

Inferencia Estadística

Examen V



Los Del DGIIM, losdeldgiim.github.io

Doble Grado en Ingeniería Informática y Matemáticas
Universidad de Granada



Esta obra está bajo una Licencia Creative Commons Atribución-NoComercial-SinDerivadas 4.0 Internacional (CC BY-NC-ND 4.0).

Eres libre de compartir y redistribuir el contenido de esta obra en cualquier medio o formato, siempre y cuando des el crédito adecuado a los autores originales y no persigas fines comerciales.

Inferencia Estadística Examen V

Los Del DGIIM, losdeldgiim.github.io

José Manuel Sánchez Varbas

Granada, 2026

Asignatura Inferencia Estadística.

Curso Académico 2021-22.

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

Descripción Examen Ordinario.

Fecha 21 de Enero de 2022.

Ejercicio 1 (2.5 puntos). Sea (X_1, \dots, X_n) una muestra aleatoria simple de una variable con función de densidad:

$$f_\theta(x) = \frac{\sqrt{\theta-1}}{2} x^{-3/2}, \quad x > \theta - 1$$

- a) Determinar para qué valores de n existe el UMVUE para θ , y encontrarlo en tales casos. Justificar la existencia en los casos que corresponda.
- b) Calcular la función de verosimilitud y encontrar el estimador máximo verosímil de θ . ¿Es insesgado? (justificar la respuesta).

Ejercicio 2 (2.25 puntos).

- a) Sea X una variable aleatoria continua con distribución en una familia regular en el sentido de Fréchet-Cramér-Rao, cuyas funciones de densidad son de la forma:

$$f_\theta(x) = \exp[Q(\theta)T(x) + D(\theta) + S(x)], \quad x \in \mathcal{X}, \quad \theta \in \Theta$$

siendo $T(X)$ un estadístico regular, insesgado en θ^3 , tal que $\text{Var}_\theta[T(X)] = \theta^2$. Calcular las funciones $Q(\theta)$ y $D(\theta)$, sabiendo que $Q(1) = 0$ y $D(0) = 1$.

- b) Sea (X_1, \dots, X_n) una muestra aleatoria simple de una variable X con función masa de probabilidad:

$$P_p[X = x] = p^2(x-1)(1-p)^{x-2}, \quad x = 2, 3, \dots; \quad 0 < p < 1$$

$$\left(\text{E}_p[X] = \frac{2}{p}, \quad \text{Var}_p[X] = \frac{2(1-p)}{p^2} \right)$$

- a) Sabiendo que la familia de distribuciones es regular, calcular la función de información asociada a la muestra. Encontrar la clase de funciones paramétricas que admiten estimador eficiente y los estimadores correspondientes.
- b) Calcular la cota para la varianza de estimadores insesgados en p^2 , regulares, y justificar si se alcanza o no dicha cota.

Ejercicio 3 (2 puntos). Sea (X_1, \dots, X_n) una muestra aleatoria simple de una variable X con distribución $\mathcal{N}(\mu, \sigma^2)$, siendo ambos parámetros desconocidos.

- a) Especificar la variable usada para hacer inferencia sobre σ^2 . Deducir su distribución, detallando y justificando exhaustivamente cada paso.
- b) Deducir, detallando el método usado, el intervalo de confianza de menor longitud esperada uniformemente para μ a nivel de confianza $1 - \alpha$.

Ejercicio 4 (1.75 puntos). El retraso medio diario (en minutos) de los trabajadores de cierta empresa tiene una distribución $U(0, \theta)$. Para contrastar la hipótesis de que el máximo retraso medio diario es de 5 minutos frente a que es de 4 minutos, se midió dicha variable durante 10 días elegidos al azar, observando que el máximo retraso medio en esos 10 días fue de 3 minutos. Obtener los tests más potentes de tamaños 0,05 y 0,15 para el problema de contraste planteado. Decidir, a la vista de lo observado, si debe rechazarse la hipótesis nula en cada uno de los casos.

Ejercicio 5 (1.5 puntos).

- a) Test de Kolmogorov–Smirnov. Comparación con χ^2 .
- b) Se recogen observaciones de la facturación (en miles de euros) de un cierto supermercado durante 25 días. Se quiere contrastar si la facturación se distribuye uniformemente.

10.1	7.3	4.2	3.1	8.4	9.2	1.1	1	8.1	6.4	10.5	1.2	
1.8	2.6	2.2	3.9	2.9	6.5	8.1	8.3	3.8	4.2	7.3	8.9	9.6

Solución.

Ejercicio 1 (2.5 puntos). Sea (X_1, \dots, X_n) una muestra aleatoria simple de una variable con función de densidad:

$$f_\theta(x) = \frac{\sqrt{\theta-1}}{2} x^{-3/2}, \quad x > \theta - 1$$

- a) Determinar para qué valores de n existe el UMVUE para θ , y encontrarlo en tales casos. Justificar la existencia 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 $T \stackrel{not}{\equiv} T(X_1, \dots, X_n)$, para una m.a.s. (X_1, \dots, X_n) con $n \in \mathbb{N}$ fijo) insesgada en $g(\theta) = \theta$, estimadora y con momento de segundo orden finito. Entonces $h(T)$ será el UMVUE.

El estadístico suficiente se calcula por medio del Teorema de Factorización de Neyman-Fisher. La función conjunta es la siguiente

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

Vemos que

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

de donde se deduce que

$$f_\theta^n(x_1, \dots, x_n) = \prod_{i=1}^n \frac{\sqrt{\theta-1}}{2} x_i^{-3/2} I_{\mathbb{R}^+}(x_{(1)} - (\theta - 1)) = \frac{(\theta - 1)^{n/2}}{2^n} \prod_{i=1}^n x_i^{-3/2} I_{\mathbb{R}^+}(x_{(1)} - (\theta - 1))$$

Tomando $T(X_1, \dots, X_n) = X_{(1)}$ y

$$h(x_1, \dots, x_n) = \frac{1}{2^n} \prod_{i=1}^n x_i^{-3/2}, \quad g_\theta(t) = (\theta - 1)^{n/2} I_{\mathbb{R}^+}(t - (\theta - 1))$$

Se cumple que

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

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 T es suficiente.

Ahora, hay que comprobar que este estadístico es completo, lo cual se hará por definición. Sabemos por teoría que la distribución del mínimo es

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

Hallamos ahora la función de distribución de X :

$$\begin{aligned} F_X(t) &= \int_{\theta-1}^t f_\theta(x) dx = \int_{\theta-1}^t \frac{\sqrt{\theta-1}}{2} x^{-3/2} dx = \frac{\sqrt{\theta-1}}{2} \int_{\theta-1}^t x^{-3/2} dx = \frac{\sqrt{\theta-1}}{2} \cdot \left[\frac{x^{-1/2}}{-1/2} \right]_{\theta-1}^t = \\ &= -\sqrt{\theta-1}(t^{-1/2} - (\theta-1)^{-1/2}) = 1 - \sqrt{\frac{\theta-1}{t}} \quad t > \theta-1 \end{aligned}$$

La función de densidad del estadístico será entonces

$$\begin{aligned} f_T(t) &= n(1 - F_X(t))^{n-1} f_\theta(t) = n \left(\sqrt{\frac{\theta-1}{t}} \right)^{n-1} \frac{\sqrt{\theta-1}}{2} t^{-3/2} = \\ &\frac{n}{2} \frac{(\sqrt{\theta-1})^n}{(\sqrt{t})^{n+2}} = \frac{n}{2} (\theta-1)^{n/2} t^{-(n+2)/2} \quad t > \theta-1 \end{aligned}$$

Sea h una función medible verificando

$$\begin{aligned} 0 &= E[h(T)] \stackrel{def}{=} \int_{\theta-1}^{+\infty} h(t) f_T(t) dt = \int_{\theta-1}^{+\infty} h(t) \frac{n}{2} (\theta-1)^{n/2} t^{-(n+2)/2} dt = \\ &\frac{n}{2} (\theta-1)^{n/2} \int_{\theta-1}^{+\infty} h(t) t^{-(n+2)/2} dt \end{aligned}$$

como $\frac{n}{2} (\theta-1)^{n/2} \neq 0 \quad \forall n \in \mathbb{N}, \quad \forall \theta > 1$, debe ser

$$\int_{\theta-1}^{+\infty} h(t) t^{-(n+2)/2} dt = 0$$

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

$$\begin{aligned} -\frac{d}{d\theta} H(\theta-1) &= 0 \iff -h(\theta-1)(\theta-1)^{-(n+2)/2} = 0 \stackrel{\theta>1}{\iff} \\ h(\theta-1) &= 0 \end{aligned}$$

