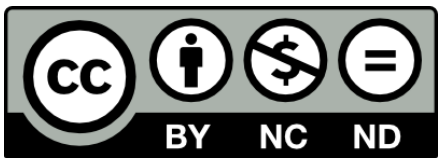
Inferencia Estadística Examen IV

Los Del DGIIM, losdeldgiim.github.io

Doble Grado en Ingeniería Informática y Matemáticas
Universidad de Granada



Inferencia Estadística Examen IV

Los Del DGIIM, `losdeldgiim.github.io`

José Manuel Sánchez Varbas

Granada, 2026

Asignatura Inferencia Estadística.

Curso Académico 2022-23.

Grado Grado en Matemáticas y Grado en Ingeniería Informática y Matemáticas.

Descripción Examen Ordinario.

Fecha 17 de Enero de 2023.

Ejercicio 1 (2,25 puntos)**.** Sea $(X_1, \ldots, X_n)$ una muestra aleatoria simple de una variable aleatoria X con distribución en una familia paramétrica.

a) Dar la definición de estadístico suficiente. Enunciar el Teorema de Factorización de Neyman–Fisher. Demostrar dicho teorema para variables discretas.

b) Si la función de distribución de X es

$$F_\theta(x) = 1 - e^{-\theta x}, \quad x > 0$$

encontrar el intervalo de confianza para θ de menor longitud media uniformemente a nivel de confianza $1 - \alpha$, basado en un estadístico suficiente.

Ejercicio 2 (2,25 puntos)**.** Sea $(X_1, \ldots, X_n)$ una muestra aleatoria simple de una variable aleatoria con función de densidad

$$f_\theta(x) = \frac{1}{2\sqrt{2\theta - 1}\sqrt{x - 1}}, \quad 1 < x < 2\theta$$

a) Sabiendo que $T = \max X_i$ es suficiente, encontrar, si existe, el UMVUE para $(2\theta - 1)^{-1}$, especificando previamente el espacio paramétrico y el espacio muestral. Justificar la no existencia del UMVUE en los casos que corresponda.

b) Calcular la función de verosimilitud y encontrar un estimador máximo verosímil de $2\theta - 1$.

Ejercicio 3 (2 puntos)**.** Sea X una variable aleatoria con distribución en una familia regular en el sentido de Fréchet–Cramér–Rao, cuyas funciones de densidad son de la forma:

$$f_\theta(x) = \exp\left\{T(x)\ln\theta - \frac{\theta^2}{2} + S(x)\right\} \quad x \in \mathbb{R}, \quad \theta \in \mathbb{R}^+$$

siendo $T(X)$ un estadístico regular.

a) Calcular la función de información asociada a X.

b) Basándose en una muestra aleatoria simple de X, $(X_1, \ldots, X_n)$, y suponiendo $T(X) > 0$, encontrar la clase de funciones paramétricas que admiten estimador eficiente y los estimadores correspondientes.

c) Bajo los supuestos del apartado b), calcular la cota inferior para la varianza de estimadores insesgados de $\ln\theta$, regulares, y justificar si se alcanza o no dicha cota.

Ejercicio 4 (2,4 puntos)**.** Contraste de hipótesis:

a) Sea $(X_1, \ldots, X_n)$ una muestra aleatoria simple de una variable aleatoria X con distribución en una familia $\{P_\theta \mid \theta \in \Theta\}$. Sea Θ_0 subconjunto arbitrario de Θ y supongamos que se pretende contrastar la hipótesis

$$H_0 : \ \theta \in \Theta_0.$$

a1) Detallar la hipótesis alternativa. Definir formalmente el concepto de test de hipótesis y dar la interpretación de sus valores.

a2) Definir el tamaño y la función de potencia de un test arbitrario para resolver el problema anterior, explicando el significado de estos conceptos en términos del rechazo de H_0.

a3) En términos del tamaño y de la función potencia, ¿qué significa que un test tiene nivel de significación α para el problema de contraste planteado? ¿Cuáles son las condiciones para que un test sea UMP a nivel de significación α?

b) Obtener un test de razón de verosimilitud de tamaño α para contrastar

$$H_0 : \theta \leq \theta_0 \qquad \text{frente a} \qquad H_1 : \theta > \theta_0,$$

basado en una observación de una variable aleatoria con la siguiente función de densidad (detallar y justificar todos los pasos para la obtención, incluyendo el estudio detallado del estadístico de contraste y su representación gráfica):

$$f_\theta(x) = \theta x^{-2} e^{-\theta/x}, \qquad x > 0.$$

¿Qué tamaños se alcanzan con dicho test?

Ejercicio 5 (1,1 puntos)**.** En cierta academia se pretende hacer un estudio sobre la efectividad de un curso intensivo de aprendizaje. Para ello, se selecciona aleatoriamente a 10 alumnos a los que se somete a una misma prueba de control antes y después del curso, obteniéndose las siguientes calificaciones:

A	7,25	7,5	7,5	8	8,5	8,75	8	8,25	7,25	8,5
D	7	8,25	8	8	8,75	9,5	8,5	8,75	7,75	8

¿Se puede concluir, a la vista de los datos, que el curso de aprendizaje no afecta a las calificaciones de los alumnos? Especificar bajo qué condiciones sobre las variables consideradas se puede resolver este problema y concretar las hipótesis nula y alternativa que se contrastan. Resolver este problema al nivel de significación 0.1 mediante el uso de diferentes tests, detallando las hipótesis necesarias para la aplicación de cada uno de ellos.

Ejercicio 1 (2,25 puntos)**.** Sea $(X_1, \ldots, X_n)$ una muestra aleatoria simple de una variable aleatoria X con distribución en una familia paramétrica.

a) Dar la definición de estadístico suficiente. Enunciar el Teorema de Factorización de Neyman–Fisher. Demostrar dicho teorema para variables discretas.

Definición 0.1 (Estadístico Suficiente)**.** Sea $(X_1, \ldots, X_n)$ una m.a.s. de $X \rightsquigarrow F \in \{F_\theta, \theta \in \Theta\}$. Un estadístico $T(X_1, \ldots, X_n)$ es *suficiente* para la familia de distribuciones considerada (o suficiente para θ) si la distribución de la muestra condicionada a cualquier valor del estadístico, $T(X_1, \ldots, X_n) = t$, es independiente de θ.

Teorema 0.1 (de Factorización de Neyman–Fisher)**.** *Sea $(X_1, \ldots, X_n)$ una m.a.s de $X \rightsquigarrow F \in \{F_\theta, \theta \in \Theta\}$. Sea f_θ la función masa de probabilidad o función de densidad de X bajo F_θ y sea f_θ^n la f.m.p. o f.d.d. de la muestra bajo F_θ. Un estadístico $T(X_1, \ldots, X_n)$ se dice que es suficiente si y solo si, para cualquier valor de $\theta \in \Theta$,*

$$f_\theta^n(x_1, \ldots, x_n) = h(x_1, \ldots, x_n) g_\theta(T(X_1, \ldots, X_n)), \quad \forall (x_1, \ldots, x_n) \in \mathcal{X}^n$$

donde h es independiente de θ y g_θ depende de $(x_1, \ldots, x_n)$ solo a través de $T(x_1, \ldots, x_n)$.

Demostración. Se demostrará por doble implicación.

$\Longrightarrow$) Supongamos que $T \equiv T(X_1, \ldots, X_n)$ es suficiente:

$$P_\theta(X_1 = x_1, \ldots, X_n = x_n) = P_\theta(X_1 = x_1, \ldots, X_n = x_n, T = T(x_1, \ldots, x_n)) =$$

$$\underbrace{P_\theta(X_1 = x_1, \ldots, X_n = x_n / T = T(x_1, \ldots, x_n))}_{h(x_1, \ldots, x_n)} \underbrace{P_\theta(T = T(x_1, \ldots, x_n))}_{g_\theta(T(x_1, \ldots, x_n))}$$

($\Longleftarrow$ Denotamos por $\mathbf{X} = (X_1, \ldots, X_n)$ y a las posibles realizaciones muestrales $\mathbf{x} = (x_1, \ldots, x_n)$. Supongamos que la función masa de probabilidad de la muestra se factoriza según se indica,

$$P_\theta(\mathbf{X} = \mathbf{x}) = h(\mathbf{x}) g_\theta(T(\mathbf{x})) \quad \forall \mathbf{x} \in \mathcal{X}^n$$

Para probar la suficiencia de T se calcula la distribución condicionada de la muestra a un valor arbitrario, $T = t$.

$$P_\theta(\mathbf{X} = \mathbf{x} / T = t) = \frac{P_\theta(\mathbf{X} = \mathbf{x}, T = t)}{P_\theta(T = t)} = \begin{cases} 0 & \text{si } T(\mathbf{x}) \neq t \\ \dfrac{P_\theta(\mathbf{X} = \mathbf{x})}{P_\theta(T = t)} & \text{si } T(\mathbf{x}) = t \end{cases}$$

