a

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ
FACULTAD DE CIENCIAS SOCIALES
ESPECIALIDAD DE ECONOMÍA

ESTADÍSTICA INFERENCIAL
PRÁCTICA CALIFICADA 4

13 de junio de 2015
Horario 522

Ejercicio 1. **(7 puntos)**

Para el ingreso mensual, X (en miles de soles), de los trabajadores en cierto sector, se propone el modelo probabilístico de Pareto siguiente: $f_X(x) = \frac{\theta\,\alpha^\theta}{x^{\theta+1}}$, $x \geq \alpha$, donde $\alpha > 0$ es una constante conocida (el ingreso mensual mínimo en este sector) y $\theta > 0$ es un parámetro por estimar a partir de una muestra aleatoria de X de tamaño, n, suficientemente grande.

a) Halle el estimador de θ por el método de máxima verosimilitud. (2 puntos)

b) A partir de la distribución asintótica del estimador de máxima verosimilitud de θ, en su versión a partir de la información de Fisher observada, deducir un intervalo del 95 % de confianza para estimar a θ. (3 puntos)

c) Al tomar una muestra aleatoria de $n = 100$, trabajadores de este sector, resultó $\sum_{j=1}^{100} Ln(X_j) = 53{,}61$. Estime θ, puntualmente y mediante un intervalo del 95 % de confianza. Asuma que $\alpha = 750$ soles. (2 puntos)

Ejercicio 2. **(6 puntos)**

Si $X \sim N(0; \sigma^2)$, donde σ^2 es desconocido. Para hacer inferencias sobre σ^2 se tomará una muestra aleatoria de tamaño $n = 20$.

a) Determine la distribución de la variable $W = \frac{20(\bar{X})^2}{\sigma^2}$ (2 puntos)

b) Use la variable W dada anteriormente y construya un intervalo del 95 % de confianza para estimar a σ^2. (2 puntos)

c) Determine el valor esperado de la longitud del intervalo anterior. (2 puntos)

Ejercicio 3. **(3 puntos)**

Sean $X \sim N(0; \sigma_1^2)$ e $Y \sim N(0; \sigma_2^2)$. Sean también dos muestras independientes de estas variables, $X_1, \ldots, X_{10}$ y $Y_1, \ldots, Y_{10}$, y S_1^2 y S_2^2 sus varianzas correspondientes. Considere la variable

$$W = \frac{\sum_{j=1}^{10} X_j^2}{\sum_{j=1}^{10} Y_j^2} \frac{\sigma_2^2}{\sigma_1^2}.$$

Halle la distribución de W para comprobar que esta es una variable base para σ_1^2/σ_2^2. Luego deduzca un intervalo del 95 % de confianza, para estimar este cociente.

$U \sim \chi^2(v)$; $E(U) = $

b.

Ejercicio 4. **(4 puntos)**

Para estimar p la proporción, en la población, que satisface cierta característica se tomará una muestra aleatoria de la variable aleatoria X que asume el valor 1, si se satisface la característica y el valor 0, si no se satisface. De este modo $f_X(x) = p^x(1-p)^{1-x}$, $x = 0; 1$. Se puede verificar que el estimador de máxima verosimilitud de p es $\bar{X}$ (la proporción, en la muestra que satisface la característica).

a) A partir de la distribución asintótica del estimador de máxima verosimilitud de p, en su versión a partir de la información de Fisher observada, deducir un intervalo del 95 % de confianza para estimar a θ. Aplique el método de la variable base. (2 puntos)

b) Una encuesta aplicada a una muestra aleatoria de 1000 electores reveló que solo 200 estaban de acuerdo con la gestión de cierta autoridad. Determine, qué se puede inferir, a partir de estos últimos resultados y con una confianza del 95 %, sobre la la proporción p de electores que aprueban la gestión de la autoridad. Use el resultado obtenido en la parte anterior. (2 puntos)

Recordatorio

Información de Fisher observada: $-H(\hat{\theta})$, $\hat{\theta}$ es el estimador de máxima verosimilitud de θ y $H(\theta) = \frac{\partial^2}{\partial \theta^2} Ln\big(\mathcal{L}(X_1, \ldots, X_n;\, \theta)\big)$, donde $\mathcal{L}(x_1, \ldots, x_n;\, \theta) = \mathcal{L}(\theta) = f_{X_1}(x_1) \cdots f_{X_1}(x_n)$: la función de verosimilitud de θ correspondiente a la muestra registrada $X_1 = x_1, \ldots, X_n = x_n$.

