ELM2401 - Métodos Estadísticos
Pauta I1

Profesora: Lorena Correa
Ayudante: Ana María Alvarado

1. Sean $X_1, \dots, X_n$ variables aleatorias i.i.d. $U(0, \theta)$. Se sugiere utilizar el siguiente estimador para el parámetro θ.

$$\hat{\theta} = max\{X_1, \dots, X_n\}$$

a.- Si se sabe que $\frac{\hat{\theta}}{\theta}$ tiene distribución $Beta(n, 1)$, y que si una variable aleatoria X tiene distribución $Beta(\alpha, \beta)$ entonces:

$$E(X) = \frac{\alpha}{\alpha + \beta} \qquad Var(X) = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}$$

Determine el ECM de $hat\theta$ en función de θ.

Solución:

$$\frac{\hat{\theta}}{\theta} \sim Beta(n, 1)$$

$$ECM(\hat{\theta}) = \mathbb{V}ar(\theta) + \{\mathbb{E}(\hat{\theta}) - \theta\}^2 \qquad (0{,}5)$$

$$\mathbb{E}(\hat{\theta}) = \theta\,\mathbb{E}(\frac{\hat{\theta}}{\theta}) = \theta\,\mathbb{E}(Beta(n, 1)) = \theta\frac{n}{n+1} \qquad (1{,}5)$$

$$\mathbb{V}ar(\hat{\theta}) = \theta^2\mathbb{V}ar(\frac{\hat{\theta}}{\theta}) = \theta^2\mathbb{V}ar(Beta(n, 1)) = \theta^2\frac{n}{(n+1)^2(n+2)} \qquad (2{,}0)$$

Así,

$$\begin{aligned} ECM(\hat{\theta}) &= \theta^2\frac{n}{(n+1)^2(n+2)} + \theta^2(\frac{n}{n+1} - 1)^2 \\ &= \theta^2\{\frac{n}{(n+1)^2(n+2)} + \frac{1}{(n+1)^2}\} \\ &= \theta^2\{\frac{n+n+2}{(n+1)^2(n+2)}\} \\ &= \frac{2\theta^2}{(n+1)(n+2)} \end{aligned}$$

b.- ¿Es $\hat{\theta}$ un estimador consistente?. ¿Qué puede decir acerca de la propiedad de insesgamiento?.

Solución:

$$\lim_{n\to\infty} ECM(\hat{\theta}) = \lim_{n\to\infty} \frac{2\theta^2}{(n+1)(n+2)} = 0 \qquad (1{,}0)$$

Luego, $\hat{\theta}$ es un estimador consistente. Además podemos decir que es un estimador asintóticamente insesgado. (1,0)

+ 1 pto base

2. Sean $X_1, \ldots, X_n$ variables aleatorias i.i.d $N(\theta, \theta),\ \theta > 0$.

a) Demuestre que el EMV de θ es raíz de la ecuación:

$$\theta^2 + \theta - W = 0, \qquad W = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

y determine $\hat{\theta}$.

Solución:

$$\begin{aligned}
L(\theta, X_1, \ldots, X_n) &= (2\pi\theta)^{-\frac{n}{2}} exp\{\frac{-1}{2\theta}\sum_{i=1}^{n}(X_i - \theta)^2\} \\
logL(\theta, X_1, \ldots, X_n) &= \frac{-n}{2} log(2\pi\theta) - \frac{1}{2\theta}\sum_{i=1}^{n}(X_i - \theta)^2 && \color{red}{(1,5)} \\
\frac{\partial logL(\theta, X_1, \ldots, X_n)}{\partial\theta} &= \frac{-n}{2\theta} + \frac{1}{2\theta^2}\sum_{i=1}^{n}(X_i - \theta)^2 + \frac{1}{\theta}(n\bar{X} - n\theta) \\
&= \frac{-n}{2\theta} - \frac{n}{2} + \frac{\sum_{i=1}^{n} X_i^2}{2\theta^2} && \color{red}{(0,5)}
\end{aligned}$$

Igualando a cero:

$$\begin{aligned}
\frac{-n}{2\theta} - \frac{n}{2} + \frac{\sum X_i^2}{2\theta^2} &= 0 \qquad \Big/ \frac{-2\theta^2}{n} \\
\theta^2 + \theta - \frac{\sum_{i=1}^{n} X_i^2}{n} &= 0 \\
\theta^2 + \theta - W &= 0
\end{aligned}$$

$$\Rightarrow \quad \hat{\theta} = \frac{-1 \pm \sqrt{1 + 4W}}{2} \qquad \color{red}{(0,5)}$$

Como θ corresponde a la varianza no puede ser negativa, luego

$$\hat{\theta} = \frac{-1 + \sqrt{1 + 4W}}{2} \qquad \color{red}{(0,5)}$$

b) Determinar la distribución asintótica de $\hat{\theta}$.