Si $T(x) = t$, entonces

$$P_\theta(\mathbf{X} = \mathbf{x}) = h(\mathbf{x}) g_\theta(T(\mathbf{x})) = h(\mathbf{x}) g_\theta(t)$$

$$P_\theta(T = t) = \sum_{\mathbf{x}' / T(\mathbf{x}') = t} P_\theta(\mathbf{X} = \mathbf{x}') = \sum_{\mathbf{x}' / T(\mathbf{x}') = t} h(\mathbf{x}') g_\theta(T(\mathbf{x}')) = g_\theta(t) \sum_{\mathbf{x}' / T(\mathbf{x}') = t} h(\mathbf{x}')$$

Así, la distribución condicionada es independiente de θ, y T es suficiente por la Definición 0.1:

$$P(\mathbf{X}=\mathbf{x}/T=t)=\begin{cases}0 & \text{si } T(\mathbf{x})\neq t\\ \dfrac{h(\mathbf{x})}{\displaystyle\sum_{\mathbf{x}'/T(\mathbf{x}')=t}h(\mathbf{x}')} & \text{si } T(\mathbf{x})=t\end{cases}$$

lo que concluye la demostración. □

b) Si la función de distribución de X es

$$F_\theta(x)=1-e^{-\theta x},\quad x>0$$

encontrar el intervalo de confianza para θ de menor longitud media uniformemente a nivel de confianza $1-\alpha$, basado en un estadístico suficiente.

Vemos que $X \rightsquigarrow \text{Exp}(\theta)$, dado que $F_\theta(x)$ es la función de distribución es la de una exponencial con parámetro $\theta\in\Theta=\mathbb{R}^+$, y $\mathcal{X}=\mathbb{R}^+$. Para hallar el estadístico suficiente, usaremos el Teorema 0.1. Primero obtenemos la función de densidad

$$f_\theta(x)=\frac{\partial F_\theta(x)}{\partial x}=\theta e^{-\theta x},\quad x>0$$

Calculamos ahora la función de densidad conjunta de la muestra

$$f_\theta^n(x_1,\dots,x_n)\overset{indep.}{=}\prod_{i=1}^n f_\theta(x_i)$$

Suponemos en este punto que $\theta>0$ (de lo contrario, $f_\theta(x_i)=0\quad\forall i=1,\dots,n$), y vemos que

$$x_i>0\quad\forall i=1,\dots,n\iff I_{\mathbb{R}^+}(x_i-0)=1\quad\forall i=1,\dots,n\iff I_{\mathbb{R}^+}(x_{(1)})=1$$

de donde se deduce que

$$f_\theta^n(x_1,\dots,x_n)=\prod_{i=1}^n\theta e^{-\theta x_i}I_{\mathbb{R}^+}(x_{(1)})=I_{\mathbb{R}^+}(x_{(1)})\theta^n e^{-\theta\left(\sum_{i=1}^n x_i\right)}$$

Tomando $S(X_1,\dots,X_n)=\displaystyle\sum_{i=1}^n X_i$ y

$$h(x_1,\dots,x_n)=I_{\mathbb{R}^+}(x_{(1)}),\quad g_\theta(t)=\theta^n e^{-\theta t}$$

Se cumple que

$$f_\theta^n(x_1,\dots,x_n)=h(x_1,\dots,x_n)g_\theta(S(x_1,\dots,x_n))\quad\forall(x_1,\dots,x_n)\in\mathcal{X}^n$$

donde h es independiente del parámetro θ y g_θ depende de la muestra solo a través del estadístico, luego, por el Teorema de Factorización de Neyman-Fisher, el estadístico S es suficiente.

Sabemos que $S = \sum_{i=1}^{n} X_i \rightsquigarrow \Gamma(n, \theta)$, ya que $X_i \rightsquigarrow \text{Exp}(\theta) \quad \forall i = 1, \ldots, n$.

Ahora encontraremos el intervalo de confianza por medio del método de la cantidad pivotal. Nos piden que el IC esté basado en un estadístico suficiente, como S. Usaremos la función de densidad de la distribución $\Gamma(p, a), p, a > 0$ conocida:

$$f(x) = \frac{a^p}{\Gamma(p)} x^{p-1} e^{-ax}, \quad x > 0$$

el Teorema de Cambio de Variable, para $Y = g(X)$

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right|$$

y la relación vista en teoría entre la distribución $\Gamma(p, a)$ y la $\chi^2(n)$:

$$X \rightsquigarrow \chi^2(n) \iff X \rightsquigarrow \Gamma\left(\frac{n}{2}, \frac{1}{2}\right) \tag{1}$$

Primero consideramos el cambio $Y = \theta S \Longrightarrow S = Y/\theta$, y entonces

$$f_Y(y) = f_S(y/\theta) \cdot \frac{1}{\theta} = \frac{\theta^n}{\Gamma(n)} \left(\frac{y}{\theta}\right)^{n-1} e^{-\theta(y/\theta)} \cdot \frac{1}{\theta} = \frac{1}{\Gamma(n)} y^{n-1} e^{-y} \rightsquigarrow \Gamma(n, 1), \quad y > 0$$

Repetimos el mismo procedimiento con el cambio $W = 2Y \Longrightarrow Y = 2/W$, y entonces

$$f_W(w) = f_Y(w/2) \cdot \frac{1}{2} = \frac{1}{\Gamma(n)} \left(\frac{w}{2}\right)^{n-1} e^{-w/2} \cdot \frac{1}{2} = \frac{(1/2)^n}{\Gamma(n)} w^{n-1} e^{-w/2} \rightsquigarrow \Gamma\left(n, \frac{1}{2}\right), \quad w > 0$$

Por (1), $W \rightsquigarrow \Gamma(n, 1/2) \iff 2\theta S = W \rightsquigarrow \chi^2(2n)$ y podemos considerar como función pivote, $T : \mathcal{X}^n \times \Theta \to \mathbb{R}$, dado por

$$T \equiv T(X_1, \ldots, X_n; \theta) = 2\theta S(X_1, \ldots, X_n) \equiv 2\theta S$$

que es una variable aleatoria cuya distribución, $\chi^2(2n)$, es independiente del parámetro θ (solo depende de $n \in \mathbb{N}$). Además:

a) T es estrictamente monótona en θ, pues

$$\frac{\partial T}{\partial \theta} = \frac{\partial}{\partial \theta}(2\theta S) = 2S > 0, \quad \forall \theta \in \Theta, \quad \forall (x_1, \ldots, x_n) \in \mathcal{X}^n = (\mathbb{R}^+)^n$$

b) $T = \lambda$ tiene solución en θ, para todo $\lambda \in \Lambda = \mathbb{R}^+$, con Λ la imagen de T, pues

$$T = \lambda \iff 2\theta S = \lambda \iff \theta = \frac{\lambda}{2S} \in \mathbb{R}^+$$

Por un teorema visto en teoría, sabemos que se puede construir un intervalo de confianza para θ a cualquier nivel de confianza $1-\alpha$, con $0<\alpha<1$. Por el método de la cantidad pivotal, buscamos $\lambda_1<\lambda_2$ verificando

$$P_\theta(\lambda_1<T<\lambda_2)=1-\alpha$$

Es decir,

$$\lambda_1<T<\lambda_2 \iff \lambda_1<2\theta S<\lambda_2 \iff \frac{\lambda_1}{2S}<\theta<\frac{\lambda_2}{2S}$$

y el intervalo tendrá longitud

$$L=\frac{\lambda_2}{2S}-\frac{\lambda_1}{2S}=\frac{\lambda_2-\lambda_1}{2S}$$

que es constante, luego coincidirá con la longitud esperada $\mathrm{E}_\theta[L]=L \quad \forall\theta\in\Theta$.