Equivalentemente

$$\forall \theta \in \Theta =]1, +\infty[\quad h(\theta-1) = 0 \iff h(t) = 0 \quad \forall t > 0$$

(tomando $t = \theta-1 \in]0, +\infty[$). Por tanto

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

y consecuentemente

$$1 \geq P[h(T) = 0] \geq P[T > 0] = 1 \implies P[h(T) = 0] = 1$$

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

$$\begin{aligned}\theta = g(\theta) = \text{E}[h(T)] &= \frac{n}{2}(\theta - 1)^{n/2} \int_{\theta-1}^{+\infty} h(t)t^{-(n+2)/2} dt \iff \\ \int_{\theta-1}^{+\infty} h(t)t^{-(n+2)/2} dt &= \frac{2\theta}{n}(\theta - 1)^{-n/2}\end{aligned}$$

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

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

y el derecho es

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

Despejamos $h(\theta - 1)$:

$$\begin{aligned}-h(\theta - 1)(\theta - 1)^{-(n+2)/2} &= \frac{2}{n}(\theta - 1)^{-n/2} - \theta(\theta - 1)^{-n/2-1} \iff \\ h(\theta - 1) &= \frac{\frac{2}{n}(\theta - 1)^{-n/2} - \theta(\theta - 1)^{-n/2-1}}{-(\theta - 1)^{-(n+2)/2}} = -\frac{2}{n}(\theta - 1) + \theta = \frac{-2(\theta - 1) + n\theta}{n} = \\ &\frac{n\theta - 2\theta + 2}{n} = \frac{(n - 2)\theta + 2}{n}\end{aligned}$$

de donde

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

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

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

$$\begin{aligned}E[h(T)^2] &\stackrel{\text{def}}{=} \int_{\theta-1}^{+\infty} h(t)^2 f_T(t) dt = \\ \int_{\theta-1}^{+\infty} \left(1 + \left(\frac{n-2}{n} \right)^2 t^2 + \frac{2(n-2)}{n}t \right) \frac{n}{2}(\theta - 1)^{n/2} t^{-(n+2)/2} dt = \\ \frac{n}{2}(\theta - 1)^{n/2} \int_{\theta-1}^{+\infty} \left(1 + \left(\frac{n-2}{n} \right)^2 t^2 + \frac{2(n-2)}{n}t \right) t^{-(n+2)/2} dt =\end{aligned}$$

$$\frac{n}{2}(\theta-1)^{n/2} \left(\int_{\theta-1}^{+\infty} t^{-(n+2)/2} dt + \left(\frac{n-2}{n} \right)^2 \int_{\theta-1}^{+\infty} t^2 t^{-(n+2)/2} dt + \frac{2(n-2)}{n} \int_{\theta-1}^{+\infty} t t^{-(n+2)/2} dt \right) =$$

$$\frac{n}{2}(\theta-1)^{n/2} \left(\left[-\frac{2}{n} t^{-n/2} \right]_{\theta-1}^{+\infty} + \left(\frac{n-2}{n} \right)^2 \left[\frac{2}{4-n} t^{(4-n)/2} \right]_{\theta-1}^{+\infty} + \frac{2(n-2)}{n} \left[\frac{1}{1-n/2} t^{1-n/2} \right]_{\theta-1}^{+\infty} \right)$$

Estudiamos ahora cada primitiva para saber cuándo converge:

- $\left[-\frac{2}{n} t^{-n/2} \right]_{\theta-1}^{+\infty} = -\frac{2}{n} \left(\lim_{m \rightarrow +\infty} m^{-n/2} - (\theta-1)^{-n/2} \right) < +\infty \iff$

$$\frac{-n}{2} < 0 \iff -n < 0 \iff n > 0$$

y ya se tenía esta condición para que $h(T)$ fuera estimador.

- $\left(\frac{n-2}{n} \right)^2 \left[\frac{2}{4-n} t^{(4-n)/2} \right]_{\theta-1}^{+\infty} =$

$$\frac{2(n-2)^2}{(4-n)n^2} \left(\lim_{m \rightarrow +\infty} m^{(4-n)/2} - (\theta-1)^{(4-n)/2} \right) < +\infty \iff$$

$$\frac{4-n}{2} < 0 \iff 4-n < 0 \iff n > 4$$

Por lo que de momento debe ser $n > 4 \iff n \geq 5$ para que el momento de segundo orden sea finito.

- $\frac{2(n-2)}{n} \left[\frac{1}{1-n/2} t^{1-n/2} \right]_{\theta-1}^{+\infty} =$

$$\frac{4(n-2)}{(2-n)n} \left(\lim_{m \rightarrow +\infty} m^{(2-n)/2} - (\theta-1)^{(2-n)/2} \right) < +\infty \iff$$

$$\frac{2-n}{2} < 0 \iff 2-n < 0 \iff n > 2$$

y ya se tenía esta condición implícitamente por la anterior.

Concluimos que $E[h(T)^2] < +\infty \iff n > 4$, ya que si $n \leq 4$, 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 $n > 4 \iff n \geq 5$.

- b) Calcular la función de verosimilitud y encontrar el estimador máximo verosímil de θ . ¿Es insesgado? (justificar la respuesta).

Como

$$f_\theta(x) = \frac{\sqrt{\theta-1}}{2} x^{-3/2}, \quad x > \theta - 1$$

y la función conjunta $f_\theta^n(x_1, \dots, x_n)$ para $(x_1, \dots, x_n) \in \mathcal{X}^n = (\theta - 1, +\infty)^n$ ya la teníamos del apartado anterior

$$\frac{(\theta-1)^{n/2}}{2^n} \prod_{i=1}^n x_i^{-3/2}$$

por definición la función de verosimilitud de θ será entonces

$$L_{x_1, \dots, x_n}(\theta) = f_\theta^n(x_1, \dots, x_n) = \frac{(\theta-1)^{n/2}}{2^n} \prod_{i=1}^n x_i^{-3/2}$$

Para hallar el EMV, usaremos las ecuaciones de verosimilitud:

$$\begin{aligned} \ln f_\theta^n(x_1, \dots, x_n) &= \frac{n}{2} \ln(\theta-1) - n \ln(2) - \frac{3}{2} \sum_{i=1}^n \ln(x_i) \implies \\ \frac{\partial \ln f_\theta^n(x_1, \dots, x_n)}{\partial \theta} &= \frac{n}{2(\theta-1)} = 0 \iff n = 0 \end{aligned}$$

lo cual no tiene sentido en este problema ya que $n \geq 1$. Vemos entonces que la función de verosimilitud es estrictamente monótona respecto de θ . De hecho, como $\theta - 1 > 0$, es estrictamente creciente. El dominio de esta función de θ teniendo en cuenta que $x_{(1)} > \theta - 1 \iff 1 < \theta < x_{(1)} + 1$, es $]1, x_{(1)} + 1[$, y como es estrictamente creciente respecto de θ , deducimos que el EMV de θ será el supremo del intervalo, es decir, $\hat{\theta} = X_{(1)} + 1$.

