Segunda Entrega

Juan Carlos Llamas Núñez

Ejercicio 1.- Sean $(X_1 \dots X_{n_1})$ e $(Y_1 \dots Y_{n_2})$ dos muestras aleatorias simples de dos poblaciones independientes con distribuciones respectivas $Exp(\lambda_1)$ y $Exp(\lambda_2)$. Hallar un intervalo de confianza al nivel $1-\alpha$ para el cociente λ_1/λ_2.

Vamos a calcular este intervalo por el método de la cantidad pivotal, para lo cuál necesitamos una función $T(X_1 \dots X_{n_1}, Y_1 \dots Y_{n_2}, \lambda_1, \lambda_2)$ cuya distribución no dependa de λ_1 ni λ_2 y de la que podamos despejar fácilmente el cociente λ_1/λ_2.

Parece razonable comenzar intentando calcular la distribución del cociente de medias muestrales $\dfrac{\bar{X}}{\bar{Y}} = \dfrac{\frac{1}{n_1}\sum_{i=1}^{n_1} X_i}{\frac{1}{n_2}\sum_{i=1}^{n_2} Y_i}$ o

simplemente de $\dfrac{\sum_{i=1}^{n_1} X_i}{\sum_{i=1}^{n_2} Y_i}$.

Si llamamos $A = \sum_{i=1}^{n_1} X_i$ y $B = \sum_{i=1}^{n_2} Y_i$ entonces

como $X \sim Exp(\lambda_1) = Gamma(\lambda_1, 1) \Rightarrow A = \sum_{i=1}^{n_1} X_i \sim Gamma(\lambda_1, n_1)$

e $Y \sim Exp(\lambda_2) = Gamma(\lambda_2, 1) \Rightarrow B = \sum_{i=1}^{n_2} Y_i \sim Gamma(\lambda_2, n_2)$.

Estamos interesados en la distribución del cociente $\frac{A}{B}$ para lo que planteamos la transformación

$$\begin{cases} U = \dfrac{A}{B} \\ V = A \end{cases} \Rightarrow \begin{cases} A = V \\ B = \dfrac{V}{U} \end{cases}$$

Así $g(u,v) = (g_1(u,v), g_2(u,v)) = \left(v, \frac{v}{u}\right)$

y $J_g = \begin{vmatrix} \frac{\partial g_1}{\partial u}(u,v) & \frac{\partial g_1}{\partial v}(u,v) \\ \frac{\partial g_2}{\partial u}(u,v) & \frac{\partial g_2}{\partial v}(u,v) \end{vmatrix} = \begin{vmatrix} 0 & -\frac{v}{u^2} \\ 1 & \frac{1}{u} \end{vmatrix}$

Por tanto $|J_g| = \frac{|V|}{u^2}$.

Como la transformación es inyectiva, entonces:

$$f_{(u,v)}(u,v) = f_{(A,B)}(g(u,v)) \cdot |J_g| \cdot I_C(g(u,v))$$

Siendo C el conjunto donde se mueven los parámetros, que en este caso es $C = (0,\infty) \times (0,\infty)$.

Por tanto,

$$f_{(u,v)}(u,v) = f_{(A,B)}\left(v, \frac{v}{u}\right) \cdot \frac{|V|}{u^2} \cdot I_C\left(v, \frac{v}{u}\right) =$$

$$\underset{\uparrow}{=} f_A(v) \cdot f_B\left(\frac{v}{u}\right) \cdot \frac{|V|}{u^2} \cdot I_{C'}(u,v)$$

$(X_1 \ldots X_{n_1}), (Y_1 \ldots Y_{n_2})$ poblaciones independientes.

Donde $C' = (0,\infty) \times (0,\infty)$ que es la imagen de C según la transformación planteada.

Como $A \sim Gamma(\lambda_1, n_1)$

$$\Rightarrow f_A(a) = \frac{\lambda_1^{n_1}}{\Gamma(n_1)} \cdot e^{-\lambda_1 a} a^{n_1 - 1} \quad a > 0 \quad \text{(para B es análogo)}$$

se tiene que

$$f_{(u,v)}(u,v) = \left(\frac{\lambda_1^{n_1}}{\Gamma(n_1)} e^{-\lambda_1 v} v^{n_1 - 1}\right) \cdot \left(\frac{\lambda_2^{n_2}}{\Gamma(n_2)} e^{-\lambda_2 \frac{v}{u}} \cdot \left(\frac{v}{u}\right)^{n_2 - 1}\right) \cdot \frac{V}{u^2} \quad \text{con } u,v > 0.$$

Como nos interesa sólo la distribución de $U = \frac{A}{B}$, marginalizamos:

$$f_u(u) = \int_0^\infty f_{(u,v)}(u,v)\, dv = \int_0^\infty \frac{\lambda_1^{n_1}}{\Gamma(n_1)} e^{-\lambda_1 v} v^{n_1 - 1} \frac{\lambda_2}{\Gamma(n_2)} e^{-\lambda_2 \frac{v}{u}} \left(\frac{v}{u}\right)^{n_2 - 1} \frac{v}{u^2}\, dv =$$

$$= \frac{\lambda_1^{n_1}}{\Gamma(n_1)} \cdot \frac{\lambda_2^{n_2}}{\Gamma(n_2)} \cdot \frac{1}{u^{n_2+1}} \cdot \frac{\Gamma(n_2+n_1)}{\left(\lambda_1 + \frac{\lambda_2}{u}\right)^{n_2+n_1}} \int_0^\infty \frac{\left(\lambda_1 + \frac{\lambda_2}{u}\right)^{n_2+n_1}}{\Gamma(n_2+n_1)} e^{-v\left(\lambda_1 + \frac{\lambda_2}{u}\right)} v^{n_1+n_2-1}\, dv =$$

$$= \frac{\lambda_1^{n_1} \lambda_2^{n_1}}{\Gamma(n_1)\Gamma(n_2)} \cdot \frac{1}{u^{n_2+1}} \cdot \frac{\Gamma(n_2+n_1)}{\left(\lambda_1 + \frac{\lambda_2}{u}\right)^{n_2+n_1}} \quad \text{con } u > 0.$$

