ESTADÍSTICA 03-05-2017

TEMA 5: ESTIMACIÓN POR INTERVALOS

1. INTERVALO PARA EL PARÁMETRO θ DESCONOCIDO DE LA POBLACIÓN

Intervalo de confianza: $(\theta_1(x_1, \dots, x_n), \theta_2(x_1, \dots x_n))$

$P(\theta \in I.C.) = 1 - \alpha$

Fijado $\alpha \in (0,1)$

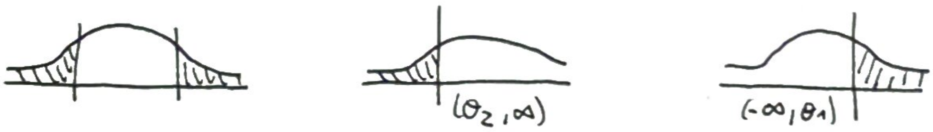
$-z_{\alpha/2} \cdot \sigma - \bar{x} \leq -\mu \leq z_{\alpha/2} \cdot \sigma \cdot \bar{x}$

$$\underbrace{\bar{x} - z_{\alpha/2} \cdot \sigma}_{\theta(x_1, \dots x_n)} \leq \mu \leq \underbrace{\bar{x} + z_{\alpha/2}\, \sigma}_{\theta_2(x_1, \dots, x_n)}$$

$(\theta_1(x_1, \dots, x_n), \theta(x_1, \dots, x_n)) \rightarrow$ INTERVALO DE CONFIANZA

$(1-\alpha)\, 100\% \rightarrow$ Nivel de confianza

$(1-\alpha) \rightarrow$ Grado de confianza



2. MÉTODOS PARA ESTIMAR EL INTERVALO

→ **Método de variable pivotal**

$T(x_1, \dots x_n)$ dado

La variable pivotal es $T(x_1, \dots x_n, \theta)$ donde $T(x_1 \dots, x_n)$ suficientes

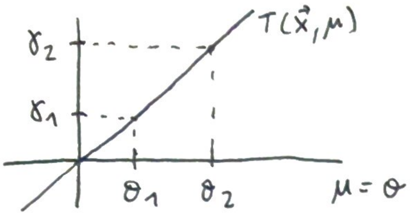
REQUISITOS

※ Ejemplo

$P(\gamma_1 \leq T(x_1, \dots x_n, \mu) \leq \gamma_2) = 1 - \alpha$ $\qquad \sigma$ conocido

$T(x_1, \dots, x_n) = \bar{x}$ $\qquad \mu$ parámetro

$T(x_1, \dots, x_n, \mu) = \dfrac{\bar{x} - \mu}{\sigma} \sim N(0,1)$



$T(\vec{x}, \theta_1) \leq T(\vec{x}, \theta)$

$\theta_1 \leq \theta_2 \qquad \theta_2 \leq \theta$

CASOS GENERALES DE $N(\mu, \sigma^2)$

(1) $h(\theta) = \mu$ $\qquad \sigma^2$ conocido

Variable puntual: $\dfrac{\bar{x} - \mu}{\sigma} = T(x_1, \dots, x_n, \mu)$

$IC(\mu) = (\bar{x} - \sigma/\sqrt{n}\; z_{\alpha/2} \;,\; \bar{x} + \sigma/\sqrt{n}\; z_{\alpha/2})$

$L = 2\, \sigma/\sqrt{n}\; z_{\alpha/n}$

Los pasos a seguir son:

→ Buscar T suficiente → $\bar{x}$

→ Buscar cantidad pivotal → $T(\vec{x}, \mu) = \dfrac{\bar{x} - \mu}{\sigma}$

→ Comprobar

↳ $T(x_1, \dots x_n, \mu) \sim$ distribución no depende del parámetro

$\dfrac{\bar{x} - \mu}{\sigma} \sim N(0,1)$

↳ como fn de μ estrictamente monótona

Fijado $\alpha \in (0,1)$

$-z_{\alpha/2}$ $\qquad z_{\alpha/2}$

$P(z \le \gamma_2(\alpha)) = 1-\alpha/2$

$z_{\alpha/2}$

$P(z \le \gamma_1(\alpha)) = \alpha/2$

$z_{1-\alpha/2} = -z_{\alpha/2}$

$\sigma = 17{,}8$ $\quad \bar{x} = 83{,}6$

$n = 64$ $\quad 1-\alpha = 0'95$

$$IC_{95\%}(\mu) = (\bar{x} - z_{\alpha/2}\, \sigma/\sqrt{n}\ ,\ \bar{x} + z_{\alpha/2}\, \sigma/\sqrt{n}) = 179239{,}87461$$

INTERVALO DE LONGITUD MÍNIMA

$L = \gamma_2(\alpha) - \gamma_1(\alpha)$ mínima

$$\underbrace{P(T \le \gamma_2)}_{\phi(\gamma_2)} - \underbrace{P(T \le \gamma_1)}_{\phi(\gamma_1)} = 1-\alpha$$

ϕ distribución $N(\mu, \sigma^2)$

Me construyo una función (que es la función a minimizar)

$$F(\gamma_1, \gamma_2) = \gamma_2(\alpha) - \gamma_1(\alpha) + \lambda[\phi(\gamma_2) - \phi(\gamma_1) - 1 + \alpha]$$

$$\frac{\partial F}{\partial \gamma_1} = -1 - \lambda \frac{\partial \phi}{\partial \gamma_1} = -1 - \lambda p(\gamma_1) = 0$$

$$\frac{\partial F}{\partial \gamma_2} = 1 + \lambda p(\gamma_2) = 0$$ donde p es la función de densidad de $N(\mu, \sigma)$

$$\frac{\partial F}{\partial \lambda} = [\phi(\gamma_2) - \phi(\gamma_1) - 1 + \alpha] = 0$$

$$\lambda = \frac{1}{p(\gamma_1)} = -\frac{1}{p(\gamma_2)} \Rightarrow p(\gamma_1) = p(\gamma_2)$$