El EMV será insesgado, por definición, si y solo si $E[\hat{\theta}] = \theta$. Obtenemos así $E[\hat{\theta}]$:

$$E[\hat{\theta}] = E[X_{(1)} + 1] = E[X_{(1)}] + 1 = E[T] + 1$$

Recordando del apartado anterior que

$$E[h(T)] = \frac{n}{2}(\theta-1)^{n/2} \int_{\theta-1}^{+\infty} h(t)t^{-(n+2)/2} dt$$

y tomando $h(T) = T$:

$$\begin{aligned} E[T] &= \frac{n}{2}(\theta-1)^{n/2} \int_{\theta-1}^{+\infty} tt^{-(n+2)/2} dt \stackrel{(*)}{=} \frac{n}{2}(\theta-1)^{n/2} \left(\frac{2}{2-n} (-(\theta-1)^{(2-n)/2}) \right) = \\ &\frac{n}{2}(\theta-1)^{n/2} \frac{2}{n-2} (\theta-1)^{(2-n)/2} = \frac{n}{n-2}(\theta-1) = \frac{n}{n-2}\theta - \frac{n}{n-2} \end{aligned}$$

donde en (*) se ha usado la convergencia de la última integral del apartado anterior, para $n > 2$. Así

$$E[\hat{\theta}] = E[T] + 1 = \frac{n}{n-2}\theta - \frac{n}{n-2} + 1 = \frac{n}{n-2}\theta - \frac{2}{n-2}$$

y como

$$E[\hat{\theta}] = \frac{n}{n-2}\theta - \frac{2}{n-2} \neq \theta \quad \forall n > 2$$

se tiene que el EMV de θ , $\hat{\theta}$, no es insesgado en θ .

Ejercicio 2 (2.25 puntos).

- a) Sea X una variable aleatoria continua con distribución en una familia regular en el sentido de Fréchet-Cramér-Rao, cuyas funciones de densidad son de la forma:

$$f_\theta(x) = \exp[Q(\theta)T(x) + D(\theta) + S(x)], \quad x \in \mathcal{X}, \quad \theta \in \Theta$$

siendo $T(X)$ un estadístico regular, insesgado en θ^3 , tal que $\text{Var}_\theta[T(X)] = \theta^2$. Calcular las funciones $Q(\theta)$ y $D(\theta)$, sabiendo que $Q(1) = 0$ y $D(0) = 1$.

Primero obtenemos relaciones con las condiciones que nos da el enunciado. Calculamos:

$$\ln f_\theta(x) = Q(\theta)T(x) + D(\theta) + S(x) \implies \frac{\partial \ln f_\theta(x)}{\partial \theta} = T(x)Q'(\theta) + D'(\theta)$$

y usamos que la familia es regular

$$\begin{aligned} E_\theta \left[\frac{\partial \ln f_\theta(x)}{\partial \theta} \right] &= 0 \iff E_\theta [T(X)Q'(\theta) + D'(\theta)] = 0 \iff \\ Q'(\theta)E_\theta [T(X)] + D'(\theta) &= 0 \iff \\ \theta^3 \stackrel{(*)}{=} E_\theta [T(X)] &= -\frac{D'(\theta)}{Q'(\theta)} \iff D'(\theta) = -\theta^3 Q'(\theta) \end{aligned} \quad (1)$$

donde en $(*)$ se ha utilizado que T es insesgado en θ^3 . Ahora, por ser T regular, se tiene que

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

en particular para $n = 1$

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

Nuevamente, como T es insesgado en θ^3 , podemos obtener el miembro izquierdo de la igualdad

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

y el miembro derecho, usando lo ya calculado al principio, es

$$\begin{aligned} E_\theta \left[T(X) \frac{\partial \ln f_\theta(X)}{\partial \theta} \right] &= E_\theta [T(X) (T(X)Q'(\theta) + D'(\theta))] = E_\theta [T(X)^2 Q'(\theta) + T(X)D'(\theta)] = \\ Q'(\theta)E_\theta [T(X)^2] + D'(\theta)E_\theta [T(X)] &\stackrel{(*)}{=} Q'(\theta)E_\theta [T(X)^2] + \theta^3 D'(\theta) \stackrel{(**)}{=} \\ Q'(\theta)(\text{Var}_\theta[T(X)] + (E_\theta[T(X)])^2) + \theta^3 D'(\theta) &= (\theta^2 + \theta^6)Q'(\theta) + \theta^3 D'(\theta) \end{aligned}$$

donde en $(*)$ se ha tenido en cuenta la insesgadez de T en θ^3 y en $(**)$ que $\text{Var}_\theta[X] = E_\theta[X^2] - (E_\theta[X])^2$.

Juntando ambos miembros

$$\frac{\partial}{\partial \theta} E_\theta[T(X)] = E_\theta \left[T(X) \frac{\partial \ln f_\theta(X)}{\partial \theta} \right] \iff 3\theta^2 = (\theta^2 + \theta^6)Q'(\theta) + \theta^3 D'(\theta) \quad (2)$$

Podemos construir un sistema de ecuaciones con (1) y (2), y obtener $Q(\theta)$ y $D(\theta)$:

$$\begin{cases} D'(\theta) = -\theta^3 Q'(\theta) \\ 3\theta^2 = (\theta^2 + \theta^6)Q'(\theta) + \theta^3 D'(\theta) \end{cases}$$

Sustituimos la primera expresión en la segunda

$$\begin{aligned} 3\theta^2 &= (\theta^2 + \theta^6)Q'(\theta) + \theta^3 D'(\theta) = (\theta^2 + \theta^6)Q'(\theta) + \theta^3(-\theta^3 Q'(\theta)) = \\ &\theta^2 Q'(\theta) + \theta^6 Q'(\theta) - \theta^6 Q'(\theta) = \theta^2 Q'(\theta) \iff \\ &Q'(\theta) = 3 \end{aligned}$$

y ahora integramos respecto de θ para hallar $Q(\theta)$:

$$Q'(\theta) = 3 \implies Q(\theta) + c_1 = \int Q'(\theta) d\theta = \int 3 d\theta = 3\theta + c_2 \implies Q(\theta) = 3\theta + c_2 - c_1$$

llamando $c_2 - c_1 \equiv c$, determinamos su valor con la condición inicial $Q(1) = 0$.

$$3 + c = Q(1) = 0 \implies c = -3 \implies Q(\theta) = 3\theta - 3 = 3(\theta - 1)$$

Como hemos visto que $Q'(\theta) = 3$, sustituimos esta expresión en (1) y obtenemos $D(\theta)$ igual que con $Q(\theta)$:

$$\begin{aligned} D'(\theta) &= -\theta^3 Q'(\theta) = -3\theta^3 \implies \\ D(\theta) + k_1 &= \int D'(\theta) d\theta = -3 \int \theta^3 d\theta = -3 \frac{\theta^4}{4} + k_2 = \frac{-3\theta^4}{4} + k_2 \implies D(\theta) = \frac{-3\theta^4}{4} + k_2 - k_1 \end{aligned}$$

llamando $k_2 - k_1 \equiv k$, determinamos su valor con la condición inicial $D(0) = 1$.

$$k = D(0) = 1 \implies D(\theta) = 1 - \frac{3\theta^4}{4}$$

En resumen, las funciones $Q(\theta)$ y $D(\theta)$ son:

$$Q(\theta) = 3(\theta - 1) \quad D(\theta) = 1 - \frac{3\theta^4}{4}$$

- b) Sea (X_1, \dots, X_n) una muestra aleatoria simple de una variable X con función masa de probabilidad:

$$\begin{aligned} P_p[X = x] &= p^2(x-1)(1-p)^{x-2}, \quad x = 2, 3, \dots; \quad 0 < p < 1 \\ \left(\mathbb{E}_p[X] = \frac{2}{p}, \quad \text{Var}_p[X] = \frac{2(1-p)}{p^2} \right) \end{aligned}$$

- a) Sabiendo que la familia de distribuciones es regular, calcular la función de información asociada a la muestra. Encontrar la clase de funciones paramétricas que admiten estimador eficiente y los estimadores correspondientes.

