Intervalos de confianza

Clase del 12/11/2024

Contenidos

Introducción

$$\bar{X} \sim N(\mu, \frac{\sigma}{\sqrt{n}}) \to \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$

$$\frac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}} \sim t_{n-1}$$

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$

Población Normal con varianza conocida

Dada una muestra $X_1, X_2, \ldots, X_n$ de una variable aleatoria normal con media μ y varianza conocida σ^2, el intervalo de confianza del $100(1-\alpha)\,\%$ para la media μ es:

$$\left(\bar{X} - z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}},\ \bar{X} + Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}\right)$$

donde:

- $\bar{X}$ es la media muestral,
- σ es la desviación estándar poblacional (conocida),
- n es el tamaño de la muestra,
- $Z_{\frac{\alpha}{2}}$ es el valor crítico de la distribución normal estándar correspondiente a un nivel de confianza $1-\alpha$.

- Este intervalo de confianza es válido cuando los datos provienen de una distribución normal y la varianza σ^2 es conocida. También deben ser independientes.
- El valor crítico $Z_{\frac{\alpha}{2}}$ se obtiene de la tabla de la distribución normal estándar para un nivel de confianza dado.
- Por regla general, interesa que la confianza sea *grande*: 95 %, 99 %, 99.9 %

IC con varianza desconocida

Dada una muestra $X_1, X_2, \ldots, X_n$ de una variable aleatoria normal con media μ y varianza desconocida, el intervalo de confianza del $100(1-\alpha)\,\%$ para la media μ es:

$$\left(\bar{X} - t_{\frac{\alpha}{2},\,n-1} \cdot \frac{S}{\sqrt{n}},\ \bar{X} + t_{\frac{\alpha}{2},\,n-1} \cdot \frac{S}{\sqrt{n}}\right)$$

donde:

- $\bar{X}$ es la media muestral, n es el tamaño de la muestra,
- S es la desviación estándar muestral, calculada como $S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}$,
- $t_{\frac{\alpha}{2},\,n-1}$ es el valor crítico de la distribución t de Student con $n-1$ grados de libertad correspondiente a un nivel de confianza $1-\alpha$.

- Este intervalo de confianza es válido cuando los datos provienen de una distribución normal y la varianza poblacional es desconocida.
- Los valores de la muestra deben ser independientes.

IC para la varianza poblacional σ^2

En esta ocasión nos basamos en la distribución chi-cuadrado (χ^2). Dada una muestra $X_1, X_2, \ldots, X_n$ de una variable aleatoria normal con media desconocida y varianza σ^2, el intervalo de confianza del $100(1-\alpha)\,\%$ para σ^2 es:

$$\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2,\, n-1}},\ \frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2},\, n-1}} \right)$$

donde:

- S^2 es la varianza muestral $= \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$,
- n es el tamaño de la muestra,
- $\chi^2_{\frac{\alpha}{2},\, n-1}$ y $\chi^2_{1-\frac{\alpha}{2},\, n-1}$ son los valores críticos de la distribución chi-cuadrado con $n-1$ grados de libertad, correspondientes a los percentiles $\frac{\alpha}{2}$ y $1-\frac{\alpha}{2}$, respectivamente.

- Este intervalo de confianza es válido cuando los datos provienen de una distribución normal.
- Los valores críticos $\chi^2_{\frac{\alpha}{2},n-1}$ y $\chi^2_{1-\frac{\alpha}{2},n-1}$ se obtienen de la tabla de la distribución chi-cuadrado.

IC para una proporción poblacional

Dada una muestra de tamaño n, en la cual se observan X éxitos (eventos de interés), se puede estimar la proporción poblacional p mediante la proporción muestral $\hat{p} = \frac{X}{n}$.
El intervalo de confianza del $100(1-\alpha)\,\%$ para la proporción poblacional p es:

$$\left(\hat{p} - z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}},\ \hat{p} + z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$$

donde:

- $\hat{p} = \frac{X}{n}$ es la proporción muestral de éxitos,
- n es el tamaño de la muestra,
- $z_{\alpha/2}$ es el valor crítico de la distribución normal estándar correspondiente a un nivel de confianza $1-\alpha$.

- Este intervalo de confianza es válido para muestras grandes (generalmente cuando $n\hat{p} \geq 5$ y $n(1-\hat{p}) \geq 5$, lo cual garantiza que la distribución binomial de X se aproxima a una distribución normal).
- El valor crítico $z_{\alpha/2}$ se obtiene de la tabla de la distribución normal estándar.

Un valor más *fino*

$$\frac{1}{1+\frac{Z_{\alpha/2}^2}{n}}\left(\hat{p}+\frac{Z_{\alpha/2}^2}{2n}\pm z_{\alpha/2}\cdot\sqrt{\frac{\hat{p}(1-\hat{p})}{n}+\frac{Z_{\alpha/2}^2}{4n^2}}\right)$$

IC en casos de 2 poblaciones normales

Cuando se dispone de dos m.a.s. de dos v.a. normales X e Y, se pueden plantear intervalos de confianza sobre la diferencia de medias o el cociente de varianzas. Dependiendo de la información de la que se dispone los *estadísticos* a utilizar tienen distribuciones de un tipo u otro y son los que condicionan los valores del intervalo. Se estudian las siguientes circunstancias:

- Estudio de la diferencia de ambas medias.
  - Caso en el que se supone que las v.a. normales tienen las varianzas son iguales (homecedasticidad).
  - Caso en el que se supone que son normales pero que no puede asumirse que las varianzas son iguales.
  - Caso general para distribuciones con muestras grandes donde la normalidad es tiene carácter asintótico.
- Intervalo para la razón de varianzas en poblaciones normales.
- Determinación del tamaño muestral.

