UNIVERSIDAD NACIONAL AGRARIA LA MOLINA
Dpto. de Estadística e Informática

Capítulo VIII.
Inferencia Estadística

Semana 13. Intervalo de confianza

Plan de aprendizaje

Inicio

- Motivación
- Logros
- Saberes previos

Desarrollo

- Intervalo de confianza para la media
- Intervalo de confianza para la varianza
- Intervalo de confianza para la proporción
- Estimación del tamaño de muestra
- Ejercicios resueltos

Cierre

- Ejercicios propuestos
- Autoevaluación (Moodle)

Motivación:

Descripción

La empresa SVS se dedica a la fabricación y distribución de contenedores que son de uso para diversas actividades. La capacidad de cada uno de estos contenedores tiene una distribución normal con un promedio de 68 m^3 con una desviación estándar de 2.8 m^3. Según la normatividad del ministerio de transporte y comunicaciones, la capacidad de los contenedores pueden tener una tolerancia de $\pm$ 7 m^3 para que la carga se transporte de forma segura. Para verificar que la carga se transporta de manera segura, se tomara una muestra de 50 contenedores.

- Defina la variable en estudio y los parámetros de la distribución.
- Calcule el contenido promedio de la muestra de contenedores.
- Calcular el margen de error para el contenido promedio con un nivel de confianza del 95%.
- Entre que valores se encontrará el contenido promedio de los contenedores con un nivel de confianza del 95%.

Logros:

Al término de la sesión, el estudiante estará en capacidad de:

- Calcular e interpretar la estimación puntual para la media, variancia y la proporción.
- Calcular e interpretar intervalo de confianza para la media, variancia y la proporción.
- Calcular el tamaño de muestra para estimar la media o la proporción.
- Resolver ejercicios propuestos.

Saberes previos:

- ¿Qué es la estimación por intervalo de confianza?
- ¿Para que sirve la estimación por intervalo de confianza?
- ¿Cómo se construye un intervalo de confianza para la media, variancia o proporción?

Contenido. Conceptos básicos

- Estimación puntual de parámetros
- Estimación por intervalo de confianza para la media
- Estimación por intervalo de confianza para la varianza
- Estimación por intervalo de confianza para la proporción
- Estimación del tamaño de muestra
- Ejemplos, ejercicios resueltos y propuestos
- Autoevaluación (Aula virtual)

Estimación puntual de los parámetros

Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de tamaño n extraída de una población con parámetro θ. Se denomina estimador puntual de θ a cualquier estadístico cuyo valor $\hat{\theta} = h(x_1, \ldots, x_n)$ dará una estimación puntual. Los estimadores puntuales para los parámetros:

La media poblacional **(μ)**: $\hat{\mu} = \bar{X} = \dfrac{\sum_{i=1}^{n} X_i}{n}$

La variancia poblacional **(σ²)**: $\hat{\sigma}^2 = S^2 = \dfrac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n-1}$

La proporción poblacional **(π)**: $\hat{\pi} = p = \dfrac{\text{Número de éxitos}}{n}$

Intervalo de confianza para la media

Si $X_1, X_2, \ldots, X_n$ es una muestra aleatoria de una población normal con media **μ** y variancia **σ²** desconocida.

Para hallar un intervalo, con un nivel de confianza del (1-α)x100% para la media μ, se tiene:

$$\left(\underbrace{\overline{X} - t_{\left(1-\alpha/2; n-1\right)} \frac{s}{\sqrt{n}}}_{a} \leq \mu \leq \underbrace{\overline{X} + t_{\left(1-\alpha/2; n-1\right)} \frac{S}{\sqrt{n}}}_{b} \right)$$

Interpretación. Se tiene una confianza del (1- α)x100% que la media poblacional μ, este contenida en el intervalo [a, b].