Tenemos que $p = \theta$, y que $f_\theta(x) = P_p[X = x]$. Para lo primero que se pide, sabemos que

$$I_{X_1, \dots, X_n}(\theta) = n I_X(\theta) = n \text{Var}_\theta \left[\frac{\partial \ln f_\theta(X)}{\partial \theta} \right]$$

Calculamos

$$\begin{aligned} \ln f_\theta(x) &= 2 \ln(\theta) + \ln(x-1) + (x-2) \ln(1-\theta) \implies \\ \frac{\partial \ln f_\theta(x)}{\partial \theta} &= \frac{2}{\theta} - \frac{x-2}{1-\theta} = \frac{2(1-\theta) - (x-2)\theta}{\theta(1-\theta)} = \frac{2-2\theta-x\theta+2\theta}{\theta(1-\theta)} = \frac{2-x\theta}{\theta(1-\theta)} = \\ &\quad \frac{2}{\theta(1-\theta)} - \frac{1}{1-\theta}x \end{aligned}$$

y

$$\begin{aligned} \text{Var}_\theta \left[\frac{\partial \ln f_\theta(X)}{\partial \theta} \right] &= \text{Var}_\theta \left[\frac{2}{\theta(1-\theta)} - \frac{1}{1-\theta}X \right] = \\ \left(\frac{1}{1-\theta} \right)^2 \text{Var}_\theta[X] &= \frac{1}{(1-\theta)^2} \frac{2(1-\theta)}{\theta^2} = \frac{2}{\theta^2(1-\theta)} \end{aligned}$$

de donde

$$I_{X_1, \dots, X_n}(\theta) = n \text{Var}_\theta \left[\frac{\partial \ln f_\theta(X)}{\partial \theta} \right] = n \frac{2}{\theta^2(1-\theta)} = \frac{2n}{\theta^2(1-\theta)}$$

Ahora, para la segunda parte, buscamos aplicar el Teorema de Caracterización de Estimadores Eficientes. Para ello, obtenemos la función conjunta de la m.a.s. de X

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

Se supondrá a partir de ahora que $x_i \in \mathbb{N} \setminus \{1\}$ y $0 < \theta < 1$. De lo contrario, $f_\theta(x_i) = 0 \quad \forall i = 1, \dots, n$.

$$f_\theta^n(x_1, \dots, x_n) \stackrel{\text{indep}}{=} \prod_{i=1}^n f_\theta(x_i) = \prod_{i=1}^n \theta^2(x_i-1)(1-\theta)^{x_i-2} = \theta^{2n} (1-\theta)^{\sum_{i=1}^n (x_i-2)} \prod_{i=1}^n (x_i-1)$$

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

$$\begin{aligned}
 & 2n \ln(\theta) + \left(\sum_{i=1}^n x_i \right) \ln(1-\theta) - 2n \ln(1-\theta) + \sum_{i=1}^n \ln(x_i - 1) \\
 & \frac{\partial \ln f_\theta^n(x_1, \dots, x_n)}{\partial \theta} = \frac{2n}{\theta} - \frac{\sum_{i=1}^n x_i}{1-\theta} + \frac{2n}{1-\theta} = \frac{2n}{\theta(1-\theta)} - \frac{\sum_{i=1}^n x_i}{1-\theta} = \\
 & -\frac{1}{1-\theta} \left(\sum_{i=1}^n x_i - \frac{2n}{\theta} \right) = -\frac{n}{1-\theta} \left(\frac{1}{n} \sum_{i=1}^n x_i - \frac{2}{\theta} \right)
 \end{aligned}$$

Ahora, supongamos que $T(X_1, \dots, X_n)$ es un estimador de $g(\theta)$ función paramétrica derivable y estrictamente monótona ($g'(\theta) \neq 0 \quad \forall \theta \in \Theta =]0, 1[$). Como el enunciado nos dice que la familia es regular, y $0 < I_X(\theta) = \frac{2}{\theta^2(1-\theta)} < +\infty \quad \forall \theta \in \Theta$, T es eficiente si y solo si $\forall \theta \in \Theta \quad \exists a(\theta) \neq 0$ tal que

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

y

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

Como

$$\begin{aligned}
 & \frac{\partial \ln f_\theta^n(X_1, \dots, X_n)}{\partial \theta} = a(\theta)[T(X_1, \dots, X_n) - g(\theta)] \iff \\
 & -\frac{n}{1-\theta} \left(\frac{1}{n} \sum_{i=1}^n x_i - \frac{2}{\theta} \right) = a(\theta)[T(X_1, \dots, X_n) - g(\theta)]
 \end{aligned}$$

claramente por comparación se obtiene que

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

Es claro que $T(X_1, \dots, X_n) = \frac{1}{n} \sum_{i=1}^n X_i$ es un estimador, pues $\Theta =]0, 1[$, $g(\Theta) =]2, +\infty[$, y $T(x_1, \dots, x_n) \geq 2 \quad \forall (x_1, \dots, x_n) \in \mathcal{X}^n = (\mathbb{N} \setminus \{1\})^n$. Además, tanto $a(\theta)$ como $g(\theta)$ verifican todas las condiciones del teorema, pues

$$g'(\theta) = 2 \cdot (\theta^{-1})' = 2 \cdot (-\theta^{-2}) = -\frac{2}{\theta^2} < 0$$

porque $0 < \theta < 1$ y $a(\theta) = -\frac{n}{1-\theta} < 0 \quad \forall n \in \mathbb{N}$, en particular, $a(\theta) \neq 0 \quad \forall \theta \in \Theta$, y usando la aditividad de la función de información de Fisher, obtenemos

$$\frac{2n}{\theta^2(1-\theta)} = I_{(X_1, \dots, X_n)}(\theta) = a(\theta)g'(\theta) = -\frac{n}{1-\theta} \left(-\frac{2}{\theta^2} \right) = \frac{2n}{\theta^2(1-\theta)}$$

Por un corolario visto en teoría, como $T(X_1, \dots, X_n)$ 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.

- b) Calcular la cota para la varianza de estimadores insesgados en p^2 , regulares, y justificar si se alcanza o no dicha cota.

En este apartado $g(\theta) = \theta^2$. Ya se han comprobado las hipótesis del Teorema de Cota de FCR en el apartado anterior, 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 $T(X_1, \dots, X_n)$:

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

Ahora, $g'(\theta) = 2\theta$, y ya hemos visto que

$$I_{(X_1, \dots, X_n)}(\theta) = \frac{2n}{\theta^2(1-\theta)}$$

luego la cota inferior para la varianza de todo estos estimadores será

$$\text{Var}_\theta[T(X_1, \dots, X_n)] \geq \frac{(2\theta)^2\theta^2(1-\theta)}{2n} = \frac{2\theta^4(1-\theta)}{n}$$

La cota no se alcanza por reducción al absurdo. Si lo hiciera, entonces tendríamos que $g(\theta) = \theta^2$ 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, $0 < I_X(\theta) = 2/\theta^2(1-\theta) < +\infty \quad \forall \theta \in \Theta$ y $g(\theta) = \theta^2$ es una función paramétrica derivable). Sin embargo, hemos visto en el apartado anterior que las únicas funciones paramétricas que admiten estimador eficiente en esta familia son las de la forma $a\frac{\theta^2}{\theta} + b$, $a, b \in \mathbb{R}$. Como $\nexists a, b \in \mathbb{R} : \theta^2 = a\frac{\theta^2}{\theta} + b$, hemos llegado a contradicción.

Ejercicio 3 (2 puntos). Sea (X_1, \dots, X_n) una muestra aleatoria simple de una variable X con distribución $\mathcal{N}(\mu, \sigma^2)$, siendo ambos parámetros desconocidos.

- a) Especificar la variable usada para hacer inferencia sobre σ^2 . Deducir su distribución, detallando y justificando exhaustivamente cada paso.

Estamos en el caso en que se quiere hacer inferencia sobre σ^2 teniendo en cuenta que la media μ es desconocida. Por teoría sabemos que la variable usada es

$$\frac{(n-1)S^2}{\sigma^2}$$

Veamos qué distribución sigue esta variable.