Podemos simplificar más aún la expresión anterior:

$$f_U(u) = \frac{\Gamma(n_2+n_1)}{\Gamma(n_1)\Gamma(n_2)} \cdot \frac{\lambda_1^{n_1}}{(\lambda_1+\frac{\lambda_2}{u})^{n_1}} \cdot \frac{\lambda_2^{n_2}}{(\lambda_1+\frac{\lambda_2}{u})^{n_2}} \cdot \frac{1}{u^{n_2}} \cdot \frac{1}{u} =$$

$$= \frac{1}{Beta(n_1,n_2)} \cdot \left(\frac{u\lambda_1}{u\lambda_1+\lambda_2}\right)^{n_1} \cdot \left(\frac{\lambda_2}{u\lambda_1+\lambda_2}\right)^{n_2} \cdot \frac{1}{u}.$$

Esta expresión hace intuir que estamos trabajando con una distribución beta, al menos haciendo cierta transformación así que si seguimos operando intentando que aparezca una beta:

$$f_U(u) = \frac{1}{Beta(n_1,n_2)} \cdot \left(\frac{u\lambda_1}{u\lambda_1+\lambda_2}\right)^{n_1-1} \left(1-\frac{u\lambda_1}{u\lambda_1+\lambda_2}\right)^{n_2-1} \cdot \frac{1}{u} \cdot \frac{u\lambda_1}{u\lambda_1+\lambda_2} \cdot \frac{\lambda_2}{u\lambda_1+\lambda_2} =$$

$$= \frac{1}{Beta(n_1,n_2)} \cdot \left(\frac{u\lambda_1}{u\lambda_1+\lambda_2}\right)^{n_1-1} \left(1-\frac{u\lambda_1}{u\lambda_1+\lambda_2}\right)^{n_2-1} \cdot \frac{\lambda_1\lambda_2}{(u\lambda_1+\lambda_2)^2}$$

Esto es claramente una distribución $Beta(n_1,n_2)$ para la variable aleatoria $W = h(U) = \frac{U\lambda_1}{U\lambda_1+\lambda_2} = \frac{1}{1+\frac{\lambda_2}{\lambda_1 U}}$.

Efectivamente, si $u \in (0,\infty)$ entonces $w = h(u) = \frac{1}{1+\frac{\lambda_2}{\lambda_1 u}} \in (0,1)$,

$$F_W(w) = P(W \le w) = P(h(U) \le w) = P(U \le h^{-1}(w)) = F_U(h^{-1}(w))$$

donde $h^{-1}(w) = \frac{\lambda_2}{\lambda_1} \frac{w}{1-w}$.

Así $f_W(w) = f_U(h^{-1}(w)) \cdot |(h^{-1}(w))'| =$

$$= \frac{1}{Beta(n_1,n_2)} \cdot (h(h^{-1}(w)))^{n_1-1} (1-h(h^{-1}(w)))^{n_2-1} \cdot \underbrace{\frac{\lambda_1\lambda_2}{\left(\frac{\lambda_2}{\lambda_1}\frac{w}{1-w}\lambda_1+\lambda_2\right)^2} \cdot \frac{\lambda_2}{\lambda_1} \cdot \frac{1-w+w}{(1-w)^2}}_{K}$$

$$= \frac{1}{Beta(n_1,n_2)} \, w^{n_1-1}(1-w)^{n_2-1} \cdot K$$

Calculamos K:

$$K = \frac{\lambda_1 \lambda_2}{\left(\frac{\lambda_2}{\lambda_1}\frac{w}{1-w}\lambda_1 + \lambda_2\right)^2} \cdot \frac{\lambda_2}{\lambda_1} \cdot \frac{1}{(1-w)^2} = \frac{\lambda_2^2}{\left(\lambda_2\left(\frac{w}{1-w}+1\right)\right)^2} \cdot \frac{1}{(1-w)^2} =$$

$$= \frac{1}{\left(\frac{w+1-w}{1-w}\right)^2} \cdot \frac{1}{(1-w)^2} = 1$$

Por tanto $f_W(w) = \frac{1}{Beta(n_1, n_2)} \cdot w^{n_1-1}(1-w)^{n_2-1}$ con $w \in (0,1)$, es decir, $W \sim Beta(n_1, n_2)$.

Esta va a ser nuestra cantidad pivotal porque recordemos que

$$W = \frac{1}{1+\frac{\lambda_2}{\lambda_1 u}} = \frac{1}{1+\frac{\lambda_2}{\lambda_1 A/B}} = \frac{1}{1+\frac{\lambda_2 B}{\lambda_1 A}} = \frac{1}{1+\frac{\lambda_2 \sum_{i=1}^{n_2} Y_i}{\lambda_1 \sum_{i=1}^{n_1} X_i}} =$$

$$= \frac{1}{1+\frac{\lambda_2 n_2 \overline{Y}}{\lambda_1 n_1 \overline{X}}}$$ y la distribución ~~de~~ de esta cantidad no depende de λ_1 ni λ_2.