Solución:

Sabemos que si $n \to \infty$ y $\hat{\theta}$ es E.M.V. de θ entonces:

$$\begin{aligned}
\hat{\theta} &\,\dot{\sim}\, N(\theta, \mathbb{V}ar(\hat{\theta})) \\
\mathbb{V}ar(\hat{\theta}) &= \Big[-\mathbb{E}\Big(\frac{\partial^2 logP_{x,\theta}}{\partial\theta^2}\Big)\Big]^{-1} && \color{red}{(1,0)}
\end{aligned}$$

$$\frac{\partial^2 logP_{x,\theta}}{\partial\theta^2} = \frac{n}{2\theta^2} - \frac{\sum_{i=1}^{n} X_i^2}{\theta^3} \qquad \color{red}{(0,5)}$$

$$
\begin{aligned}
\mathbb{E}\Big(\frac{\partial^2 log P_{x,\theta}}{\partial\theta^2}\Big) &= \frac{n}{2\theta^2} - \frac{1}{\theta^3} n\mathbb{E}(X_i^2) \\
&= \frac{n}{2\theta^2} - \frac{1}{\theta^3} n[\mathbb{V}ar(X_i) + \mathbb{E}^2(X_i)] \\
&= \frac{n}{2\theta^2} - \frac{1}{\theta^3} n[\theta + \theta^2] \\
&= \frac{-n - 2n\theta}{2\theta^2} && (0{,}5)
\end{aligned}
$$

Luego,

$$
\mathbb{V}ar(\hat{\theta}) = \frac{2\theta^2}{n + 2n\theta}
$$

Así,

$$
\hat{\theta} \;\dot{\sim}\; N(\theta, \frac{2\theta^2}{n + 2n\theta}) \qquad (1{,}0)
$$

También pueden calcular la cota de Cramer Rao (ó $\mathbb{V}ar(\hat{\theta})$) como:

$$
\mathbb{V}ar(\hat{\theta}) = [nI(\theta)]^{-1}
$$

donde $I(\theta) = -\mathbb{E}\Big(\frac{\partial^2 log f(x)}{\partial\theta^2}\Big)$

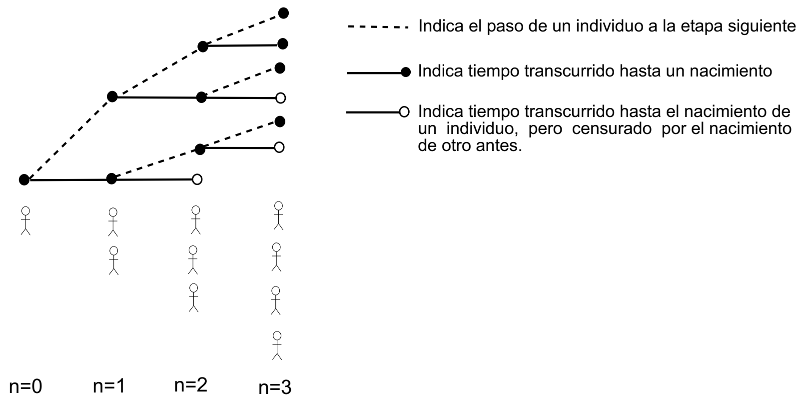
+ 1 pto base

3. En el tiempo $t = 0$, hay un individuo vivo en cierta población. Entonces se presenta un ”proceso de nacimiento puro”. El tiempo que transcurre hasta el primer nacimiento está distribuido exponencialmente con parámetro λ. Después del primer nacimiento, hay dos individuos vivos. El tiempo hasta que el primero de ellos vulve a dar a luz es exponencial con parámetro λ y del mismo modo para el segundo individo.

a.- Demuestre que el tiempo transcurrido hasta que nace el n-ésimo individuo es exponencial de parámetro $n\lambda$.

Solución:

Sea T_n:Tiempo transcurrido hasta que nace el n-ésimo individuo.



Así, $\quad T_n = min\{X_1, X_2, \ldots, X_n\} \quad$ con $\quad X_i \sim exp(\lambda)$ (0,5)

$$\begin{aligned}
F_{T_n}(t) = P(T_n \leq t) &= 1 - P(T_n > t) \\
&= 1 - P(min\{X_1, X_2, \ldots, X_n\} > t) \\
&= 1 - P(X_1 > t, X_2 > t, \ldots, X_n > t) \\
&= 1 - [P(X_i > t)]^n \\
&= 1 - [e^{-\lambda t}]^n
\end{aligned}$$ (1,5)

$$f_{T_n}(t) = \lambda n e^{-n\lambda t} \qquad t > 0 \qquad \Rightarrow \qquad T_n \sim exp(n\lambda)$$

b) ¿Pertenece la distribución del tiempo transcurrido hasta que nace el n-ésimo individuo, a la familia exponencial?. Determine el estadístico suficiente para estimar λ