Límite inferior: $LI(\mu) = \overline{X} - t_{\left(1-\alpha/2; n-1\right)} \frac{s}{\sqrt{n}}$

Límite superior: $LS(\mu) = \overline{X} + t_{\left(1-\alpha/2; n-1\right)} \frac{s}{\sqrt{n}}$

Intervalo de confianza para la proporción

Para hallar un intervalo con un nivel de confianza del (1- α)x100% para la proporción **π**, se tiene:

$$\underbrace{p - Z_{\left(1-\frac{\alpha}{2}\right)}\sqrt{\frac{p(1-p)}{n}}}_{a} \leq \pi \leq \underbrace{p + Z_{\left(1-\frac{\alpha}{2}\right)}\sqrt{\frac{p(1-p)}{n}}}_{b}$$

Interpretación. Se tiene una confianza del (1- α)x100% que la proporción poblacional π, esta contenida en el intervalo [a, b].

Intervalo de confianza para la varianza

Un intervalo de confianza del (1- α)x100% para la variancia **σ**² se halla:

$$\underbrace{\frac{(n-1)s^2}{\chi^2_{\left(1-\frac{\alpha}{2},n-1\right)}}}_{a} \leq \sigma^2 \leq \underbrace{\frac{(n-1)s^2}{\chi^2_{\left(\frac{\alpha}{2},n-1\right)}}}_{b}$$

Interpretación. Se tiene una confianza del (1- α)x100% que la variancia poblacional σ**²**, este contenida en el intervalo [a, b].

Para la desviación estándar:

$$\underbrace{\sqrt{\frac{(n-1)s^2}{\chi^2_{\left(1-\frac{\alpha}{2},n-1\right)}}}}_{a} \leq \sigma \leq \underbrace{\sqrt{\frac{(n-1)s^2}{\chi^2_{\left(\frac{\alpha}{2},n-1\right)}}}}_{b}$$

Ejemplo 1

En una granja de cuyes, se desea evaluar el nuevo alimento para el engorde de cuyes. Con este fin se extrae una muestra de 30 cuyes, registrando la ganancia de peso (grs) y cambio de pelo.

Medida	Ganancia de peso
$\bar{x}$	125.5
S	35.5
Nr. cambiaron el pelo	24

a. Halle un intervalo de confianza de 90% para la verdadera ganancia media de peso de los cuyes.

Tabla t: $\alpha = 0.10 \rightarrow t_{(1-\alpha/2,n-1)} = t_{(0.95,29)=1.699}$

$$\bar{X} - t_{(1-\alpha/2,n-1)} x \frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + t_{(1-\alpha/2,n-1)} x \frac{S}{\sqrt{n}}$$

$$125.5 - 1.699 x \frac{35.5}{\sqrt{30}} \leq \mu \leq 125.5 + 1.699 x \frac{35.5}{\sqrt{30}} \Longrightarrow [114.49, 136.51]$$

Interpretación. Se tiene una confianza de 90% que la ganancia media del peso, está en el intervalo [114.49, 136.51] grs.

b. Halle un intervalo de confianza de 98% para la verdadera variancia de la ganancia de peso de los cuyes

$$\textbf{Tabla Chi – Cuadrado}: \alpha = 0.02 \begin{cases} \chi^2_{\left(1-\alpha/_2, n-1\right)} = \chi^2_{(0.99,29)=49.588} \\ \chi^2_{\left(\alpha/_2, n-1\right)} = \chi^2_{(0.01,29)=14.256} \end{cases}$$

$$\frac{(n-1)S^2}{\chi^2_{\left(1-\alpha/_2, n-1\right)}} \le \sigma^2 \le \frac{(n-1)S^2}{\chi^2_{\left(\alpha/_2, n-1\right)}} \Rightarrow \frac{29x35.5^2}{49.588} \le \sigma^2 \le \frac{29x35.5^2}{14.256} \Rightarrow [737.02, 2563.64]$$

Interpretación: Con un nivel de confianza de 98% la variancia de la ganancia del peso, está en el intervalo [737.02, 2563.64] gr²