Por tanto nos podemos construir nuestro intervalo de nivel de confianza $1-\alpha$ como

$$P(a \leq W \leq b) = 1-\alpha$$

$$a \leq W \leq b \Leftrightarrow a \leq \frac{1}{1+\frac{\lambda_2 n_2 \overline{Y}}{\lambda_1 n_1 \overline{X}}} \leq b \Leftrightarrow \frac{1}{a} \geq 1 + \frac{\lambda_2 n_2 \overline{Y}}{\lambda_1 n_1 \overline{X}} \geq \frac{1}{b} \Leftrightarrow$$

$$\Leftrightarrow \frac{1}{a} - 1 \geq \frac{\lambda_2 n_2 \overline{Y}}{\lambda_1 n_1 \overline{X}} \geq \frac{1}{b} - 1 \Leftrightarrow \frac{1-a}{a} \geq \frac{\lambda_2 n_2 \overline{Y}}{\lambda_1 n_1 \overline{X}} \geq \frac{1-b}{b} \Leftrightarrow$$

$$\Leftrightarrow \frac{a}{1-a} \leq \frac{\lambda_1 n_1 \overline{X}}{\lambda_2 n_2 \overline{Y}} \leq \frac{b}{1-b} \Leftrightarrow \frac{a}{1-a} \cdot \frac{n_2 \overline{Y}}{n_1 \overline{X}} \leq \frac{\lambda_1}{\lambda_2} \leq \frac{b}{1-b} \cdot \frac{n_2 \overline{Y}}{n_1 \overline{X}}$$

Por tanto si a y b son los puntos críticos:

$$IC_{1-\alpha}\left(\frac{\lambda_1}{\lambda_2}\right) = \left(\frac{a\, n_2 \overline{Y}}{(1-a)\, n_1 \overline{X}} , \frac{b\, n_2 \overline{Y}}{(1-b)\, n_1 \overline{X}}\right)$$

Si particularizamos para el caso de probabilidad de colas iguales entonces $a = \beta_{n_1,n_2;1-\alpha/2}$ y $b = \beta_{n_1,n_2;\alpha/2}$ con

$$F_{Beta(n_1,n_2)}(\beta_{n_1,n_2;1-\alpha/2}) = \alpha/2 \quad y \quad F_{Beta(n_1,n_2)}(\beta_{n_1,n_2;\alpha/2}) = 1-\frac{\alpha}{2}.$$

En este caso el intervalo de confianza queda:

$$IC_{1-\alpha}\left(\frac{\lambda_1}{\lambda_2}\right) = \left(\frac{\beta_{n_1,n_2;1-\alpha/2}\cdot n_2 \overline{Y}}{(1-\beta_{n_1,n_2;1-\alpha/2})\, n_1 \overline{X}} , \frac{\beta_{n_1,n_2;\alpha/2}\cdot n_2 \overline{Y}}{(1-\beta_{n_1,n_2;\alpha/2})\, n_1 \overline{X}}\right)$$

Ejercicio 2: Sea $(X_1 -- X_n)$ una muestra aleatoria simple de una población con función de densidad $f_\theta(x) = \frac{2}{\theta^2}(\theta-x)\, I_{(0,\theta)}(x)$. Hallar una cantidad pivotal basada en el estadístico $T = X_{(n)}$ y utilizarla para encontrar un intervalo de confianza con probabilidad de colas iguales para θ al nivel de confianza $1-\alpha$.

Primero vamos a calcular la distribución de $X_{(n)}$.

$$F_{X_{(n)}}(y) = P\{X_{(n)} \le y\} = P\{X_1 \le y, X_2 \le y, --, X_n \le y\} = \prod_{i=1}^{n} P\{X_i \le y\} =$$

$$= F_X(y)^n$$

Así $f_{X_{(n)}}(y) = n\, F_X(y)^{n-1} \cdot f_X(y)$

Tenemos que calcular ahora $F_X(x)$ y como tenemos $f_\theta(x)$ entonces

$$F_X(x) = \begin{cases} 0 & si\ x<0 \\ \int_0^x f_\theta(t)dt & si\ x\in[0,\theta) \\ 1 & si\ x \ge \theta. \end{cases} \quad ; \int_0^x \frac{2}{\theta^2}(\theta-t)dt = \frac{2}{\theta}x - \frac{2x^2}{2\theta^2}$$

Por tanto

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ \dfrac{2\theta x - x^2}{\theta^2} & \text{si } x \in [0,\theta) \\ 1 & \text{si } x \geq \theta \end{cases}$$

Entonces la función de densidad de $X_{(n)}$ será

$$f_{X_{(n)}}(y) = n\, F_X(y)^{n-1} f_X(y) = n \cdot \left(\frac{2\theta y - y^2}{\theta^2}\right)^{n-1} \cdot \frac{2}{\theta^2} \cdot (\theta - y) \cdot I_{(0,\theta)}(y) =$$

$$= \frac{n}{\theta^{2n}} \left(2\theta y - y^2\right)^{n-1} \cdot (2\theta - 2y) \cdot I_{(0,\theta)}(y).$$

Está distribución depende de θ por lo que vamos a hacer una transformación para intentar encontrar una cantidad pivotal. Como $2\theta - 2y$ es la derivada de $2\theta y - y^2$ respecto a y podemos probar calculando la distribución de $g(X_{(n)}) = W = 2\theta X_{(n)} - X_{(n)}^2$

$$F_W(w) = P\{W \leq w\} = P\{2\theta X_{(n)} - X_{(n)}^2 \leq w\} = P\{X_{(n)}^2 - 2\theta X_{(n)} + w \geq 0\} =$$

$$= \begin{cases} 1, & \text{si } (2\theta)^2 - 4w < 0 \\ P\left\{X_{(n)} \in \left(-\infty, \dfrac{2\theta - \sqrt{(2\theta)^2 - 4w}}{2}\right] \cup \left[\dfrac{2\theta + \sqrt{(2\theta)^2 - 4w}}{2}, \infty\right)\right\} & \text{c.c.} \end{cases} =$$

$$= \begin{cases} 1 & \text{si } \theta^2 < w \\ P\left\{X_{(n)} \leq \dfrac{2\theta - \sqrt{4\theta^2 - 4w}}{2}\right\} + P\left\{\dfrac{2\theta + \sqrt{4\theta^2 - 4w}}{2} \leq X_{(n)}\right\} & \text{c.c} \end{cases} =$$

$$= \begin{cases} 1 & \text{si } \theta^2 < w \\ F_{X_{(n)}}\left(\theta - \sqrt{\theta^2 - w}\right) + 1 - F_{X_{(n)}}\left(\theta + \sqrt{\theta^2 - w}\right) & \text{si } \theta^2 \geq w \end{cases}$$