Distribución asintótica del estimador de máxima verosimilitud:

$\sqrt{-H(\widehat{\theta})}\ (\widehat{\theta} - \theta) \overset{aprox.}{\sim} N(0;\,1)$ (versión con la información de Fisher observada).

$X \sim N(\mu;\, \sigma^2):\quad \frac{X-\mu}{\sigma} \sim N(0;\,1);\quad \frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \sim N(0;\,1);\quad \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1).$

$Z \sim N(0;\,1) \Rightarrow Z^2 \sim \chi^2(1).$ $\bar{X} \sim N(\mu;\, \frac{\sigma^2}{n})$

$U \sim \chi^2(\nu):\ E(X) = \nu/2.$

$W_1 \sim \chi^2(\nu_1), W_2 \sim \chi^2(\nu_2)$, W_1 y W_2 independientes $\Rightarrow W := \frac{W_1/\nu_1}{W_2/\nu_2} \sim F(\nu_1,\, \nu_2).$

J.F. San Miguel, 13 de junio de 2015.

PC4 2015-1

1. $$f_X(x) = \frac{\theta \alpha^{\theta}}{x^{\theta+1}} \quad ; \quad x \geq \alpha \quad ; \quad \begin{matrix} \alpha > 0 \\ \theta > 0 \end{matrix}$$

a) Hallar estimación por máx verosimilitud

$$\mathcal{L}(\theta) = f_{X_1}(x_1) \times \dots \times f_{X_n}(x_n)$$

$$\mathcal{L}(\theta) = \frac{\theta \alpha^{\theta}}{(x_1)^{\theta+1}} \times \dots \times \frac{\theta \alpha^{\theta}}{(x_n)^{\theta+1}} = \frac{\theta^n \alpha^{\theta n}}{\left(\prod_{j=1}^{n} x_j\right)^{\theta+1}}$$

$$\ln(\mathcal{L}(\theta)) = n \ln(\theta) + \theta n \ln(\alpha) - (\theta+1) \ln\left(\prod_{j=1}^{n} x_j\right)$$

$$\ln(\mathcal{L}(\theta)) = n \ln(\theta) + \theta n \ln(\alpha) - (\theta+1) \sum_{j=1}^{n} \ln(x_j)$$

$$\frac{\partial \ln(\mathcal{L}(\theta))}{\partial \theta} = \frac{n}{\theta} + n \ln(\alpha) - \sum_{j=1}^{n} \ln(x_j) = 0$$

$$\boxed{\hat{\theta}_{MV} = \frac{n}{\sum_{j=1}^{n} \ln(x_j) - n \ln(\alpha)}}$$

b) A partir distribución asintótica del estimador MV de θ con info Fischer observada deducir intervalo 95% confianza

$$\sqrt{-H(\hat{\theta})}\,(\hat{\theta} - \theta) \sim N(0,1)$$

$$\ln(\mathcal{L}(\theta)) = n \ln(\theta) + \theta n \ln(\alpha) - (\theta+1) \sum_{j=1}^{n} \ln(x_j)$$

$$H(\hat{\theta}) = \frac{\partial \ln(\mathcal{L}(\theta))}{\partial \theta^2} = -\frac{n}{\hat{\theta}^2}$$

$$-1{,}96 \leq \sqrt{\frac{n}{\hat{\theta}^2}}\,(\hat{\theta} - \theta) \leq 1{,}96$$

$$\frac{-1{,}96}{\sqrt{n}} \leq \left(1 - \frac{\theta}{\hat{\theta}}\right) \leq \frac{1{,}96}{\sqrt{n}}$$

$$\frac{-1{,}96}{\sqrt{n}} - 1 \leq -\frac{\theta}{\hat{\theta}} \leq \frac{1{,}96}{\sqrt{n}} - 1$$