Para la desviación estándar:

$$\sqrt{737.02} \le \sigma \le \sqrt{2563.64} \Rightarrow [27.15, 50.63]$$

Interpretación: Con un nivel de confianza de 98% la desviación estándar de la ganancia del peso, está en el intervalo [27.15, 50.63] gr

c. Halle un intervalo de confianza de 95% para la proporción de cuyes que cambiaron de color de pelo..

$$\textbf{Tabla Z}:\ \alpha = 0.05 \rightarrow Z_{\left(1-{}^{\alpha}/_{2}\right)} = Z_{(0.975)=1.96},\qquad p = \frac{24}{30} = 0.8$$

$$p - Z_{\left(1-{}^{\alpha}/_{2}\right)}\mathrm{x}\sqrt{\frac{\mathrm{px}(1-\mathrm{p})}{\mathrm{n}}} \leq \pi \leq p + Z_{\left(1-{}^{\alpha}/_{2}\right)}\mathrm{x}\sqrt{\frac{\mathrm{px}(1-\mathrm{p})}{\mathrm{n}}}$$

$$0.8 - 1.96\mathrm{x}\sqrt{\frac{0.8\mathrm{x}0.2}{30}} \leq \pi \leq 0.8 + 1.96\mathrm{x}\sqrt{\frac{0.8\mathrm{x}0.2}{30}} \Longrightarrow [0.657, 0.943]$$

Interpretación: Se estima con una confianza de 95% que la proporción de cuyes que cambiaron de pelo, se encuentra en el intervalo [0.657, 0.934]

Ejercicio

Un inspector de calidad está evaluando si el contenido de fruta por lata es el adecuado. Se supone que el contenido de fruta se distribuye como una normal. El inspector escoge al azar 35 latas y encuentra que el peso promedio es de 29.2 onzas, la variancia es de 4 onzas^2 y que 5 no tienen un peso adecuado.

a. Halle e interprete un intervalo de confianza del 95% para el peso promedio poblacional.
b. Halle e interprete un intervalo de confianza del 95% para la variancia y desviación estándar poblacional del contenido de fruta en las latas.
c. Halle e interprete un intervalo de confianza del 99% para la proporción poblacional de latas que presentan un peso adecuado

Solución:

a. Halle e interprete un intervalo de confianza del 95% para el peso promedio poblacional.

μ = Peso promedio del contenido de latas de fruta

$$\text{Se tiene: n} = 35 \text{ latas}, \bar{\text{X}} = 29.2 \text{ onzas}, \text{S}^2 = 4 \text{ onzas}^2$$

$$\textbf{Nivel: 95\%} \Longrightarrow \alpha = 0.05 \rightarrow \text{Tablat: } t_{\left(1-\alpha/_{2},n-1\right)} = t_{(0.975,34)} = 2.032$$

$$\bar{X} - t_{\left(1-\alpha/_{2},n-1\right)} x \frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + t_{\left(1-\alpha/_{2},n-1\right)} x \frac{S}{\sqrt{n}}$$

$$29.2 - 2.032x\frac{2}{\sqrt{35}} \leq \mu \leq 29.2 + 2.032x\frac{2}{\sqrt{35}} \Longrightarrow [28.51, 29.89]$$

Interpretación. Con una confianza del 95%, se estima que el peso promedio del contenido de fruta se encuentra en el intervalo [28.51, 29.89] onzas

b. Halle e interprete un intervalo de confianza del 95% para la variancia y desviación estándar poblacional del contenido de fruta en las latas.

