TEMA 5

ESTIMACIÓN POR INTERVALOS

5.1. *MÉTODO DE ESTIMACIÓN DE MÁXIMA VEROSIMILITUD*

- MV y MM construyen buenos estimadores θ^*
- Pero $\theta^*(X)$ puede tomar un valor erróneo
- La estimación puede mejorar acompañando de un intervalo donde confiamos se encuentre θ

  $\theta \in [\underline{\theta}, \overline{\theta}]$ ó IC_θ

- Ejemplo: m.a.s. n=100, B(1,θ)

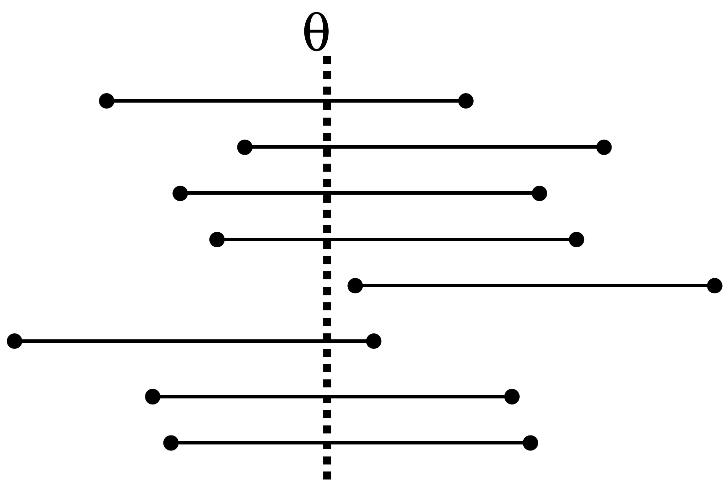
  $\theta^* = \overline{x}$

  Posibles medias muestrales: 0.63, 0.42...

- Nivel de confianza $\gamma=1-\alpha$

  $P[\underline{\theta}(X) \leq \theta \leq \overline{\theta}(X)] = \gamma = 1-\alpha$

  $P[T_1(X) \leq \theta \leq T_2(X)] = \gamma = 1-\alpha$



  (1-α)% de los IC contiene el parámetro

  α = nivel de significación

- Interpretación:
  - SI: P(intervalo contenga a θ)
  - NO: P(θ pertenezca al intervalo)

  Porque:
  - θ siempre desconocida
  - las variables aleatorias son $\underline{\theta}(X)$ y $\overline{\theta}(X)$, no θ

5.2. *MÉTODOS DE CONSTRUCCIÓN DE INTERVALOS DE CONFIANZA*

- Determinación de $\underline{\theta}(\mathrm{X})$ y $\overline{\theta}(\mathrm{X})$

MÉTODO DE LA CANTIDAD PIVOTAL

- Definición:
  - Si X viene de m.a.s. de $\xi \rightarrow f(x,\theta)$, $\theta \in \Theta$
  - Entonces $T(\mathrm{X},\theta)$ es:
    - Cantidad pivotal o pivote
    - Función de probabilidad que no dependa de θ
    - θ^*
- Teorema
  - Si la cantidad pivotal es una función monótona de θ, es posible determinar IC_θ.
- Si se puede elegir $k_1(\alpha)$ y $k_2(\alpha)$ de $T(\mathrm{X},\theta)$ tal que:

  $P\,[k_1(\alpha) \leq T(\mathrm{X},\theta) \leq k_2(\alpha)] = 1-\alpha$

  Entonces:

  $P[\underline{\theta}(\mathrm{X}) \leq \theta \leq \overline{\theta}(\mathrm{X})] = \gamma = 1-\alpha$

MÉTODO GENERAL

- Menos restrictivo: la distribución de $T(\mathrm{X})$ puede depender de θ = $T(\mathrm{X},\theta)$
- Ejemplo:
  - $\overline{x} \rightarrow N(E(\xi), \sqrt{\frac{V(\xi)}{n}})$
  - $T(\mathrm{X},\theta) \rightarrow \dfrac{\overline{x}-\mu}{\sqrt{\frac{V(\xi)}{n}}} \rightarrow N(0,1)$
- Procedimiento:
  - Población: $\xi \rightarrow f(x,\theta)$
  - $\theta^* \rightarrow g(\theta^*,\theta)$
  - $P\,[k_1(\alpha,\theta) \leq \theta^* \leq k_2(\alpha,\theta)] = 1-\alpha$

    $$\int_{k_1}^{k_2} g(\theta^*,\theta)\, d\theta^* = 1-\alpha$$

$$\left.\begin{array}{l}\int_{-\infty}^{k_1} g(\theta^*,\theta)d\theta^* = \alpha_1 \\ \int_{k_1}^{\infty} g(\theta^*,\theta)d\theta^* = \alpha_2\end{array}\right\}\alpha_1+\alpha_2=\alpha$$