$N(0,1)$

$$\underbrace{\frac{1}{\sqrt{2\pi}} e^{-1/2\, \gamma_1^2} = \frac{1}{\sqrt{2\pi}} e^{-1/2\, \gamma_2^2}}$$

* Ejemplo

04-05-2017

$N(\mu, \sigma^2)$

$h(\sigma) = \sigma^2$

¿ IC (σ^2)

$(1-\alpha) = 100\%$

μ desconocida

Método de la variable pivotal

$T(x_1 \dots x_n; \sigma)$ → como función $\bar{x}$, no depende de σ función σ, estrictamente creciente o decreciente

→ $S^2 = \frac{1}{n-1} \sum (x_i - \mu)^2$

$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$ variable pivotal quiero que sea mínima

$P(\gamma_1 \leq \underbrace{T(x_1, \dots, x_n, \sigma)}_{\chi^2_{n-1}} \leq \gamma_2) = 1-\alpha$

$L = \gamma_2 - \gamma_1 \rightarrow$ mini

$\phi(\gamma_2) - \phi(\gamma_1) = 1-\alpha$

ϕ distribución χ^2_{n-1}, f. densidad de χ^2

$F(\gamma_1, \gamma_2, \lambda) = \gamma_2 - \gamma_1 + \lambda(\phi_2(\gamma_2) - \phi_1(\gamma_1) - 1 + \alpha)$

$\frac{\partial F}{\partial \gamma_1} = -1 - \lambda f(\gamma_1) = 0$

$\frac{\partial F}{\partial \gamma_2} = 1 + \lambda f(\gamma_2) = 0$

→ $\lambda = -1/f(\gamma_1)$, $\lambda = -1/f(\gamma_2)$ > $f(\gamma_1) = f(\gamma_2)$

Lo igualamos y nos sale $\gamma_2^n e^{-1/2\gamma_2} = e^{-1/2\gamma_1} \gamma_1^n$

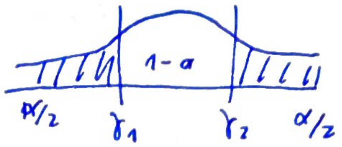
Para ... se ...

... $= \alpha/2$

$P(\gamma_1 \leq \gamma_1) = \alpha/2$

Ahora hacemos $\gamma_2 = \chi^2_{n-1,\, \alpha/2}$

$\gamma_1 = \chi^2_{n-1,\, 1-\alpha/2}$



Cogemos la de $\gamma_2 \longrightarrow \frac{(n-1)\, S^2}{\sigma^2} \leq \chi^2_{n-1,\, \alpha/2}$

$\sigma^2 \geq (n-1)\, S^2 / \chi^2_{(n-1)\, \alpha/2}$

Cogemos ahora la de γ_1

$\sigma^2 \leq (n-1)\, S^2 / \chi^2_{n-1,\, 1-\alpha/2}$

$$IC(\sigma^2)_{(1-\alpha)/100\%} = \left((n-1) \frac{S^2}{\chi^2_{n-1,\alpha/2}} ,\ (n-1)\, S^2 / \chi^2_{n-1,\, 1-\alpha/2} \right)$$

\# Ejercicio

m.a.s. $n = 64$, $\bar{x} = 66'3$, μ desconocido

$IC = 99\%$ para μ

$IC = 95\%$ para μ

Extraer conclusiones

De modo que en los casos generales

(1) $N(\mu\, \sigma)$ σ^2 varianza

$h(\theta) = \mu$ σ conocido

(2) $N(\mu, \sigma)$ σ^2 desconocido

$$T(x_1, \dots x_n, \theta) = \sqrt{n}\, \frac{\bar{x} - \mu}{S} \sim t_{n-1}$$

(3) $N(\mu, \sigma)$ σ^2 desconocido, muestra grande

n grande

(4) $N(\mu, \theta)$ μ conocido

$h(\theta) = \sigma^2$

$$T(x_1, \dots x_n, \theta) = \frac{1}{\sigma^2} \sum (x_i - \mu)^2 \sim \chi$$

(5) $N(\mu, \theta)$ μ desconocido, $h(\theta) = \sigma^2$

(6) $N(\mu, \theta)$

$h(\theta) = \mu_1 - \mu_2$

σ_1^2, σ_2^2 conocidos

$$\frac{\bar{x} - \bar{y} - (\mu_x - \mu_y)}{\sqrt{\frac{\sigma_x^2}{n_x} - \frac{\sigma_y^2}{n_y}}} \longrightarrow N(0,1)$$

(7) $N(\mu, \theta)$

$h(\theta) = \mu_1 - \mu_2$ $\sigma_1^2 = \sigma_2^2$ y desconocidos

$$\frac{\bar{x} - \bar{y} - (\mu_x - \mu_y)}{\sqrt{S_x^2/n_x + S_y^2/n_y}}$$

(8) $N(\mu, \theta)$

$h(\theta) = \frac{\sigma_1^2}{\sigma_2^2} \sim F_{n,m}$

- Si nos da $X \sim f_\theta(x)$ localización ✓

$(x \to y = x - \theta)$

la variable pivotal es $y = x \pm n\theta$

TEOREMA ⚠ Importante

$X_1, \dots X_n$ m.a.s. X $\quad P_\theta(\{F_\theta : \theta \in \Theta\})$

F_θ continua y estrictamente monótona, entonces

$T = -2 \sum_{i=1}^{n} \ln F_\theta$ es la cantidad pivotal siempre y $T \sim \chi^2_{2n}$