Solución:

$$\begin{aligned}
L(\lambda, X_1, X_2, \ldots, X_k) &= \prod_{i=1}^{k} n\lambda\, e^{-n\lambda X_i} \\
&= (n\lambda)^k e^{-n\lambda \sum_{i=1}^{k} X_i} \\
&= exp\{\underbrace{\,k\, log(n\lambda)}_{d(\theta)} \;-\; \underbrace{n\,\lambda}_{c_1(\theta)} \;\; \underbrace{\sum_{i=1}^{k} X_i}_{t_1(X)}\}
\end{aligned}$$

Luego, la distribución pertenece a la familia exponencial.

El estadístico suficiente corresponde a $\quad t_1(X) = \sum_{i=1}^{k} X_i$ (1,0) c.- Suponga que el proceso se observa hasta que ha ocurrido el sexto nacimiento y los tiempos sucesivos de nacimiento son: 25.2 , 41.7 , 51.2 , 55.5 , 55.5 , 59.5 , 61.8. Derive el EMV de λ y obtenga el valor para la muestra obtenida.

Solución:

$$\begin{aligned} logL(\lambda) &= t\,log(n\lambda) - n\lambda\sum_{i=1}^{k} X_i \\ \frac{\partial logL(\lambda)}{\partial\lambda} &= \frac{k}{n\lambda}n - n\sum_{i=1}^{k} X_i \\ & \frac{k}{\hat{\lambda}} - n\sum_{i=1}^{k} X_i = 0 \qquad \Rightarrow \qquad \hat{\lambda} = \frac{k}{n\sum_{i=1} X_i} = \frac{1}{n\bar{X}} \end{aligned}$$ (1,0)

Luego $\hat{\lambda}_{EMV} = \dfrac{1}{n\bar{x}}$

Evaluando en la muestra dada $\hat{\lambda}_{EMV} = 0{,}0034$ (0,5)

d.- Determine la distribución asintótica de $\hat{\lambda}$.

Solución:

$$\begin{aligned} \frac{\partial^2 l(\lambda)}{\partial\lambda^2} &= \frac{-k}{\lambda^2} \\ E\Big(-\frac{\partial^2 l(\lambda)}{\partial\lambda^2}\Big) &= \frac{k}{\lambda^2} \\ Var(\hat{\lambda}) &= \Big[E\Big(-\frac{\partial^2 l(\lambda)}{\partial\lambda^2}\Big)\Big]^{-1} = \frac{\lambda^2}{k} \end{aligned}$$

Luego,

$$\hat{\lambda} \ \dot{\sim}\ N(\lambda, \frac{\lambda^2}{k})$$ (0,5)

e.- Obtenga el estimador de mínimos cuadrados y obtenga su error cuadrático medio.

Solución:

$$\begin{aligned} S^2 &= \sum_{i=1}^{k}(X_i - \mathbb{E}(X_i))^2 = \sum_{i=1}^{k}(X_i - (n\lambda)^{-1})^2 \qquad \text{Donde} \lambda \sim exp(n\lambda) \\ \frac{\partial S^2}{\partial\lambda} &= 2\sum_{i=1}^{k}\Big(X_i - \frac{1}{n\lambda}\Big)\Big(-\frac{1}{n\lambda^2}\Big) \qquad \Rightarrow \qquad \sum_{i=1}^{k} X_i = \frac{k}{n\hat{\lambda}} \end{aligned}$$

$$\hat{\lambda} = \frac{k}{n\sum_{i=1}^{k} X_i} = \frac{1}{n\bar{X}}$$ (0,5)

$$ECM(\hat{\lambda}) = \mathbb{V}ar(\hat{\lambda}) + [\mathbb{E}(\hat{\lambda}) - \lambda]^2$$

$$\mathbb{E}(\hat{\lambda}) = \mathbb{E}\Big(\frac{1}{n\bar{X}}\Big) = \frac{1}{n}\mathbb{E}\Big(\frac{1}{\bar{X}}\Big)$$
$$\mathbb{V}ar(\hat{\lambda}) = \mathbb{V}ar\Big(\frac{1}{n\bar{X}}\Big) = \frac{1}{n^2}\mathbb{V}ar\Big(\frac{1}{\bar{X}}\Big)$$ (0,5)

Luego, es necesario determinar $\mathbb{E}\left(\frac{1}{\bar{X}}\right)$ y $\mathbb{V}ar\left(\frac{1}{\bar{X}}\right)$

+ 1 pto base

Sea $T = \frac{1}{\bar{X}} = \frac{k}{\sum X_i} = \frac{k}{Z}$ donde $Z \sim Gamma(k, n\lambda)$