- Nota: con distribuciones discretas
  - P $[k_1(\alpha,\theta) \le \theta^* \le k_2(\alpha,\theta)] \ge 1-\alpha$

INTERVALOS DE LONGITUD MÍNIMA

- min $\overline{\theta}(X)-\underline{\theta}(X)=D$
  - Método de multiplicadores de Lagrange
  - Convenio $\alpha_1=\alpha_2=\frac{\alpha}{2}$

5.3. *INTERVALOS DE CONFIANZA EN POBLACIONES NORMALES*

- Población: $\xi \to N(\mu,\sigma)$

INTERVALO DE CONFIANZA PARA LA MEDIA, CON VARIANZA CONOCIDA

- $\overline{x} \to N(\theta, \sqrt{\frac{\sigma^2}{n}})$
- $T(X,\theta) \to \frac{\overline{x}-\theta}{\sqrt{\frac{\sigma^2}{n}}} \to N(0,1) \Rightarrow$ DESPEJAR θ
- P $[k_1(\alpha,\theta) \le \theta^* \le k_2(\alpha,\theta)]=1-\alpha$
- P $[k_1(\alpha,\theta) \le g(\theta^*,\theta) \le k_2(\alpha,\theta)]=1-\alpha$
- P $[k_1 \le N(0,1) \le k_2]=1-\alpha$
- P $[k_1 \le \frac{\overline{x}-\theta}{\sqrt{\frac{\sigma^2}{n}}} \le k_2]=1-\alpha$
- P $[k_1\sqrt{\frac{\sigma^2}{n}} \le \overline{x}-\theta \le k_2\sqrt{\frac{\sigma^2}{n}}]=1-\alpha$
- P $[-\overline{x}+k_1\sqrt{\frac{\sigma^2}{n}} \le -\theta \le -\overline{x}+k_2\sqrt{\frac{\sigma^2}{n}}]=1-\alpha$
- P $[\overline{x}-k_2\sqrt{\frac{\sigma^2}{n}} \le \theta \le \overline{x}-k_1\sqrt{\frac{\sigma^2}{n}}]=1-\alpha$
- $k=k_2=-k_1$
- P $[\overline{x}-k\sqrt{\frac{\sigma^2}{n}} \le \theta \le \overline{x}+k\sqrt{\frac{\sigma^2}{n}}]=1-\alpha$

- $P[\theta^*-k\,V(\theta^*)\leq\theta\leq\theta^*+k\,V(\theta^*)]=1-\alpha$
- $k\rightarrow N(0,1)$; $k=z_{\alpha/2}$
- $\theta\in[\bar{x}-k\sqrt{\frac{\sigma^2}{n}},\bar{x}+k\sqrt{\frac{\sigma^2}{n}}]$
- $\theta\in[\bar{x}\pm k\sqrt{\frac{\sigma^2}{n}}]$
- $\theta\in[\bar{x}\pm\text{error}]$
- $\theta\in[\theta^*\pm k\,V(\theta^*)]$
- $\theta\in[\theta^*-k\,V(\theta^*),\theta^*+k\,V(\theta^*)]$
- Ejemplo:
  - Población: $\xi\rightarrow N(\mu,6)$
  - Muestra: m.a.s $n=100$, $\bar{x}=14.35$
  - Estadístico: $\mu^*=\bar{x}\rightarrow N(\theta,\sqrt{\frac{6^2}{100}})=N(\theta,0.6)$
  - Confianza: $\gamma=0.99\Rightarrow\alpha=0.01\Rightarrow z_{0.005}=2.576$
  - Intervalo: $\theta\in[14.35\pm2.576(0.6)]\Rightarrow$