Sea ahora F la función de distribución de la distribución χ^2 de Pearson $\chi^2(2n)$, y sea f la correspondiente función de densidad. La restricción es

$$1-\alpha=P_\theta(\lambda_1<T<\lambda_2)=F(\lambda_2)-F(\lambda_1)$$

Minimizamos $\lambda_2-\lambda_1$ con el método de los multiplicadores de Lagrange visto en teoría

$$H(\lambda_1,\lambda_2)=\lambda_2-\lambda_1-\lambda[F(\lambda_2)-F(\lambda_1)-(1-\alpha)]$$

donde no se ha considerado la parte constante de la longitud media por no afectar al procedimiento de minimización. Buscamos aquellos λ_1,λ_2 que minimicen $H(\lambda_1,\lambda_2)$, luego obtenemos sus derivadas parciales

$$\frac{\partial H}{\partial\lambda_1}=-1+\lambda f(\lambda_1) \quad \frac{\partial H}{\partial\lambda_2}=1-\lambda f(\lambda_2)$$

Ahora, igualamos ambas parciales a 0 y despejamos λ:

$$\left.\begin{array}{l} 0=\dfrac{\partial H}{\partial\lambda_1}=-1+\lambda f(\lambda_1) \\ 0=\dfrac{\partial H}{\partial\lambda_2}=1-\lambda f(\lambda_2) \end{array}\right\} \Longrightarrow \left.\begin{array}{l} \lambda=\dfrac{1}{f(\lambda_1)} \\ \lambda=\dfrac{1}{f(\lambda_2)} \end{array}\right\} \Longrightarrow f(\lambda_1)=f(\lambda_2)$$

En este caso, la asimetría de la distribución hace que los valores λ_1 y λ_2 no sean los de colas iguales. Sin embargo, la diferencia no es suficientemente importante (sobre todo para grandes muestras), y en la práctica se usa el intervalo de colas iguales:

$$\lambda_1=\chi^2_{2n;1-\alpha/2} \quad \lambda_2=\chi^2_{2n;\alpha/2}$$

siendo $P[\chi^2_{2n}>\chi^2_{2n;\alpha/2}]=\alpha/2$. En conclusión, el intervalo de confianza para θ de menor longitud media uniformemente a nivel de confianza $1-\alpha$, basado en un estadístico suficiente S es:

$$\left]\frac{\lambda_1}{2S},\frac{\lambda_2}{2S}\right[=\left]\frac{\chi^2_{2n;1-\alpha/2}}{2S},\frac{\chi^2_{2n;\alpha/2}}{2S}\right[$$

Ejercicio 2 (2,25 puntos)**.** Sea $(X_1, \ldots, X_n)$ una muestra aleatoria simple de una variable aleatoria con función de densidad

$$f_\theta(x) = \frac{1}{2\sqrt{2\theta-1}\sqrt{x-1}}, \quad 1 < x < 2\theta$$

a) Sabiendo que $T = \max X_i$ es suficiente, encontrar, si existe, el UMVUE para $(2\theta-1)^{-1}$, especificando previamente el espacio paramétrico y el espacio muestral. Justificar la no existencia del UMVUE en los casos que corresponda.

Buscamos obtener el UMVUE mediante el método alternativo visto en teoría. Para ello, en primer lugar hay que encontrar un estadístico suficiente y completo T, y luego una función del estadístico $h(T)$ (denotaremos indistintamente $T \overset{not}{\equiv} T(X_1, \ldots, X_n)$, para una m.a.s. $(X_1, \ldots, X_n)$ con $n \in \mathbb{N}$ fijo) insesgada en $g(\theta) = (2\theta-1)^{-1}$, estimadora y con momento de segundo orden finito. Entonces $h(T)$ será el UMVUE.

Ya sabemos que $T = X_{(n)}$ es suficiente por el enunciado. Ahora, hay que comprobar que este estadístico es completo, lo cual se hará por definición. Sabemos por teoría que la distribución del máximo es

$$F_T(t) = (F_X(t))^n \Longrightarrow f_T(t) = n(F_X(t))^{n-1} f_\theta(t)$$

Hallamos ahora la función de distribución de X:

$$F_X(t) = \int_1^t f_\theta(x)dx = \int_1^t \frac{1}{2\sqrt{2\theta-1}\sqrt{x-1}}dx = \frac{1}{2\sqrt{2\theta-1}}\int_1^t \frac{1}{\sqrt{x-1}}dx =$$

$$\frac{1}{2\sqrt{2\theta-1}} \cdot \left[2\sqrt{x-1}\right]_1^t = \frac{2\sqrt{t-1}}{2\sqrt{2\theta-1}} = \frac{\sqrt{t-1}}{\sqrt{2\theta-1}} \quad 1 < t < 2\theta$$

La función de densidad del estadístico será entonces

$$f_T(t) = n(F_X(t))^{n-1} f_\theta(t) = n\left(\frac{\sqrt{t-1}}{\sqrt{2\theta-1}}\right)^{n-1} \frac{1}{2\sqrt{2\theta-1}\sqrt{t-1}} =$$

$$\frac{(\sqrt{t-1})^{n-1}}{(\sqrt{2\theta-1})^{n-1}} \frac{n}{2\sqrt{2\theta-1}\sqrt{t-1}} = \frac{n(\sqrt{t-1})^{n-2}}{2(\sqrt{2\theta-1})^n} \quad 1 < t < 2\theta$$

Sea h una función medible verificando

$$0 = \mathrm{E}[h(T)] \overset{def}{=} \int_1^{2\theta} h(t) f_T(t)dt = \int_1^{2\theta} h(t)\frac{n(\sqrt{t-1})^{n-2}}{2(\sqrt{2\theta-1})^n}dt =$$

$$\frac{n}{2(\sqrt{2\theta-1})^n}\int_1^{2\theta} h(t)(\sqrt{t-1})^{n-2}dt$$

como $\dfrac{n}{2(\sqrt{2\theta-1})^n} \neq 0 \quad \forall n \in \mathbb{N}, \quad \forall \theta \in \Theta = \left]\dfrac{1}{2}, +\infty\right[$, debe ser

$$\int_1^{2\theta} h(t)(\sqrt{t-1})^{n-2}dt = 0$$

Por el Teorema Fundamental del Cálculo, podemos considerar una primitiva $H(t)$ del integrando $h(t)(\sqrt{t-1})^{n-2}$, y esta cumple, por la Regla de Barrow, que $H(2\theta) - H(1) = 0 \quad \forall\theta \in \Theta$. Derivando respecto de θ, se obtiene que

$$\frac{d}{d\theta}H(2\theta) = 0 \iff 2h(2\theta)(\sqrt{2\theta-1})^{n-2} = 0 \overset{(*)}{\iff} h(2\theta) = 0$$

donde en $(*)$ se ha usado que $2(\sqrt{2\theta-1})^{n-2} \neq 0$ por ser $\theta \in \Theta$ y $2 > 0$. Equivalentemente,

$$\forall\theta \in \Theta \quad h(2\theta) = 0 \iff h(t) = 0 \quad \forall t > 1$$

(tomando $t = 2\theta \in]1, +\infty[$). Por tanto

$$]1, +\infty[\subseteq \{t : h(t) = 0\}$$

y consecuentemente

$$1 \geqslant P[h(T) = 0] \geqslant P[T > 1] = 1 \Longrightarrow P[h(T) = 0] = 1$$

y entonces por definición concluimos que T es un estadístico completo. Tenemos entonces en este punto que T es un estadístico suficiente y completo.

Ahora hay que buscar un estimador insesgado en $g(\theta)$ y de segundo orden finito. Sea h (independiente de la anterior) función medible tal que

$$(2\theta-1)^{-1} = g(\theta) = \mathrm{E}[h(T)] = \frac{n}{2(\sqrt{2\theta-1})^n}\int_1^{2\theta} h(t)(\sqrt{t-1})^{n-2}dt \iff$$

$$\int_1^{2\theta} h(t)(\sqrt{t-1})^{n-2}dt = \frac{2(\sqrt{2\theta-1})^n}{n}(2\theta-1)^{-1} = \frac{2(2\theta-1)^{\frac{n}{2}-1}}{n}$$

Derivamos respecto de θ a ambos lados e igualamos. El miembro izquierdo ya lo tenemos por el apartado anterior:

$$2h(2\theta)(\sqrt{2\theta-1})^{n-2}$$

y el derecho es

$$\frac{2}{n}\left(\frac{n}{2}-1\right)(2\theta-1)^{\frac{n}{2}-2}\cdot 2 = \left(2-\frac{4}{n}\right)(2\theta-1)^{\frac{n}{2}-2} = \frac{2n-4}{n}(2\theta-1)^{\frac{n}{2}-2}$$

Despejamos $h(2\theta)$:

$$2h(2\theta)(\sqrt{2\theta-1})^{n-2} = \frac{2n-4}{n}(2\theta-1)^{\frac{n}{2}-2} \iff$$

$$h(2\theta) = \frac{2n-4}{2n}\frac{(2\theta-1)^{\frac{n}{2}-2}}{(\sqrt{2\theta-1})^{n-2}} = \frac{n-2}{n}\frac{(2\theta-1)^{\frac{n}{2}-2}}{(2\theta-1)^{(n-2)/2}} = \frac{n-2}{n}\frac{1}{2\theta-1}$$

de donde

$$h(t) = \frac{n-2}{n}\frac{1}{t-1}$$

Por construcción $h(T)$ es insesgada en $g(\theta)$. Vemos que $h(T)$ también es estimador de $g(\theta)$, pues $\Theta =]1/2, +\infty[$, y $g(\theta) = \frac{1}{2\theta-1} \Longrightarrow g(\Theta) =]0, +\infty[$. Como $T = X_{(n)} > 1$, entonces $T - 1 > 0$, y $\frac{n-2}{n} > 0$ si $n \geqslant 3$, luego $h(T) > 0$ si $n \geqslant 3$. Queda comprobar que tiene momento de segundo orden finito.