Ejemplo

$f_\theta(x) = \theta x^{\theta-1} I_{(0,1)}(x) \quad \theta \geq 0 \qquad I(\theta) \mid (1-\alpha) \quad 10\%$

$F_\theta(x) = \int_0^x \theta x^{\theta-1} dx = x^\theta \Big]_0^x = x^\theta \qquad x \in (0,1)$

$\theta \geq 0$

Sabemos como se distribuye x_i, luego podemos saber como se distribuye $F(x_i)$, $\sum F(x_i)$, $2\sum \ln F(x_i)$, ...

Tengo $x_i \xrightarrow{g} -\ln x_i = y_i$

$x_i = e^{-y_i} \xleftarrow[g^{-1}]{} y_i$

Por tanto

$f_{y_i}(y_i) = f_{y_i}(e^{y_i}) \cdot e^{-y_i} = \theta e^{-y_i(\theta-1)} e^{-y_i}$

$= \theta e^{-y_i\theta} \sim \exp(\theta) = \text{gamma}(\overset{p}{1}, \overset{a}{\theta})$

$= \text{gama}(\underset{a}{\theta}, \underset{p}{1})$

De modo que $\sum y_i = -\sum \ln x_i$

$\sim \text{gamma}(\theta, n)$ pq es reproductiva respecto a p

Por tanto $-\theta \sum \ln F(x_i) = \text{gamma}(1, n)$

3. MÉTODO DE NEYMAN (general)

08-05-2017

NOTA
ECM: estimador cuadrático medio
ECMV: estimador centrado de mínima varianza

Buscaremos E.M.V (estimador máxima verosimilitud)

Ejemplo:

$X_1 \dots X_n$ m.a.s. población $X \sim U(0,\theta)$

I C (θ)
$(1-\alpha|100\%)$ nivel de significación

$0 \quad X_{(1)} \quad X_{(n)} \quad \theta$ $\qquad \theta \in [X_n, \alpha)$

$\hat{\theta} = X_{(n)}$ EMV

$F_{X(n)}(y) = (F(y))^n = \left(\frac{y}{\theta}\right)^n \qquad 0 < y < \theta$

$P\{\gamma_1(\alpha) \le \underset{\parallel}{\overset{X(n)}{T}} \le \gamma_2(\alpha)\} = 1-\alpha$

$T = X_{(n)}$

(1) Despejo γ_1, γ_2 como función del parámetro

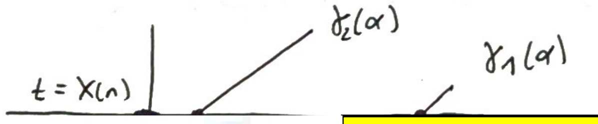
$\underbrace{P(\gamma_1(\alpha) \le X_{(n)} \le \gamma_2(\alpha))} = 1-\alpha = 1-\alpha_1-\alpha_2 \qquad \alpha \in (0,1)$
$\alpha = \alpha_1 + \alpha_2$

$F_{X(n)}(\gamma_2(\alpha)) - F_{X(n)}(\gamma_1(\alpha)) = 1-\alpha_1-\alpha_2$

$P(X_{(n)} \le \gamma_2) = F_{X(n)}(\gamma_2(\alpha)) = 1-\alpha_2 = \left(\frac{\gamma_2}{\theta}\right)^n \Rightarrow \gamma_2 = \theta(1-\alpha+\alpha_1)^{1/n}$

$P(X_{(n)} \le \gamma_1) = F_{X(n)}(\gamma_1(\alpha)) = \alpha_1 = \left(\frac{\gamma_1}{\theta}\right)^n \Rightarrow \gamma_1 = \theta(\alpha_1)^{1/n}$

(2) γ_1 y γ_2 son estrictamente monótonos como función del parámetro $\gamma_2(\alpha)$



$\gamma_1 = \theta(\alpha_1)^{1/n} = X_{(n)}$

(3) Que se pueda resolver

$$X_{(n)} = \theta \cdot \alpha_1^{1/n} \Rightarrow \theta_2 = X_{(n)} / \alpha_1^{1/n}$$

$$X_{(n)} = \theta (1-\alpha+\alpha_1)^{1/n} \Rightarrow \theta_1 = \frac{X_{(n)}}{(1-\alpha+\alpha_1)^{1/n}}$$

$$\underset{(1-\alpha)100\%}{I.C} (\theta) = (\theta_1, \theta_2) = \left(\frac{X_{(n)}}{(1-\alpha+\alpha_1)^{1/n}}, \frac{X_{(n)}}{\alpha_1^{1/n}} \right)$$

4. INTERVALOS DE CONFIANZA BASADOS EN DISTRIBUCIONES ASINTÓTICAS

Sea $\{T_n\}_{n \in \mathbb{N}}$ una sucesión de estimadores de θ asintóticamente normal

$$\frac{\sqrt{n}\,(T_n(\vec{x}) - \theta)}{\sqrt{V(\theta)}} \xrightarrow[n\to\infty]{\mathcal{L}} N(0,1)$$

