Clase 6

jueves, 4 de agosto de 2022 18:59

Intervalos de confianza

Ejercicio 6

Se arroja 50 veces una moneda con probabilidad p de salir cara. Hallar un intervalo de confianza asintótico de nivel 0.95 para p basado en la observación x=50.

$$X_i = \begin{cases} 1 & \text{si sale cara en el experimento } i \\ 0 & \text{" " ceca " " " "} \end{cases} \qquad X_i \sim Be(p)$$

↳ Parámetro a estimar

Estimador de p: $\overline{X} = \dfrac{\sum_{i=1}^{50} X_i}{50} = \hat{p}$ $\qquad \sum_{i=1}^{50} X_i \sim Bi(50, p)$

$\underline{X} = \{X_1, X_2, \dots, X_n\}$ m.a.,

si n es grande: $\dfrac{\overline{X} - E[\overline{X}]}{\sqrt{V(\overline{X})}} \overset{(a)}{\sim} N(0,1) \rightarrow$

$$U = \frac{\overline{X} - p}{\sqrt{p\cdot(1-p)}} \cdot \sqrt{50} \overset{(a)}{\sim} N(0,1)$$

$$P(a < U < b) \simeq 0{,}95 \Rightarrow P\left(-z_{1-\frac{\alpha}{2}} < \frac{\overline{X}-p}{\sqrt{p(1-p)}} \cdot \sqrt{50} < z_{1-\frac{\alpha}{2}}\right) = 0{,}95$$

$$U = \frac{\overline{X} - p}{\sqrt{p(1-p)}} \cdot \sqrt{50} \overset{(a)}{\sim} N(0,1)$$

$p \rightarrow \overline{X}$

↳ Por teorema de slutsky

$$I(\overline{x}) = \left(\overline{X} - z_{1-\frac{\alpha}{2}} \sqrt{\frac{\overline{x}(1-\overline{x})}{50}} ; \overline{X} + z_{1-\frac{\alpha}{2}} \sqrt{\frac{\overline{x}(1-\overline{x})}{50}}\right)$$

↳ 0,975 ↳ 0,975.

De la muestra $n = 50$, $x = 50 \Rightarrow \overline{x} = 1$

$$IC(p) = \left(1 - z_{1-\frac{\alpha}{2}} \sqrt{\frac{1(1-1)}{50}} ; 1 + z_{1-\frac{\alpha}{2}} \sqrt{\frac{1(1-1)}{50}}\right)$$

$$IC(p) = (1, 1)$$

De un experimento en los efectos de un medicamento para la ansiedad se midió el puntaje en un test de memoria antes y después de tomar el medicamento. A partir de los datos que se encuentran en el archivo Islander_data.csv hallar un IC para la media del tiempo de respuesta después de consumir el medicamento.

X_i: puntaje del tiempo de respuesta después de consumir el medicamento

Estimador de μ: $\overline{X} \Rightarrow U = \dfrac{\overline{X} - \mu}{\sigma} \overset{(a)}{\sim} N(0,1)$

Estimador de μ: $\bar{X} \Rightarrow U = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \overset{(a)}{\sim} N(0,1)$

Usamos s (estimado) porque σ no se conoce.

$$IC(\mu) = \left(\bar{x} - z_{1-\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}} , \bar{x} + z_{1-\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}}\right)$$

$$IC(\mu) = \left(60{,}92 - 1{,}96 \cdot \frac{18{,}13}{\sqrt{198}} , 60{,}92 + 1{,}96 \cdot \frac{18{,}13}{\sqrt{198}}\right)$$

$z_{0,975}$

$IC(\mu) = (58{,}39 , 63{,}45)$ con 95% de confianza.

Test de Hipótesis

Ejercicio 1

Supongamos que en las especificaciones de procedimientos de una planta de energía nuclear se establece que la resistencia media de soldadura debe superar 100lb/plg. Supongamos que somos el director del equipo de inspección del ente regulador estatal que debe determinar si la planta cumple con las especificaciones. e planea seleccionar una muestra al azar de soldaduras y realizar pruebas en cada una de ellas.

X_i: resistencia de la soldadura i (lb/plg.)

1. ¿Cuáles son las hipótesis a testear?

$H_0: \mu \leq 100$ $\quad\quad$ $H_1: \mu > 100$

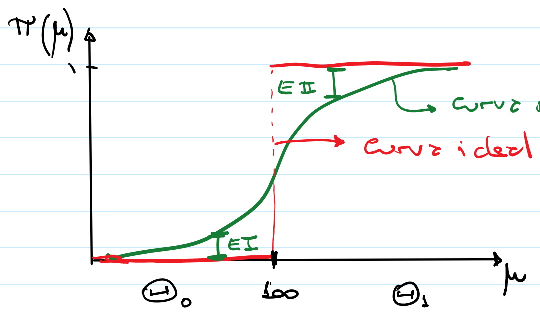
resistencia media de las soldaduras.

2. Explicar que significa en este contexto el error de tipo I y el de tipo II, y discutir cuáles son las consecuencias de cometer cada tipo de error

E I: las resistencias son buenas ($\mu > 100$) cuando en realidad no lo son

E II: las resistencias son malas ($\mu \leq 100$) cuando en realidad no lo son.

CURVA DE POTENCIA



$\pi(\mu) = P_\mu(\text{rechazar } H_0)$

Ejercicio 2

Se tiene una m.a. de tamaño n de una población uniforme en el intervalo $(0,\theta)$.

1. Diseñar un test de hipótesis para decidir si θ es mayor a 2.5 con un nivel de significación de 0.05.
2. Suponer θ=3, n=20, simular la m.a. y decidir en base a ella.
3. Hallar el p-valor.