Demostración.

$$\sum_{i=1}^n (X_i - \mu)^2 = \sum_{i=1}^n (\overbrace{X_i - \bar{X}}^a + \overbrace{\bar{X} - \mu}^b)^2$$

y aplicamos el cuadrado de un binomio $(a+b)^2 = a^2 + b^2 + 2ab$:

$$\begin{aligned} \sum_{i=1}^n (X_i - \mu)^2 &= \sum_{i=1}^n ((X_i - \bar{X})^2 + (\bar{X} - \mu)^2 + 2(X_i - \bar{X})(\bar{X} - \mu)) = \\ &= \sum_{i=1}^n (X_i - \bar{X})^2 + \sum_{i=1}^n (\bar{X} - \mu)^2 + \sum_{i=1}^n 2(X_i - \bar{X})(\bar{X} - \mu) = \\ &\quad \xrightarrow{0} (n-1)S^2 + n(\bar{X} - \mu)^2 \end{aligned}$$

Dividimos a cada lado de los extremos de la cadena de igualdades por $\sigma^2 > 0$:

$$\underbrace{\frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma^2}}_A = \underbrace{\frac{(n-1)S^2}{\sigma^2}}_B + \underbrace{\frac{n(\bar{X} - \mu)^2}{\sigma^2}}_C$$

Hallamos la distribución de A , C , y la que nos interesa, B :

- A. El enunciado nos dice que $X \rightsquigarrow \mathcal{N}(\mu, \sigma^2)$, luego, como (X_1, \dots, X_n) es una m.a.s. de X , se tiene que

$$X_i \rightsquigarrow \mathcal{N}(\mu, \sigma^2) \quad \forall i = 1, \dots, n$$

Tipificando para cada $i = 1, \dots, n$:

$$\frac{X_i - \mu}{\sigma} \rightsquigarrow \mathcal{N}(0, 1) \xrightarrow{(*)} \frac{(X_i - \mu)^2}{\sigma^2} \rightsquigarrow \chi^2(1)$$

donde en $(*)$ se ha aplicado una propiedad de la $\chi^2(n)$. Nuevamente, como (X_1, \dots, X_n) es una m.a.s. de X , las variables de la m.a.s. son independientes, y se puede usar la reproductividad de la $\chi^2(n)$:

$$\frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma^2} \rightsquigarrow \chi^2(n)$$

- $C \rightsquigarrow \chi^2(1)$, porque sabemos que $\bar{X} \rightsquigarrow \mathcal{N}(\mu, \sigma^2/n)$, y tipificando:

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \rightsquigarrow \mathcal{N}(0, 1)$$

y aplicando la misma propiedad (*) que antes

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \rightsquigarrow \mathcal{N}(0, 1) \xrightarrow{(*)} \frac{(\bar{X} - \mu)^2}{\sigma^2/n} = \frac{n(\bar{X} - \mu)^2}{\sigma^2} \rightsquigarrow \chi^2(1)$$

Ahora, como $A = B + C$, y B y C son independientes porque son funciones medibles de variables independientes (B es función medible de S^2 y C función medible de \bar{X} , independientes por el Lema de Fisher), podemos utilizar el teorema de multiplicación de esperanzas:

$$M_A(t) = E[e^{tA}] = E[e^{t(B+C)}] = E[e^{tB}e^{tC}] \stackrel{TME}{=} E[e^{tB}]E[e^{tC}] = M_B(t)M_C(t)$$

La función generatriz de momentos de la χ^2 es conocida,

$$M(t) = \frac{1}{(1-2t)^{n/2}}, \quad t < 1/2$$

por lo que basta despejar $M_B(t)$, con $A \rightsquigarrow \chi^2(n)$ y $C \rightsquigarrow \chi^2(1)$:

$$\begin{aligned} \frac{1}{(1-2t)^{n/2}} &= M_A(t) = M_B(t)M_C(t) = M_B(t)\frac{1}{(1-2t)^{1/2}} \iff \\ M_B(t) &= \frac{(1-2t)^{1/2}}{(1-2t)^{n/2}} = (1-2t)^{1/2-n/2} = (1-2t)^{(1-n)/2} = \frac{1}{(1-2t)^{(n-1)/2}} \quad t < 1/2 \end{aligned}$$

y por el Teorema de Unicidad de la Función Generatriz de Momentos,

$$\frac{(n-1)S^2}{\sigma^2} = B \rightsquigarrow \chi^2(n-1)$$

lo que concluye la prueba. □

- b) Deducir, detallando el método usado, el intervalo de confianza de menor longitud esperada uniformemente para μ a nivel de confianza $1 - \alpha$.

Usaremos el método de la cantidad pivotal. Sabemos que un pivote para μ , siendo σ^2 desconocida, es, considerando $\theta = \mu$ y $\Theta = \mathbb{R}$, $T : \mathcal{X}^n \times \Theta \rightarrow \mathbb{R}$, dado por

$$T(X_1, \dots, X_n; \theta) = \frac{\bar{X} - \theta}{S/\sqrt{n}} \rightsquigarrow t(n-1)$$

variable aleatoria cuya distribución, $t(n-1)$, no depende de θ . Además:

- a) T es estrictamente monótona en θ , pues

$$\frac{\partial T}{\partial \theta} = \frac{\partial}{\partial \theta} \left(\frac{\bar{X} - \theta}{S/\sqrt{n}} \right) = -\frac{1}{S/\sqrt{n}} < 0, \quad \forall \theta \in \Theta$$

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

$$T = \lambda \iff \frac{\bar{X} - \theta}{S/\sqrt{n}} = \lambda \iff \bar{X} - \theta = \lambda(S/\sqrt{n}) \iff \theta = \bar{X} - \lambda(S/\sqrt{n}) \in \Theta = \mathbb{R}$$

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,

$$\begin{aligned} \lambda_1 < T < \lambda_2 &\iff \lambda_1 < \frac{\bar{X} - \theta}{S/\sqrt{n}} < \lambda_2 \iff \lambda_1(S/\sqrt{n}) < \bar{X} - \theta < \lambda_2(S/\sqrt{n}) \iff \\ &\lambda_1(S/\sqrt{n}) - \bar{X} < -\theta < \lambda_2(S/\sqrt{n}) - \bar{X} \iff \\ &\bar{X} - \lambda_2(S/\sqrt{n}) < \theta < \bar{X} - \lambda_1(S/\sqrt{n}) \end{aligned}$$

y el intervalo tendrá longitud $L = (\bar{X} - \lambda_1(S/\sqrt{n})) - (\bar{X} - \lambda_2(S/\sqrt{n})) = (S/\sqrt{n})(\lambda_2 - \lambda_1)$, que es constante, luego coincidirá con la longitud esperada $E_\theta[L] = L \quad \forall \theta \in \Theta$.

Sea ahora F la función de distribución de la distribución t de Student $t(n-1)$, 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 :

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

Como la distribución t de Student tiene las mismas propiedades de simetría con respecto al origen que la $\mathcal{N}(0, 1)$, las únicas posibilidades son o bien $\lambda_1 = \lambda_2$ o bien $\lambda_1 = -\lambda_2$. La primera de las dos opciones se descarta ya que no verifica la restricción ($F(\lambda_2) - F(\lambda_1) = 0 \neq 1 - \alpha$), por lo que tiene que ser $\lambda_1 = -\lambda_2$. Teniendo nuevamente en cuenta la restricción $F(\lambda_2) - F(\lambda_1) = 1 - \alpha$, estos quedan determinados por:

$$\lambda_2 = t_{n-1; \alpha/2}, \quad \lambda_1 = -\lambda_2 = -t_{n-1; \alpha/2}$$

siendo $P[T > t_{n-1; \alpha/2}] = \alpha/2$ y $T \rightsquigarrow t(n-1)$. En conclusión, el intervalo de confianza para θ de mínima longitud esperada uniformemente a nivel de confianza $1 - \alpha$ es:

$$[\bar{X} - \lambda_2(S/\sqrt{n}), \bar{X} - \lambda_1(S/\sqrt{n})] = [\bar{X} - t_{n-1; \alpha/2}(S/\sqrt{n}), \bar{X} + t_{n-1; \alpha/2}(S/\sqrt{n})]$$

Ejercicio 4 (1.75 puntos). El retraso medio diario (en minutos) de los trabajadores de cierta empresa tiene una distribución $U(0, \theta)$. Para contrastar la hipótesis de que el máximo retraso medio diario es de 5 minutos frente a que es de 4 minutos, se midió dicha variable durante 10 días elegidos al azar, observando que el máximo retraso medio en esos 10 días fue de 3 minutos. Obtener los tests más potentes de tamaños 0,05 y 0,15 para el problema de contraste planteado. Decidir, a la vista de lo observado, si debe rechazarse la hipótesis nula en cada uno de los casos.