(1) Si $V(\theta)$ (varianza de θ) es conocida ; $\alpha \in [0,1]$ prefijado

$$P\left(c_1(\alpha) \le \sqrt{n}\,\frac{T(\vec{x}) - \theta}{\sqrt{V(\theta)}} \le c_2(\alpha)\right) = 1-\alpha$$

$$\underset{(1-\alpha)100\%}{I.C.} (\theta) = \left(T(\vec{x}) - c_2(\alpha)\sqrt{\frac{V(\theta)}{n}}, T(\vec{x}) - c_1(\alpha)\sqrt{\frac{V(\theta)}{n}} \right)$$

$$c_2(\alpha) = -c_1(\alpha) = z_{\alpha/2}$$

(2) Si la $V(\theta)$ es desconocida

Si $\{S_n\}_{n\in\mathbb{N}}$ suc. tal que $\{S_n\} \xrightarrow{P} V(\theta)$

$$\sqrt{n}\left(\frac{T_n(\vec{x}) - \theta}{\sqrt{V(\theta)}}\right) \xrightarrow{\mathcal{L}} Z \sim N(0,1)$$

$$\underset{1-\alpha\,100\%}{I.C} (\theta) = \left(T(\vec{x}) + z_{\alpha/2}\sqrt{\frac{S_n(x)}{n}}\right)$$

$T(\vec{x})$ ...

$(1-\alpha)100\%$

$$\frac{\sqrt{n}\,(T_n - \theta)}{S_n} \xrightarrow{\mathcal{L}} Z \sim N(0,1)$$

PROPIEDAD EMV = NORMALIDAD ASINTÓTICA

$\Theta \subset \mathbb{R}$ $\quad \{f_\theta(x) : \theta \in \Theta\}$

H1: • $\forall \theta \subset A \subset \Theta$ $\quad$ A intervalo abierto de Θ

$\exists \ \dfrac{\partial^i}{\partial \theta^i} \ln f_\theta(\vec{x}) \quad i = 1, 2, 3$ $\quad$ Que existan las 3 derivadas

• $\left| \dfrac{\partial^3}{\partial \theta^3} \ln f_\theta(x) \right| < M(x)$ $\quad$ Que la tercera derivada esté acotada

• $E[M(x)] < k_\theta$ $\quad$ Que la esp. de la cota sea finita

H2: $\exists \theta_0 \in \Theta$ $\quad$ • $E\left[\dfrac{\partial}{\partial \theta} \ln f_{\theta_0}(\vec{x})\right] = 0$ $\quad$ La esperanza de la 1ª deriv se anula en un pto

• $E\left[\dfrac{1}{f_{\theta_0}(x)} \cdot \dfrac{\partial^2 f_{\theta_0}}{\partial \theta^2}(\vec{x})\right] = 0$

• $0 < I_1(\theta_0) < \infty$ $\quad$ Información de Fisher

→ Entonces

$$\sqrt{n}\,(\hat{\theta}_n - \theta_0) \xrightarrow{\mathcal{L}} N\left(0, \frac{1}{\sqrt{I(\theta)}}\right)$$

PROBLEMAS TEMA 5

1. Calcular un intervalo creíble de probabilidad $1-\alpha$ para la probabilidad de éxito θ en una distribución de Bernoulli. Suponer que la distribución inicial queda recogida por una distribución Beta (1,1) y que en 10 repeticiones del experimento se han observado 5 éxitos.

$$\left.\begin{array}{l} X \sim \text{Bernoulli}(\theta) \\ \theta \sim \text{Beta}(1,1) \\ n = 10 \\ \sum_{1}^{10} x_i = 5 \end{array}\right\}$$

$$\pi(\theta|\vec{x}) = \frac{f_\theta(x_1,\dots,x_n)\,\pi(\theta)}{\int_\theta f_\theta(x_1,\dots,x_n)\,\pi(\theta)\,d\theta}$$

$\pi(\theta|\vec{x}) \sim \text{Beta}(p_1, q_1)$ $\Longrightarrow$ $\pi(\theta|\vec{x}) \sim \text{Beta}(6,6)$

$$\begin{cases} p_1 = p_0 + \sum_{i=1}^{n} x_i = 1+5 = 6 \\ q_1 = q_0 + n - \sum_{i=1}^{n} x_i = 1+10-5 = 6 \end{cases}$$

$$E[\pi(\theta|\vec{x})] = \frac{p_1}{p_1+q_1} = \frac{6}{12} = \frac{1}{2}$$ Estadístico bayesiano

¿Fórmula?

$$V[\pi(\theta|\vec{x})] = \frac{p_1 q_1}{(p_1+q_1)^2(p_1+q_1+1)} = \frac{36}{12^2 \cdot 13}$$ P.F.E

$$P(a \le \theta \le b) = 1-\alpha$$

Beta(6,6) Probabilidad de colas iguales

$$\begin{cases} b = z_{\alpha/2} \\ a = z_{1-\alpha/2} \end{cases}$$

2. Calcular un intervalo creíble de probabilidad $1-\alpha$ para el parámetro θ de una distribución Poisson(θ) cuando la información inicial viene dada por la distribución Gamma(a,p).