Por tanto la función de densidad de $W = 2\theta X_{(n)} - X_{(n)}^2$ es

$$f_W(w) = \left(f_{X_{(n)}}\left(\theta - \sqrt{\theta^2 - w}\right) \cdot \frac{1}{2\sqrt{\theta^2 - w}} + f_{X_{(n)}}\left(\theta + \sqrt{\theta^2 - w}\right) \cdot \frac{1}{2\sqrt{\theta^2 - w}} \right) \cdot I_{(-\infty, \theta^2]}(w) =$$

Si calculamos por separado cada uno de los términos

$$f_{X_{(n)}}\left(\theta - \sqrt{\theta^2 - w}\right) = \frac{n}{\theta^{2n}} \left(2\theta\left(\theta - \sqrt{\theta^2 - w}\right) - \left(\theta - \sqrt{\theta^2 - w}\right)^2\right)^{n-1} \left(2\theta - 2\left(\theta - \sqrt{\theta^2 - w}\right)\right) \cdot I_{(0,\theta)}\left(\theta - \sqrt{\theta^2 - w}\right) =$$

$$= \frac{n}{\theta^{2n}} \left(2\theta^2 - 2\theta\sqrt{\theta^2 - w} - \theta^2 + 2\theta\sqrt{\theta^2 - w} - \theta^2 + w\right)^{n-1} \left(2\theta - 2\theta + 2\sqrt{\theta^2 - w}\right) \cdot I_{(0,\theta)}\left(\theta - \sqrt{\theta^2 - w}\right) =$$

$$= \frac{n}{\theta^{2n}} w^{n-1} \cdot 2\sqrt{\theta^2 - w} \cdot I_{(0,\theta)}\left(\theta - \sqrt{\theta^2 - w}\right).$$

La condición $0 < \theta - \sqrt{\theta^2 - w} < \theta$ equivale a $\sqrt{\theta^2 - w} < \theta \Leftrightarrow$

$\Leftrightarrow \theta^2 - w < \theta^2 \Leftrightarrow w > 0$.

Por tanto $f_{X_{(n)}}\left(\theta - \sqrt{\theta^2 - w}\right) = \frac{n}{\theta^{2n}} w^{n-1} 2\sqrt{\theta^2 - w} \cdot I_{(0,\infty)}(w)$.

El segundo caso es mucho más fácil ya que

$$f_{X_{(n)}}\left(\theta + \sqrt{\theta^2 - w}\right) = 0 \quad \text{porque} \quad \theta + \sqrt{\theta^2 - w} \notin (0, \theta).$$

En conclusión

$$f_W(w) = f_{X_{(n)}}\left(\theta - \sqrt{\theta^2 - w}\right) \frac{1}{2\sqrt{\theta^2 - w}} \cdot I_{(-\infty, \theta^2]}(w) = \frac{n}{\theta^{2n}} w^{n-1} \frac{2\sqrt{\theta^2 - w}}{2\sqrt{\theta^2 - w}} \cdot I_{(0,\theta^2)}(w) =$$

$$= \frac{n}{\theta^{2n}} w^{n-1} I_{(0,\theta^2)}(w) = \frac{n}{\theta^2} \left(\frac{w}{\theta^2}\right)^{n-1} I_{(0,\theta^2)}(w).$$

Esta distribución es mucho más manejable pero aún depende de θ por lo que consideramos la distribución de $S = \frac{W}{\theta^2}$

$$F_S(s) = P\{S \leq s\} = P\left\{\frac{W}{\theta^2} \leq s\right\} = P\{W \leq s\theta^2\} = F_W(s\theta^2).$$

La función de densidad de S será:

$$F_S(s) = \theta^2 \cdot f_W(\theta^2 s) = \theta^2 \cdot \frac{n}{\theta^2} \cdot \left(\frac{\theta^2 s}{\theta^2}\right)^{n-1} \cdot I_{(0,\theta^2)}(\theta^2 s) =$$

$$= n\, s^{n-1}\, I_{(0,\theta^2)}(\theta^2 s) = n\, s^{n-1}\, I_{(0,1)}(s)$$ ya que la condición

$\theta^2 s \in (0,\theta^2)$ equivale a $s \in (0,1)$.

Esta es ya una distribución que no depende de θ y además es conocida (es una distribución beta de parámetros $\alpha = n$ y $\beta = 1$).