    $\theta\in[14.35\pm1.55]\Rightarrow\theta\in[12.80,15.90]$

    Si $\bar{x}=15.83$, $\theta\in[15.83\pm1.55]\Rightarrow\theta\in[14.28,17.38]$

INTERVALO DE CONFIANZA PARA LA MEDIA, CON VARIANZA DESCONOCIDA

- $T(X,\theta)\rightarrow\frac{\bar{x}-\theta}{\sqrt{\frac{s^2}{n-1}}}\rightarrow t_{n-1}\Rightarrow$ DESPEJAR θ
- $P[k_1(\alpha,\theta)\leq\theta^*\leq k_2(\alpha,\theta)]=1-\alpha$
- $P[k_1(\alpha,\theta)\leq g(\theta^*,\theta)\leq k_2(\alpha,\theta)]=1-\alpha$
- $P[k_1\leq\frac{\bar{x}-\theta}{\sqrt{\frac{s^2}{n-1}}}\leq k_2]=1-\alpha$
- Si $-k_1=k_2=t_{n-1,\alpha/2}$
  - $k=t_{n-1,\alpha/2}$
  - $\theta^*=\bar{x}$
  - $V(\theta^*)=\frac{s^2}{n-1}$
  - Error$=t_{n-1,\alpha/2}\sqrt{\frac{s^2}{n-1}}$

- Ejemplo:
  - Población: $\xi \rightarrow N(\mu,\sigma)$
  - Muestra: m.a.s n=9, $\overline{x}$=166, s=2.4037
  - Estadístico: $\mu^*=\overline{x}$ porque $T(X,\theta) \rightarrow \frac{\overline{x}-\mu}{\sqrt{\frac{s^2}{n-1}}} \rightarrow t_{n-1}$
  - Confianza: $\gamma=0.99 \Rightarrow \alpha=0.01 \Rightarrow t_{9-1,0.005}=2.896$
  - Intervalo: $\theta \in [166 \pm 2.896(\frac{2.4037}{\sqrt{9-1}})] \Rightarrow$

    $\theta \in [166 \pm 2.4611] \Rightarrow \theta \in [163.54,168.46]$

INTERVALO DE CONFIANZA PARA LA VARIANZA, DADA POBLACIÓN NORMAL

- $T(X,\theta) \rightarrow \frac{ns^2}{\sigma^2} \rightarrow \chi^2_{n-1} \Rightarrow$ DESPEJAR σ^2
- $P[k_1 \leq \frac{ns^2}{\sigma^2} \leq k_2]=1-\alpha$
- $P[\frac{ns^2}{k_2} \leq \sigma^2 \leq \frac{ns^2}{k_1}]=1-\alpha$
- $P[\frac{ns^2}{\chi^2_{n-1,\alpha/2}} \leq \sigma^2 \leq \frac{ns^2}{\chi^2_{n-1,1-\alpha/2}}]=1-\alpha$
- Ejemplo:
  - Población: $\xi \rightarrow N(\mu,\sigma)$
  - Muestra: m.a.s n=7, $\overline{x}$=2.72, s^2=0.06
  - Estadístico: $T(X,\theta) \rightarrow \frac{ns^2}{\sigma^2} \rightarrow \chi^2_{n-1}$
  - Confianza: $\gamma=0.98 \Rightarrow \alpha=0.02 \Rightarrow \chi^2_{7-1,0.01}=16.812$, $\chi^2_{7-1,0.99}=0.872$
  - Intervalo: $\sigma^2 \in [\frac{7*0.06}{16.812},\frac{7*0.06}{0.872}] \Rightarrow$

    $\theta \in [0.025,0.481]$

5.4. *INTERVALOS DE CONFIANZA EN POBLACIONES NO NORMALES*

- Población: $\xi \rightarrow$NO Normal, pero con $E(\xi)=\mu$ y $\sqrt{V(\xi)}=\sigma$ conocidas

CASO GENERAL PARA MEDIAS

- Solución: Desigualdad de Chebichev da una cota
- $P[|\xi-E(\xi)| \leq k\sqrt{V(\xi)}]=1-\frac{1}{k^2}=1-\alpha$