Sea $X \equiv$ “Retraso medio diario de los trabajadores de la empresa”. Nos dicen que $X \sim U(0, \theta)$. Al considerar (X_1, \dots, X_n) m.a.s. de X con $n = 10$, se observó que $X_{(n)} = 3$. El contraste que se quiere resolver es de la forma

$$\begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta = \theta_1 \end{cases}$$

con $4 = \theta_1 < \theta_0 = 5$. Como el contraste es de hipótesis simple frente a hipótesis simple, sabemos por el Lema de Neyman-Pearson que el Test de Neyman-Pearson será el más potente de tamaño α , de la forma

$$\varphi(X_1, \dots, X_n) = \begin{cases} 1 & \text{si } \lambda(X_1, \dots, X_n) > k \\ \gamma & \text{si } \lambda(X_1, \dots, X_n) = k \\ 0 & \text{si } \lambda(X_1, \dots, X_n) < k \end{cases}$$

para ciertos $\gamma \in [0, 1]$, $k \in \mathbb{R}$ y

$$\lambda(X_1, \dots, X_n) = \frac{f_1^n(X_1, \dots, X_n)}{f_0^n(X_1, \dots, X_n)}$$

Definimos el espacio muestral y el espacio paramétrico, en ambos casos dependiente del parámetro θ . Si estamos en H_0 , entonces $\mathcal{X}_0 = [0, \theta_0]$, y $\Theta_0 = \{\theta_0\}$. Análogamente, si estamos en H_1 , entonces $\mathcal{X}_1 = [0, \theta_1]$, y $\Theta_1 = \{\theta_1\}$. Tenemos entonces que

$$\Theta = \Theta_0 \cup \Theta_1 = \{\theta_0, \theta_1\}$$

y como $\theta_1 < \theta_0$, entonces $[0, \theta_1] \subset [0, \theta_0]$ y

$$\mathcal{X} = \mathcal{X}_0 \cup \mathcal{X}_1 = [0, \theta_0] \cup [0, \theta_1] = [0, \theta_0]$$

Consecuentemente

$$\begin{aligned} \mathcal{X}^n &= \{(x_1, \dots, x_n) \in \mathbb{R}^n : 0 \leq x_i \leq \theta_0 \quad \forall i = 1, \dots, n\} = \\ &\quad \{(x_1, \dots, x_n) \in \mathbb{R}^n : 0 \leq x_{(1)} \wedge x_{(n)} \leq \theta_0\} \end{aligned}$$

Podemos considerar entonces $(x_1, \dots, x_n) \in \mathcal{X}^n$, y obtener la función conjunta $f_\theta^n(x_1, \dots, x_n)$, que es

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

Y vemos que

$$x \leq \theta \iff I_{\mathbb{R}_0^-}(x - \theta) = 1$$

luego

$$x_i \leq \theta \quad \forall i = 1, \dots, n \iff x_{(n)} \leq \theta \iff I_{\mathbb{R}_0^-}(x_{(n)} - \theta) = 1$$

Por tanto

$$\begin{aligned} f_\theta^n(x_1, \dots, x_n) &= \prod_{i=1}^n \frac{1}{\theta} I_{\mathbb{R}_0^-}(x_{(n)} - \theta) = \\ &\theta^{-n} I_{\mathbb{R}_0^-}(x_{(n)} - \theta) \end{aligned}$$

Se tiene entonces que

$$f_0^n(x_1, \dots, x_n) = \theta_0^{-n} I_{\mathbb{R}_0^-}(x_{(n)} - \theta_0) = \theta_0^{-n} \quad \forall (x_1, \dots, x_n) \in \mathcal{X}^n$$

$$f_1^n(x_1, \dots, x_n) = \theta_1^{-n} I_{\mathbb{R}_0^-}(x_{(n)} - \theta_1) \quad \forall (x_1, \dots, x_n) \in \mathcal{X}^n$$

y podemos obtener $\lambda(x_1, \dots, x_n)$, con $(x_1, \dots, x_n) \in \mathcal{X}^n$:

$$\begin{aligned} \lambda(x_1, \dots, x_n) &= \frac{f_1^n(x_1, \dots, x_n)}{f_0^n(x_1, \dots, x_n)} = \\ \frac{\theta_1^{-n}}{\theta_0^{-n}} I_{\mathbb{R}_0^-}(x_{(n)} - \theta_1) &= \frac{\begin{cases} \theta_1^{-n} & \text{si } x_{(n)} \leq \theta_1 \\ 0 & \text{si } x_{(n)} > \theta_1 \end{cases}}{\theta_0^{-n}} \end{aligned}$$

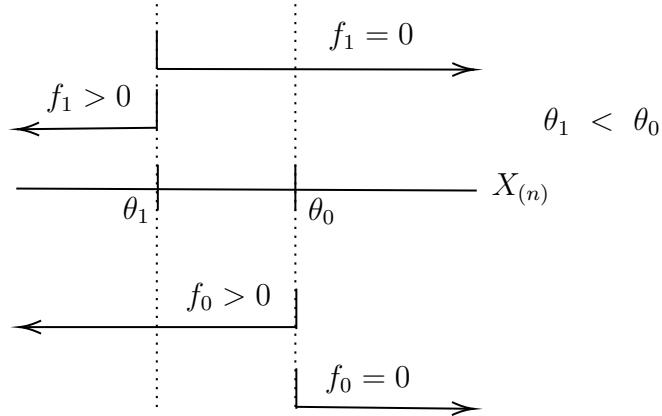
Ahora, si $x_{(n)} \leq \theta_1$, simplificamos

$$\frac{\theta_1^{-n}}{\theta_0^{-n}} = \frac{\theta_0^n}{\theta_1^n} = \left(\frac{\theta_0}{\theta_1}\right)^n$$

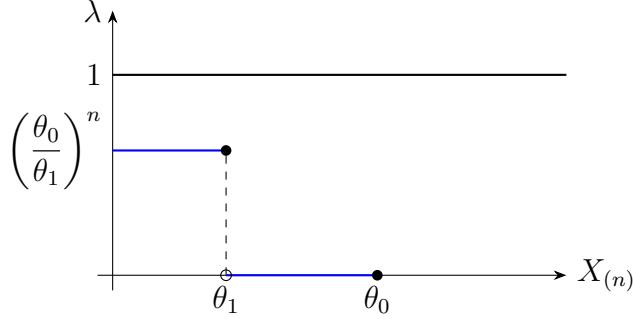
de donde

$$\lambda(x_1, \dots, x_n) = \frac{f_1^n(x_1, \dots, x_n)}{f_0^n(x_1, \dots, x_n)} = \begin{cases} \left(\frac{\theta_0}{\theta_1}\right)^n & \text{si } x_{(n)} \leq \theta_1 \\ 0 & \text{si } x_{(n)} > \theta_1 \end{cases} \quad \forall (x_1, \dots, x_n) \in \mathcal{X}^n$$

Gráficamente, como $\theta_1 < \theta_0$:



La semirrecta en que $f_0 = 0$, que se corresponde con la condición $x_{(n)} > \theta_0$, no nos interesa ya que $(x_1, \dots, x_n) \notin \mathcal{X}^n$. Gráficamente, la situación en la que estamos es la siguiente:



Tenemos que $k \in \left\{ 0, \left(\frac{\theta_0}{\theta_1}\right)^n \right\}$, luego distinguimos entre estos dos casos.