Diferencia de medias de dos normales de varianza común

Se consideran dos v.a. normales $X \sim N(\mu_1, \sigma)$, $Y \sim N(\mu_2, \sigma)$. El intervalo de confianza al $100(1-\alpha)\,\%$ es:

$$I_{\mu_1-\mu_2}^{1-\alpha} = \left((\bar{X} - \bar{Y}) \pm t_{\frac{\alpha}{2}, n_1+n_2-2} \cdot S_T \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right)$$

- $\bar{X}, \bar{Y}$ son las medias muestrales respectivas.
- $\hat{S_T}^2 = \frac{(n_1-1)\cdot \hat{s}_1^2 + (n_2-1)\cdot \hat{s}_2^2}{n_1+n_2-2}$
- n_1, n_2 son los tamaños muestrales respectivos,
- $t_{\frac{\alpha}{2}}$ es el valor crítico de la distribución t de Student correspondiente a cola a derecha de $\frac{\alpha}{2}$.

Diferencia de medias de dos normales de varianza distinta

Se consideran dos v.a. normales $X \sim N(\mu_1, \sigma_1)$, $Y \sim N(\mu_2, \sigma_2)$ $(\sigma_1 \neq \sigma_2)$. El intervalo de confianza al $100(1-\alpha)\,\%$ es:

$$I^{1-\alpha}_{\mu_1-\mu_2} = \left((\bar{X} - \bar{Y}) \pm t_{\frac{\alpha}{2},g} \cdot \sqrt{\frac{\hat{s}_1^2}{n_1} + \frac{\hat{s}_2^2}{n_2}} \right)$$

- $g = n_1 + n_2 - 2 - \Delta$, gr. de libertad ,
- $\Delta =$ es el entero más próximo a $\frac{((n_2-1)S_1-(n_1-1)S_2)^2}{(n_2-1)S_1^2+(n_1-1)S_2^2}$
- $S_i = \frac{\hat{s}_i{}^2}{n_i}, i = 1, 2$

IC de de la diferencia de medias de dos poblaciones normales, caso general

Para casos en los que n es *grande* sabemos que asintóticamente las v.a. pueden ser consideradas normales. El intervalo de confianza del $100(1-\alpha)\,\%$ para la diferencia de medias es:

$$I_{\mu_1-\mu_2}^{1-\alpha} = \left((\bar{X} - \bar{Y}) \pm z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{s}_1^2}{n_1} + \frac{\hat{s}_2^2}{n_2}} \right)$$

- $\hat{s}_i^2$ Son las quasi-varianzas muestrales respectivas,
- $Z_{\frac{\alpha}{2}}$ es el valor crítico de la distribución normal estándar correspondiente a cola a derecha de $\frac{\alpha}{2}$.

IC de de la razón de varianzas en poblaciones normales

El intervalo de confianza del $100(1-\alpha)\,\%$ para el cociente de varianzas $\frac{\sigma_1^2}{\sigma_2^2}$ es:

$$I^{1-\alpha}_{\frac{\sigma_1^2}{\sigma_2^2}} = \left(F_a \cdot \frac{\hat{s}_1^2}{\hat{s}_2^2}, F_b \cdot \frac{\hat{s}_1^2}{\hat{s}_2^2} \right)$$

- $\hat{s}_i^2$ Son las quasi-varianzas muestrales respectivas,
- F_a y F_b cumplen que $P(F_a < F < F_b) = 1-\alpha$, donde F es una distribución *F de Snedecor* con (n_2-1) y $n_1-1)$ grados de libertad. O sea:
- F_b es el valor que deja $1-\frac{\alpha}{2}$ a derecha de la F y F_a el que deja a izquierda el mismo valor y de la misma F.

Determinación del tamaño muestral

Si se plantea determinar el valor del tamaño muestral n para que el intervalo tenga una amplitud determinada L en los casos donde el intervalo se plantea como ($centro \pm amplitud$), se nos plantearía una ecuación $L = amplitud$. Como n interviene en la expresión de la amplitud, de ahí podríamos obtener un valor estimado o una cota de n.

1. Media de población normal y varianza conocida: $amplitud = z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$
2. Proporción de una población: $amplitud = z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

1. $L = z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \rightarrow n = z^2_{\frac{\alpha}{2}} \cdot \frac{\sigma^2}{L^2}$
2. $L = z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \rightarrow n = z^2_{\frac{\alpha}{2}} \cdot \frac{\hat{p}(1-\hat{p})}{L^2}$

Estimadores de máxima verosimilitud

Clase del 12/11/2024

Contenidos

Introducción

Por regla general nos enfrentaremos a problemas donde se quiere averiguar la característica de una determinada población contando con información acerca de una muestra de esa población. Este proceso es lo que se conoce como inferencia.

Inferir

La R.A.E. lo define como: Deducir algo o sacarlo como conclusión de otra cosa.

Definición

Definición

Los estimadores de máxima verosimilitud (MLE, *maximum likelihood estimator*) son métodos para estimar los parámetros de una distribución probabilística que maximizan la probabilidad de observar los datos dados.