Ello se cumplirá en caso de que $E[h(T)^2] < +\infty$:

$$E[h(T)^2] \overset{def}{=} \int_1^{2\theta} h(t)^2 f_T(t)dt = \int_1^{2\theta} \left(\frac{n-2}{n}\right)^2 \frac{1}{(t-1)^2} \frac{n(\sqrt{t-1})^{n-2}}{2(\sqrt{2\theta-1})^n} dt =$$

$$\frac{(n-2)^2}{2n(2\theta-1)^{n/2}} \int_1^{2\theta} (t-1)^{(n-6)/2} dt = \frac{(n-2)^2}{2n(2\theta-1)^{n/2}} \frac{2}{(n-4)} \left[(t-1)^{(n-4)/2}\right]_1^{2\theta} =$$

$$\frac{(n-2)^2}{n(n-4)} \frac{(2\theta-1)^{(n-4)/2}}{(2\theta-1)^{n/2}} = \frac{(n-2)^2}{n(n-4)} \frac{1}{(2\theta-1)^2}$$

Y vemos que $E[h(T)^2] < +\infty \iff n > 4$, ya que si $n \leqslant 4$, el momento de segundo orden no es finito. Por tanto, por el Teorema de Lehmann-Scheffé, $E[h(T)/T] = h(T)$ es el UMVUE para $g(\theta)$, y existe siempre y cuando

$$n > 4 \iff n \geqslant 5$$

b) Calcular la función de verosimilitud y encontrar un estimador máximo verosímil de $2\theta - 1$.

Como

$$f_\theta(x) = \frac{1}{2\sqrt{2\theta-1}\sqrt{x-1}}, \quad 1 < x < 2\theta$$

Entonces

$$f_\theta^n(x_1, \ldots, x_n) \overset{indep}{=} \prod_{i=1}^n f_\theta(x_i)$$

A partir de aquí asumimos que $x_i > 1 \quad \forall i = 1, \ldots, n \iff x_{(1)} > 1$. En otro caso, $f_\theta(x_i) = 0 \quad \forall i = 1, \ldots, n$. Tenemos que

$$x_i < 2\theta \quad \forall i = 1, \ldots, n \iff x_{(n)} < 2\theta \iff I_{\mathbb{R}^-}(x_{(n)} - 2\theta) = 1$$

Entonces

$$f_\theta^n(x_1, \ldots, x_n) \overset{indep}{=} \prod_{i=1}^n f_\theta(x_i) = \prod_{i=1}^n \frac{1}{2\sqrt{2\theta-1}\sqrt{x_i-1}} I_{\mathbb{R}^-}(x_{(n)} - 2\theta) =$$

$$\frac{1}{(2\sqrt{2\theta-1})^n} \prod_{i=1}^n \frac{1}{\sqrt{x_i-1}} I_{\mathbb{R}^-}(x_{(n)} - 2\theta)$$

La función de verosimilitud de θ será entonces

$$L_{x_1,\ldots,x_n}(\theta) = f_\theta^n(x_1, \ldots, x_n) = \frac{1}{(2\sqrt{2\theta-1})^n} \prod_{i=1}^n \frac{1}{\sqrt{x_i-1}} I_{\mathbb{R}^-}(x_{(n)} - 2\theta)$$

Consideramos la función paramétrica $\lambda = g(\theta) = 2\theta - 1$, con $g : \Theta \to \Lambda$. Por definición de función de verosimilitud de una función paramétrica, considerando una realización muestral $(x_1, \ldots, x_n) \in \mathcal{X}^n$, con $\mathcal{X} =]1, +\infty[$, obtenido del apartado anterior:

$$M_{x_1,\ldots,x_n}(\lambda) = \sup_{\theta \in g^{-1}(\lambda)} L_{x_1,\ldots,x_n}(\theta)$$

Ahora expresamos θ explícitamente en función de λ, usando que

$$\lambda = 2\theta - 1 \iff \lambda + 1 = 2\theta \iff \theta = \frac{\lambda + 1}{2}$$

Así pues, la función de verosimilitud de la función paramétrica $g(\theta)$ es

$$M_{x_1,\ldots,x_n}(\lambda) = L_{x_1,\ldots,x_n}\left(\frac{\lambda+1}{2}\right) = \frac{1}{(2\sqrt{\lambda})^n} \prod_{i=1}^{n} \frac{1}{\sqrt{x_i - 1}} I_{\mathbb{R}^-}(x_{(n)} - (\lambda + 1))$$

Ahora, puede razonarse que, expresando la función de verosimilitud en función de lambda, debe ser $x_{(n)} < \lambda + 1 \iff \lambda > x_{(n)} - 1$ y como $\prod_{i=1}^{n} \frac{1}{\sqrt{x_i - 1}} > 0$ no influye en el estudio de la monotonía, la monotonía la determina $\dfrac{1}{(2\sqrt{\lambda})^n} = (2\sqrt{\lambda})^{-n}$, que es estrictamente decreciente como función de λ, hecho del cual se deduce que el máximo se alcanza en el ínfimo de $]x_{(n)} - 1, +\infty[$. Por tanto, $\hat{\lambda}(x_1, \ldots, x_n) \overset{not}{\equiv} \hat{\lambda} = x_{(n)} - 1$ es el EMV de $\lambda = g(\theta) = 2\theta - 1$.

Ejercicio 3 (2 puntos)**.** Sea X una variable aleatoria con distribución en una familia regular en el sentido de Fréchet–Cramér–Rao, cuyas funciones de densidad son de la forma:

$$f_\theta(x) = \exp\left\{T(x)\ln\theta - \frac{\theta^2}{2} + S(x)\right\} \quad x \in \mathbb{R}, \quad \theta \in \mathbb{R}^+$$

siendo $T(X)$ un estadístico regular.

a) Calcular la función de información asociada a X.

Por el enunciado se deduce que $(X_1, \ldots, X_n)$ m.a.s de X es de tamaño $n = 1$, y que solo hay una observación de X. Por ser T regular, se cumple que

$$\frac{\partial}{\partial\theta}E_\theta[T(X_1, \ldots, X_n)] = E_\theta\left[T(X_1, \ldots, X_n)\frac{\partial \ln f_\theta^n(X_1, \ldots, X_n)}{\partial\theta}\right]$$

en particular para $n = 1$

$$\frac{\partial}{\partial\theta}E_\theta[T(X)] = E_\theta\left[T(X)\frac{\partial \ln f_\theta(X)}{\partial\theta}\right]$$

Primero usamos que la familia es regular

$$\mathrm{E}_\theta\left[\frac{\partial \ln f_\theta(X)}{\partial\theta}\right] = 0 \iff \mathrm{E}_\theta\left[\frac{T(X)}{\theta} - \theta\right] = 0 \iff \frac{1}{\theta}E_\theta[T(X)] - \theta = 0 \iff$$

$$E_\theta[T(X)] = \theta^2$$

y entonces

$$\frac{\partial}{\partial\theta}E_\theta[T(X)] = \frac{\partial}{\partial\theta}(\theta^2) = 2\theta$$

Calculando

$$\ln f_\theta(X) = T(X)\ln\theta - \frac{\theta^2}{2} + S(X)$$

$$\frac{\partial \ln f_\theta(X)}{\partial\theta} = \frac{T(X)}{\theta} - \theta$$

Como

$$E_\theta\left[T(X)\frac{\partial \ln f_\theta(X)}{\partial\theta}\right] = E_\theta\left[T(X)\left(\frac{T(X)}{\theta} - \theta\right)\right] = \frac{1}{\theta}E_\theta[T(X)^2] - \theta E_\theta[T(X)] \overset{E_\theta[T(X)]=\theta^2}{=}$$

$$\frac{1}{\theta}E_\theta[T(X)^2] - \theta^3$$

Tenemos que

$$2\theta = \frac{\partial}{\partial\theta}E_\theta[T(X)] = E_\theta\left[T(X)\frac{\partial \ln f_\theta(X)}{\partial\theta}\right] = \frac{1}{\theta}E_\theta[T(X)^2] - \theta^3 \iff$$

$$2\theta = \frac{1}{\theta}E_\theta[T(X)^2] - \theta^3 \iff E_\theta[T(X)^2] = \theta(2\theta + \theta^3) = 2\theta^2 + \theta^4$$

Ahora, sabemos que la función de información asociada a X puede calcularse usando

$$I_X(\theta) = E_\theta\left[\left(\frac{\partial \ln f_\theta(X)}{\partial\theta}\right)^2\right] \quad \forall\theta\in\Theta$$

Desarrollando

$$I_X(\theta) = E_\theta\left[\left(\frac{\partial \ln f_\theta(X)}{\partial\theta}\right)^2\right] = E_\theta\left[\left(\frac{T(X)}{\theta}-\theta\right)^2\right] = E_\theta\left[\left(\frac{T(X)}{\theta}\right)^2 - 2\cdot T(X) + \theta^2\right] =$$

$$\frac{1}{\theta^2}E_\theta[T(X)^2] - 2E_\theta[T(X)] + E_\theta[\theta^2] \overset{E_\theta[T(X)]=\theta^2}{=} \frac{1}{\theta^2}E_\theta[T(X)^2] - 2\theta^2 + \theta^2 \overset{E_\theta[T(X)^2]=2\theta^2+\theta^4}{=}$$

$$\frac{1}{\theta^2}(2\theta^2+\theta^4) - \theta^2 = 2 + \theta^2 - \theta^2 = 2$$

b) Basándose en una muestra aleatoria simple de X, $(X_1,\ldots,X_n)$, y suponiendo $T(X) > 0$, encontrar la clase de funciones paramétricas que admiten estimador eficiente y los estimadores correspondientes.