Entonces $S = \frac{W}{\theta^2} = \frac{2\theta X_{(n)} - X_{(n)}^2}{\theta^2} \sim Beta(\alpha = n, \beta = 1)$

Podemos tomar S como cantidad pivotal y si queremos calcular ahora un intervalo de confianza para θ con probabilidad de colas iguales al nivel de confianza $1-\alpha$ prodecemos de la siguiente forma:

$$P\{a \leq S \leq b\} = 1-\alpha.$$

Como nos piden que las probabilidades de las colas sean iguales

$a = \beta_{n,1:1-\alpha/2}$ con $F_{Beta(n,1)}(\beta_{n,1:1-\alpha/2}) = \frac{\alpha}{2}$ y

$b = \beta_{n,1:\alpha/2}$ con $F_{Beta(n,1)}(\beta_{n,1:\alpha/2}) = 1-\alpha/2$

Ahora tenemos que construir el intervalo para θ:

$$a \leq S \leq b \Leftrightarrow a \leq \frac{2\theta X_{(n)} - X_{(n)}^2}{\theta^2} \leq b \Leftrightarrow a\theta^2 \leq 2\theta X_{(n)} - X_{(n)}^2 \leq b\theta^2.$$

$$\Leftrightarrow a\theta^2 - 2X_{(n)}\theta + X_{(n)}^2 \leq 0 \quad (1)$$

y

$$b\theta^2 - 2X_{(n)}\theta + X_{(n)}^2 \geq 0. \quad (2)$$

Ahora, la desigualdad (1) no se cumple para ningún θ si el discriminante es negativo y la desigualdad (2) se cumple para todo θ si el discriminante es positivo.

$$\Delta_1 = 4X_{(n)}^2 - 4aX_{(n)}^2 = 4X_{(n)}^2(1-a)$$

$$\Delta_2 = 4X_{(n)}^2 - 4bX_{(n)}^2 = 4X_{(n)}^2(1-b)$$

Como a y b son valores del dominio de la distribución beta, entonces $a, b \in (0,1)$ y $\Delta_1 > 0$, $\Delta_2 > 0$.

Esto quiere decir que la desigualdad (2) es cierta $\forall\theta$ por lo que sólo consideramos la restricción dada por la desigualdad (1) que equivale a que $\theta \in \left(\frac{2X_{(n)} - \sqrt{4X_{(n)}^2(1-a)}}{2a}, \frac{2X_{(n)} + \sqrt{4X_{(n)}^2(1-a)}}{2a} \right)$

o simplificando $\theta \in \left(X_{(n)} \frac{1-\sqrt{1-a}}{a}, X_{(n)} \frac{1+\sqrt{1-a}}{a} \right)$

y sustituyendo a por $\beta_{n,1:1-\alpha/2}$, el intervalo de confianza de probabilidad de colas iguales queda:

$$\boxed{IC_{1-\alpha}(\theta) = \left(X_{(n)} \frac{1-\sqrt{1-\beta_{n,1:1-\alpha/2}}}{\beta_{n,1:1-\alpha/2}}, X_{(n)} \frac{1+\sqrt{1-\beta_{n,1:1-\alpha/2}}}{\beta_{n,1:1-\alpha/2}} \right)}$$

Ejercicio 3.- Sea $(X_1, \dots X_n)$ una muestra aleatoria simple de $X \sim f_\theta(x) = \theta x^{\theta-1} I_{(0,1)}(x)$ con $\theta > 0$. Encontrar un intervalo de confianza de longitud mínima para θ, al nivel de confianza $1-\alpha$.

La función de distribución de X es:

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ \int_0^x f_\theta(t)dt & \text{si } x \in [0,1) \\ 1 & \text{si } x \geq 1 \end{cases} \quad \text{y} \int_0^x \theta t^{\theta-1} dt = x^\theta$$

$$\Rightarrow F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ x^\theta & \text{si } x \in [0,1) \\ 1 & \text{si } x \geq 1 \end{cases}$$

Como la función de distribución, en función de θ, es continua y monótona, entonces,

$$T(X_1 \cdots X_n, \theta) = -2 \sum_{i=1}^{n} Ln(F_\theta(X_i)) \sim \chi^2_{2n}$$

Por tanto $T = -2 \sum_{i=1}^{n} Ln(X_i^\theta) = -2\theta \sum_{i=1}^{n} Ln(X_i) \sim \chi^2_{2n}$.

Podemos escribir

$$P\{a \leq -2\theta \sum_{i=1}^{n} Ln(X_i) \leq b\} = 1-\alpha$$

Es decir, $F_{\chi^2_{2n}}(b) - F_{\chi^2_{2n}}(a) = 1-\alpha$

$$a \leq -2\theta \sum_{i=1}^{n} Ln(X_i) \leq b \iff -\frac{a}{2\sum_{i=1}^{n} Ln(X_i)} \leq \theta \leq -\frac{b}{2\sum_{i=1}^{n} Ln(X_i)}$$

$-\sum Ln X_i > 0$ porque $X_i \in (0,1)$

Como este intervalo tiene que ser de longitud mínima, hay que minimizar la función $L(a,b) = -\frac{b}{2\sum_{i=1}^{n} Ln(X_i)} + \frac{a}{2\sum Ln(X_i)} =$

$$= \frac{1}{2\sum_{i=1}^{n} Ln(X_i)} (a-b).$$

Además tenemos la restricción $F_{\chi^2_{2n}}(b) - F_{\chi^2_{2n}}(a) = 1-\alpha$, es decir, $F_{\chi^2_{2n}}(b) - F_{\chi^2_{2n}}(a) - 1 + \alpha = 0$.

Si consideramos la función $g(a,b) = F_{\chi^2_{2n}}(b) - F_{\chi^2_{2n}}(a) - 1 + \alpha$ podemos aplicar el Teorema de los multiplicadores de Lagrange a la función $L(a,b)$ con la restricción $g(a,b) = 0$.