Nos enfrentaremos a situaciones en las que

- Se tiene una *m.a.s.* $\{x_1, x_2, ..., x_n\}$ de una v.a.X. (Observaciones)
- Dicha variable aleatoria X tendrá una función de probabilidad o de densidad de probabilidad que dependerá de determinado/s parámetro/s.
- Se planteará el objetivo de obtener una estimación (o inferencia) para dicho/s parámetro/s.

Estimador insesgado

Así se define a aquel estimador cuyo valor esperado es el del propio parámetro.

$$\mathbb{E}[\hat{\theta}] = \theta$$

Distribución Binomial

- **Función de masa de probabilidad (fmp):**

$$f(k; n, p) = \binom{n}{k} p^k (1-p)^{n-k}$$

- **Parámetro a estimar:** Probabilidad de éxito p
- **MLE de p:**

$$\hat{p} = \frac{k}{n}$$

donde k es el número de éxitos observados y n es el número total de ensayos.

Distribución Binomial

- También podríamos tener una muestra de K variables aleatorias binomiales $X_i \sim B(n_i, p)$ $i = 1, 2, ..., K$. La muestra se representaría como: $\{x_1, x_2, ...x_k\}$
- **MLE de** p**:**

$$\hat{p} = \frac{\sum_{i=1}^{K} x_i}{\sum_{i=1}^{K} n_i}$$

Distribución de Poisson

- **Función de masa de probabilidad (fmp):**

$$f(k;\lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$$

- **Parámetro a estimar:** Tasa de eventos λ
- **MLE de λ:**

$$\hat{\lambda} = \frac{\sum_{i=1}^{n} k_i}{n}$$

donde k_i es el número de eventos en la i-ésima observación y n es el número total de observaciones.

Distribución Exponencial

- **Función de densidad de probabilidad (fdp):**

$$f(x;\lambda) = \lambda e^{-\lambda x} \quad \text{para } x \geq 0$$

- **Parámetro a estimar:** Tasa λ
- **MLE de λ:**

$$\hat{\lambda} = \frac{1}{\bar{x}}$$

donde $\bar{x}$ es el promedio de los datos observados.

Distribución Normal

- **Función de densidad de probabilidad (fdp):**

$$f(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

- **Parámetros a estimar:** Media μ y varianza σ^2
- **MLE de μ:**

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

- **MLE de σ^2:**

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{\mu})^2$$

Distribución Uniforme $U(0, b)$

- **Función de densidad de probabilidad (fdp):**

$$f(x; b) = \frac{1}{b} \quad \text{para } 0 \leq x \leq b$$

- **Parámetro a estimar:** Límite superior b
- **MLE de b:**

$$\hat{b} = \text{máx}(x_1, x_2, \ldots, x_n)$$

- Se puede demostrar que no es un estimador insesgado pues

$$\mathbb{E}[\hat{b}] = \frac{n}{n+1} \cdot b$$

Estimador insesgado de mínima varianza

Existencia de varios estimadores insesgados

En ocasiones podría resultar que tenemos dos estimadores insesgados. En esa ocasión podría ser interesante considerar aquel que tenga mínima varianza. Lo denominaríamos **Estimador insesgado de mínima varianza**

Distribución Uniforme $U(0, b)$

- El estimador $\mathbb{E}[\hat{b}] = \frac{n}{n+1} \cdot b$ sí es *insesgado*
- $\hat{b}_2 = 2\overline{X}$ también es un estimador insesgado. Pero este es el que tiene la varianza menor.

Estimador Insesgado de la Media

Para una distribución cualquiera, los estimadores insesgados de la media y la varianza se obtienen basándose en una muestra de datos. Dada una muestra $X_1, X_2, \ldots, X_n$ de una variable aleatoria con media poblacional μ, el estimador insesgado de la media es simplemente el promedio muestral:

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

Este estimador es insesgado porque su esperanza matemática es igual a la media poblacional, es decir, $\mathbb{E}[\hat{\mu}] = \mu$.

Estimador Insesgado de la Varianza

Dada una muestra $X_1, X_2, \ldots, X_n$ de una variable aleatoria con varianza poblacional σ^2, el estimador insesgado de la varianza es:

$$\hat{\sigma}^2 = S_{n-1} = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \hat{\mu})^2$$

donde $\hat{\mu}$ es el promedio muestral. Este estimador es insesgado porque su esperanza matemática es igual a la varianza poblacional, es decir, $\mathbb{E}[\hat{\sigma}^2] = \sigma^2$.
Este ajuste de $\frac{1}{n-1}$ en lugar de $\frac{1}{n}$ es necesario para compensar el sesgo que se introduce al utilizar la media muestral $\hat{\mu}$ en lugar de la media poblacional μ en el cálculo de la varianza.

Contrastes de hipótesis

Clase del 19/11/2024

Contenidos

Ejemplo

En una población normal con varianza conocida $\sigma^2 = 4$, se extrajo una muestra de tamaño $n = 16$, obteniéndose una media muestral de $\bar{x} = 6$. Queremos contrastar la hipótesis de que la media poblacional es $\mu = 5$ frente a la hipótesis alternativa de que $\mu > 5$, utilizando un nivel de significación de $\alpha = 0.05$.