σ^2 = Variancia del peso del contenido de latas de fruta

$$\text{Nivel: } 95\% \Rightarrow \alpha = 0.05 \Rightarrow \text{Tabla Chi} - \text{Cuadrado:} \begin{cases} \chi^2_{\left(1-\alpha/2,n-1\right)} = \chi^2_{(0.975,34)} = 51.966 \\ \chi^2_{\left(\alpha/2,n-1\right)} = \chi^2_{(0.025,34)} = 19.806 \end{cases}$$

$$\frac{(n-1)S^2}{\chi^2_{\left(1-\alpha/2,n-1\right)}} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi^2_{\left(\alpha/2,n-1\right)}} \Rightarrow \frac{34x4}{51.966} \leq \sigma^2 \leq \frac{34x4}{19.806} \Rightarrow [2.62, 6.87]$$

Interpretación. Con una confianza del 95%, se estima que la variancia del peso del contenido de fruta se encuentra en el intervalo [2.62, 6.87] onzas²

Desviación estándar: $\sqrt{2.62} \leq \sigma \leq \sqrt{6.87} \Rightarrow [1.62, 2.62]$

Interpretación. Con una confianza del 95%, se estima que la desviación estándar del peso se encuentra en el intervalo [1.62, 2.62] onzas

c. Halle e interprete un intervalo de confianza del 99% para la proporción poblacional de latas que presentan un peso adecuado.

π = Proporción de latas que tienen el peso adecuado

$$\textbf{Nivel: } \mathbf{99\% \Rightarrow \alpha = 0.01 \Rightarrow} \textbf{ Tabla Z: } Z_{(1-\alpha/_2)} = Z_{(0.995)} = 2.58$$

$$\text{Estimar p: 30 tienen peso adecuado } \ p = \frac{30}{35} = 0.86$$

$$p - Z_{(1-\alpha/_2)} x \sqrt{\frac{px(1-p)}{n}} \le \pi \le p + Z_{(1-\alpha/_2)} x \sqrt{\frac{px(1-p)}{n}}$$

$$0.86 - 2.586x\sqrt{\frac{0.86x0.14}{35}} \le \pi \le 0.86 - 2.586x\sqrt{\frac{0.86x0.14}{35}} \Rightarrow [0.709, 1.011]$$

Interpretación: Se estima con una confianza de 99% que la proporción de latas que presentan el peso adecuado, se encuentra en el intervalo [0.709, 1.011]

Determinación del tamaño de muestra

El tamaño de muestra cuando se estima a la media poblacional μ o cuando se estima a la proporción poblacional π.

Tamaño de muestra para estimar a la media μ

Se sabe que: $t = \dfrac{\overline{x} - \mu}{s / \sqrt{n}}$

Despejando el tamaño de muestra: $\boxed{n = \dfrac{t^2 s^2}{e^2}}$

Donde:

e =Es el margen de error que se quiere aceptar.

t = El valor tabular del nivel de confianza requerido.

Los G.L. se busca en la tabla con el máximo valor ($\rightarrow \propto$: última fila).

S^2=La variancia se obtiene de una muestra piloto.

Tamaño de muestra para estimar a la proporción π

Se sabe que: $Z = \dfrac{p-\pi}{\sqrt{\dfrac{\pi(1-\pi)}{n}}}$

Despejando el tamaño de muestra:

$$\boxed{n = \frac{Z^2\pi\left(1-\pi\right)}{e^2}}$$

Donde:

e =Es el margen de error que se quiere aceptar.

Z =El valor tabular del nivel de confianza requerido.

Π =Se estima con una muestra piloto. Se puede considerar el valor conservador π=0.5.

Ejemplo

Una cadena de hoteles desea estimar el número promedio de habitaciones ocupadas cada noche en sus sucursales con una confianza del 99%. ¿Cuántas noches deben incluirse en la muestra si se puede tolerar un error de 10 habitaciones y en una muestra piloto se halló una desviación estándar de 60 habitaciones?

Se tiene:

Margen de error: e=10 habitaciones

Nivel de confianza 99%: $\Rightarrow \alpha = 0.01 \Rightarrow \textbf{Tabla t}: t_{\left(1-\alpha/2, \infty\right)} = t_{(0.995,\infty)} = 2.577$

Desviación estándar: S=60 habitaciones

$$n = \frac{t^2 S^2}{e^2} = \frac{2.577^2 60^2}{10^2} = 239.07 \approx 239 \text{ noches}$$

Se debe considerar 239 noches como tamaño de muestra.

