Métodos Estadísticos
Tarea II

Rubén Pérez Palacios Lic. Computación Matemática
Profesora: Dra. Eloísa Díaz Francés Murguía

22 de marzo de 2021

Problemas

1. Considera una muestra de n variables aleatorias independientes $X_1, \cdots, X_n$ idénticamente distribuidas como exponencial con valor esperado θ. Da las expresiones de una estadística suficientes t unidimensional para θ. Usala para expresar la logverosimilitud $l\,(\theta;t)$, la función Score $Sc(\theta;t)$ el emv de θ, denotado como $\hat{\theta}$, la información observada de Fisher $I_{\hat{\theta}}$ y finlmente la verosimilitud relativa de θ, denotada como $R(\theta;t)$.

   Recordemos que la función de densidad de una $Exp(\theta)$ es

   $$f(x;\theta) = \frac{1}{\theta}\exp\left(-\frac{x}{\theta}\right)\mathbb{1}_{[0,\infty)}(x),$$

   por lo que función de distribución conjunto es

   $$f(\vec{x};\theta) = \prod_{i=1}^{n} f(x_i;\theta) = \prod_{i=1}^{n}\frac{1}{\theta}\exp\left(-\frac{x_i}{\theta}\right)\mathbb{1}_{[0,\infty)}(x_i) = \left(\prod_{i=1}^{n}\mathbb{1}_{[0,\infty)}(x_i)\right)\frac{1}{\theta^n}\exp\left(-\frac{\sum_{i=1}^{n} x_i}{\theta}\right),$$

   por el Teorema de la factorización de Fisher tenemos que $t = \sum_{i=1}^{n} x_i$ es una estadística suficiente de θ.

   Luego

   $$L_U(\theta;t) = \frac{1}{\theta^n}\exp\left(-\frac{t}{\theta}\right),$$

   por lo que

   $$l(\theta;t) = \log\left(\frac{1}{\theta^n}\right) - \frac{t}{\theta},$$

   entonces

   $$Sc(\theta;t) = -\frac{n}{\theta} + \frac{t}{\theta^2},$$

   igualando a cero la función Score obtenemos

$$0 = -\frac{n}{\hat{\theta}} + \frac{t}{\hat{\theta}^2},$$

por lo que

$$\hat{\theta} = \frac{t}{n} = \overline{X}.$$

La información de Fisher es

$$I_{\hat{\theta}} = \left(\frac{n}{\theta^2} - \frac{2t}{\theta^3} \right) \bigg|_{\theta=\hat{\theta}} = \frac{n^3}{t^2}.$$

Por último la verosimilitud relativa es

$$R(\theta; t) = \frac{\frac{1}{\theta^n} \exp\left(-\frac{t}{n}\right)}{\frac{1}{\hat{\theta}^n} \exp\left(-\frac{t}{\hat{\theta}}\right)} = \left(\frac{t}{n\theta} \right)^n \exp\left(\frac{1}{n} - \frac{t}{\theta} \right).$$

2. Da un ejemplo de una distribución que pertenezca a la familia de distribuciones de valores extremos y también a la de localización y escala.

   Es la distribución Gumbel. Sabemos que esta pertence a la familia de distribuciones de valores extremos, ahora demostraremos que pertenece a la de localización y escala.

   Demostración. Recordemos que la función de densidad de un Gumbel (a, b) es

$$f(x; a, b) = \frac{1}{b} \exp\left\{ -\frac{x-a}{b} - \exp\left[-\left(\frac{x-a}{b} \right) \right] \right\} \mathbb{1}_{(-\infty,\infty)}(x).$$

   Ahora podemos reexpresar esta como

$$f(x; a, b) = \frac{1}{b} f_0 \left(\frac{x-a}{b} \right),$$

   donde

$$f_0(x) = \exp\left[-x - \exp(-x)\right] \mathbb{1}_{(-\infty,\infty)}(x).$$

   Podemos ver que su soporte no depende de ningún parametro desconocido. Además si $Y = \frac{x-a}{b}$ entonces por el Teorema de Cambio de Variable su función de densidad es $f_0(x)$ la cual no depende de ningún parametro desconocido, por lo tanto concluimos que la Gumbel pertenece a la familia de localización y escala. □

3. La Densidad Gamma con parámetros de forma α y escala β está dada como

$$f(x; \alpha, \beta) = \frac{x^{\alpha-1}}{\Gamma(\alpha)\, \beta^\alpha} \exp\left(-\frac{x}{\beta} \right) \mathbb{1}_{(0,\infty)}(x).$$

   a) Demuestra que la distirbución Gamma pertence a la familia exponencial de distribuciones.