Formulación de las Hipótesis

$$H_0 : \mu = 5 \quad \text{(la media poblacional es 5).}$$

$$H_a : \mu > 5 \quad \text{(la media poblacional es mayor que 5).}$$

Estadístico de Prueba

Dado que la población sigue una distribución normal y se conoce la varianza, utilizamos la z-prueba:

$$d = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \sim N(0, 1)$$

Sustituyendo los valores:

$$\bar{x} = 6, \quad \mu_0 = 5, \quad \sigma = \sqrt{4} = 2, \quad n = 16,$$

obtenemos:

$$d = \frac{6 - 5}{2/\sqrt{16}} = \frac{6 - 5}{2/4} = \frac{1}{0.5} = 2.$$

Nivel de Significación y Región Crítica

Para un nivel de significación de $\alpha = 0.05$
(Probabilidad de rechazar H_0 cuando es verdadera)
el valor crítico $z_{\text{crítico}}$ se obtiene de la tabla de la distribución normal estándar: $z_{\text{crítico}} = 1.645$
La regla de decisión es la siguiente:

- Si $d \geq z_{\text{crítico}}$, rechazamos H_0.
- Si $d < z_{\text{crítico}}$, no rechazamos H_0.

Comparación del Estadístico Calculado con el Valor Crítico

El estadístico calculado es: $d = 2$. Por tanto como $d = 2 > z_{\text{crítico}} = 1.645$ entonces **rechazamos** H_0. Equivalentemente, rechazamos H_0 si:

$$\bar{x} \geq \frac{\sigma}{\sqrt{n}} \cdot z_{crtico} + \mu_0$$

Alternativamente podemos razonar calculando el **nivel crítico**:

$$P(d > 2|N(0,1) = 0.023$$

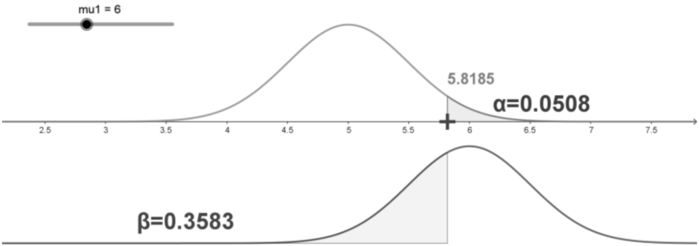
Como el nivel de significación (0.05) es mayor que el nivel crítico (0.023) , entonces rechazamos la hipótesis.

error de tipo II

El error de tipo II β viene dado por la probabilidad de no rechazar H_0 cuando es falsa. Supongamos, por ejemplo, que $\mu = 6$.
Tendríamos que $\bar{X} \sim N(6, 0.5)$.
$P(d \leq 1,65|\mu = 6) = P(z \leq -0.35) = 0.363$

Resultado del contraste

Con un nivel de significación de $\alpha = 0.05$, los datos proporcionan suficiente evidencia para concluir que la media poblacional es mayor que 5.


mu1 = 6
5.8185
α=0.0508
2.5
3
3.5
4
4.5
5
5.5
6
6.5
7
7.5
β=0.3583

Introducción

El contraste de hipótesis es una herramienta esencial en la inferencia estadística para tomar decisiones sobre una población basándose en información muestral. Podríamos establecer el paralelismo con la celebración de un juicio. Hay que establecer la culpabilidad de alguien y hay que aportar suficientes evidencias para determinar si es así o si nos quedamos con la hipótesis inicial de la inocencia.

Hipótesis Nula (H_0) y Alternativa (H_a)

- **Hipótesis nula (H_0):** Representa la afirmación inicial o el estado "por defecto". Usualmente indica ausencia de efecto o diferencia (por ejemplo, $\mu = 0$, $p = 0.5$).
- **Hipótesis alternativa (H_a):** Es la afirmación que se desea probar. Puede ser:
    - *Unilateral*: Por ejemplo, $\mu > 0$ o $\mu < 0$.
    - *Bilateral*: Por ejemplo, $\mu \neq 0$.

Estadístico de Prueba

Es una función de los datos muestrales que resume la información necesaria para el contraste. Su forma depende del tipo de prueba empleada, como t-prueba, z-prueba o χ^2-prueba. Este estadístico se compara con una distribución teórica bajo la suposición de que H_0 es verdadera.

Nivel de Significación (α)

Es la probabilidad máxima tolerada de cometer un **error tipo I** (rechazar H_0 cuando es verdadera). Los valores más comunes son $\alpha = 0.05$ o $\alpha = 0.01$.

Regla de Decisión

- **Valor p:** Es la probabilidad de obtener un resultado tan extremo como el observado, bajo la suposición de H_0. Si $p < \alpha$, se rechaza H_0.
- **Valor Crítico:** Se define un rango crítico en función de α y la distribución del estadístico. Si el estadístico cae dentro del rango crítico, se rechaza H_0.

Errores en la Decisión

- **Error tipo I (α):** Rechazar H_0 cuando es verdadera.
- **Error tipo II (β):** No rechazar H_0 cuando es falsa. verdadera.

Distribución bajo la Hipótesis Nula

Es la distribución teórica del estadístico de prueba cuando H_0 es verdadera. Por ejemplo, para pruebas sobre medias:

- z-distribución (si la varianza poblacional es conocida).
- t-distribución (si la varianza poblacional es desconocida y la muestra es pequeña).