Ejemplo

Halle el tamaño de muestra, con el fin de estimar la proporción de estudiantes que tienen correo electrónico. En un estudio previo, se halló que de 150 estudiantes, 120 tenían correo electrónico. Se desea tener un nivel de confianza del 95% y un error del 3% en las estimaciones.

Se tiene:

Margen de error: e=0.03

Nivel de confianza 95%: $\Rightarrow \alpha = 0.05 \Rightarrow$ **Tabla Z**: $Z_{(1-\alpha/2)} = Z_{(0.975)} = 1.96$

Estimación de π : p=120/150=0.8

$$n = \frac{Z^2 p(1-p)}{e^2} = \frac{1.96^2 0.8x0.2}{0.03^2} = 682.95 \approx 683 \text{ estudiantes}$$

Se debe considerar un tamaño de muestra 683 estudiantes

Ejercicio

¿Cuál debe ser el tamaño de muestra para estimar el nivel medio del consumo de proteínas de los adultos de una zona minera, si se desea tener un margen de error de 0.45 g/dl y un nivel de confianza del 95%?. De una muestra piloto se obtuvo que el consumo de proteínas tiene una desviación estándar de 2.5 g/dl.

Solución:

Margen de error: e= 0.45 g/dl

Nivel de confianza 95%: $\alpha = 0.05 \Rightarrow$ **Tabla t**: $t_{(1-\alpha/2,\infty)} = t_{(0.975,\infty)} = 1.96$

Desviación estándar: S=2.5 g/dl

$$n = \frac{t^2 S^2}{e^2} = \frac{1.96^2 2.5^2}{0.45^2} = 118.57 \approx 119 \text{ adultos}$$

Ejercicio

El director comercial de cierta compañía que realiza ventas por correo electrónico, desea precisar con mucho cuidado su política de crédito. Si el director desea tener un intervalo de confianza del 99% para la proporción de clientes que están al día en sus pagos. ¿Qué tamaño de muestra debe usar si se desea tener un margen de error del 4.5%?. Suponga que en una muestra piloto de 35 clientes se halló que 8 clientes están al día en sus pagos.

Solución:

Margen de error: e=0.045

Nivel de confianza 99%: $\Rightarrow \alpha = 0.01 \Rightarrow$ **Tabla Z**: $Z_{(1-\alpha/2)} = Z_{(0.995)} = 2.58$

Estimación de π : p=8/35=0.23

$$n = \frac{Z^2 p(1-p)}{e^2} = \frac{2.58^2 0.23x0.77}{0.045^2} = 582.15 \approx 582 \text{ clientes}$$

Ejercicios propuestos.

Un fabricante de fibras sintéticas diseña un experimento para estimar la tensión de ruptura media de una fibra, para lo cual observa las tensiones de ruptura (en libras) de una muestra aleatoria de 12 hilos. Los datos obtenidos fueron:

19.3	20.2	21.4	18.3	18.6	19.4	22.5	20.8	19.6	21.3	18.5	22.4

a) Con una confianza del 98%, halle e interprete el intervalo de confianza para la tensión de ruptura media de las fibras sintéticas.

b) Halle e interprete un intervalo del 90% de confianza para la variancia de la tensión de ruptura de las fibras sintéticas.

c) En una muestra de 80 hilos, se halló que el 15% tienen una tensión de ruptura es menor a 20.5 libras. Hallar un intervalo de confianza del 95% para la proporción de hilos cuya tensión de ruptura es menor a 20.5 libras.

Referencias bibliográficas

- Anderson D., Sweendy D., Williams T. (2016) Estadística para Administración y Economía. 12ª. Edición. México. Cengage Learning Editores. Capítulo 8 y 9.
- Newbold, P. y Carlson, W. y Thorne, B. (2008). Estadística para Administración y Economía (6ta. ed.) Madrid: Pearson Education. Prentice Hall