Si $\boxed{k = 0}$, entonces el test será:

$$\varphi(X_1, \dots, X_n) = \begin{cases} 1 & \text{si } \lambda(X_1, \dots, X_n) > k \iff X_{(n)} \leq \theta_1 \\ \gamma & \text{si } \lambda(X_1, \dots, X_n) = k \iff X_{(n)} > \theta_1 \\ 0 & \text{si } \lambda(X_1, \dots, X_n) < k \text{ nunca} \end{cases}$$

es decir:

$$\varphi(X_1, \dots, X_n) = \begin{cases} 1 & \text{si } X_{(n)} \leq \theta_1 \\ \gamma & \text{si } X_{(n)} > \theta_1 \end{cases}$$

Determinamos γ imponiendo tamaño α :

$$\begin{aligned} \alpha &\stackrel{def}{=} \sup_{\theta \in \Theta_0} \beta_\varphi(\theta) = \sup_{\theta \in \Theta_0} E_\theta[\varphi(X_1, \dots, X_n)] = E_{\theta_0}[\varphi(X_1, \dots, X_n)] = \\ &= 1 \cdot P_{\theta_0}[X_{(n)} \leq \theta_1] + \gamma P_{\theta_0}[X_{(n)} > \theta_1] \end{aligned}$$

Para calcular las probabilidades, obtenemos la función de distribución $F_X(t)$:

$$F_X(t) \stackrel{def}{=} \int_0^t \frac{1}{\theta} dx = \frac{1}{\theta} \int_0^t dx = \frac{t}{\theta}$$

Sabemos por teoría que, para $T = X_{(n)}$, la distribución del máximo verifica

$$P_\theta[T < t] = F_T(t) = (F_X(t))^n = \left(\frac{t}{\theta}\right)^n$$

Por lo tanto,

$$P_{\theta_0}[T \leq \theta_1] = \left(\frac{\theta_1}{\theta_0}\right)^n$$

y

$$P_{\theta_0}[T > \theta_1] = 1 - P_{\theta_0}[T \leq \theta_1] = 1 - \left(\frac{\theta_1}{\theta_0}\right)^n$$

de donde

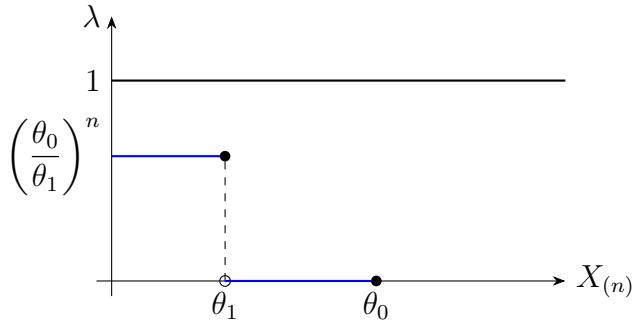
$$\alpha = 1 \cdot P_{\theta_0}[X_{(n)} \leq \theta_1] + \gamma P_{\theta_0}[X_{(n)} > \theta_1] = \left(\frac{\theta_1}{\theta_0}\right)^n + \gamma \left(1 - \left(\frac{\theta_1}{\theta_0}\right)^n\right) \iff$$

$$0 \leq \gamma = \frac{\alpha - \left(\frac{\theta_1}{\theta_0}\right)^n}{1 - \left(\frac{\theta_1}{\theta_0}\right)^n} \leq 1 \iff \alpha \geq \left(\frac{\theta_1}{\theta_0}\right)^n$$

y el test resultante es:

$$\varphi(X_1, \dots, X_n) = \begin{cases} 1 & \text{si } X_{(n)} \leq \theta_1 \\ \frac{\alpha - \left(\frac{\theta_1}{\theta_0}\right)^n}{1 - \left(\frac{\theta_1}{\theta_0}\right)^n} & \text{si } X_{(n)} > \theta_1 \end{cases}$$

Si $k = \left(\frac{\theta_0}{\theta_1}\right)^n$, el test será, mirando nuevamente la misma gráfica



el que sigue:

$$\varphi(X_1, \dots, X_n) = \begin{cases} 1 & \text{si } \lambda(X_1, \dots, X_n) > k \text{ nunca} \\ \gamma & \text{si } \lambda(X_1, \dots, X_n) = k \iff X_{(n)} \leq \theta_1 \\ 0 & \text{si } \lambda(X_1, \dots, X_n) < k \iff X_{(n)} > \theta_1 \end{cases}$$

es decir:

$$\varphi(X_1, \dots, X_n) = \begin{cases} \gamma & \text{si } X_{(n)} \leq \theta_1 \\ 0 & \text{si } X_{(n)} > \theta_1 \end{cases}$$

Determinamos γ igual que antes imponiendo tamaño α :

$$\alpha \stackrel{def}{=} \sup_{\theta \in \Theta_0} \beta_\varphi(\theta) = \sup_{\theta \in \Theta_0} E_\theta[\varphi(X_1, \dots, X_n)] = E_{\theta_0}[\varphi(X_1, \dots, X_n)] =$$

$$\gamma \cdot P_{\theta_0}[X_{(n)} \leq \theta_1] + 0 \cdot P_{\theta_0}[X_{(n)} > \theta_1] = \gamma \cdot P_{\theta_0}[X_{(n)} \leq \theta_1]$$

Ya sabemos del caso anterior que

$$P_{\theta_0}[T \leq \theta_1] = \left(\frac{\theta_1}{\theta_0}\right)^n$$

luego

$$\alpha = \gamma \cdot P_{\theta_0}[X_{(n)} \leq \theta_1] = \gamma \left(\frac{\theta_1}{\theta_0} \right)^n \iff 0 \leq \gamma = \frac{\alpha}{\left(\frac{\theta_1}{\theta_0} \right)^n} = \alpha \left(\frac{\theta_0}{\theta_1} \right)^n \leq 1 \iff \alpha \leq \left(\frac{\theta_1}{\theta_0} \right)^n$$

y el test resultante es:

$$\varphi(X_1, \dots, X_n) = \begin{cases} \alpha \left(\frac{\theta_0}{\theta_1} \right)^n & \text{si } X_{(n)} \leq \theta_1 \\ 0 & \text{si } X_{(n)} > \theta_1 \end{cases}$$

Sustituimos ahora los datos del enunciado, $\theta_0 = 5$, $\theta_1 = 4$, $n = 10$, y $X_{(n)} = 3$. Distinguimos según el tamaño pedido:

- $\alpha = 0,05$. Como

$$0,05 = \alpha \leq \left(\frac{\theta_1}{\theta_0} \right)^n = \left(\frac{4}{5} \right)^{10} \approx 0,1074$$

el test más potente en este caso es

$$\varphi(X_1, \dots, X_n) = \begin{cases} \alpha \left(\frac{\theta_0}{\theta_1} \right)^n & \text{si } X_{(n)} \leq \theta_1 \\ 0 & \text{si } X_{(n)} > \theta_1 \end{cases}$$

y sustituyendo

$$\varphi(X_1, \dots, X_n) = \begin{cases} 0,05 \left(\frac{5}{4} \right)^{10} & \text{si } X_{(n)} \leq 4 \\ 0 & \text{si } X_{(n)} > 4 \end{cases}$$

como $3 = X_{(n)} \leq 4$, se rechaza H_0 con probabilidad:

$$0,05 \left(\frac{5}{4} \right)^{10} \approx 0,4657$$

- $\alpha = 0,15$. Como

$$0,15 = \alpha \geq \left(\frac{\theta_1}{\theta_0} \right)^n = \left(\frac{4}{5} \right)^{10} \approx 0,1074$$

el test más potente en este caso es

$$\varphi(X_1, \dots, X_n) = \begin{cases} 1 & \text{si } X_{(n)} \leq \theta_1 \\ \frac{\alpha - \left(\frac{\theta_1}{\theta_0} \right)^n}{1 - \left(\frac{\theta_1}{\theta_0} \right)^n} & \text{si } X_{(n)} > \theta_1 \end{cases}$$

y sustituyendo

$$\varphi(X_1, \dots, X_n) = \begin{cases} 1 & \text{si } X_{(n)} \leq 4 \\ \frac{0,15 - \left(\frac{5}{4}\right)^{10}}{1 - \left(\frac{5}{4}\right)^{10}} & \text{si } X_{(n)} > 4 \end{cases}$$

como $3 = X_{(n)} \leq 4$, se rechaza H_0 con probabilidad 1.

Ejercicio 5 (1.5 puntos).

- a) Test de Kolmogorov–Smirnov. Comparación con χ^2 .