Buscamos aplicar el Teorema de Caracterización de Estimadores Eficientes. Para ello, obtenemos la función conjunta de la m.a.s. de X

$$f_\theta^n(x_1,\ldots,x_n) \overset{indep}{=} \prod_{i=1}^n f_\theta(x_i)$$

Se supondrá a partir de ahora que $x_i \in \mathbb{R} \quad \forall i = 1,\ldots,n$ y $\theta\in\mathbb{R}^+$. De lo contrario, $f_\theta(x) = 0$.

$$f_\theta^n(x_1,\ldots,x_n) \overset{indep}{=} \prod_{i=1}^n f_\theta(x_i) = \prod_{i=1}^n e^{\left(T(x_i)\ln\theta - \frac{\theta^2}{2} + S(x)\right)} = e^{\left(\ln\theta(\sum_{i=1}^n T(x_i)) - \frac{n\theta^2}{2} + \sum_{i=1}^n S(x)\right)} =$$

$$\exp\left\{\ln\theta\left(\sum_{i=1}^n T(x_i)\right) - \frac{n\theta^2}{2} + \sum_{i=1}^n S(x)\right\}$$

$$\ln f_\theta^n(x_1,\ldots,x_n) = \ln\theta\left(\sum_{i=1}^n T(x_i)\right) - \frac{n\theta^2}{2} + \sum_{i=1}^n S(x)$$

$$\frac{\partial \ln f_\theta^n(x_1,\ldots,x_n)}{\partial\theta} = \frac{1}{\theta}\left(\sum_{i=1}^n T(x_i)\right) - n\theta = \frac{1}{\theta}\left(\sum_{i=1}^n T(X_i) - n\theta^2\right)$$

Ahora, supongamos que $T(X_1,\ldots,X_n)$ es un estimador de $g(\theta)$ función paramétrica derivable y estrictamente monótona ($g'(\theta)\neq 0 \quad \forall\theta\in\Theta=\mathbb{R}^+$). Como el enunciado nos dice que la familia es regular, y $0 < I_X(\theta) = 2 < +\infty \quad \forall\theta\in\Theta$, T es eficiente si y solo si $\forall\theta\in\Theta \quad \exists a(\theta)\neq 0$ tal que

$$P_\theta\left[\frac{\partial \ln f_\theta^n(X_1,\ldots,X_n)}{\partial\theta} = a(\theta)[T(X_1,\ldots,X_n) - g(\theta)]\right] = 1$$

y

$$I_{(X_1,\ldots,X_n)}(\theta) = a(\theta)g'(\theta)$$

Como

$$\frac{\partial \ln f_\theta^n(X_1,\ldots,X_n)}{\partial\theta} = a(\theta)[T(X_1,\ldots,X_n) - g(\theta)] \iff$$
$$\frac{1}{\theta}\left(\sum_{i=1}^n T(X_i) - n\theta^2\right) = a(\theta)[T(X_1,\ldots,X_n) - g(\theta)]$$

claramente por comparación se obtiene que

$$T(X_1,\ldots,X_n) = \sum_{i=1}^n T(X_i), \quad g(\theta) = n\theta^2, \quad a(\theta) = \frac{1}{\theta}$$

y tanto $a(\theta)$ como $g(\theta)$ verifican todas las condiciones del teorema, pues $g'(\theta) = 2n\theta > 0$, porque $\theta > 0$ y $2n \geqslant 2$, y usando la aditividad de la función de información de Fisher

$$2n = nI_X(\theta) = I_{(X_1,\ldots,X_n)}(\theta) = a(\theta)g'(\theta) = \frac{1}{\theta}2n\theta = 2n$$

Por un corolario visto en teoría, como $T(X_1,\ldots,X_n)$ es un estimador eficiente para $g(\theta)$, con $g'(\theta) \neq 0$, sabemos que las únicas funciones paramétricas que admiten estimadores eficientes son las de la forma $ag(\theta) + b$ y los correspondientes estimadores eficientes son $aT + b$, con probabilidad 1, bajo todas las distribuciones de la familia.

c) Bajo los supuestos del apartado b), calcular la cota inferior para la varianza de estimadores insesgados de $\ln\theta$, regulares, y justificar si se alcanza o no dicha cota.

En este apartado $g(\theta) = \ln\theta$. Ya se han comprobado las hipótesis del Teorema de Cota de FCR en el apartado b), y supuesto que los estimadores insesgados sean además de segundo orden (si no, la varianza no será finita), se tiene que la cota inferior de la varianza es, tomando como uno de estos estimadores $T(X_1,\ldots,X_n)$:

$$\mathrm{Var}_\theta[T(X_1,\ldots,X_n)] \geqslant \frac{(g'(\theta))^2}{I_{(X_1,\ldots,X_n)}(\theta)} \quad \forall\theta \in \Theta$$

Ahora, $g'(\theta) = 1/\theta$, y ya hemos visto que $I_{(X_1,\ldots,X_n)}(\theta) = nI_X(\theta) = 2n$, luego la cota inferior para la varianza de todo estos estimadores será

$$\mathrm{Var}_\theta[T(X_1,\ldots,X_n)] \geqslant \frac{(1/\theta)^2}{2n} = \frac{1}{2n\theta^2}$$

La cota no se alcanza por reducción al absurdo. Si lo hiciera, entonces tendríamos que $g(\theta) = \ln\theta$ admitiría un estimador regular, insesgado y cuya varianza alcanza la cota de FCR para cualquier valor $\theta \in \Theta$, es decir, un estimador eficiente (dado que $\{F_\theta : \theta \in \Theta = \mathbb{R}^+\}$ es una familia regular, $0 < I_X(\theta) = 2 < +\infty \quad \forall\theta \in \Theta$ y $g(\theta) = \ln\theta$ es una función paramétrica derivable). Sin embargo, hemos visto en b) que las únicas funciones paramétricas que admiten estimador eficiente en esta familia son las de la forma $an\theta^2 + b$, $a, b \in \mathbb{R}$. Como $\nexists a, b \in \mathbb{R} : \ln\theta = an\theta^2 + b$, hemos llegado a contradicción.

Ejercicio 4 (2,4 puntos)**.** Contraste de hipótesis:

a) Sea $(X_1, \ldots, X_n)$ una muestra aleatoria simple de una variable aleatoria X con distribución en una familia $\{P_\theta \mid \theta \in \Theta\}$. Sea Θ_0 subconjunto arbitrario de Θ y supongamos que se pretende contrastar la hipótesis

$$H_0 : \theta \in \Theta_0$$

a1) Detallar la hipótesis alternativa. Definir formalmente el concepto de test de hipótesis y dar la interpretación de sus valores.

La hipótesis alternativa es la negación de la hipótesis nula H_0, y se denota por

$$H_1 : \theta \in \Theta_1 = \Theta - \Theta_0$$

El problema de contraste de hipótesis, en general, es el siguiente: se constrasta la hipótesis nula, $H_0 : \theta \in \Theta_0$, frente a la hipótesis alternativa, $H_1 : \theta \in \Theta_1$. La hipótesis sobre el parámetro establece una partición sobre el espacio paramétrico, pues $\Theta = \Theta_0 \cup \Theta_1$ y $\Theta_0 \cap \Theta_1 = \emptyset$, y además sobre la familia de distribuciones, ya que

$$\{F_\theta, \theta \in \Theta\} = \{F_\theta, \theta \in \Theta_0\} \cup \{F_\theta, \theta \in \Theta_1\}$$

La definición formal del concepto de test de hipótesis es la siguiente:

Definición 0.2 (Test de Hipótesis)**.** Un *test de hipótesis* es un estadístico, $\varphi(X_1, \ldots, X_n)$, con valores en $[0, 1]$, que especifica la probabilidad de rechazar H_0 para cada realización muestral.

