

Ejercicios sobre estimaciones e intervalos.

FUNDAMENTOS:

Los intervalos de confianza para una media tienen expresión formal

$$\bar{x} \pm z_{1-\alpha/2} S / \sqrt{n}$$

cuando se tiene una muestra.

Sino, se usa la expresión análoga

$$\bar{x} \pm z_{1-\alpha/2} \sigma_{\bar{\mu}} / \sqrt{n}$$

con

$$\sigma_{\bar{\mu}}$$

la desviación del estimador (o una aproximación a esta). Si lo que se tiene es la varianza

$$\sigma_{\bar{\mu}}^2$$

del estimador recordar extraerle raíz cuadrada para obtener su desviación. Y, a veces se tiene una estimación de la varianza

$$\sigma^2$$

(o desviación

$$\sigma$$

) poblacional y puede ser usada para obtener lo que va en lugar de S o

$$\sigma_{\bar{\mu}}$$

en los intervalos.

En general el procedimiento matemático para intervalos de confianza es

$$\bar{\theta} \pm c_{1-\frac{\alpha}{2}} \cdot \frac{\sigma_{\bar{\theta}}}{\sqrt{n}}$$

donde

$$c_{1-\frac{\alpha}{2}}$$

es el cuantil que consigue la probabilidad

$$P(L \leq \theta \leq U) = 1 - \alpha$$

para el parámetro de interés  $\theta$ , se usa un cuantil normal

$$z_{1-\frac{\alpha}{2}}$$

cuando n es mayor a 31 y uno T-student

$$t_{1-\frac{\alpha}{2}}$$

con  $n-1$  grados de libertad en los otros casos.

En el caso de ser una estimación de diferencia de parámetros el formato matemático de intervalo es

$$\bar{D} \pm c_{1-\frac{\alpha}{2}} \cdot (S_{\bullet}) \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

, donde

$$\bar{D} = \bar{\theta}_1 - \bar{\theta}_2$$

es el estimador de la diferencia de parámetros de interés,

$$S_{\bullet}^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

es el estimador ponderado de

$$\sigma^2$$

, donde esta varianza poblacional se estima aproximadamente igual para ambas

muestras independientes

de tamaños

$$n_1, n_2$$

respectivamente, y el cuantil sigue una distribución T con

$$n_1 + n_2 - 2$$

grados de libertad si ambas muestras son inferiores a 32 y normal en otro caso.

Ejercicios

Ejercicio: (A) Se desea estimar los milímetros promedio del lluvia en una región para determinar si es factible para cierto cultivo. Se midieron las cantidades en

5

precipitaciones y se registraron los milímetros. La media muestral junto con su desviación estándar muestral debe ser usada para estimar la media real con un intervalo de confianza del

96%

. Muestra: [181.2, 220.3, 215, 185.3, 200] (B) A un nivel de confianza del 99%, esta finca desea estimar el rendimiento promedio de la cosecha de maíz. Sobre una muestra de 600 ha se obtuvo una media 98.7 y una desviación del estimador de 1.25, construya un intervalo de confianza con estos datos.

Ejercicio: Una empresa desea evaluar el rendimiento de un algoritmo de clasificación en la detección de spam en correos electrónicos. Se seleccionan

500

correos electrónicos, de los cuales 100 son spam y 400 no son spam. El algoritmo clasifica correctamente 93 de los correos spam y 383 de los correos no spam. Estime (son dos intervalos de confianza) las proporciones reales con la información de esta muestra con un intervalo de confianza del

91%

para los "spam" y el otro al

94%

Ejercicio En general, nuevas costureras de una máquina, requieren de capacitación para aprender a elaborar un producto. Se diseñó un nuevo proceso de capacitación, el que se desea comparar con el usual. Se estudiaron dos grupos de cinco obreras, el primer grupo siguió el método usual y el otro el nuevo procedimiento. Los resultados de la medición de los horas de capacitaciones invertidos

para lograr la elaboración del producto fueron	usual	33	37	45	31	43
	nuevo	35	31	24	34	32