Demostración. Podemos reexpresarla a la densidad de la distribución Gamma de la siguiente forma

$$f(x;\alpha,\beta)=\left(\frac{1}{\Gamma\alpha\beta^{\alpha}}\right)\left(\frac{1}{x}\right)exp\left(\alpha\log(x)-\frac{x}{\beta}\right)\mathbb{1}_{(0,\infty)}(x).$$

Donde su soporte no depende de parámetros desconocidos y $A(\theta)=\frac{1}{\Gamma\alpha\beta^{\alpha}}$, $B(x)=\frac{1}{x}$, $C_1(\theta)=\alpha$, $D_1(x)=log(x)$, $C_2(\theta)=\frac{1}{\beta}$ y $D(x)=x$, por lo tanto la distirbución Gamma pertenece a la familia exponencial. □

b) Considera una muestra de n variables $X_1,\cdots,X_n$ independientes e idénticamente distribuidas como Gamma (a,b). Demuestra que las medias arimética y geométrica,

$$t_1=\frac{1}{n}\sum_{i=1}^{n}x_i,\ \text{y}\ t_2=\left(\prod_{i=1}^{n}x_i\right)^{\frac{1}{n}}.$$

constituyen un vector de estadísticas suficientes para α y β. También demuestra que son suficientes para (α,β).

Demostración. Veamos que la densidad conjunta de la muestra es

$$f(\vec{x};\vec{\theta})=\prod_{i=1}^{n}f(x_i;\vec{\theta})=\left(\prod_{i=1}^{n}\mathbb{1}_{(0,\infty)}(x)\right)\left(\frac{1}{\Gamma(\alpha)\beta}\right)^{n}(t_2)^{-n}\exp\left(\alpha\log(t_2^n)-\frac{nt}{\beta}\right).$$

Por el Teorema de la factorización de Fisher concluimos que el vector $\vec{t}=(t_1,t_2)$ es un vector de estadísticas suficientes de (α,β). □

c) Reparametriza en términos de (μ,α) donde la media es $\mu=\alpha\beta$ y da la espresión de la logverosimilitud de μ,α.

Solución 1 – *La reparametrización de la densidad es*

$$f(x;\alpha,\beta)=\left(\frac{\alpha^{\alpha}}{\Gamma\alpha\mu^{\alpha}}\right)\left(\frac{1}{x}\right)exp\left(\alpha\left(\log(x)-\frac{x}{\mu}\right)\right)\mathbb{1}_{(0,\infty)}(x),$$

por lo que

$$f(\vec{x};\vec{\theta})=\prod_{i=1}^{n}f(x_i;\vec{\theta})=\left(\prod_{i=1}^{n}\mathbb{1}_{(0,\infty)}(x)\right)\left(\frac{\alpha^{\alpha}}{\Gamma(\alpha)\mu}\right)^{n}(t_2)^{-n}\exp\left(\alpha\left(\log(t_2^n)-\frac{nt}{\mu}\right)\right),$$

por el Teorema de la factorización de Fisher concluimos que el vector $\vec{t}=(t_1,t_2)$ *es un vector de estadísticas suficientes de* (μ,α). *Entonces*

$$L_u(\vec{\theta};\vec{t})=\left(\frac{\alpha^{\alpha}}{\Gamma(\alpha)\mu}\right)^{n}\exp\left(\alpha\left(\log(t_2^n)-\frac{nt}{\mu}\right)\right),$$

por lo tanto

$$l(\vec{\theta,t})=n\alpha\left(\log\left(\frac{\sigma}{\mu}\right)+(\log(t_2))-\frac{t_1}{\mu}\right)-n\log(\Gamma\alpha).$$

Demuestra que el estimador de momentos de μ denotado como $\tilde{\mu}$ es igual al estimador de máxima verosimilitud de μ, denotado por $\hat{\mu}$. Demuestra que el estimador de máxima verosimilitud restringido de μ dado α el cual es denotado por $\hat{\mu}(\alpha)$, es igual al emv de μ.