1. $\underline{X} = \{X_1, X_2, X_3, \dots, X_n\} \rightarrow$ m.a. $\quad X_i \sim U(0,\theta) \quad i=1,2,\dots,n$

Estimador de máxima verosimilitud de θ : $\max\{X_i\}_{i=1}^{n} = U$

$$F_U(u) = P(U \leq u) = P(X_1 \leq u, X_2 \leq u, \dots, X_n \leq u) = \left(\frac{u}{\theta}\right)^n$$

$$F_{X_1}(x) = \frac{x}{\theta}\,\mathbb{1}\{0 < x < \theta\}$$

$$H_0 : \theta \leq 2,5 \qquad H_1 : \theta > 2,5$$

Propongo como estadístico de prueba: $T = \dfrac{\overbrace{\max(X_i)}^{U}}{\theta}$

$$F_T(t) = P(T \leq t) = P\left(\frac{U}{\theta} \leq t\right) = P(U \leq \theta t) =$$

$$= F_U(\theta t) = \left(\frac{\theta t}{\cancel{\theta}}\right)^n$$

$$F_T(t) = t^n\,\mathbb{1}\{0 < t < 1\}$$

$$\delta(\underline{X}) = \begin{cases} 1 & \text{si } \dfrac{\max(X_i)}{\theta} > k_\alpha \\ 0 & \text{en caso contrario.} \end{cases}$$

Bajo Ho verdadera.

$$\alpha = \max_{\theta \in \Theta_0} P(\delta(\underline{X}) = 1) \;, \quad P_{\theta=2,5}\left(\frac{\max(X_i)}{\theta} > k_{0,05}\right) = 0,05 \Rightarrow$$

$$\Rightarrow 1 - t^n = 0,05 \Rightarrow t = \sqrt[n]{0,95} = k_{0,05}$$

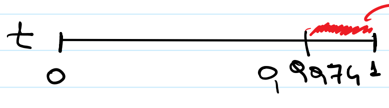
$$\delta(\underline{X}) = \mathbb{1}\left\{\frac{\max(X_i)}{2,5} > \sqrt[n]{0,95}\right\}$$

2. Para $n = 20$:

$$\delta(\underline{X}) = \mathbb{1}\left\{\frac{\max(X_i)}{2,5} > \underline{0,9974}\right\}$$

De la muestra $n=20$ para $\theta = 3$ se obtuvo $\max\{x_i\}_{i=1}^{20} = 2,9291$

entonces $\dfrac{\max\{x_i\}}{2,5} = \dfrac{2,9291}{2,5} = 1,1716 > 0,9974 \Rightarrow$ se rechaza Ho.



$\alpha = \max\limits_{\theta \in \Theta_0} P(EI) = 0,05$

$$p\text{-valor} = P_{\theta=2,5}\left(\frac{\max(X_i)}{2,5} > 1,1716\right) = 0$$

- Otro ejemplo: se toma una muestra de $n=20$ y $\dfrac{\max(x_i)}{2,5} = 0,9998$

$$p\text{-valor} = P_{\theta=2,5}\left(\frac{\max(X_i)}{2,5} > 0,9998\right) = 1 - F_T(0,9998) =$$

$$= 1 - 0,9998^{20} = \underline{0,004}$$

Ejercicio 3

Se realiza un estudio para estimar la proporción de residentes en una ciudad que están a favor de la construcción de una planta de energía nuclear. En una muestra de 100 personas se observó que la proporción de individuos que se encuentran a favor fue de 0.62

1. ¿Puede decirse con un nivel de significación de 0.01 que la mayor parte de la población está a favor de la construcción de la planta nuclear?
2. Hallar el p-valor.

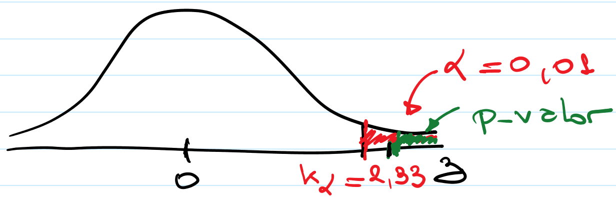
$$X_i = \begin{cases} 1 & \text{si el residente } i \text{ está a favor de la planta} \\ 0 & \text{" " " " no está " " " " "} \end{cases} \qquad X_i \sim Be(p)$$

$$H_0: p \leq \frac{1}{2} \qquad H_1: p > \frac{1}{2}$$

Estimador de p: $\hat{p} = \overline{X} = \dfrac{\sum_{i=1}^{100} X_i}{100}$

Por T.C.L. $\dfrac{\overline{X} - p}{\sqrt{p(1-p)}} \sqrt{100} \overset{(a)}{\sim} N(0;1)$

$$\delta(\underline{X}) = \begin{cases} 1 & \text{si } \dfrac{\overline{X} - 0{,}5}{\sqrt{0{,}5 \cdot 0{,}5}} \sqrt{100} > k_\alpha = 2{,}33 \\ 0 & \text{en otro caso} \end{cases}$$



De la muestra de 100 se obtuvo $\overline{x} = 0{,}62 \Rightarrow$

$$\Rightarrow \frac{0{,}62 - 0{,}5}{0{,}5} \cdot \sqrt{100} = 2{,}4 > 2{,}33 \text{ se rechaza } H_0$$

$$\text{p-valor} = P_{p=0,5}\left(\frac{\overline{X} - 0{,}5}{0{,}5} \sqrt{100} > 2{,}4 \right) = 1 - 0{,}9918 = 0{,}0082$$

- Si $p = 0{,}6$, calcular $P(\text{E II})$

$$\alpha = P\left(\frac{\overline{X} - 0{,}5}{0{,}5} \cdot 10 > 2{,}33 \right) = 0{,}05$$

$$P_{p=0,6}(\delta(\underline{X}) = 0) =$$