcula como:

$$\text{Error estándar} = \frac{\sigma}{\sqrt{n}} = \frac{5}{\sqrt{40}}$$

Paso 3: Calcular el margen de error

El margen de error se calcula como:

$$\text{Margen de error} = z \times \text{Error estándar} = 1.96 \times \frac{5}{\sqrt{40}}$$

Paso 4: Calcular el intervalo de confianza

El intervalo de confianza se calcula como:

$$\text{Intervalo de confianza} = \bar{x} \pm \text{Margen de error}$$

Ahora, realicemos los cálculos:

Cálculo del error estándar de la media

$$\text{Error estándar} = \frac{5}{\sqrt{40}} \approx 0.7906$$

Cálculo del margen de error

$$\text{Margen de error} = 1.96 \times 0.7906 \approx 1.5496$$

Cálculo del intervalo de confianza

$$\text{Intervalo de confianza} = 170 \pm 1.5496 = (168.4504, 171.5496)$$

Resultado

El intervalo de confianza del 95% para la media poblacional es aproximadamente $(168.45, 171.55)$. Esto significa que con un 95% de confianza, podemos afirmar que la verdadera media poblacional de las alturas se encuentra dentro de este rango.

Ejercicio 6: Varianza Desconocida

Pregunta: En un estudio de satisfacción laboral, se toma una muestra de 25 empleados y se encuentra una media muestral de 78 puntos con una desviación estándar muestral de 12 puntos. Calcula el intervalo de confianza del 95% para la media poblacional.

Solución:

Para calcular el intervalo de confianza del 95% para la media poblacional, usaremos la media y la desviación estándar muestral. Dado que la muestra es pequeña (n < 30) y la desviación estándar poblacional no es conocida, utilizaremos la distribución t de Student.

Datos Proporcionados

- Tamaño de la muestra (n): 25
- Media muestral ($\bar{x}$): 78 puntos
- Desviación estándar muestral (s): 12 puntos
- Nivel de confianza: 95%

Paso 1: Determinar el valor crítico t

Para un nivel de confianza del 95% y $n-1$ grados de libertad, necesitamos el valor crítico t. Con $n-1 = 25-1 = 24$ grados de libertad, el valor crítico t ($t_{0.025}$) se puede encontrar en una tabla de la distribución t o usando software estadístico. Aproximadamente, $t_{0.025,24} \approx 2.064$.

Paso 2: Calcular el error estándar de la media

El error estándar de la media se calcula como:

$$\text{Error estándar} = \frac{s}{\sqrt{n}} = \frac{12}{\sqrt{25}} = 2.4$$

Paso 3: Calcular el margen de error

El margen de error se calcula como:

$$\text{Margen de error} = t \times \text{Error estándar} = 2.064 \times 2.4 \approx 4.9536$$

Paso 4: Calcular el intervalo de confianza

El intervalo de confianza se calcula como:

$$\text{Intervalo de confianza} = \bar{x} \pm \text{Margen de error}$$

Cálculo del intervalo de confianza

$$\text{Intervalo de confianza} = 78 \pm 4.9536 = (73.0464, 82.9536)$$

Resultado

El intervalo de confianza del 95% para la media poblacional es aproximadamente $(73.05, 82.95)$. Esto significa que con un 95% de confianza, podemos afirmar que la verdadera media poblacional de la satisfacción laboral se encuentra dentro de este rango.

Ejercicio 7: Muestra Grande

Pregunta: Una encuesta a 200 estudiantes universitarios encuentra que el gasto promedio mensual en libros es de $100 con una desviación estándar de $20. Calcula el intervalo de confianza del 95% para la media del gasto mensual en libros para todos los estudiantes universitarios.

Solución:

Para calcular el intervalo de confianza del 95% para la media del gasto mensual en libros para todos los estudiantes universitarios, utilizaremos la media y la desviación estándar muestral. Dado que la muestra es grande ($n > 30$), podemos utilizar la distribución normal estándar.