Demostración. Empezaremos por calcular cada uno de los estimadores. El de momentos esta dado por

$$\tilde{\mu} = E[X] = t_1.$$

Puesto que

$$\frac{\partial l}{\partial \mu} = n\alpha\left(t_1 - \mu\right),$$

por lo tanto el emv de μ es

$$\hat{\mu} = t_1.$$

Por último como $\hat{\mu}$ es independiente de α entonces

$$\hat{\mu} = \mu$$

□

Da la logverosimilitud perfil del parámetro de forma α

$$l(\alpha, \hat{\mu}(\alpha); \vec{t}) = n\alpha\left(\log\left(\frac{\alpha}{t_1}\right) + \log(t_2) - 1\right) - n\log\left(\Gamma\left(\alpha\right)\right).$$

Da la expresión de la matriz de información de Fisher para (μ, α). Calcula las inormaciones individuales de μ y α.

$$I_{\hat{\mu}\hat{\alpha}} = \begin{pmatrix} \frac{n\hat{\alpha}}{t_1^2} & 0 \\ 0 & n\left(\left(\frac{\Gamma(\hat{\alpha})\Gamma^{(2)}(\hat{\alpha}) - \left(\Gamma'(\hat{\alpha})\right)^2}{\left(\Gamma(\hat{\alpha})\right)^2}\right) - \frac{1}{\alpha}\right)\end{pmatrix},$$

al ser una matriz diagonal la matriz inversa $I^{-1}_{\hat{\mu}\hat{\alpha}}$ solo tendra como entrada (i,i) el inverso de la entrada (i,i) de la matriz $I_{\hat{\mu}\hat{\alpha}}$, luego como las informaciones individuales están dadas por el inverso de la entrada (i,i) de la matriz $I^{-1}_{\hat{\mu}\hat{\alpha}}$ entonces concluimos que las informaciones individuales están dadas por las entradas (i,i) de la matriz $I_{\hat{\mu}\hat{\alpha}}$, es decir

$$I_{\hat{\mu}} = \frac{n\hat{\alpha}}{t_1^2},$$

$$I_{\hat{\mu}} = n\left(\left(\frac{\Gamma(\hat{\alpha})\Gamma^{(2)}\left(\hat{\alpha}\right) - \left(\Gamma'\left(\hat{\alpha}\right)\right)^2}{\left(\Gamma\left(\hat{\alpha}\right)\right)^2}\right) - \frac{1}{\alpha}\right).$$

4. Considera una muestra de n variables aleatorias continuas $X_1, \cdots, X_n$ que son independientes e idénticamente distribuidas como Uniformes en $(0, \theta)$ con $\theta > 0$ desconocido. Da un estimador de momentos, el emv y la verosimilitud relativa de θ, el intervalo de nivel $0 \leq c \leq 1$. Di si existe o no la información de fisher.

   El estimador de momentos esta dado por

$$\tilde{\theta} = E\left[X\right] = \frac{\theta}{2}.$$

Puesto que la densidad conjunta de las X_i esta dada por (demostrado tarea 1)

$$f(\vec{x};\theta) = \left(\frac{1}{\theta}\right)^n \mathbb{1}_{(0,\theta)}(\text{máx}(\vec{x})),$$

entonces con $t = \text{máx}(\vec{x})$

$$L_U(\theta;T) = \left(\frac{1}{\theta}\right)^n \mathbb{1}_{(0,\theta)}(t),$$

puesto que esto es una función decreciente para θ enotnces esta se maximizara en el menor valor que θ pueda tomar. Ahora hay que recordar que $\hat{\theta}$ es solo una aproximación al emv que alguna veces resulta serlo pero no es cierto que esto se cumple, ya que por definición

$$\hat{\theta} = \max_{\theta} L_u(\theta;t),$$

entonces $\hat{\theta}$ no tiene que vivir en el espacio parametral. Como $\hat{\theta} = \sup\{\theta : \theta > t\}$ concluimos que $\hat{\theta} = t$.

Ahora (no se si este bien porque la indicadora del denominador es cero, pero...)

$$R(\theta;t) = \frac{\left(\frac{1}{\theta}\right)^n \mathbb{1}_{(0,\theta)}(t)}{\left(\frac{1}{t}\right)^n \mathbb{1}_{(0,t)}(t)} = \left(\frac{1}{\theta t}\right)^n \mathbb{1}_{(0,\theta)}(t),$$

luego $\left(\frac{1}{\theta t}\right)^n \mathbb{1}_{(0,\theta)}(t) \geq c$ si y sólo si $\theta > t$ y $\theta = \frac{1}{\sqrt[n]{c}t}$ por lo tanto

$$IV(c) = \left\{\theta : \theta = \frac{1}{\sqrt[n]{c}t} \text{ y } \theta > t\right\}.$$

Por úlimto la Información de Fisher es (no se si este bien porque la indicadora es cero)

$$I_{\hat{\theta}} = \left.\frac{n}{\theta^2}\right|_{\theta=\hat{\theta}} = \frac{n}{t^2}.$$