Datos Muestrales

Los datos son la base para calcular el estadístico de prueba. Deben cumplir ciertos supuestos, como normalidad, independencia y homogeneidad de varianzas (dependiendo del contraste).

Interpretación del Resultado

- Si se **rechaza** H_0, se concluye que los datos ofrecen evidencia suficiente para respaldar H_a.
- Si **no se rechaza** H_0, no hay evidencia suficiente para descartar H_0, pero esto no implica que sea verdadera.

Contraste sobre proporciones: Formulación de las hipótesis

Caso bilateral

$$H_0 : p = p_0$$

$$H_a : p \neq p_0$$

Estadístico del contraste a un nivel de significación α

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Donde:

- $\hat{p} = \frac{x}{n}$: Proporción muestral, con x el número de éxitos en la muestra y n el tamaño muestral.
- p_0: Proporción teórica según H_0.
- n: Tamaño de la muestra.

Región crítica o regla de decisión

Se rechaza H_0 si $|Z| > Z_{\alpha/2}$

Contraste unilateral sobre proporciones

$$H_0 : p \leq p_0$$

$$H_a : p > p_0$$

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Se rechaza H_0 si $Z > Z_{\alpha/2}$

Contraste unilateral 2 sobre proporciones

$$H_0 : p \geq p_0$$

$$H_a : p < p_0$$

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Se rechaza H_0 si $Z < Z_{\alpha/2}$

Contraste sobre diferencia de proporciones

En esta ocasión se dispone de dos muestras de sendas poblaciones. A partir de ahí se pretende estudiar si la proporción de una determinada característica de una población es igual, mayor o menor que la de la otra población.

Contraste sobre diferencia de proporciones

$$\left.\begin{array}{lcl} H_0 & : & p_x = p_y \\ H_a & : & p_x \neq p_y \end{array}\right\} \equiv$$

Se rechaza H_0 al nivel de significación α si

$$\left|\frac{\hat{p}_x - \hat{p}_y}{\sqrt{p(1-p)(\frac{1}{n_x}+\frac{1}{n_y})}}\right| > Z_\alpha$$

$$p = \frac{n_x p_x + n_y p_y}{n_x + n_y}$$

Contraste sobre diferencia de proporciones (unilateral)

$$\left.\begin{array}{lcl} H_0 & : & p_x \leq p_y \\ H_a & : & p_x > p_y \end{array}\right\}$$

Se rechaza H_0 al nivel de significación α si

$$\frac{\hat{p}_x - \hat{p}_y}{\sqrt{p(1-p)(\frac{1}{n_x} + \frac{1}{n_y})}} > Z_\alpha$$

CH para la media de poblaciones normales. Varianza conocida

$$\left.\begin{array}{lll} H_0 & : & \mu = \mu_0 \\ H_a & : & \mu \neq \mu_0 \end{array}\right\}$$

Se rechaza H_0 al nivel de significación α si

$$\left|\frac{\sqrt{n}(\bar{X} - \mu_0)}{\sigma}\right| > Z_{\frac{\alpha}{2}}$$

CH para la media de poblaciones normales.Varianza conocida. (Unilateral)

$$\left.\begin{array}{lcl} H_0 & : & \mu \leq \mu_0 \\ H_a & : & \mu > \mu_0 \end{array}\right\}$$

Se rechaza H_0 al nivel de significación α si

$$\frac{\sqrt{n}(\bar{X} - \mu_0)}{\sigma} > Z_{\frac{\alpha}{2}}$$

CH para la media de poblaciones normales.Varianza conocida. (Unilateral 2)

$$\left.\begin{array}{lcl} H_0 & : & \mu \geq \mu_0 \\ H_a & : & \mu < \mu_0 \end{array}\right\}$$

Se rechaza H_0 al nivel de significación α si

$$\frac{\sqrt{n}(\bar{X} - \mu_0)}{\sigma} < -Z_{\frac{\alpha}{2}}$$

CH para la media de poblaciones normales. Varianza desconocida

$$\left.\begin{array}{lcl} H_0 & : & \mu = \mu_0 \\ H_a & : & \mu \neq \mu_0 \end{array}\right\}$$

Se rechaza H_0 al nivel de significación α si

$$\left|\frac{\sqrt{n}(\bar{X}-\mu_0)}{\S}\right| > t_{n-1,\frac{\alpha}{2}}$$

CH para la media de poblaciones normales.Varianza desconocida. (Unilateral)

$$\left.\begin{array}{lcl} H_0 & : & \mu \leq \mu_0 \\ H_a & : & \mu > \mu_0 \end{array}\right\}$$

Se rechaza H_0 al nivel de significación α si

$$\frac{\sqrt{n}(\bar{X} - \mu_0)}{\S} > t_{n-1,\frac{\alpha}{2}}$$

CH para la media de poblaciones normales.Varianza desconocida. (Unilateral 2)

$$\left.\begin{array}{lcl} H_0 & : & \mu \geq \mu_0 \\ H_a & : & \mu < \mu_0 \end{array}\right\}$$

Se rechaza H_0 al nivel de significación α si

$$\frac{\sqrt{n}(\bar{X}-\mu_0)}{\S} < t_{n-1,\frac{\alpha}{2}}$$

Contraste para la varianza

$$\left.\begin{array}{lcl} H_0 & : & \sigma^2 = \sigma_0^2 \\ H_a & : & \sigma^2 \neq \sigma_0^2 \end{array}\right\}$$