$\pi(\theta|\vec{x}) = \text{Gamma}(a_1, p_1)$

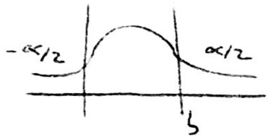
$$\begin{cases} a_1 = a + \sum_{i=1}^{n} x_i \\ p_1 = n + p \end{cases}$$

$E[\pi(\theta|\vec{x})] = \frac{a_1}{p_1}$ $\underline{\underline{E.B}}$

$V[\pi(\theta|\vec{x})] = \frac{a_1}{p_1^2}$ $\underline{\underline{P.F.E}}$

$Pr(a \leq \theta \leq b) = 1-\alpha$

$\quad\quad\quad\quad\quad \gamma$
gamma(a_1, p_1)



$P(\theta \leq b) = 1 - \alpha/2$

$P(\theta \leq a) = \alpha/2$

Sabemos que gamma$(1/2, n/2) = \chi^2_n$, luego tendremos que mirar en la tabla de χ^2:

$\theta \sim \text{gamma}(a_1, p_1) \equiv 2a_1\theta \sim \text{gamma}(\overset{1/2}{a_1}, p_1) \sim \chi^2_{2p_1}$:

$$\theta \xrightarrow{g} 2a_1\theta = y$$

$$\theta = \frac{y}{2a_1} \xleftarrow{g^{-1}}$$

$$\left(\frac{d\theta}{dy}\right) = \frac{1}{2a_1}$$

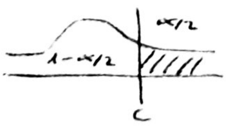
$$\Rightarrow f_Y(y) = f_X\left(\frac{y}{2a_1}\right) \cdot \frac{1}{2a_1} = \frac{\left(\frac{1}{2}\right)^{p_1}}{\Gamma(p_1)} e^{-\frac{1}{2}y} y^{p_1 - 1} \sim \text{Gamma}\left(\frac{1}{2}, p_1\right)$$

$\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad \gamma$
$\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad \chi^2_{2p_1}$

$\rightarrow P(2a_1 a \leq 2a_1\theta \leq \underbrace{2a_1 b}_{c}) = 1-\alpha$

$P(\chi^2_{2p_1} \leq c) = 1 - \alpha/2$

$c = \chi^2_{2p_1, \alpha/2}$



6. Para una m.a.s de tamaño n de una población con función de densidad

$$f(x|\theta) = \frac{5x^4}{\theta^5} I_{(0,\theta]}(x) \ , \ \theta > 0$$

a) Hallar el estimador de máxima verosimilitud T.

(No está acotado por el parámetro)

$$f_\theta(\vec{x}) = \frac{5^n \prod_1^n x_i^4}{\theta^{5n}} \prod_1^n I_{(0,\theta]}(x_i)$$

$0 \quad x_{(1)} \quad x_{(n)} \quad \theta$

$\theta \in (x_{(n)}, \infty)$

$\Longrightarrow \boxed{\hat{\theta} = X_{(n)}}$ es el EMV

b) Determinar la región de confianza de grado 0'95 para θ, de la forma $(\lambda T, \infty)$ para λ conveniente.

$P(\lambda X_{(n)} \leq \theta) = 1-\alpha \ , \ \forall \theta \in \Theta \ , \ X_{(n)} \leq \theta/\lambda$

$F_{X_{(n)}}(y) = (F(y))^n = \left(\frac{y}{\theta}\right)^{5n}$

$F(y) = \int_0^y \frac{5x^4}{\theta^5}\,dx = \left(\frac{y}{\theta}\right)^5$

$\underbrace{F_{X_{(n)}}(\theta/\lambda)}_{\left(\frac{\theta}{\lambda}\cdot\frac{1}{\theta}\right)^{5n}} = 1-\alpha \qquad \Longrightarrow \lambda = \left(\frac{1}{1-\alpha}\right)^{1/5n}$

$$\boxed{I.C(\theta) = \left(\frac{1}{(1-\alpha)^{5n}} X_{(n)} \ , \ \infty\right)}$$

$1-\alpha = 0'95$

c) Encuéntrese dos variables pivotales.

Se trata de una función de escala (ejercicio de la entrega), luego su distribución no depende del parámetro y por tanto,

3. A una persona se le pasa un test de inteligencia, cuyo resultado X se supone que sigue una distribución Normal$(\theta, \sigma=10)$, donde θ es su nivel de inteligencia real. Supongamos que en el colectivo al que pertenece la persona, la inteligencia θ se distribuye según Normal$(100, \sigma_0=15)$. Determinar un intervalo creíble de probabilidad 0'95 para su nivel de inteligencia cuando el resultado del test ha sido 110.

$$\underbrace{P(a \le \theta \le b)}_{N(\mu_1, \sigma_1^2)} = 1-\alpha = 0'95$$

$$\begin{cases} X \sim N(\theta, 10) \\ \theta \sim N(100, 15) \\ \overline{X} = 110 \end{cases}$$

¿$n=1$?

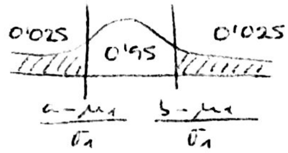
$$\mu_1 = \frac{\frac{\mu_0}{\sigma_0^2} + \overline{X}\frac{n}{\sigma^2}}{\frac{1}{\sigma_0^2} + \frac{n}{\sigma^2}} = 106'923$$

$$\sigma_1^2 = \frac{1}{\frac{1}{\sigma_0^2} + \frac{n}{\sigma^2}} \longrightarrow \sigma_1 = \sqrt{69'23}$$

$$P\left(\frac{a-\mu_1}{\sigma_1} \le \underbrace{\frac{\theta-\mu_1}{\sigma_1}}_{N(0,1)} \le \frac{b-\mu_1}{\sigma_1}\right) = 1-\alpha = 0'95$$

$$P\left(N(0,1) \le \frac{b-\mu_1}{\sigma_1}\right) = 1-\alpha/2 \Rightarrow \frac{b-\mu_1}{\sigma_1} = 1'96$$

$$P\left(N(0,1) \ge \frac{a-\mu_1}{\sigma_1}\right) = \frac{\alpha}{2} \Rightarrow \frac{a-\mu_1}{\sigma_1} = -1'96$$