Como $\hat{\theta} < 0$ para : $\alpha = 0,75$

$$\left[\frac{+1,96}{\sqrt{n}} - 1\right](-\hat{\theta}) \leq \theta \leq \left[\frac{1,96}{\sqrt{n}} + 1\right][+\hat{\theta}]$$ con 95% confianza

dado que $$\hat{\theta} = \frac{n}{\sum_{i=1}^{n} \ln(x_i) - n\ln(\alpha)}$$

Finalmente

$$\boxed{\hat{\theta}\left[\frac{-1,96}{\sqrt{n}} + 1\right] \leq \theta \leq \hat{\theta}\left[\frac{+1,96}{\sqrt{n}} + 1\right]}$$

c) Al tomar muestra $n=100$. Estimar puntualmente θ mediante intervalo 95%

$\alpha = 750$

7

$$\hat{\theta} = \frac{100}{53,61 - 100\ln(0,750)} \cong 1,21 \ ;$$

6 ~~[illegible]~~ → intervalo con 95% confianza

$$0,97 \leq \theta \leq 1,44$$

(2) $X \sim N(0, \sigma^2)$ donde σ^2 es desconocido : Se tomará muestra $n=20$

a) determinar distribución $$W = \frac{20(\bar{X})^2}{\sigma^2}$$

Dado $$\bar{X} \sim N\left(0; \frac{\sigma^2}{n}\right)$$

estandarizando : $$\frac{\bar{X} - 0}{\sqrt{\frac{\sigma^2}{n}}} \sim N(0,1)$$

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

como $n=20$: $$\frac{\sqrt{20}\,\bar{X}}{\sigma} \sim N(0,1)$$

Por formula : $Z \sim N(0,1) \rightarrow Z^2 \sim \chi^2(1)$

Como $Z = \dfrac{\sqrt{20}\ \bar{x}}{\sigma} \sim N(0,1)$

$$\boxed{Z^2 = \frac{20(\bar{x})^2}{\sigma^2} \sim \chi^2(1)}$$

Como $W = Z^2$ esta es la distribución de $W = \dfrac{20(\bar{x})^2}{\sigma^2}$

b) Usar W para construir intervalo 95% confianza para σ^2

$F_W(a) = 0{,}025$ $\qquad F_W(b) = 0{,}975$

$$a \leq W \leq b$$

$$a \leq \frac{20(\bar{x})^2}{\sigma^2} \leq b$$

$$b \leq \frac{\sigma^2}{20(\bar{x})^2} \leq a \rightarrow 20\left(\frac{1}{b}\right)(\bar{x})^2 \leq \sigma^2 \leq 20\left(\frac{1}{a}\right)(\bar{x})^2$$

6

Como $a = 0{,}0010$; $b = 5{,}0234$

$$\boxed{3{,}98\,(\bar{x})^2 \leq \sigma^2 \leq 20000\,(\bar{x})^2}$$

c) Hallar Esperado longitud intervalo

$$E\left[20000(\bar{x})^2 - 3{,}98(\bar{x})^2\right] = E\left(1{,}9996{,}02(\bar{x})^2\right)$$

$$= 1{,}9996\ E(\bar{x}^2)$$

$$= 1{,}9996\left[V(\bar{x}) + E^2(\bar{x})\right]$$

$$= 1{,}996\left[\frac{V(x)}{n} + E^2(x)\right]$$

$$\boxed{\text{esperado longitud intervalo} = 1{,}996\left[\frac{\sigma^2}{n}\right]}$$

(3) $X \sim N(0; \sigma_1^2)$; $Y \sim N(0, \sigma_2^2)$

Sean muestras $\left\{ \begin{array}{l} X_1; \dots X \\ Y_1 \dots Y \end{array} \right.$ con S_1^2 y S_2^2 varianzas

$$\hookrightarrow \quad W = \frac{\sum_{i=1} x_i^2}{\sum_{i=1} y_i^2} \cdot \frac{\sigma_2^2}{\sigma_1^2}$$

(a) Halla distribución W para comprobar que es variable base para $\frac{\sigma_1^2}{\sigma_2^2}$ y deducir intervalo 95% confianza

Estandarizando X y Y:
$$Z_x = \frac{x}{\sigma_1} \quad ; \quad Z_y = \frac{Y}{\sigma_2}$$