Datos Proporcionados

- Tamaño de la muestra (n): 200
- Media muestral ($\bar{x}$): \$100
- Desviación estándar muestral (σ): \$20
- Nivel de confianza: 95%

Paso 1: Determinar el valor crítico

Para un nivel de confianza del 95%, el valor crítico z (para una distribución normal estándar) es aproximadamente 1.96.

Paso 2: Calcular el error estándar de la media

El error estándar de la media se calcula como:

$$\text{Error estándar} = \frac{\sigma}{\sqrt{n}} = \frac{20}{\sqrt{200}}$$

Paso 3: Calcular el margen de error

El margen de error se calcula como:

$$\text{Margen de error} = z \times \text{Error estándar} = 1.96 \times \frac{20}{\sqrt{200}}$$

Paso 4: Calcular el intervalo de confianza

El intervalo de confianza se calcula como:

$$\text{Intervalo de confianza} = \bar{x} \pm \text{Margen de error}$$

Ahora, realicemos los cálculos:

Cálculo del error estándar de la media

$$\text{Error estándar} = \frac{20}{\sqrt{200}} \approx 1.4142$$

Cálculo del margen de error

$$\text{Margen de error} = 1.96 \times 1.4142 \approx 2.7714$$

Cálculo del intervalo de confianza

$$\text{Intervalo de confianza} = 100 \pm 2.7714 = (97.2286, 102.7714)$$

Resultado

El intervalo de confianza del 95% para la media del gasto mensual en libros para todos los estudiantes universitarios es aproximadamente $(97.23, 102.77)$. Esto significa que con un 95% de confianza, podemos afirmar que la verdadera media del gasto mensual en libros para todos los estudiantes universitarios se encuentra dentro de este rango.

Ejercicio 8: Proporción

Pregunta: En una muestra de 150 votantes, 90 indican que planean votar por el candidato A. Calcula el intervalo de confianza del 95% para la proporción de votantes que planean votar por el candidato A.

Solución:

Para calcular el intervalo de confianza del 95% para la proporción de votantes que planean votar por el candidato A, utilizamos la fórmula para el intervalo de confianza de una proporción.

Datos Proporcionados

- Tamaño de la muestra (n): 150
- Número de éxitos (x): 90
- Nivel de confianza: 95%

Paso 1: Calcular la proporción muestral

La proporción muestral ($\hat{p}$) se calcula como:

$$\hat{p} = \frac{x}{n} = \frac{90}{150} = 0.6$$

Paso 2: Determinar el valor crítico

Para un nivel de confianza del 95%, el valor crítico z (para una distribución normal estándar) es aproximadamente 1.96.

Paso 3: Calcular el error estándar de la proporción

El error estándar de la proporción se calcula como:

$$\text{Error estándar} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0.6 \times (1-0.6)}{150}} = 0.04$$

Paso 4: Calcular el margen de error

El margen de error se calcula como:

$$\text{Margen de error} = z \times \text{Error estándar} = 1.96 \times 0.04 = 0.0784$$

Paso 5: Calcular el intervalo de confianza

El intervalo de confianza se calcula como:

$$\text{Intervalo de confianza} = \hat{p} \pm \text{Margen de error}$$

Cálculo del intervalo de confianza

$$\text{Intervalo de confianza} = 0.6 \pm 0.0784 = (0.5216, 0.6784)$$

Resultado

El intervalo de confianza del 95% para la proporción de votantes que planean votar por el candidato A es aproximadamente $(0.5216, 0.6784)$. Esto significa que con un 95% de confianza, podemos afirmar que la verdadera proporción de votantes que planean votar por el candidato A se encuentra dentro de este rango.

Ejercicio 9: Comparación de Proporciones

Pregunta: Dos empresas están interesadas en la proporción de empleados que están satisfechos con su trabajo. En una muestra de 100 empleados de la empresa A, 60 se declaran satisfechos. En una muestra de 120 empleados de la empresa B, 90 se declaran satisfechos. Calcula el intervalo de confianza del 95% para la diferencia en las proporciones de empleados satisfechos entre las dos empresas.

Solución:

Para calcular el intervalo de confianza del 95% para la diferencia en las proporciones de empleados satisfechos entre las dos empresas, utilizamos la fórmula para el intervalo de confianza para la diferencia entre dos proporciones.