$$P\left(-1'96 \le \frac{\theta-\mu_1}{\sigma_1} \le 1'96\right) = 0'95$$

$$-1'96\,\sigma_1 + \mu_1 \le \theta \le 1'96\,\sigma_1 + \mu_1$$

$$\Rightarrow \boxed{IC_{95\%} = (90'9149, 123'837)}$$

Hay que comprobar que la distribución no depende del parámetro:

$$F_{X_{(n)}}(y) = \left(\frac{y}{\theta}\right)^{5n}$$

$$f_{X_{(n)}}(y) = \frac{5}{\theta^5} y^4$$

$$y \xrightarrow{g} y/\theta = z$$

$$y = z\theta \xleftarrow{g^{-1}} z$$

$$\Rightarrow f_z(z) = f_y(z\theta)\theta = \frac{5}{\theta^5}(z\theta)^4\theta = 5z^4$$

$z \in [0,1]$
No depende del parámetro

$\Rightarrow \frac{X_{(n)}}{\theta}$ no depende del parámetro θ. Como función de θ es estrictamente monótona $\Rightarrow$ Es cantidad pivotal

$\rightarrow$ ¿Cómo construiríamos el intervalo de confianza?

$$P\Big(a \leq \underbrace{\frac{X_{(n)}}{\theta}}_{\text{Tenemos su distribución}} \leq b\Big) = 1-\alpha$$

Despejamos a y b:

$$a \leq \frac{X_{(n)}}{\theta} \Rightarrow \theta \leq \frac{X_{(n)}}{a}$$

$$\Rightarrow \boxed{I.C_{(1-\alpha)100\%}(\theta) = \left(\frac{X_{(n)}}{b}, \frac{X_{(n)}}{a}\right)}$$

$\rightarrow$ La otra cantidad pivotal:

$\boxed{T = -2\sum_{1}^{n} \ln F_\theta(x_i)}$ será cantidad pivotal si $F_\theta(x)$ es continua en θ y estrictamente monótona (Habría que demostrarlo)

$\rightarrow$ Veamos el I.C:

$$F_\theta(x) = \left(\frac{x}{\theta}\right)^5$$

$$P\left(a \leq -2\sum_{1}^{n} \ln\left(\frac{x_i}{\theta}\right)^5 \leq b\right) = 1-\alpha$$

$$\sim \chi^2_{2n}$$

$$b = \chi^2_{2n,\alpha/2}$$

$$a = \chi^2_{2n,1-\alpha/2}$$

$$\Rightarrow \chi^2_{2n,1-\alpha/2} \leq -2\sum_{i=1}^{n} \ln\left(\frac{x_i}{\theta}\right)^5 \leq \chi^2_{2n,\alpha/2}$$

d) Constrúyase un intervalo de confianza mediante el método de Neyman.

$$P(\gamma_1 \leq X_{(n)} \leq \gamma_2) = 1-\alpha$$
$$\shortparallel$$
$$F_{X_{(n)}}(\gamma_2) - F_{X_{(n)}}(\gamma_1) = 1-\alpha = 1-\alpha_1-\alpha_2$$

$$F_{X_{(n)}}(\gamma_2) = 1-\alpha_1 = \left(\frac{\gamma_2}{\theta}\right)^{5n}$$

$$F_{X_{(n)}}(\gamma_1) = \alpha_2 = \left(\frac{\gamma_1}{\theta}\right)^{5n}$$

Ahora despejamos θ:

$$\theta_1 = \frac{\gamma_2}{(1-\alpha_1)^{1/5n}}$$

$$\theta_2 = \frac{\gamma_1}{\alpha_2^{1/5n}}$$

Veamos que son estrictamente monótonas:

$$X_{(n)} = \gamma_2 = \theta_1 (1-\alpha_1)^{1/5n}$$

$$X_{(n)} = \gamma_1 = \theta_2 \alpha_2^{1/5n} \leftarrow \text{Por el dibujo de } t^{a}$$

Y este va a ser el intervalo γ_1, γ_2.

7. Para una m.a.s de tamaño n de una población con función de densidad
$$f(x|\theta) = (\theta+1)x^{\theta} I_{(0,1)}(x), \quad \theta > 0$$

Construir una región de confianza de grado $1-\alpha$ basada en la variable pivotal $T(\vec{x},\theta) = -\sum_{i=1}^{n}(\theta+1)\ln x_i$ tomando probabilidad de colas iguales.

Veamos cómo se distribuye:

$$x_i \xrightarrow{g} -(\theta+1)\ln x_i = y_i$$
$$x_i = e^{-\frac{y_i}{\theta+1}} \xleftarrow[g^{-1}]{} y_i$$

$$f_\theta(x_i) = f_\theta\left(e^{-\frac{y_i}{\theta+1}}\right) = (\theta+1)e^{\frac{-y_i\theta}{\theta+1}}$$

$$\Rightarrow f_Y(y) = e^{-\frac{y_i}{\theta+1}} \left|\frac{1}{\theta+1}\right| (\theta+1) e^{\frac{-y_i\theta}{\theta+1}} = e^{\frac{-y_i}{\theta+1}} e^{\frac{-y_i\theta}{\theta+1}} = e^{\frac{-y_i-y_i\theta}{\theta+1}} = e^{\frac{-y_i(1+\theta)}{\theta+1}} = e^{-y_i}$$

independiente

$$y_i = \frac{z_i}{2} \xleftarrow[g^{-1}]{} z_i$$

$$\Rightarrow f_Z(z_i) = f_Y\left(\frac{z_i}{2}\right) \cdot \frac{1}{2} = \frac{1}{2}e^{-\frac{z_i}{2}} \sim \text{Gamma}\left(\frac{1}{2}, 1\right)$$

$$\Longrightarrow \sum_{i=1}^{n} t_i \sim \text{gamma}\left(\tfrac{1}{2}, n\right) \underset{\wr}{} \; \chi^2_{2n}$$

$$\sum_{1}^{n} t_i = \sum_{1}^{n} 2y_i = -2(\theta+1)\sum_{1}^{n} \ln x_i \sim \chi^2_{2n}$$