Determine un intervalo de confianza para la diferencia de tiempos medios de estos métodos trabajando con una confianza del 93

Ejercicio: (A) Un proceso químico ha producido, en promedio, 600 toneladas de productos químicos al día en base a una muestra de 200 días. Estime la producción media diaria, con un intervalo al

96%

de confianza si se conoce una estimación de 43.5 para

$\sigma$

poblacional. (B) Esta misma industria ha registrado los días en que ha sobrepasado lo permitido en desechos químicos, resultando en una muestra [0,1,1,1,0,1,0]. Calcule un intervalo de confianza para esta proporción al

94%

de confianza.

Ejercicio: (A) Una nueva planta de riego fue instalada en seis fincas y se midieron las producciones en miles de quintales de arroz :

31.05, 29.95, 29.15, 29.85, 30.85, 28.37.

Estime la verdadera producción promedio con un intervalo del 99(B) Esta misma industria ha registrado los días en que se han sobrecalentado ciertos componentes del sistema de riego, resultando en una muestra [0,1,1,1,0,0,1]. Calcule un intervalo de confianza para esta proporción al

94%

de confianza.

Ejercicio En general, nuevos obreros, requieren de capacitación para aprender a elaborar un producto. Se diseñó un nuevo proceso de capacitación, el que se desea comparar con el usual. Se estudiaron dos grupos de cinco obreros, el primer grupo siguió el método usual y el otro el nuevo procedimiento. Los resultados de la medición de porcentajes obtenidos al evaluar la elaboración del

producto fueron 

usual	93	97	65	81	83
nuevo	75	71	84	94	92

 Determine un intervalo de confianza para la diferencia de porcentajes de estos métodos trabajando con una confianza del 93

Ejercicio: Se desea evaluar el rendimiento de un algoritmo de clasificación en la detección de transacciones fraudulentas en una plataforma de pagos en línea. Se seleccionan

1000

transacciones, de las cuales 50 son fraudulentas y 950 no son fraudulentas. El algoritmo clasifica correctamente 45 de las transacciones fraudulentas y 900 de las transacciones no fraudulentas. Estime las proporciones reales (son dos intervalos) con la información de esta muestra con un intervalo de confianza del

95%

para las fraudulentas y del

91%

para las otras.

Ejercicio: Se desea estimar la diferencia de medias del tiempo (minutos) que los clientes esperan ser atendidos en dos sucursales de una empresa. Se seleccionan 50 clientes de una y 60 de la otra, y se registra el tiempo que cada uno pasó en la sala de espera. Los registros arrojaron medias de 19.5 y 18.8 minutos, junto con varianzas del estimador (muestrales) de 3.4 y 2.2 minutos cuadrados respectivamente. Basado en estos datos estime la diferencia con un intervalo de confianza del

99%

Ejercicio: (A) Se desea estimar la media del tiempo que los usuarios pasan en una página web antes de abandonarla. Se seleccionan

4

usuarios y se registra el tiempo que cada uno pasó en la página web. La media muestral junto con su desviación estándar muestral debe ser usada para Estimar la media real con un intervalo de confianza del

97%

4

. Muestra: [12, 8, 9, 11] (B) La misma empresa de software desea estimar el número medio de clics sobre determinado anuncio en una aplicación para dispositivos móviles. Se toma una muestra de 300 dispositivos y se obtiene una media de 4 clics al día y una varianza del estimador de 0.8.

Ejercicio: Se desea estimar la diferencia de proporciones de clientes de determinada marca que lo recomiendan después de haber comprado. Una sucursal está en un centro comercial céntrico y la otra más alejada del centro de una ciudad. Se seleccionan

150

clientes de la primera y 200 de la segunda, de los cuales 121 fueron recomendadores de la primera y 180 de la segunda. Estime la diferencia real con la información de estas muestras con un intervalo de confianza del

95%