No se rechaza H_0 si:

$$\chi^2_{n-1,1-\frac{\alpha}{2}} < \frac{(n-1)S^2}{\sigma_0^2} < \chi^2_{n-1,\frac{\alpha}{2}}$$

Contraste para la varianza (unilateral1)

$$\left.\begin{array}{lcl} H_0 & : & \sigma^2 \leq \sigma_0^2 \\ H_a & : & \sigma^2 > \sigma_0^2 \end{array}\right\}$$

No se rechaza H_0 si:

$$\frac{(n-1)S^2}{\sigma_0^2} < \chi^2_{n-1,\alpha}$$

Contraste para la varianza (unilateral 2)

$$\left.\begin{array}{lcl} H_0 & : & \sigma^2 \geq \sigma_0^2 \\ H_a & : & \sigma^2 < \sigma_0^2 \end{array}\right\}$$

No se rechaza H_0 si:

$$\chi^2_{n-1,1-\alpha} < \frac{(n-1)S^2}{\sigma_0^2}$$

Contraste para el cociente de varianzas de dos poblaciones normales

$$\left.\begin{array}{lcl} H_0 & : & \sigma_x^2 = \sigma_y^2 \\ H_a & : & \sigma_x^2 \neq \sigma_y^2 \end{array}\right\}$$

No se rechaza H_0 si:

$$F_{n_x-1,n_y-1,1-\frac{\alpha}{2}} < \frac{S_x^2}{S_y^2} < F_{n_x-1,n_y-1,\frac{\alpha}{2}}$$

Contraste sobre diferencia de medias en dos poblaciones normales. Varianzas conocidas

$$\left.\begin{array}{lcl} H_0 & : & \mu_x = \mu_y \\ H_a & : & \mu_x \neq \mu_y \end{array}\right\} \equiv$$

Se rechaza H_0 al nivel de significación α si

$$\left| \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}} \right| > Z_{\frac{\alpha}{2}}$$

Contraste sobre diferencia de medias en dos poblaciones normales. Varianzas conocidas.

$$\left.\begin{array}{lcl} H_0 & : & \mu_x \leq \mu_y \\ H_a & : & \mu_x > \mu_y \end{array}\right\} \equiv$$

Se rechaza H_0 al nivel de significación α si

$$\frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}} > Z_\alpha$$

Contraste sobre diferencia de medias en dos poblaciones normales. Varianzas conocidas.

$$\left. \begin{array}{lcl} H_0 & : & \mu_x \geq \mu_y \\ H_a & : & \mu_x < \mu_y \end{array} \right\} \equiv$$

Se rechaza H_0 al nivel de significación α si

$$\frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}} < -Z_\alpha$$

Contraste sobre diferencia de medias en dos poblaciones normales. Varianzas desconocidas e iguales

$$\left.\begin{array}{lcl} H_0 & : & \mu_x = \mu_y \\ H_a & : & \mu_x \neq \mu_y \end{array}\right\} \equiv$$

Se rechaza H_0 al nivel de significación α si

$$\left| \frac{\bar{X} - \bar{Y}}{S_d \sqrt{\frac{1}{n_x} + \frac{1}{n_y}}} \right| > t_{n_x + n_y - 2, \frac{\alpha}{2}}$$

$$S_d^2 = \frac{(n_x - 1) \cdot S_x^2 + (n_y - 1) \cdot S_y^2}{n_x + n_y - 2}$$

Contraste sobre diferencia de medias en dos poblaciones normales. Varianzas desconocidas e iguales

$$\left.\begin{array}{lcl} H_0 & : & \mu_x \leq \mu_y \\ H_a & : & \mu_x > \mu_y \end{array}\right\} \equiv$$

Se rechaza H_0 al nivel de significación α si

$$\frac{\bar{X} - \bar{Y}}{S_d \sqrt{\frac{1}{n_x} + \frac{1}{n_y}}} > t_{n_x + n_y - 2, \alpha}$$

Contraste sobre diferencia de medias en dos poblaciones normales. Varianzas desconocidas e iguales

$$\left.\begin{array}{lll} H_0 & : & \mu_x \geq \mu_y \\ H_a & : & \mu_x < \mu_y \end{array}\right\} \equiv$$

Se rechaza H_0 al nivel de significación α si

$$\frac{\bar{X}-\bar{Y}}{S_d\sqrt{\frac{1}{n_x}+\frac{1}{n_y}}} < -t_{n_x+n_y-2,\alpha}$$

Contraste sobre diferencia de medias en dos poblaciones normales. Varianzas desconocidas y distintas

$$\left.\begin{array}{lcl} H_0 & : & \mu_x = \mu_y \\ H_a & : & \mu_x \neq \mu_y \end{array}\right\} \equiv$$

Se rechaza H_0 al nivel de significación α si

$$\left| \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{X_x^2}{n_x} + \frac{S_y^2}{n_y}}} \right| > t_{g, \frac{\alpha}{2}}$$

$g = n_1 + n_2 - 2 - \Delta$, gr. de libertad
$\Delta =$ es el entero más próximo a $\frac{((n_y-1)S_x - (n_x-1)S_y)^2}{(n_y-1)S_x^2 + (n_x-1)S_y^2}$