Comenzamos describiendo el Test de Kolmogorov–Smirnov. Sea una función de distribución específica F_0 , y sea (X_1, \dots, X_n) una m.a.s. de una v.a. X continua que se distribuye según una función de distribución F que es completamente desconocida. El contraste a resolver es

$$\begin{cases} H_0 : F = F_0 \\ H_1 : F \neq F_0 \end{cases}$$

El estadístico que se usa para resolver el problema es el estadístico de Kolmogorov–Smirnov

$$D(X_1, \dots, X_n) = \sup_{x \in \mathbb{R}} |F_{X_1, \dots, X_n}^*(x) - F_0(x)|$$

Intuitivamente, este mide la máxima distancia vertical entre la función de distribución empírica F_{X_1, \dots, X_n}^* y la función de distribución que se está contrastando como teórica F_0 . Más teóricamente, este test se basa en el teorema de Glivenko-Cantelli:

Teorema 0.1 (de Glivenko-Cantelli). *Sea $\{X_n\}_{n \in \mathbb{N}}$ una sucesión de v.a.i.i.d. con función de distribución común F . Si F_{X_1, \dots, X_n}^* es la función de distribución muestral asociada a la m.a.s. (X_1, \dots, X_n) , se verifica que F_{X_1, \dots, X_n}^* converge casi seguramente y uniformemente a la función de distribución de X , F .*

$$P \left\{ \lim_{n \rightarrow +\infty} \sup_{x \in \mathbb{R}} |F_{X_1, \dots, X_n}^*(x) - F(x)| = 0 \right\} = 1$$

Por otro lado, el test χ^2 de Pearson se basa en comparar frecuencias observadas y esperadas. Para ello, requiere una agrupación en k grupos distintos (al menos 5 en la práctica). Resuelve el mismo contraste que el test de Kolmogorov–Smirnov si la hipótesis nula es simple. Si fuera compuesta, deben estimarse los parámetros de los que depende F_0 (normalmente por máxima verosimilitud) y luego aplicar el test χ^2 de Pearson a esa adaptación. El estadístico de contraste es

$$\chi^2 = \sum_{i=1}^k \frac{(N_i - E_i)^2}{E_i}$$

La aproximación que hace este test es asintótica y en la práctica se exigen frecuencias esperadas $E_i \geq 5$ para cada $i \in \{1, \dots, k\}$.

Algunas ventajas del test de Kolmogorov–Smirnov respecto al test χ^2 de Pearson son las siguientes:

- El test de Kolmogorov-Smirnov no necesita hacer particiones de los datos, mientras que el test χ^2 de Pearson sí.
- Para v.a. continuas, es más apropiado usar el test de Kolmogorov-Smirnov que el test χ^2 de Pearson.
- Bajo H_0 , si F_0 es continua, la distribución del estadístico de Kolmogorov-Smirnov no depende de F_0 , mientras que el test χ^2 de Pearson es asintótico.

De igual manera, algunas ventajas del test χ^2 de Pearson respecto al test de Kolmogorov–Smirnov son las siguientes:

- El test χ^2 puede aplicarse tanto a variables aleatorias discretas como continuas (estas últimas tras agrupar en clases), mientras que el test de Kolmogorov–Smirnov está pensado fundamentalmente para distribuciones continuas.
 - El test χ^2 permite dar mayor peso a determinadas regiones (por ejemplo, las colas) mediante una elección adecuada de las particiones, mientras que el test de Kolmogorov–Smirnov se basa únicamente en la distancia máxima y no distingue zonas.
- b) Se recogen observaciones de la facturación (en miles de euros) de un cierto supermercado durante 25 días. Se quiere contrastar si la facturación se distribuye uniformemente.

10.1	7.3	4.2	3.1	8.4	9.2	1.1	1	8.1	6.4	10.5	1.2
1.8	2.6	2.2	3.9	2.9	6.5	8.1	8.3	3.8	4.2	7.3	8.9

En este caso, aunque la variable aleatoria es continua, dado que tenemos frecuencias, es más apropiado usar el test χ^2 de Pearson. Sea

$$X \equiv \text{“Facturación del supermercado en un día”}$$

El contraste a resolver es

$$\begin{cases} H_0 : X \rightsquigarrow U(a, b) & a, b \in \mathbb{R} : a < b \\ H_1 : X \not\rightsquigarrow U(a, b) \end{cases}$$

Vemos que la hipótesis nula es compuesta, luego primeramente debemos estimar el valor de los parámetros a y b . Sabemos que el EMV de a es $\hat{a} = X_{(1)} = 1$ y el EMV de b es $\hat{b} = X_{(n)} = 10,5$. El contraste adaptado sería

$$\begin{cases} H_0 : X \rightsquigarrow U(1, 10.5) \\ H_1 : X \not\rightsquigarrow U(1, 10.5) \end{cases}$$

Denotemos por N_1, \dots, N_k las frecuencias observadas en las k clases consideradas, y por

$$\hat{p}_i = P_{\hat{a}, \hat{b}}(X \in A_i), \quad i = 1, \dots, k$$

las probabilidades teóricas bajo H_0 con los parámetros a y b estimados por \hat{a} y \hat{b} , respectivamente.

El estadístico de contraste viene dado por

$$\hat{\chi}(N_1, \dots, N_k) = \sum_{i=1}^k \frac{(N_i - n\hat{p}_i)^2}{n\hat{p}_i}$$

Como el parámetro se ha estimado a partir de los mismos datos, la distribución asintótica bajo H_0 es

$$\hat{\chi}(N_1, \dots, N_k) \rightsquigarrow_{H_0} \chi^2(k - q - 1)$$

con $q = 2$ el número de parámetros estimados.

Por teoría, para poder aplicar el test hay que verificar que

$$E_i^* = n\hat{p}_i \geq 5 \quad \forall i = 1, \dots, k, \quad n = 25, \quad \hat{p}_i = \int_{c_{i-1}}^{c_i} \frac{1}{\hat{b} - \hat{a}} dx = \frac{c_i - c_{i-1}}{\hat{b} - \hat{a}}$$

donde $A_i = [c_{i-1}, c_i]$, y $\hat{a} \leq c_i \leq c_{j-1} < c_j \leq \hat{b} \quad \forall i < j = 1, \dots, k$. Como

$$E_i = n\hat{p}_i = n \frac{c_i - c_{i-1}}{\hat{b} - \hat{a}} = 25 \frac{c_i - c_{i-1}}{10,5 - 1} = 25 \frac{c_i - c_{i-1}}{9,5}$$

podemos agrupar por medio de los percentiles. Se tomarán k clases con $\hat{p}_i = 1/k$, de tal manera que $E_i = n/k$, para cada $i = 1, \dots, k$. Como $n = 25$, será $E_i \geq 5$ si $k \leq 5$. Para que sean al menos 5 grupos, debe ser $k = 5$, y

$$E_i = 5 \quad \forall i = 1, \dots, k$$

Los extremos de los grupos (intervalos), en este caso los quintiles, vienen dados por $c_i = 1 + 9,5 \frac{i}{5}$:

c_1	c_2	c_3	c_4	c_5
1	2.9	4.8	6.7	8.6

Tenemos entonces la siguiente partición de $k = 5$ intervalos

$$A_i = [c_{i-1}, c_i] \quad i = 1, \dots, k-1 \quad A_k = [c_k, 10,5]$$

Ahora hay que contar las frecuencias observadas N_i de cada quintil, con $i = 1, \dots, k$. Vemos que

N_1	N_2	N_3	N_4	N_5
6	6	2	6	5

y obtenemos

$$\begin{aligned}\chi^2_{exp} &= \sum_{i=1}^5 \frac{(N_i - E_i)^2}{E_i} = \frac{(6-5)^2}{5} + \frac{(6-5)^2}{5} + \frac{(2-5)^2}{5} + \frac{(4-5)^2}{5} + \\ &\quad \frac{(5-5)^2}{5} = \frac{12}{5} = 2,4\end{aligned}$$

El test asintótico de tamaño α es

$$\varphi(X_1, \dots, X_n) = \begin{cases} 1 & \text{si } \hat{\chi}^2(N_1, \dots, N_k) \geq \chi^2_{k-1;\alpha} \\ 0 & \text{si } \hat{\chi}^2(N_1, \dots, N_k) < \chi^2_{k-1;\alpha} \end{cases}$$

con

$$p - valor = P_{H_0}[\hat{\chi}^2(N_1, \dots, N_k) \geq \chi^2_{exp}] \approx_{n \rightarrow +\infty} P[\chi^2(k-q-1) \geq \chi^2_{exp}]$$

y χ^2_{exp} el valor del estadístico obtenido con la muestra observada. Usando que $k-q-1 = 2$, obtenemos

$$p - valor \approx P[\chi^2(2) \geq 2,4] \approx 0,3$$

Como el $p - valor$ es grande (respecto a los niveles habituales de significación), no se rechaza H_0 , por lo que puede suponerse que la facturación del supermercado en un día se distribuye uniformemente.