Es decir,

$$f_Z(z) = \frac{(n\lambda)^k}{\Gamma(k)} z^{k-1} e^{-n\lambda z} \qquad z > 0$$

$$F_T(t) = 1 - F_Z(\frac{k}{t}) \qquad \Rightarrow \qquad f_T(t) = \frac{k}{t^2} f_Z(\frac{k}{t}) \qquad t > 0$$

Así,

$$f_T(t) = \frac{k}{t^2} \frac{(n\lambda)^k}{\Gamma(k)} \left(\frac{k}{t}\right)^{k-1} e^{-n\lambda \frac{k}{t}}$$

$$f_T(t) = \frac{(n\lambda k)^k e^{-(n\lambda k)\frac{1}{t}}}{\Gamma(k)} \left(\frac{1}{t}\right)^{k+1}$$

Sea $\alpha = k$ y $\beta = n\lambda k$

Luego,

$$f_T(t) = \frac{\beta^\alpha}{\Gamma(\alpha)} t^{-(\alpha+1)} e^{-\frac{\beta}{t}} \qquad \alpha, \beta, t > 0$$

$T \sim Gamma - Inversa(\alpha, \beta)$

$$\begin{aligned} \Longrightarrow \quad \mathbb{E}(T) &= \frac{\beta}{\alpha - 1} = \frac{n\lambda k}{k-1} \\ \mathbb{V}ar(T) &= \frac{\beta^2}{(\alpha-1)^2(\alpha-2)} = \frac{(n\lambda k)^2}{(k-1)^2(k-2)} \end{aligned}$$

Luego,

$$\begin{aligned} ECM(\hat{\lambda}) &= \mathbb{V}ar(\hat{\lambda}) + [\mathbb{E}(\hat{\lambda}) - \lambda]^2 \\ &= \frac{1}{n^2}\mathbb{V}ar(T) + \left[\frac{1}{n}\mathbb{E}(T) - \lambda\right]^2 \\ &= \frac{1}{n^2}\frac{(n\lambda k)^2}{(k-1)^2(k-2)} + \left[\frac{1}{n}\frac{n\lambda k}{k-1} - \lambda\right]^2 = \lambda^2\left[\frac{k+2}{(k-1)(k-2)}\right] \end{aligned}$$

4. Sean $X_1, \ldots, X_n$ variables aleatorias i.i.d. con densidad de probabilidades dada por:

$$f(x) = \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} \qquad x \geq 0 \qquad \alpha, \beta \geq 0$$

Determine el estimador de momentos de α y β, utilizando $\frac{1}{n}\sum_{i=1}^n X_i$ y $\frac{1}{n}\sum_{i=1}^n log(X_i)$

Solución:

i) $E(X_i) = \bar{X}$ donde $E(X_i) = \frac{\alpha}{\beta}$

Así $\bar{X} = \dfrac{\alpha}{\beta}$ (2,0)

ii) $E(logX_i) = \frac{1}{n}\sum_{i=1}^n log(X_i)$ (0,5)

Para obtener $E(logX_i)$ utilizaremos la extensión de Taylor en torno a $\frac{\alpha}{\beta}$, y aplicaremos el valor esperado.

Asi,

$$g(X_i) = g(\tfrac{\alpha}{\beta}) + g'(\tfrac{\alpha}{\beta})(X_i - \tfrac{\alpha}{\beta}) + \tfrac{1}{2}g''(\tfrac{\alpha}{\beta})(X_i - \tfrac{\alpha}{\beta})^2$$

$$g(X_i) = log(X_i) \implies \frac{\partial g(X_i)}{\partial X_i} = \frac{1}{X_i} \qquad \frac{\partial^2 g(X_i)}{\partial X_i^2} = -\frac{1}{X_i^2}$$ (1,5)

Así, la aproximación para la función $log(X_i)$ está dada por

$$log(X_i) \approx log(\frac{\alpha}{\beta}) + \frac{(X_i - \frac{\alpha}{\beta})}{(\frac{\alpha}{\beta})} - \frac{(X_i - \frac{\alpha}{\beta})^2}{2(\frac{\alpha}{\beta})^2}$$ (1,0)

Aplicando Esperanza

$$E(log(X_i)) = log(\frac{\alpha}{\beta}) - \frac{1}{2\alpha} \implies log(\frac{\alpha}{\beta}) - \frac{1}{2\alpha} = \frac{1}{n}\sum_{i=1}^n log(X_i)$$ (0,5)

Despejando las ecuaciones de estimación se obtiene que:

$$\hat{\alpha} = \frac{1}{2[log(\bar{X}) - \frac{1}{n}\sum_{i=1}^n log(X_i)]}$$

$$\hat{\beta} = \frac{1}{2\bar{X}[log(\bar{X}) - \frac{1}{n}\sum_{i=1}^n log(X_i)]}$$ (0,5)

+ 1 pto base