Probabilidad de colas iguales:

$$P(\gamma_1 \le \underset{\chi^2_{2n}}{\underset{\wr}{2T}} \le \gamma_2) = 1-\alpha$$

$$P(\chi^2_{2n} \le \gamma_2) = \alpha/2 \qquad \longrightarrow \qquad \gamma_2 = \chi^2_{2n,\alpha/2}$$

$$P(\chi^2_{2n} \ge \gamma_1) = 1-\alpha/2 \qquad\qquad \gamma_1 = \chi^2_{2n,1-\alpha/2}$$

Sustituimos:

$$\left(\chi^2_{2n,1-\alpha/2} \le -2(\theta+1)\sum_{i=1}^{n} \ln x_i \le \chi^2_{2n,\alpha/2}\right) \iff 1 + \frac{\chi^2_{2n,1-\alpha/2}}{2\sum_{i=1}^{n}\ln x_i} \le -\theta \le 1 + \frac{\chi^2_{2n,\alpha/2}}{2\sum_{i=1}^{n}\ln x_i}$$

$$\Longrightarrow \boxed{\underset{(1-\alpha)\,100\%}{\text{I. Confianza}} = \left(-1-\frac{\chi^2_{2n,\alpha/2}}{2\sum_{i=1}^{n}\ln x_i},\; -1-\frac{\chi^2_{2n,\alpha/2}}{2\sum_{i=1}^{n}\ln x_i}\right)}$$

¿Qué ocurriría si me diera un punto, es decir, $1-\alpha = 0'95$? ¿Cómo lo ponemos?

9. Para una m.a.s de tamaño n de una población con función de densidad

$$f(x|\theta) = e^{-(x-\theta)} I_{[\theta,\infty)}(x), \quad \theta \in \mathbb{R}$$

Determinar un intervalo de confianza de grado $1-\alpha$ mediante el método de la variable pivotal con probabilidad de colas iguales.

Hay tres maneras de hallar un I.C
- → Método de la variable pivotal (como antes porque es de localización)
- → Cogiendo un estimador suficiente
- → Método de Neyman

$$f_\theta(\vec{x}) = e^{-n\bar{x}+n\theta} I_{(-\infty, x_{(1)})}(\theta)$$

$X_{(1)}$ es suficiente

$X_{(1)}$ es el EMV

Hacemos el cambio:

$$y \xrightarrow{g} z = y - \theta$$

$$y = z + \theta \xleftarrow[g^{-1}]{} z$$

$$\left|\frac{dy}{dz}\right| = |1|$$

$$\Longrightarrow f_z(z) = f_y(z+\theta)\cdot 1 = n e^{(-z-\theta+\theta)n} \cdot 1 = n e^{-nz}$$

De modo que nuestro $\underline{\underline{T = X_{(1)} - \theta}}$ porque la distribución no depende de θ

$$P(\gamma_1 \le T \le \gamma_2) = 1 - \alpha$$

$$\phi(\gamma_2) - \phi(\gamma_1) = 1 - \alpha = 1 - \alpha_1 - \alpha_2$$

ϕ distribución T

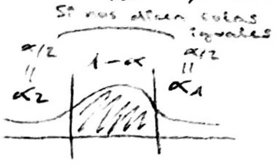
función de densidad de T

$$\phi(\gamma_2) = 1 - \alpha/2$$

$$\Longrightarrow 1 - \alpha/2 = \phi(\gamma_2) = \int_0^{\gamma_2} -n e^{-nz}\,dz = e^{-n\gamma_2} + 1$$

¿+?

(Por el teorema del cambio de variable)



$$\alpha/2 = \phi(\gamma_1) = \int_0^{\gamma_1} -n e^{-nz}\,dz = e^{-n\gamma_1} + 1$$

$$\ln(1-\alpha/2) = -n\gamma_2 \Longrightarrow \gamma_2 = -\frac{1}{n}\ln(1-\alpha/2)$$

$$\ln(\alpha/2) = -n\gamma_1 \Longrightarrow \gamma_1 = -\frac{1}{n}\ln(\alpha/2)$$

$$I.C = (\gamma_1, \gamma_2) = \boxed{\left(-\frac{1}{n}\ln(\alpha/2), -\frac{1}{n}\ln(1-\alpha/2)\right)}$$

⑩ H5

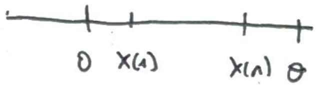
16-05-2017

Para una m.a.s de tamaño n de una población con función de densidad $f(x|\theta) = 3x^2\theta^{-3} I_{(0,\theta)}(x)$

Determinar un intervalo de confianza de grado $1-\alpha$ mediante el método de Neyman con probabilidad de colas iguales

$$f_\theta(\vec{x}) = \underbrace{3^n \prod_{i=1}^{n} x_i^2 \cdot \theta^{-3n}}\ \underbrace{\prod_{i=1}^{n} I_{(0,\theta)}\, x_i}$$

Como función de θ es máximo cuando θ es lo más pequeño posible, por tanto CMV $\hat{\theta} = X_{(n)}$