$\nabla g(a,b) = \left(-f_{\chi^2_{2n}}(a), f_{\chi^2_{2n}}(b)\right)$ que es linealmente independiente y

$$\nabla L(a,b) = \left(\frac{1}{2\sum_{i=1}^{n} Ln(X_i)}, -\frac{1}{2\sum_{i=1}^{n} Ln(X_i)} \right)$$

$\Rightarrow \nabla L(a,b) = \lambda \nabla g(a,b)$, esto es:

$$\begin{cases} +\dfrac{1}{2\sum_{i=1}^{n} Ln(X_i)} = -\lambda f_{\chi^2_{2n}}(a) \\ -\dfrac{1}{2\sum_{i=1}^{n} Ln(X_i)} = \lambda f_{\chi^2_{2n}}(b) \\ F_{\chi^2_{2n}}(b) - F_{\chi^2_{2n}}(a) - 1 + \alpha = 0 \end{cases} \Rightarrow \begin{matrix} f_{\chi^2_{2n}}(a) = f_{\chi^2_{2n}}(b) \\ y \\ F_{\chi^2_{2n}}(b) - F_{\chi^2_{2n}}(a) - 1 + \alpha = 0 \end{matrix}$$

Por tanto para un α dado existirán unos únicos $a_0, b_0 > 0$ que verifiquen $f_{\chi^2_{2n}}(a_0) = f_{\chi^2_{2n}}(b_0)$ y $F_{\chi^2_{2n}}(a_0) - F_{\chi^2_{2n}}(b_0) = 1 - \alpha$.
Estos a_0 y b_0 determinan el intervalo de confianza de longitud mínima que será:

$$\boxed{IC_{1-\alpha}(\theta) = \left(-\frac{a_0}{2\sum_{i=1}^{n} Ln(X_i)}, -\frac{b_0}{2\sum_{i=1}^{n} Ln(X_i)} \right)}$$

Ejercicio 4.- Sea $(X_1 -- X_n)$ una muestra aleatoria simple de $X = U^{\beta}$ donde $U \sim U(0,1)$ y $\beta > 0$ es un parámetro desconocido.. Obtener un intervalo de confianza al nivel de confianza $1-\alpha$ basado en el ECUMV para β.

Veamos cuál es la distribución de X.

$$F_X(x) = P\{X \leq x\} = P\{U^{\beta} \leq x\} = P\{U \leq x^{1/\beta}\} = F_U(x^{1/\beta})$$

$$f_X(x) = f_U(x^{1/\beta}) \cdot \frac{1}{\beta} \cdot x^{\frac{1}{\beta}-1} = \frac{1}{\beta} x^{\frac{1}{\beta}-1} \cdot I_{(0,1)}(x)$$, que es una ~~d~~función de densidad de una distribución $Beta(\frac{1}{\beta}, 1)$.

Calculemos ahora el ECUMV.

$$f(x_1 -- x_n | \beta) = \prod_{i=1}^{n} f(x_i|\beta) = \prod_{i=1}^{n} \frac{1}{\beta} x_i^{\frac{1}{\beta}-1} = \frac{1}{\beta^n} \left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{\beta}-1} =$$

$$= \frac{1}{\beta^n} e^{(\frac{1}{\beta}-1)\sum_{i=1}^{n} Ln x_i}$$

Por el teorema de factorización, $S(X_1 -- X_n) = \sum_{i=1}^{n} Ln X_i$ es un estadístico suficiente. Además, la imagen de

$$g : (0,\infty) \longrightarrow \mathbb{R}, \quad \beta \longrightarrow g(\beta) = \frac{1}{\beta} - 1$$

contiene un rectángulo abierto de $\mathbb{R}$, por lo que S es además completo (estamos en la familia exponencial uniparamétrica).

Para construir el ECUMV podemos intentar calcular primero la esperanza de $S(X_1 -- X_n) = \sum_{i=1}^{n} Ln X_i$

$$E[S] = E\left[\sum_{i=1}^{n} Ln X_i\right] = \sum_{i=1}^{n} E[Ln X_i] = n E[Ln X_i].$$

Ahora,

$$E[\ln X] = \int_0^1 \ln x \cdot f(x|\beta)\,dx = \int_0^1 \ln x \cdot \frac{1}{\beta} x^{\frac{1}{\beta}-1} dx =$$

$$= \ln x \cdot x^{1/\beta}\Big]_0^1 - \int_0^1 x^{1/\beta} \frac{1}{x}\,dx =$$

$\ln x = u \quad \frac{1}{x}dx = du$

$\frac{1}{\beta} x^{\frac{1}{\beta}-1} dx = dv$

$x^{1/\beta} = v$

$$= x^{1/\beta} \ln x\Big]_0^1 - \beta x^{1/\beta}\Big]_0^1 =$$

$$= 0 - \lim_{x\to 0} x^{1/\beta}\ln x - \beta + 0 = -\beta$$

Por tanto $E\left[\sum_{i=1}^n \ln X_i\right] = n \cdot E[\ln X] = -n\beta$

Si llamamos T a $T(X_1 \dots X_n) = -\frac{1}{n}\sum_{i=1}^n \ln X_i$ entonces

$E[T] = -\frac{1}{n} E\left[\sum_{i=1}^n \ln X_i\right] = \beta$, es decir, T es insesgado para β.

Además, T es función de S (un estadístico suficiente y completo).

Podemos concluir entonces que $T(X_1 \dots X_n) = -\frac{1}{n}\sum_{i=1}^n \ln X_i$ es el ECUMV.