Distinguimos dos tipos de test de hipótesis, los no aleatorizados y los aleatorizados.

En los no aleatorizados, se considera la región crítica o de rechazo $\mathcal{C} \subset \mathcal{X}^n$, que es aquella en la que se rechaza con probabilidad 1 la hipótesis nula H_0 y el test se puede escribir como $\varphi : \mathcal{X}^n \to \{0, 1\}$, dado por

$$\varphi(X_1, \ldots, X_n) = \begin{cases} 1 & \text{si } (X_1, \ldots, X_n) \in \mathcal{C} \\ 0 & \text{si } (X_1, \ldots, X_n) \notin \mathcal{C} \end{cases}$$

En los aleatorizados, el codominio del test se extiende a $[0, 1]$, y el test es $\varphi : \mathcal{X}^n \to [0, 1]$, donde $\varphi(X_1, \ldots, X_n)$ es la probabilidad de rechazar la hipótesis nula H_0 en cada realización muestral, al igual que en el caso de los tests no aleatorizados.

a2) Definir el tamaño y la función de potencia de un test arbitrario para resolver el problema anterior, explicando el significado de estos conceptos en términos del rechazo de H_0.

Se define la función de potencia de un test arbitrario como sigue:

Definición 0.3 (Función de Potencia)**.** Para un test φ, se define la *función de potencia* como la función que asocia a cada parámetro, θ, la probabilidad media de rechazar H_0 cuando el (verdadero) valor del parámetro es θ

$$\begin{array}{rccl} \beta_\varphi: & \Theta & \longrightarrow & [0,1] \\ & \theta & \longmapsto & B_\varphi(\theta) = E_\theta[\varphi(X_1,\ldots,X_n)] \end{array}$$

En el caso de un test no aleatorizado, se puede definir como

$$\begin{array}{rccl} \beta_\varphi: & \Theta & \longrightarrow & [0,1] \\ & \theta & \longmapsto & B_\varphi(\theta) = P_\theta[(X_1,\ldots,X_n) \in \mathcal{C}] \end{array}$$

Asimismo, se define el tamaño de un test arbitrario de la siguiente forma, recordando que un error de tipo I consiste en rechazar la hipótesis nula H_0 siendo esta verdadera:

Definición 0.4 (Tamaño de un Test)**.** Para un test φ, se define el *tamaño del test* como la máxima probabilidad media de cometer un error de tipo I con dicho test, es decir

$$\sup_{\theta\in\Theta_0} E_\theta[\varphi(X_1,\ldots,X_n)] = \sup_{\theta\in\Theta_0} B_\varphi(\theta)$$

En el caso de un test φ no aleatorizado, se puede definir como

$$\sup_{\theta\in\Theta_0} P_\theta[(X_1,\ldots,X_n) \in \mathcal{C}]$$

a3) En términos del tamaño y de la función potencia, ¿qué significa que un test tiene nivel de significación α para el problema de contraste planteado? ¿Cuáles son las condiciones para que un test sea UMP a nivel de significación α?

El nivel de significación se define por:

Definición 0.5 (Nivel de Significación)**.** Se dice que φ es un test de hipótesis con *nivel de significación* $\alpha \in [0,1]$ si su tamaño es menor o igual que α (actuando α como cota superior de las probabilidades medias de cometer un error de tipo I), es decir:

$$\forall\theta \in \Theta_0, \quad \beta_\varphi(\theta) = E_\theta[\varphi(X_1,\ldots,X_n)] \leqslant \alpha$$

En el caso de un test φ no aleatorizado, se puede definir como

$$\forall\theta \in \Theta_0, \quad P_\theta[(X_1,\ldots,X_n) \in \mathcal{C}] \leqslant \alpha$$

Por lo tanto, un test tiene nivel de significación $\alpha \in [0,1]$ si su tamaño, es decir, el supremo de la función de potencia en el espacio paramétrico de la hipótesis nula, es menor o igual que α.

Se dice que un test es uniformemente más potente a nivel de significación α si verifica la Definición 0.6:

Definición 0.6 (Test Uniformemente Más Potente (UMP))**.** Dado un nivel de significación α fijo, se dice que un test φ es *uniformemente más potente* (UMP), a ese nivel, si se cumplen las siguientes dos condiciones:

1) φ tiene nivel de significación α, es decir

$$E_\theta[\varphi(X_1,\ldots,X_n)] \leqslant \alpha \quad \forall\theta \in \Theta_0$$

2) Para cualquier otro test φ^*, con nivel de significación α, se cumple que la función de potencia en φ es mayor que en φ^*, para cualquier $\theta \in \Theta_1$. Equivalentemente

$$B_\varphi(\theta) \geqslant \beta_{\varphi^*}(\theta) \quad \forall\theta \in \Theta_1$$

El hecho de considerar la maximización solo en Θ_1 y no en Θ_0 la función de potencia se debe a que se busca minimizar el error de tipo II, que consiste en no rechazar H_0 siendo H_0 falsa.

En general, no tiene porqué existir un test UMP para un problema dado.

b) Obtener un test de razón de verosimilitud de tamaño α para contrastar

$$H_0:\ \theta \leq \theta_0 \qquad \text{frente a} \qquad H_1:\ \theta > \theta_0,$$

basado en una observación de una variable aleatoria con la siguiente función de densidad (detallar y justificar todos los pasos para la obtención, incluyendo el estudio detallado del estadístico de contraste y su representación gráfica):

$$f_\theta(x) = \theta x^{-2} e^{-\theta/x}, \qquad x > 0.$$

¿Qué tamaños se alcanzan con dicho test?

Dado que las hipótesis no son simples, no podemos aplicar el Lema de Neyman-Pearson, por lo que habrá que utilizar el Test de la Razón de Verosimilitudes (TRV). Se tiene que $\mathcal{X} = \mathbb{R}^+$ y la muestra se compone de una única observación (pongamos que x_0 es el valor obtenido en la realización muestral). Obtenemos ahora el EMV de θ.

$$L_{x_0}(\theta) = f_\theta(x_0) = \theta x_0^{-2} e^{-\theta/x_0} \Longrightarrow \ln f_\theta(x_0) = \ln\theta - 2\ln x_0 - \theta/x_0 \Longrightarrow$$

$$\frac{\partial \ln f_\theta(x_0)}{\partial\theta} = \frac{1}{\theta} - \frac{1}{x_0} = \frac{x_0 - \theta}{x_0\theta}$$

Resolviendo la ecuación de verosimilitud

$$\frac{\partial \ln f_\theta(x_0)}{\partial \theta} = 0 \iff \frac{x_0 - \theta}{x_0 \theta} = 0 \iff \theta = x_0$$

Como

$$\frac{\partial^2 \ln f_\theta(x_0)}{\partial^2 \theta} = -\frac{1}{\theta^2} < 0$$

se tiene que $\hat{\theta} = x_0$ es el EMV de θ. La razón de verosimilitudes es entonces

$$\lambda(x_0) = \frac{\sup_{\theta \in \Theta_0} L_{x_0}(\theta)}{\sup_{\theta \in \Theta} L_{x_0}(\theta)}, \quad x_0 \in \mathcal{X}$$

Como $\hat{\theta}$ es el EMV de x_0, entonces $\sup_{\theta \in \Theta} L_{x_0}(\theta) = L_{x_0}(x_0)$. Ahora, estudiamos la monotonía respecto de θ de la función de verosimilitud

$$L_{x_0}(\theta) = \theta x_0^{-2} e^{-\theta/x_0} \Longrightarrow \frac{\partial L_{x_0}(\theta)}{\partial \theta} = x_0^{-2}\left(e^{-\theta/x_0} + \theta e^{-\theta/x_0}\frac{-1}{x_0}\right) = x_0^{-2} e^{-\theta/x_0}\left(1 - \frac{\theta}{x_0}\right)$$

Como $x_0^{-2} e^{-\theta/x_0} > 0$ el signo lo determinará el factor $1 - \theta/x_0$. Hay tres posibilidades:

a) Si $1 - \theta/x_0 < 0 \iff 1 < \theta/x_0 \iff x_0 < \theta$, entonces $L_{x_0}(\theta)$ es decreciente.

b) Si $1 - \theta/x_0 > 0 \iff 1 > \theta/x_0 \iff x_0 > \theta$, entonces $L_{x_0}(\theta)$ es creciente.

c) Si $1 - \theta/x_0 = 0 \iff 1 = \theta/x_0 \iff \theta = x_0$, sabemos que se llega al EMV de θ.