Datos Proporcionados

- Tamaño de la muestra de la empresa A (n_A): 100
- Número de empleados satisfechos en la empresa A (x_A): 60
- Tamaño de la muestra de la empresa B (n_B): 120
- Número de empleados satisfechos en la empresa B (x_B): 90

Paso 1: Calcular las proporciones muestrales

La proporción muestral para la empresa A ($\hat{p}_A$) es:

$$\hat{p}_A = \frac{x_A}{n_A} = \frac{60}{100} = 0.60$$

La proporción muestral para la empresa B ($\hat{p}_B$) es:

$$\hat{p}_B = \frac{x_B}{n_B} = \frac{90}{120} = 0.75$$

Paso 2: Determinar el valor crítico

Para un nivel de confianza del 95%, el valor crítico z (para una distribución normal estándar) es aproximadamente 1.96.

Paso 3: Calcular el error estándar de la diferencia entre proporciones

El error estándar de la diferencia entre proporciones se calcula como:

$$\text{Error estándar} = \sqrt{\left(\frac{\hat{p}_A(1-\hat{p}_A)}{n_A}\right) + \left(\frac{\hat{p}_B(1-\hat{p}_B)}{n_B}\right)}$$

Cálculo del error estándar

$$\text{Error estándar} = \sqrt{\left(\frac{0.60 \times (1-0.60)}{100}\right) + \left(\frac{0.75 \times (1-0.75)}{120}\right)} \approx 0.06296$$

Paso 4: Calcular el margen de error

El margen de error se calcula como:

$$\text{Margen de error} = z \times \text{Error estándar} = 1.96 \times 0.06296 \approx 0.1234$$

Paso 5: Calcular el intervalo de confianza

La diferencia entre las proporciones muestrales ($\hat{p}_A - \hat{p}_B$) es:

$$\hat{p}_A - \hat{p}_B = 0.60 - 0.75 = -0.15$$

El intervalo de confianza se calcula como:

$$\text{Intervalo de confianza} = (\hat{p}_A - \hat{p}_B) \pm \text{Margen de error}$$

Cálculo del intervalo de confianza

$$\text{Intervalo de confianza} = -0.15 \pm 0.1234 = (-0.2734, -0.0266)$$

Resultado

El intervalo de confianza del 95% para la diferencia en las proporciones de empleados satisfechos entre las dos empresas es aproximadamente $(-0.2734, -0.0266)$. Esto significa que con un 95% de confianza, podemos afirmar que la proporción de empleados satisfechos es mayor en la empresa B que en la empresa A, y la diferencia en las proporciones se encuentra en el rango de -0.2734 a -0.0266.

Ejercicio 10: Interpretación Crítica

Pregunta: Un informe estadístico presenta el intervalo de confianza del 95% para la media de una variable como (75, 85). El informe concluye que la media poblacional es 80. ¿Es esta conclusión válida? Explica por qué o por qué no, utilizando conceptos de la interpretación de intervalos de confianza.

Solución:

La conclusión del informe no es necesariamente válida, y aquí está la razón:

El intervalo de confianza del 95% para la media de una variable proporciona un rango dentro del cual es probable que se encuentre la verdadera media poblacional con un nivel de confianza del 95%. En este caso, el intervalo de confianza es (75, 85), lo que significa que hay un 95% de probabilidad de que la verdadera media poblacional se encuentre en este rango.

Dado que el valor de 80 se encuentra dentro del intervalo de confianza, no podemos rechazar la posibilidad de que la media poblacional sea 80. Sin embargo, la conclusión de que la media poblacional es exactamente 80 no es necesariamente válida. El intervalo de confianza proporciona un rango de valores plausibles para la media poblacional, pero no nos dice con certeza cuál es el valor exacto.

Por lo tanto, la conclusión del informe de que la media poblacional es 80 podría ser una interpretación razonable dada la información proporcionada en el intervalo de confianza, pero debemos tener en cuenta que otros valores dentro del intervalo también son posibles.