- Como $\theta^*=\bar{x}$, distribución de $\bar{x}$ en vez de ξ:
    - P $[|\bar{x}-E(\bar{x})|\leq k\sqrt{V(\bar{x})}]=1-\frac{1}{k^2}=1-\alpha$
- Como $E(\bar{x})=E(\xi)$ y $\sqrt{V(\bar{x})}=\sqrt{\frac{V(\xi)}{n}}$
    - P $[|\bar{x}-E(\xi)|\leq k\sqrt{\frac{V(\xi)}{n}}]=1-\frac{1}{k^2}=1-\alpha$
    - P $[|\bar{x}-\mu|\leq k\frac{\sigma}{\sqrt{n}}]=1-\frac{1}{k^2}=1-\alpha$
- Entonces:
    - P $[-k\frac{\sigma}{\sqrt{n}}\leq\bar{x}-\mu\leq k\frac{\sigma}{\sqrt{n}}]=1-\frac{1}{k^2}=1-\alpha$
    - P $[\bar{x}-k\frac{\sigma}{\sqrt{n}}\leq\mu\leq\bar{x}+k\frac{\sigma}{\sqrt{n}}]=1-\frac{1}{k^2}=1-\alpha$
- $k=\sqrt{\frac{1}{\alpha}}$
- Ejemplo:
    - Población: $\xi\rightarrow\mu=?,\sigma=4$
    - Muestra: m.a.s $n=10$, $\bar{x}=1006$
    - Estadístico: $\bar{x}$
    - Confianza: $\gamma=0.95\Rightarrow\alpha=0.05\Rightarrow k=\sqrt{\frac{1}{0.05}}=4.47$
    - Intervalo: $\mu\in[1006\pm4.47\frac{4}{\sqrt{10}}]\Rightarrow\mu\in[1006\pm5.66]$

      $\mu\in[1000.34,1011.66]$
    - Nota: Si $\xi\rightarrow N$ $\mu\in[1006\pm1.96\frac{4}{\sqrt{10}}]\Rightarrow\mu\in[1003.52,1008.48]$

INTERVALOS DE CONFIANZA BASADOS EN GRANDES MUESTRAS

- INTERVALOS PARA UN PARÁMETRO θ
    - $\theta^*_{MV}\rightarrow N(\theta,\sqrt{V(\theta^*)})$

      Asintóticamente normal, insesgado y eficiente
    - $T(X,\theta)\rightarrow\frac{\theta^*-\theta}{\sqrt{V(\theta^*)}}\rightarrow N(0,1)\Rightarrow$ DESPEJAR θ
    - P $[\theta^*-k\,V(\theta^*)\leq\theta\leq\theta^*+k\,V(\theta^*)]=1-\alpha$, para $k\rightarrow N(0,1)$; $k=z_{\alpha/2}$

- INTERVALOS PARA UN θ=p de B(1,p)
  - $p^*_{MV}=\bar{x}$
    - $E(p^*)=E(\xi)=p$
    - $E(p^*)=E(\xi)/n=p(1-p)/n$
  - Ejemplo:
    - $\bar{x}=0.20$, n=100, $\alpha=0.05$
    - $p\in [0.20 \pm 1.96\sqrt{\frac{0.20*0.80}{100}}] \Rightarrow$

$p\in [0.20 \pm 0.0784] \Rightarrow p\in [0.1216, 0.2784]$

EXTRA: COMPARACIÓN ENTRE MUESTRAS

INTERVALOS DE CONFIANZA DE LA DIFERENCIA DE MEDIAS CON VARIANZAS CONOCIDAS

- $\xi_1 \rightarrow Normal(\mu_1,\sigma_1)$ ➔m.a.s n_1➔ $\bar{x}, s_1^2$
- $\xi_2 \rightarrow Normal(\mu_2,\sigma_2)$ ➔m.a.s n_2➔ $\bar{y}, s_2^2$
- $\bar{x} \rightarrow N(\mu_1, \frac{\sigma_1}{\sqrt{n_1}})$
- $\bar{y} \rightarrow N(\mu_2, \frac{\sigma_2}{\sqrt{n_2}})$
- $\bar{x}-\bar{y} \rightarrow N(\mu_1-\mu_2, \sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}})$
- $IC_{\mu_1-\mu_2} \in (\bar{x}-\bar{y}) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$