Por lo tanto, $L_{x_0}(\theta)$ es creciente en $]0, x_0[$ y decreciente en $]x_0, +\infty[$. Usando esto, se deduce que

$$\sup_{\theta \in \Theta_0} L_{x_0}(\theta) = \begin{cases} L_{x_0}(x_0) & \text{si } x_0 \leqslant \theta_0 \\ L_{x_0}(\theta_0) & \text{si } x_0 > \theta_0 \end{cases}$$

y como

$$\frac{L_{x_0}(\theta_0)}{L_{x_0}(x_0)} = \frac{\theta_0 x_0^{-2} e^{-\theta_0/x_0}}{x_0 x_0^{-2} e^{-x_0/x_0}} = \frac{\theta_0}{x_0} e^{1-\theta_0/x_0}$$

la razón de verosimilitudes se escribe de la siguiente forma

$$\lambda(x_0) = \begin{cases} \dfrac{L_{x_0}(x_0)}{L_{x_0}(x_0)} = 1 & \text{si } x_0 \leqslant \theta_0 \\ \dfrac{L_{x_0}(\theta_0)}{L_{x_0}(x_0)} = \dfrac{\theta_0}{x_0} e^{1-\theta_0/x_0} & \text{si } x_0 > \theta_0 \end{cases}$$

Por construcción $0 \leqslant \lambda(x_0) \leqslant 1$, y el test de razón de verosimilitud (TRV) es:

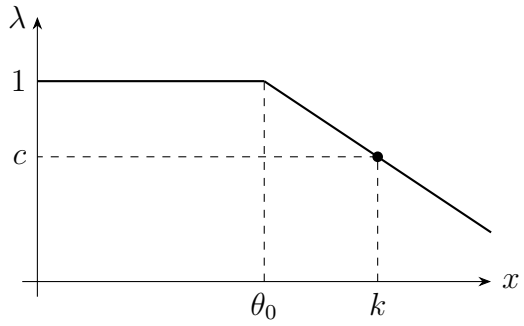
$$\varphi(X) = \begin{cases} 1 & \text{si } \lambda(X) < c \\ 0 & \text{si } \lambda(X) \geqslant c \end{cases}$$

donde $c \in]0,1]$ es constante y se determina imponiendo el tamaño o nivel de significación $\alpha \in [0,1]$.

Estudiamos ahora la monotonía de λ. Si $x > \theta_0$ consideramos el cambio de variable $t = \theta_0/x \in]0,1[$ (como $f_\theta(x)$ es densidad, debe ser $\theta > 0$ para que pueda integrar a 1, y $x \in \mathcal{X} = \mathbb{R}^+$), de donde $\lambda(x) = te^{1-t} = g(t)$. Como

$$g'(t) = e^{1-t} + te^{1-t} \cdot (-1) = e^{1-t}(1-t) > 0 \quad \forall t \in]0,1[$$

vemos que g es creciente en t, pero $t = \theta_0/x$ crece si x decrece, luego $\lambda(x)$ es estrictamente decreciente en $]\theta_0, +\infty[$. Además $\lambda(\theta_0) = 1$ y $\lim\limits_{x\to\infty} \lambda(x) = 0$. Gráficamente:



Luego

$$\left.\begin{array}{l} \lambda < c \\ \lambda \geqslant c \end{array}\right\} \Longleftrightarrow \left\{\begin{array}{l} x > k \\ x \leqslant k \end{array}\right. \qquad \text{con } k \geqslant \theta_0$$

Por lo que el test a considerar será:

$$\varphi(X) = \begin{cases} 1 & \text{si } X > k \\ 0 & \text{si } X \leqslant k \end{cases}$$

Imponemos tamaño α:

$$\alpha = \sup_{\theta\in\Theta_0} \mathrm{E}_\theta[\varphi(X)] = \sup_{\theta\leqslant\theta_0} P_\theta[X > k]$$

Como

$$f_\theta(x) = \theta x^{-2} e^{-\theta/x}$$

entonces considerando el cambio $u = \theta/t \Longrightarrow t = \theta/u \Longrightarrow dt = -\dfrac{\theta}{u^2}$, y $t^{-2} = (u/\theta)^2$, sustituyendo

$$\theta t^{-2} e^{-\theta/t} dt = \theta \left(\frac{u}{\theta}\right)^2 e^{-u} \left(-\frac{\theta}{u^2} du\right) = -e^{-u} du$$

y viendo que si $t \to 0^+ \Longrightarrow u = \theta/t \to +\infty$ y si $t = x \Longrightarrow u = \theta/x$, se tiene que

$$F_\theta(x) = \int_0^t f_\theta(t)dt = \int_0^t \theta x^{-2} e^{-\theta/x} dt =$$

$$\int_{+\infty}^{\theta/x} -e^{-u}du = \int_{\theta/x}^{+\infty} e^{-u}du = [-e^{-u}]_{\theta/x}^{+\infty} = e^{-\theta/x}$$

Así pues, para $x \in \mathcal{X}$, $F_\theta(x) = e^{-\theta/x}$, de donde

$$P_\theta[X > k] = 1 - F_\theta(k) = 1 - e^{-\theta/k}$$

Entonces el tamaño queda

$$\alpha = \sup_{\theta \leqslant \theta_0} P_\theta[X > k] = P_{\theta_0}[X > k] = 1 - e^{-\theta_0/k}$$

y k queda determinado por

$$\alpha = 1 - e^{-\theta_0/k} \iff \alpha - 1 = -e^{-\theta_0/k} \iff 1 - \alpha = e^{-\theta_0/k} \iff \ln(1-\alpha) = -\theta_0/k \iff$$

$$k = -\frac{\theta_0}{\ln(1-\alpha)}$$

y el TRV no aleatorizado es

$$\varphi(X) = \begin{cases} 1 & \text{si } X > -\dfrac{\theta_0}{\ln(1-\alpha)} \\ \\ 0 & \text{si } X \leqslant -\dfrac{\theta_0}{\ln(1-\alpha)} \end{cases}$$

Los tamaños que se alcanzan con este test pueden hallarse considerando el tamaño como función de k

$$\alpha(k) = 1 - e^{-\theta/k}, \quad k > 0$$

Vemos que

$$\lim_{k \to +\infty} \alpha(k) = \lim_{k \to +\infty} (1 - e^{-\theta/k}) =$$

$$1 - \lim_{k \to +\infty} e^{-\theta/k} = 1 - e^0 = 0$$

y que el menor k se obtiene cuando $k = \theta_0$, con $\alpha(\theta_0) = 1 - e^{-\theta_0/\theta_0} = 1 - e^{-1}$. Por lo tanto, los tamaños que pueden obtenerse con el TRV no aleatorizado son

$$\alpha \in [0, 1 - e^{-1}]$$

Para obtener cualquier tamaño $\alpha \in [0, 1]$ hay que aleatorizar el TRV, es decir, considerar un test de la forma

$$\varphi(X) = \begin{cases} 1 & \text{si } \lambda(X) < c \\ \gamma & \text{si } \lambda(X) = c \quad \gamma \in [0,1] \\ 0 & \text{si } \lambda(X) > c \end{cases}$$

Y como $\lambda(x) = 1 \quad \forall x \leqslant \theta_0$ y $\lambda(x) < 1 \quad \forall x > \theta_0$, lo natural es tomar $c = 1$, y entonces

$$\varphi(X) = \begin{cases} 1 & \text{si } \lambda(X) < 1 \iff X > \theta_0 \\ \gamma & \text{si } \lambda(X) = 1 \iff X \leqslant \theta_0 \quad \gamma \in [0,1] \\ 0 & \text{si } \lambda(X) > 1 \text{ nunca} \end{cases}$$

es decir

$$\varphi(X) = \begin{cases} 1 & \text{si } X > \theta_0 \\ \gamma & \text{si } X \leqslant \theta_0 \end{cases} \quad \gamma \in [0,1]$$

Si $X > \theta_0$ se rechaza siempre, y si $X \leqslant \theta_0$ se rechaza con probabilidad γ. Imponiendo tamaño α:

$$\alpha = \sup_{\theta \in \Theta_0} \mathrm{E}_\theta[\varphi(X)] = E_{\theta_0}[\varphi(X)] = P_{\theta_0}[X > \theta_0] + \gamma P_{\theta_0}[X \leqslant \theta_0]$$

y como

$$P_{\theta_0}[X > \theta_0] = 1 - e^{-1} \quad P_{\theta_0}[X \leqslant \theta_0] = e^{-1}$$

sacamos

$$\alpha = (1 - e^{-1}) + \gamma e^{-1}$$

De esta manera, dado un tamaño $\alpha \in [1 - e^{-1}, 1]$, se obtiene que

$$\alpha = (1-e^{-1}) + \gamma e^{-1} \iff \alpha - (1-e^{-1}) = \gamma e^{-1} \iff \gamma = \frac{\alpha - (1 - e^{-1})}{e^{-1}} \in [0,1]$$

y el TRV aleatorizado es

$$\varphi(X) = \begin{cases} 1 & \text{si } X > \theta_0 \\ \dfrac{\alpha - (1 - e^{-1})}{e^{-1}} & \text{si } X \leqslant \theta_0 \end{cases}$$