$\theta \in [X_{(n)}, \infty]$

$$P(\gamma_2 \leq X_{(n)} \leq \gamma_1) = 1-\alpha$$

$$1-\alpha = P(\gamma_2 \leq X_{(n)} \leq \gamma_1) = \underbrace{P(X_{(n)} \leq \gamma_1)}_{F_{X_{(n)}}(\gamma_1)} - \overbrace{P(X_{(n)} \leq \gamma_2)}^{F_{X_{(n)}}(\gamma_2)}$$

$$P(X_{(n)} \leq \gamma_1) = F_{X_{(n)}}(\gamma_1) = 1 - \alpha/2 \quad ⊛ = \left(\frac{\gamma_1}{\theta}\right)^{3n} \Rightarrow \theta\left(1 - \frac{\alpha/2}{2}\right)^{1/3n} = \gamma_1 \quad ⊛$$

$$P(X_{(n)} \leq \gamma_2) = F_{X_{(n)}}(\gamma_2) = \alpha/2 \quad ⊛ = \left(\frac{\gamma_2}{\theta}\right)^{3n} \Rightarrow \theta\left(\frac{\alpha/2}{2}\right)^{1/3n} = \gamma_2$$

$$f_\theta(x) = \frac{3x^2}{\theta^3} I_{(0,\theta)}(x)$$

$$F_\theta(x) = \int_0^x \frac{3x^2}{\theta^3}\,dx = \frac{3x^3}{3\theta^3}\Big]_0^x = \left(\frac{x}{\theta}\right)^3 \quad \text{si } 0 < x < \theta$$

monótonas)

$$\gamma_1 = \theta\left(1-\frac{\alpha}{2}\right)^{1/3n} = X_{(n)} \Rightarrow \theta_1 = \frac{X_{(n)}}{(1-\alpha/2)^{1/3n}}$$

$$\gamma_2 = \theta\,(\alpha/2)^{1/3n} = X_{(n)} \Rightarrow \theta_2 = \frac{X_{(n)}}{(\alpha/2)^{1/3n}}$$

$$\underset{(1-\alpha/100\%)}{I.C.\ (\theta)} = \left(X_{(n)}\left(1-\frac{\alpha}{2}\right)^{-1/3n},\ X_{(n)}\,(\alpha/2)^{-1/3n}\right)$$

→ ¿Para amplitud mínima?

Hoja 5

(4) Se estudia la cantidad de lluvia caída en dos regiones diferentes A y B a partir de m.a.s. de ambas poblaciones, que se suponen normales e independientes. Si se ha obtenido $\bar{x}_1 = 93,43$, $s_1^2 = 420,06$, $\bar{x}_2 = 85,24$, $s_2^2 = 421,414$

$N(\mu_1, \sigma_1^2)$ A $\bar{X}_1 = 93,43$ $\quad s_1^2 = 420.009$ $\quad \boxed{n = 10}$ ← no está en el enunciado

$N(\mu_1, \sigma_2^2)$ B $\bar{X}_2 = 85,24$ $\quad s_2^2 = 421,414$

(a) $Ic(\sigma_1^2) = 95\%$ → $1-\alpha = 0,95$; $\alpha = 0,05$; $\alpha/2 = 0,025$

$(n-1)\dfrac{s^2}{\sigma^2} \sim \chi^2_{n-1}$ TEOREMA DE FISHER ¡¡SABER!!

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

E distribución χ^2_{n-1}

$F(b) = 1 - \alpha/2 = 0{,}975 \implies b = \chi^2_{n-1,\,\alpha/2} = 19{,}02$

$F(a) = \alpha/2 = 0{,}025 \implies a = \chi^2_{n-1,\,\alpha/2} = 2{,}7$

Por tanto

$$a \leq (n-1)\frac{s^2}{\sigma^2} \leq b$$

$$\frac{(n-1)s^2}{b} \leq \sigma^2 \leq \frac{(n-1)s^2}{a}$$

$IC(\sigma) = (198{,}48 \,;\, 1400{,}23)$

$s^2,\ a,\ b$ $\qquad s^2 = 420{,}269$

$$\frac{1}{(n-1)} \sum_{1}^{n} (x_i - \bar{x})^2$$

(b) Suponiendo que las dos varianzas poblaciones son desconocidas pero iguales, determinar un intervalo de confianza para grado 0,95 para la diferencia de medias poblacionales

$\Delta\mu = \mu_2 - \mu_1$

$$\frac{(\bar{x} - \bar{y}) - (\mu_2 - \mu_1)}{\sqrt{\left(\frac{1}{n} + \frac{1}{m}\right)\underbrace{\left(\frac{1}{n+m-2}(n-1)s_1^2 + (m-1)s_2^2\right)}_{sp}}} \sim t_{n+m-2}$$

t simétrica, colas iguales, $n = m = 10$

$P(a \leq t_{n+m-2} \leq b) = 1 - \alpha$

$\begin{cases} \phi(b) - \phi(a) = 1 - \alpha & \phi \text{ dist. } t_{n+m-2} \\ \phi(b) = 1 - \frac{\alpha}{2} = 0{,}975 \longrightarrow b = 2'101 \\ \phi(a) = \frac{\alpha}{2} = 0'025 \longrightarrow a = -2'101 \end{cases}$

$$\implies IC_{95\%}(\Delta\mu) = \left(\bar{x} - \bar{y} - t_{n+m-2,\alpha/2}\, sp\sqrt{\frac{1}{n} + \frac{1}{m}} \,;\, \bar{x} - \bar{y} + t_{n+m-2,\alpha/2}\, sp\sqrt{\frac{1}{n} + \frac{1}{m}}\right)$$