Contraste sobre diferencia de medias en dos poblaciones normales. Varianzas desconocidas y distintas

$$\left.\begin{array}{lcl} H_0 & : & \mu_x \leq \mu_y \\ H_a & : & \mu_x > \mu_y \end{array}\right\} \equiv$$

Se rechaza H_0 al nivel de significación α si

$$\frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_x^2}{n_x} + \frac{S_y^2}{n_y}}} > t_{g,\alpha}$$

Contraste sobre diferencia de medias en dos poblaciones normales. Varianzas desconocidas y distintas

$$\left.\begin{array}{lcl} H_0 & : & \mu_x \geq \mu_y \\ H_a & : & \mu_x < \mu_y \end{array}\right\} \equiv$$

Se rechaza H_0 al nivel de significación α si

$$\frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_x^2}{n_x} + \frac{S_y^2}{n_y}}} < -t_{g,\alpha}$$

En los contrastes de hipótesis sobre dos poblaciones normales, se asume que son independientes. Si las variables aleatorias son dependientes o apareadas, se considera la variable aleatoria $D = X - Y \sim N(\mu_x - \mu_y, \sigma_D)$

Estado Finalizado
Comenzado   domingo, 24 de noviembre de 2024, 13:42
Completado   domingo, 24 de noviembre de 2024, 14:14
Duración   31 minutos 17 segundos
Puntos 6,00/7,00
Calificación   8,57 de 10,00 (85,71%)
Pregunta 1
Correcta
Se puntúa 1,00 sobre 1,00
Marcar pregunta
Enunciado de la pregunta
Dame un intervalo de confianza al 92
% para la media de la siguiente muestra aleatoria simple de una Variable Aleatoria Normal de varianza σ^2=(85)^2
.

Muestra=[4.74,1.51,3.93,3.06,0.47,0.21,1.77,0.6,0.63,4.2,1.13,0.02,2.23,1.18,2.9,2.05]

IC92=[a,b]

a=
1.2141005

Tu respuesta fue interpretado como: 1.2141005

b=
2.614649

Tu respuesta fue interpretado como: 2.614649

Retroalimentación
Respuesta correcta, bien hecho.

Perfecto

Respuesta correcta, bien hecho.

Observa que se trata de un ejercicio para calcular el intervalo de confianza de una v.a. normal de varianza conocida.

Encontrarás el intervalo en las transparencias relativas a los intervalos de confianza.

The answer 1.2141005715
, which can be typed as 1.2141005715, would be correct.

The answer 2.6146494285
, which can be typed as 2.6146494285, would be correct.

Pregunta 2

Correcta
Se puntúa 1,00 sobre 1,00
Marcar pregunta
Enunciado de la pregunta
Calcula un intervalo de confianza al 97
% para la varianza de una v.a. Normal a partir de la siguiente muestra aleatoria simple:

Muestra=[4.25,0.09,1.22,2.27,3.35,3.42,0.04,4.06,0.34,4.75,0.57,3.52,0.37,2.02,3.45]

IC97=[a,b]

a=
1.4312343

Tu respuesta fue interpretado como: 1.4312343

b=
7.87517094

Tu respuesta fue interpretado como: 7.87517094

Retroalimentación
Respuesta correcta, bien hecho.

El valor menor del intervalo está bien calculado.

Respuesta correcta, bien hecho.

El segundo está bien calculado.

Observa que se trata de un ejercicio para calcular el intervalo de confianza de la varianza de una v.a. normal.

Encontrarás el intervalo en las transparencias relativas a los intervalos de confianza.

The answer 1.43123438554
, which can be typed as 1.43123438554, would be correct.

The answer 7.87517094199
, which can be typed as 7.87517094199, would be correct.

Pregunta 3
Correcta
Se puntúa 1,00 sobre 1,00
Marcar pregunta
Enunciado de la pregunta
Dame un intervalo de confianza al 91
% para la media de la siguiente muestra aleatoria simple de una Variable

Aleatoria Normal de varianza desconocida .

Muestra=[4.98,1.91,2.36,3.99,1.6,1.21,0.84,2.16,2.53,0.51,0.46,1.28,3.78,3.52]

IC91=[a,b]

a=
1.54058971017

Tu respuesta fue interpretado como: 1.54058971017

b=
2.90655314

Tu respuesta fue interpretado como: 2.90655314

Retroalimentación
Respuesta correcta, bien hecho.

Perfecto

Respuesta correcta, bien hecho.

Observa que se trata de un ejercicio para calcular el intervalo de confianza de una v.a. normal de varianza conocida.

Encontrarás el intervalo en las transparencias relativas a los intervalos de confianza.

The answer 1.54058975345
, which can be typed as 1.54058975345, would be correct.

The answer 2.90655310369
, which can be typed as 2.90655310369, would be correct.

Pregunta 4
Correcta
Se puntúa 1,00 sobre 1,00
Marcar pregunta
Enunciado de la pregunta
Calcula un intervalo de confianza al 99
% para la diferencia de medias dos V. A. Normales independientes de varianza común cuyas dos muestras aleatorias simples son:

Muestra1=[0.82,0.22,2.65,3.72,0.84,4.18,4.74,2.82,1.07,1.26,1.56,4.88,4.11,0.7,4.27,0.12,1.54,3.59]

Muestra2=[3.17,3.06,4.4,1.48,2.54,3.68,0.7,4.84,1.77,3.29,0.11,1.07,0.75,1.46]

IC99=[a,b]

a=
-1.34298947

Tu respuesta fue interpretado como: −1.34298947

b=
1.51362439

Tu respuesta fue interpretado como: 1.51362439

Retroalimentación
Respuesta correcta, bien hecho.