INTERVALOS DE CONFIANZA DE LA DIFERENCIA DE MEDIAS CON VARIANZAS DESCONOCIDAS

- $\xi_1 \rightarrow Normal(\mu_1,\sigma_1)$ ➔m.a.s n_1➔ $\bar{x}, s_1^2$
- $\xi_2 \rightarrow Normal(\mu_2,\sigma_2)$ ➔m.a.s n_2➔ $\bar{y}, s_2^2$
- $\frac{(\bar{x}-\bar{y})-(\mu_1-\mu_2)}{\sqrt{\frac{n_1 s_1^2+n_2 s_2^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \longrightarrow t_{n1+n2-2}$
- $IC_{\mu_1-\mu_2} \in (\bar{x}-\bar{y}) + t_{\alpha/2,n1+n2-2}\sqrt{\frac{n_1 s_1^2+n_2 s_2^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$

INTERVALOS DE CONFIANZA DE LA DIFERENCIA DE PROPORCIONES

- $\xi_1 \to B(1,\pi_1)$ ➔m.a.s n_1➔p_1
- $\xi_2 \to B(1,\pi_2)$ ➔m.a.s n_1➔p_2
- $IC\pi_1-\pi_2 \in (p_1-p_2)+z_{\alpha/2}\sqrt{\frac{p_1(1-p_1)}{n_1}+\frac{p_2(1-p_2)}{n_2}}$

INTERVALOS DE CONFIANZA DEL COCIENTE DE VARIANZAS DE DISTRIBUCIONES NORMALES

- $\xi_1 \to Normal(\mu_x,\sigma_x)$ ➔m.a.s n_1➔ $\bar{x}$, s_x^2
- $\xi_2 \to Normal(\mu_y,\sigma_y)$ ➔m.a.s n_2➔ $\bar{y}$, s_y^2
- $\frac{\sigma_y^2}{\sigma_x^2}\frac{s_{1x}^2}{s_{1y}^2} \longrightarrow F_{n1-1,n2-1}$
- $IC\sigma_x^2/\sigma_y^2 \in \left[\frac{1}{F_{n1-1,n2-1,\alpha/2}}\frac{s_{1x}^2}{s_{1y}^2}, \frac{1}{F_{n1-1,n2-1,1-\alpha/2}}\frac{s_{1x}^2}{s_{1y}^2}\right]$
- Ejemplo:
  - $n_1=21$, $s_{1x}^2=10$
  - $n_2=21$, $s_{1y}^2=20$
  - $\alpha=0.05$
  - $F_{n1-1,n2-1,\alpha/2} = F_{20,20,0.025}$ =2.46
  - $F_{n1-1,n2-1,\alpha/2} = F_{20,20,0.975}$ =1/2.46=0.41
  - Como $s_{1y}^2 > s_{1x}^2$, s_{1y}^2 en el numerador
    - $IC\sigma_x^2/\sigma_y^2 \in \left[\frac{1}{2.46}\frac{20}{10}, \frac{1}{0.41}\frac{20}{10}\right]$
    - $IC\sigma_x^2/\sigma_y^2 \in [0.81, 4.93]$

5.5. *DETERMINACIÓN DEL TAMAÑO DE LA MUESTRA*

- Se fija el error máximo e

  $e=k\sqrt{V(\theta^*)}$

INTERVALOS DE CONFIANZA DE LA MEDIA CON VARIANZA CONOCIDA

- $\xi \to Normal(\mu,\sigma)$ ➔m.a.s n➔ $\bar{x}$, s^2
- $\bar{x} \to N(\mu, \frac{\sigma}{\sqrt{n}})$
- $IC\mu \in \bar{x}$ +error = $\bar{x} + z_{\alpha/2}\sqrt{\frac{\sigma^2}{n}}$

- error = $z_{\alpha/2}\sqrt{\dfrac{\sigma^2}{n}}$
- $n=\dfrac{z^2\sigma^2}{e^2}$
- Ejemplo: Error=e=0.20, N(θ,2), γ=0.95
  - $n=\dfrac{1.96^2 * 2^2}{0.2^2} \geq 384.16$

INTERVALOS DE CONFIANZA DE LA MEDIA CON VARIANZA DESCONOCIDA

- $n=\dfrac{{t_{n-1}}^2\sigma^2}{e^2}+1$

INTERVALOS DE CONFIANZA DE LA PROPORCIÓN

- $n=\dfrac{z^2 p^*\left(1-p^*\right)}{e^2}$
- Ejemplo: Error=e=0.10
  - $n=\dfrac{1.96^2 * 0.5 * (1-0.5)}{0.1^2} \geq 26.04$
  - $n=\dfrac{3^2 * 0.5 * (1-0.5)}{0.1^2} \geq 225$