Veamos cuál es la distribución de T. Primero calculamos la distribución de $Y = \ln\left(\frac{1}{X}\right)$

$$F_Y(y) = P\{Y \le y\} = P\{\ln\left(\frac{1}{X}\right) \le y\} = P\{\frac{1}{X} \le e^y\} = P\{e^{-y} \le X\} =$$

$$= 1 - P\{X \le e^{-y}\} = 1 - F_X(e^{-y})$$

$$f_Y(y) = + f_X(e^{-y}) \cdot e^{-y} = \frac{1}{\beta} \cdot (e^{-y})^{\frac{1}{\beta}-1} \cdot e^{-y} \cdot I_{(0,1)}(e^{-y}) =$$

$$= \frac{1}{\beta} e^{-\frac{y}{\beta}} \cdot I_{(0,\infty)}(y)$$ ya que $0 < e^{-y} < 1 \Leftrightarrow -y < 0 \Leftrightarrow y > 0$

Por tanto $Y \sim Gamma\left(\frac{1}{\lambda}, 1\right)$

Además, $-\sum_{i=1}^{n} Ln(X_i) = \sum_{i=1}^{n} Ln\left(\frac{1}{X_i}\right) = \sum_{i=1}^{n} Y_i \sim Gamma\left(\frac{1}{\lambda}, n\right)$

nT

Por último $\frac{2}{\lambda} nT \sim Gamma\left(\frac{1}{\lambda} \cdot \frac{\lambda}{2}, \frac{2n}{2}\right) = Gamma\left(\frac{1}{2}, \frac{2n}{2}\right) = \chi^2_{2n}$

Por tanto $\frac{2}{\lambda} nT$ es una cantidad pivotal ya que su distribución es una χ^2_{2n} que no depende de λ y, además, la χ^2_{2n} está tabulada.

$$\Rightarrow P\left\{a \leq \frac{2}{\lambda} nT \leq b\right\} = 1-\alpha$$

$$a \leq \frac{2}{\lambda} nT \leq b \iff \frac{1}{a} \geq \frac{\lambda}{2nT} \geq \frac{1}{b} \iff$$

$$\iff \frac{2nT}{a} \geq \lambda \geq \frac{2nT}{b} \iff \frac{2nT}{b} \leq \lambda \leq \frac{2nT}{a}.$$

En general, a y b verifican que $F_{\chi^2_{2n}}(b) - F_{\chi^2_{2n}}(a) = 1-\alpha$ pero podemos particularizar para el caso concreto de probabilidad de colas iguales. En este caso

$a = \chi^2_{2n:1-\alpha/2}$ con $F_{\chi^2_{2n}}\left(\chi^2_{2n:1-\alpha/2}\right) = \alpha/2$ y

$b = \chi^2_{2n:\alpha/2}$ con $F_{\chi^2_{2n}}\left(\chi^2_{2n:\alpha/2}\right) = 1-\alpha/2$.

$$\Rightarrow \boxed{IC_{1-\alpha}(\lambda) = \left(\frac{2nT}{\chi^2_{2n:\alpha/2}}, \frac{2nT}{\chi^2_{2n:\alpha/2-1}}\right)}$$ donde recordemos que

$T(X_1 \dots X_n) = -\frac{1}{n}\sum_{i=1}^{n} Ln(X_i)$ es el ECUMV.

Ejercicio 5.- Sea $(X_1 - X_n)$ una muestra aleatoria simple de $X \sim f_\theta(x) = \theta e^{-\theta x} \cdot I_{(0,\infty)}(x)$, $\theta > 0$. Construir un intervalo de confianza de longitud mínima al nivel de confianza $1-\alpha$ para la media poblacional.

Como $X \sim Exp(\theta) = Gamma(\theta, 1) \Rightarrow E[X] = \frac{1}{\theta}$.

Vamos a construir una cantidad pivotal.

$$\sum_{i=1}^{n} X_i \sim Gamma(\theta, n) \quad y \quad 2\theta \sum_{i=1}^{n} X_i \sim Gamma\left(\frac{1}{2}, \frac{2n}{2}\right) = \chi^2_{2n}$$

Así, $P\{a \leq 2\theta n \bar{X} \leq b\} = 1-\alpha$

$$a \leq 2\theta n \bar{X} \leq b \iff \frac{1}{a} \geq \frac{1}{2\theta n \bar{X}} \geq \frac{1}{b} \iff$$

$$\iff \frac{2n\bar{X}}{b} \leq \frac{1}{\theta} \leq \frac{2n\bar{X}}{a}$$

Como nos piden que la longitud del intervalo sea mínima, hay que minimizar la función $L(a,b) = \frac{2n\bar{X}}{a} - \frac{2n\bar{X}}{b}$ con la restricción $g(a,b) = 0$ siendo $g(a,b) = F_{\chi^2_{2n}}(b) - F_{\chi^2_{2n}}(a) - 1 + \alpha$ ya que la condición $g(a,b) = 0$ equivale a $F_{\chi^2_{2n}}(b) - F_{\chi^2_{2n}}(a) = 1-\alpha$.

$$L(a,b) = 2n\bar{X}\left(\frac{1}{a} - \frac{1}{b}\right)$$

$$\nabla L(a,b) = \left(-\frac{2n\bar{X}}{a^2}, \frac{2n\bar{X}}{b^2}\right) \quad y \quad \nabla g(a,b) = \left(-f_{\chi^2_{2n}}(a), f_{\chi^2_{2n}}(b)\right)$$

que es linealmente independiente.

Podemos aplicar el Teorema de los multiplicadores de Lagrange

$$\Rightarrow \nabla L(a,b) = \lambda \nabla g(a,b)$$

$$\Rightarrow \left\{ \begin{array}{l} -\dfrac{2n\bar{X}}{a^2} = -\lambda f_{\chi^2_{2n}}(a) \\ \dfrac{2n\bar{X}}{b^2} = \lambda f_{\chi^2_{2n}}(b) \\ F_{\chi^2_{2n}}(b) - F_{\chi^2_{2n}}(a) = 1-\alpha \end{array} \right. \Rightarrow b^2 f_{\chi^2_{2n}}(b) = a^2 f_{\chi^2_{2n}}(a)$$

Por tanto, $L(a,b)$ alcanza un mínimo para aquellos valores $a_0, b_0 > 0$ que verifiquen que $F_{\chi^2_{2n}}(b_0) - F_{\chi^2_{2n}}(a_0) = 1-\alpha$ y $a_0^2 f_{\chi^2_{2n}}(a_0) = b_0^2 f_{\chi^2_{2n}}(b_0)$.