Perfecto

Respuesta correcta, bien hecho.

Observa que se trata de un ejercicio para calcular el intervalo de confianza de la diferencia entre dos poblaciones normales de varianza desconocida.

Encontrarás el intervalo en las transparencias relativas a los intervalos de confianza.

The answer −1.34298947862
, which can be typed as -1.34298947862, would be correct.

The answer 1.51362439926
, which can be typed as 1.51362439926, would be correct.

Pregunta 5
Incorrecta
Se puntúa 0,00 sobre 1,00
Marcar pregunta
Enunciado de la pregunta
Calcula un intervalo de confianza al 98
% para la diferencia de medias dos V. A. Normales independientes de varianza común cuyas dos muestras aleatorias simples son:

Muestra1=[0.23,4.83,3.49,3.4,0.83,3.14,0.37,4.94,1.56,2.1,0.7,2.52,0.23]

Muestra2=[4.91,3.79,3.24,1.15,4.7,1.1,4.73,1.79,4.85,4.62,0.23,0.16,2.8,0.65,2.12]

IC98=[a,b]

a=
-2.1809587

Tu respuesta fue interpretado como: −2.1809587

b=
1.09562541

Tu respuesta fue interpretado como: 1.09562541

Retroalimentación
Respuesta incorrecta.

El valor inferior del intervalo parece haber fallado.

Para plantear el correspondiente intervalo, primero deberemos calcular el centro:
$\bar{X}_1=2.18$

$\bar{X}_2=2.72266666667$

El centro del intervalo viene dado por $\bar{X}_1-\bar{X}_2$
. Ahora obtendremos la Quasi-varianza muestral. Observa que
$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{X})^2$

$S^2_1=2.84721666667$ $S^2_2=3.26270666667$

Para obtener la amplitud del intervalo, necesitaremos :

$S_T^2=3.07094205128$

Por tanto $S_T=1.75241035471$
. Finalmente, obtenemos el valor crítico de la t-student:

$t_{\frac{\alpha}{2},n_1+n_2-2}=2.47862973255$

Y ya podríamos plantear la amplitud como $Amplitud=1.64592171751$
.

El intervalo buscado sería:

$(-2.18858838418,1.10325505084)$

Respuesta incorrecta.

El valor superior del intervalo ha fallado.

Para plantear el correspondiente intervalo, primero deberemos calcular el centro:
$\bar{X}_1=2.18$

$\bar{X}_2=2.72266666667$

El centro del intervalo viene dado por $\bar{X}_1-\bar{X}_2$
. Ahora obtendremos la Quasi-varianza muestral. Observa que
$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{X})^2$

$S21=2.84721666667S22=3.26270666667$

Para obtener la amplitud del intervalo, necesitaremos :

$ST2=3.07094205128$

Por tanto $ST=1.75241035471$
. Finalmente, obtenemos el valor crítico de la t-student:

$t\alpha 2,n1+n2-2=2.47862973255$

Y ya podríamos plantear la amplitud como $Amplitud=1.64592171751$
.

El intervalo buscado sería:

$(-2.18858838418,1.10325505084)$

Observa que se trata de un ejercicio para calcular el intervalo de confianza de la diferencia entre dos poblaciones normales de varianza desconocida.

Encontrarás el intervalo en las transparencias relativas a los intervalos de confianza.

The answer -2.18858838418
, which can be typed as -2.18858838418, would be correct.

The answer 1.10325505084
, which can be typed as 1.10325505084, would be correct.

Pregunta 6
Correcta
Se puntúa 1,00 sobre 1,00
Marcar pregunta
Enunciado de la pregunta
Obtén un intervalo de confianza al 93
% para la proporción de trilingues en el alumnado de la UA sabiendo que hemos encontrado 155
en una muestra de tamaño 171
:

$IC93=[a,b]$

a=
0.86608043

Tu respuesta fue interpretado como: 0.86608043

b=
0.94678506

Tu respuesta fue interpretado como: 0.94678506

Retroalimentación
Respuesta correcta, bien hecho.

Muy bien con el valor inferior del intervalo.

Respuesta correcta, bien hecho.

...y el valor superior del intervalo está bien calculado.

Observa que se trata de un ejercicio para calcular el intervalo de confianza de una v.a. normal de varianza conocida.

Encontrarás el intervalo en las transparencias relativas a los intervalos de confianza.

The answer 0.866080433269
, which can be typed as 0.866080433269, would be correct.

The answer 0.946785063807
, which can be typed as 0.946785063807, would be correct.

Pregunta 7
Correcta
Se puntúa 1,00 sobre 1,00
Marcar pregunta
Enunciado de la pregunta
Se pretende tomar una muestra para identificar la proporción de ordenadores infectados por un nuevo virus. Si pretendemos dar un intervalo de confianza al 94
% que tenga una radio 0.111
, ¿Qué tamaño debería tener dicha muestra?

n=
72

Tu respuesta fue interpretado como: 72

Retroalimentación
Respuesta correcta, bien hecho.

Bien calculado.

]

The answer 72
, which can be typed as 72, would be correct.