Transformando a χ^2:
$$Z_x^2 = \frac{x^2}{\sigma_1^2} \sim \chi^2(1) \; ; \; Z_y^2 = \frac{Y^2}{\sigma_2^2} \sim \chi^2(1)$$

propiedad
$$\sum_{i=1}^{n} \chi^2(1) \sim \chi^2(n)$$

Como $n=10$ para X y para Y:

$$\boxed{W_1 = \frac{\sum_{i=1}^{9} x_i^2}{\sigma_1^2} \sim \chi^2(9) \quad W_2 = \frac{\sum_{i=1}^{9} Y_i^2}{\sigma_2^2} \sim \chi^2(9)}$$

$$\hookrightarrow \quad W = \frac{W_1 / v_1}{W_2 / v_2} \sim F(v_1; v_2)$$

$$W = \frac{\dfrac{\sum_{i=1}^{9} X_i^2}{\sigma_1^2 (9)}}{\dfrac{\sum Y_i^2}{\sigma_2^2 (9)}} \sim F(9; 9)$$

$$\boxed{W = \frac{\sum_{i=1}^{9} x_i^2 \cdot \sigma_2^2}{\sum_{i=1}^{9} Y_i^2 \cdot \sigma_1^2} \sim F(9; 9)}$$

Para hallar intervalo 95%. $F_W(a) = 0,025$; $F_W(b) = 0,975$

$\Rightarrow a = 0,25$; $b = 4,03$

$$0,25 \le \frac{\sum_{i=1}^{9} x_i^2 \cdot \sigma_2^2}{\sum_{i=1}^{9} y_i^2 \cdot \sigma_1^2} \le 4,03$$

3

intervalo 95% confianza para σ_1^2 / σ_2^2

$$\boxed{\frac{\sum_{i=1}^{9} x_i^2}{\sum_{i=1}^{9} y_i^2} \cdot \left(\frac{1}{4,03}\right) \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{\sum_{i=1}^{9} x_i^2}{\sum_{i=1}^{9} y_i^2} \cdot \left(\frac{1}{0,25}\right)}$$

4) $f_X(x) = p^x (1-p)^{1-x}$; $x = 0, 1$

a) A partir distribución asintótica estimador máx verosimilitud, hallar intervalo 95% confianza para θ

$$\mathcal{L}(p) = f_{X_1}(x_1) \dots f_{X_n}(x_n) = p^{x_1}(1-p)^{1-x_1} \times \dots \times p^{x_n}(1-p)^{1-x_n}$$

$$\mathcal{L}(p) = p^{\sum_{i=1}^{n} x_i} (1-p)^{n - \sum_{i=1}^{n} x_i}$$

$$\ln(\mathcal{L}(p)) = \sum_{i=1}^{n} x_i \ln(p) + \left(n - \sum_{i=1}^{n} x_i\right) \ln(1-p)$$

$$\frac{d \ln(\mathcal{L}(p))}{dp} = \frac{\sum_{i=1}^{n} x_i}{p} + \left(\frac{n - \sum_{i=1}^{n} x_i}{1-p}\right)(-1) = 0$$

$$(1-p)\left(\sum x_i\right) = pn - p\sum x_i$$

$$\sum x_i - \cancel{p\sum x_i} = pn - \cancel{p\sum x_i}$$

$$\boxed{\bar{x} = \hat{p}}$$

$$H(\hat{p}) = \frac{d^2}{d} \ln(\mathcal{L}) = \frac{-\left(\sum^{n} x_i\right)}{\hat{p}^2} + \frac{\left(n - \sum^{n} x_i\right)}{(1-\hat{p})^2} = 0$$

$$\sqrt{-H(\hat{p})} = \sqrt{\frac{\left(\sum^{n} x_i\right)}{\hat{p}^2} - \frac{\left(n - \sum^{n} x_i\right)}{(1-\hat{p})^2}}$$

$$-1{,}96 \leq \sqrt{-H(\hat{p})}\,(\hat{p} - p) \leq 1{,}96$$

$$\frac{-1{,}96}{\sqrt{-H(\hat{p})}} - \hat{p} \leq -p \leq \frac{1{,}96}{\sqrt{-H(\hat{p})}} - \hat{p}$$

$$\hat{p} - \frac{1{,}96}{\sqrt{H(\hat{p})}} \leq p \leq \hat{p} + \frac{1{,}96}{\sqrt{-H(\hat{p})}}$$

$$\boxed{\hat{p} - \frac{1{,}96}{\sqrt{-H(\hat{p})}} \leq p \leq \hat{p} + \frac{1{,}96}{\sqrt{-H(\hat{p})}}}$$ intervalo 95 % confianza

b) $n = 1000$

$\sum_{i=1} x_i = 200 \rightarrow$ $\hat{p} = 1/5 \rightarrow \sqrt{-H(\hat{p})} = 61{,}237$

$$0{,}167 \leq p \leq 0{,}232$$ → intervalo 95 % confianza

c)