Entonces el intervalo de confianza de longitud mínima al nivel de confianza $1-\alpha$ para $\frac{1}{\theta}$ es:

$$\boxed{IC_{1-\alpha}\left(\frac{1}{\theta}\right) = \left(\frac{2n\bar{X}}{b_0}, \frac{2n\bar{X}}{a_0}\right)}$$

Ejercicio 6.- Sea $(X_1 -- X_n)$ una muestra aleatoria simple de $X \sim f_\theta(x) = e^{-(x-\theta)} I_{(\theta,\infty)}(x)$. $\theta > 0$. Encontrar el intervalo Bayesiano de máxima densidad a posteriori al nivel $1-\alpha$ si la distribución a priori es $\pi(\theta) = e^{-\theta} I_{(0,\infty)}(\theta)$

La distribución a posteriori de θ es

$$\pi(\theta | x_1 -- x_n) = \frac{\pi(\theta) \cdot f(x_1 -- x_n | \theta)}{\int_{\Theta} \pi(\theta) \cdot f(x_1 -- x_n | \theta)\, d\theta}$$ donde

$$f(x_1 -- x_n | \theta) = \prod_{i=1}^{n} f(x_i | \theta) = \prod_{i=1}^{n} e^{-(x_i-\theta)} I_{(\theta,\infty)}(x_i) = e^{n\theta} \cdot e^{-\sum_{i=1}^{n} x_i} \cdot I_{(\theta,\infty)}(x_{(1)})$$

$$= e^{n\theta} \cdot e^{-\sum_{i=1}^{n} x_i} I_{(-\infty, x_{(1)})}(\theta)$$

De esta manera

$$\pi(\theta|x_1 \dots x_n) = \frac{e^{-\theta} I_{(0,\infty)}(\theta) \cdot e^{n\theta} \cdot \cancel{e^{-\sum_{i=1}^n x_i}} \cdot I_{(-\infty, x_{(1)})}(\theta)}{\int_0^\infty e^{-\theta} e^{n\theta} \cancel{e^{-\sum_{i=1}^n x_i}} I_{(-\infty, x_{(1)})}(\theta)\, d\theta} =$$

$$= \frac{e^{-(1-n)\theta} \quad I_{(0,x_{(1)})}(\theta)}{\int_0^{x_{(1)}} e^{-(1-n)\theta}\, d\theta} \underset{n\neq 1}{\overset{\uparrow}{=}} \frac{e^{-(1-n)\theta}}{\left.\dfrac{e^{-(1-n)\theta}}{n-1}\right]_{\theta=0}^{\theta=x_{(1)}}} \cdot I_{(0,x_{(1)})}(\theta) =$$

$$= \frac{e^{(n-1)\theta}}{e^{-(1-n)x_{(1)}} - e^0}(n-1) \cdot I_{(0,x_{(1)})}(\theta) = \frac{e^{(n-1)\theta} \quad (n-1)}{e^{(n-1)x_{(1)}} - 1} \cdot I_{(0,x_{(1)})}(\theta)$$

$$\left(\text{Si } n=1 \Rightarrow \pi(\theta|x_1 \dots x_n) = \frac{1}{x_{(1)}} I_{(0,x_{(1)})}(\theta)\right)$$

Para el caso $n=1$, cualquier intervalo $(a,b) \subset [0, x_{(1)}]$ y que cumpla que $F_\theta(b|x_1 \dots x_n) - F_\theta(a|x_1 \dots x_n) = 1-\alpha$ tienen la misma longitud porque estamos tratando con una v.a. $\theta \sim \mathcal{U}(0, x_{(1)})$.

Si $n>1$, como la función de densidad a posteriori de θ es monótona creciente, el intervalo Bayesiano de máxima densidad (a,b) tiene que cumplir que $b = x_{(1)}$ y $F_\theta(a|x_1 \dots x_n) = \alpha$

Si calculamos explicitamente a:

$$F_\theta(a|x_1 \dots x_n) = \int_0^a \frac{e^{(n-1)\theta} \quad (n-1)}{e^{(n-1)x_{(1)}} - 1}\, d\theta = \frac{n-1}{e^{(n-1)x_{(1)}} - 1} \cdot \left.\frac{e^{(n-1)\theta}}{\cancel{n-1}}\right]_0^a =$$

$$= \frac{e^{(n-1)a} - 1}{e^{(n-1)x_{(1)}} - 1} = \alpha.$$

Si despejamos a de la ecuación anterior

$$e^{(n-1)a} = \alpha\left(e^{(n-1)x_{(1)}}-1\right)+1$$

$$\Updownarrow$$

$$(n-1)a = Ln\left(\alpha\left(e^{(n-1)x_{(1)}}-1\right)+1\right)$$

$$\Updownarrow$$

$$a = \frac{1}{n-1} Ln\left(\alpha\left(e^{(n-1)x_{(1)}}-1\right)+1\right).$$

Así, el intervalo de confianza Bayesiano al nivel de confianza $1-\alpha$ es

$$\boxed{IC_{1-\alpha}(\theta) = \left(\frac{1}{n-1} Ln\left(\alpha\left(e^{(n-1)x_{(1)}}-1\right)+1\right),\ x_{(1)}\right)}$$

y como estamos tomando una aproximación bayesiana aquí sí podemos decir que

$$P\left\{\theta\in\left(\frac{1}{n-1} Ln\left(\alpha\left(e^{(n-1)x_{(1)}}-1\right)+1\right),\ x_{(1)}\right)\right\} = 1-\alpha$$

porque θ es una variable aleatoria.