Y ya podemos cubrir cualquier tamaño $\alpha \in [0,1]$:

a) Si $\alpha \in [0, 1 - e^{-1}]$, el TRV que se toma es el no aleatorizado:

$$\varphi(X) = \begin{cases} 1 & \text{si } X > -\dfrac{\theta_0}{\ln(1-\alpha)} \\ \\ 0 & \text{si } X \leqslant -\dfrac{\theta_0}{\ln(1-\alpha)} \end{cases}$$

b) Si $\alpha \in [1 - e^{-1}, 1]$, el TRV que se toma es el aleatorizado:

$$\varphi(X) = \begin{cases} 1 & \text{si } X > \theta_0 \\ \dfrac{\alpha - (1 - e^{-1})}{e^{-1}} & \text{si } X \leqslant \theta_0 \end{cases}$$

Si $\alpha = 1 - e^{-1}$:

$$\varphi(X) = \begin{cases} 1 & \text{si } X > -\dfrac{\theta_0}{\ln(1-\alpha)} \\ \\ 0 & \text{si } X \leqslant -\dfrac{\theta_0}{\ln(1-\alpha)} \end{cases} \iff \begin{cases} 1 & \text{si } X > \theta_0 \\ \dfrac{\alpha - (1 - e^{-1})}{e^{-1}} = 0 & \text{si } X \leqslant \theta_0 \end{cases}$$

por lo que el TRV no aleatorizado y el TRV aleatorizado coinciden.

Ejercicio 5 (1,1 puntos)**.** En cierta academia se pretende hacer un estudio sobre la efectividad de un curso intensivo de aprendizaje. Para ello, se selecciona aleatoriamente a 10 alumnos a los que se somete a una misma prueba de control antes y después del curso, obteniéndose las siguientes calificaciones:

A	7,25	7,5	7,5	8	8,5	8,75	8	8,25	7,25	8,5
D	7	8,25	8	8	8,75	9,5	8,5	8,75	7,75	8

¿Se puede concluir, a la vista de los datos, que el curso de aprendizaje no afecta a las calificaciones de los alumnos? Especificar bajo qué condiciones sobre las variables consideradas se puede resolver este problema y concretar las hipótesis nula y alternativa que se contrastan. Resolver este problema al nivel de significación 0.1 mediante el uso de diferentes tests, detallando las hipótesis necesarias para la aplicación de cada uno de ellos.

Para estudiar si el curso de aprendizaje "no afecta" a las calificaciones de los alumnos, trabajamos con la variable $X = D - A$, y con la m.a.s. asociada $X_i = D_i - A_i \quad \forall i = 1, \ldots, n$, con $n = 10$. De esta manera, no afectar es equivalente a que la distribución de X esté centrada en 0, para lo cual usaremos el constraste de la mediana M_X. La hipótesis bilateral será

$$\begin{cases} H_0 : M_X = 0 \\ H_1 : M_X \neq 0 \end{cases}$$

La m.a.s. de X quedaría

$$(-0.25, 0.75, 0.5, 0, 0.25, 0{,}75, 0.5, 0.5, 0.5, -0.5)$$

Por teoría sabemos que, como hay un 0, se elimina ese dato, y se reajusta el tamaño de la muestra a $n = 9$. La muestra resultante es entonces

$$(-0.25, 0.75, 0.5, 0.25, 0.75, 0.5, 0.5, 0.5, -0.5)$$

En primer lugar aplicaremos el test de los signos.

Test de los Signos de Fisher

Se supondrá que X tiene una distribución continua. Hay un total de 2 signos negativos, y $9 - 2 = 7$ positivos. El estadístico es

$$T = \text{Número de Signos i} : X_i > 0 \rightsquigarrow_{H_0} B(n, 1/2)$$

Entonces $T_{exp} = 7$ y ya sabíamos que $n = 9$. El p-valor sabemos que

$$p - valor = \begin{cases} 2P_{H_0}[T(X_1, \ldots, X_n) \leqslant T_{exp}] & \text{si } T_{exp} \leqslant n/2 \\ 2P_{H_0}[T(X_1, \ldots, X_n) \geqslant T_{exp}] & \text{si } T_{exp} \geqslant n/2 \end{cases}$$

En este caso $7 = T_{exp} \geqslant n/2 = 4{,}5$. Obtenemos esta probabilidad usando que $T \rightsquigarrow_{H_0} B(9, 1/2)$, luego

$$P_{H_0}[T = k] = \binom{9}{k} \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{9-k} = \binom{9}{k} \left(\frac{1}{2}\right)^9 = \frac{\binom{9}{k}}{2^9}$$

y

$$P_{H_0}[T(X_1,\ldots,X_n) \geqslant 7] = P_{H_0}[T(X_1,\ldots,X_n) = 7] + P_{H_0}[T(X_1,\ldots,X_n) = 8] +$$

$$P_{H_0}[T(X_1,\ldots,X_n) = 9] = \frac{\binom{9}{7} + \binom{9}{8} + \binom{9}{9}}{2^9} = \frac{36+9+1}{512} = \frac{46}{512} \approx 0{,}0898$$

Entonces $p - valor = 2 \cdot P_{H_0}[T(X_1,\ldots,X_n) \geqslant 7] \approx 2 \cdot 0{,}0898 = 0{,}1796$. Como el nivel de significación que nos dan es $\alpha = 0{,}1$, concluimos que no se rechaza H_0 con el test de los signos.

Test de los Rangos Signados de Wilcoxon

Se supondrá que además de ser X continua, como en el caso anterior, es además simétrica (alrededor de la mediana). Consideramos $D_i = X_i - 0 = X_i, \quad \forall i = 1,\ldots,n = 9$. Ordenamos crecientemente los valores $|D_i|$, y luego asignamos el lugar que ocupan o rango $r(|D_i|)$ con $i = 1,\ldots,n$. Sabemos que si hay empates, entonces se le asigna el promedio de los rangos a cada uno de los $|D_i|$ que producen el empate. Hay en total tres $|D_i|$, que son 0.25, 0.5 y 0.75. Además

1. 0.25 aparece 2 veces, con rangos 1 y 2, luego a cada uno le asignamos el rango promedio 1,5.
2. 0.5 aparece 5 veces, con rangos $3, 4, 5, 6$ y 7, luego a cada uno le asignamos el rango promedio $(3+4+5+6+7)/5 = 5$.
3. 0.75 aparece 2 veces, con rangos 8 y 9, luego a cada uno le asignamos el rango 8,5

El estadístico de Wilcoxon es:

$$T^+(X_1,\ldots,X_n) = \text{Suma de los Rangos de los } D_i > 0$$

Obtenemos su valor experimental:

$$T^+_{exp} = T(x_1,\ldots,x_n) = 1{,}5 + 4 \cdot 5 + 2 \cdot 8{,}5 = 38{,}5$$

La distribución de T^+ es simétrica alrededor de $n(n+1)/4$ y está tabulada para $9 = n \leqslant 15$. En este caso $n(n+1)/4 = 9 \cdot 10/4 = 22{,}5$. Como $H_1 : M_X \neq 0$, sabemos que se rechaza H_0 para un nivel de significación α si

$$T^+_{exp} \leqslant k \quad \text{ó} \quad T^+_{exp} \geqslant \frac{n(n+1)}{2} - k$$

con k aquel valor tal que $P[T^+(X_1,\ldots,X_n) \leqslant k] \leqslant \alpha/2$. La cola izquierda será para valores $t \leqslant 22$, y la cola derecha para valores $t \geqslant 23$. En este caso, $n(n+1)/2 = 9(10)/2 = 45$, y buscamos en la cola izquierda el mayor k tal que:

$$P[T^+(X_1,\ldots,X_n) \leqslant k] \leqslant \alpha/2 = 0{,}05$$

Usando la tabla de Wilcoxon para $n = 9$:

$$P[T^+(X_1,\ldots,X_n) \leqslant 8] = 0{,}049 \leqslant 0{,}05$$

$$P[T^+(X_1,\ldots,X_n) \leqslant 9] = 0{,}064 > 0{,}05$$

Por tanto, $k = 8$. Como

$$38{,}5 = T^+_{exp} \geqslant \frac{n(n+1)}{2} - k = 45 - 8 = 37$$

se rechaza H_0 a nivel de significación $\alpha = 0{,}1$.