

Inferencia Estadística

Examen I

FACULTAD
DE
CIENCIAS
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 I

Los Del DGIIM, losdeldgiim.github.io

José Manuel Sánchez Varbas

Granada, 2025

Asignatura Inferencia Estadística.

Curso Académico 2024-25.

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

Descripción Examen Ordinario.

Fecha 23 de Enero de 2025.

Ejercicio 1. Sean (X_1, \dots, X_{n_1}) y (Y_1, \dots, Y_{n_2}) m.a.s. de X e Y , variables que siguen $N(\mu_1, 4)$ y $N(\mu_2, 5)$ respectivamente.

- a) Si $\mu_1 = 2$, $\mu_2 = 3$ y sean (X_1, \dots, X_8) y (Y_1, \dots, Y_{10}) dos muestras de tamaño 8 y 10 respectivamente con medias muestrales \bar{X}, \bar{Y} , calcular el percentil 99 de

$$V = \frac{\bar{X} - \bar{Y} + 1}{\sqrt{\frac{\sum_{i=1}^8 (X_i - \bar{X})^2}{4} + \frac{\sum_{i=1}^{10} (Y_i - \bar{Y})^2}{5}}}. \quad (1)$$

- b) Calcular el intervalo de confianza para $\mu_1 - \mu_2$ de menor longitud esperada uniformemente a nivel de confianza $1 - \alpha$. ¿Cómo sería el intervalo si las varianzas fueran desconocidas pero iguales?

Ejercicio 2.

- a) Sea X una v.a. con función de densidad

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

Calcular el UMVUE y determinar para qué valores de n existe. ¿Es eficiente?

- b) Sea

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

en una familia regular según Fréchet–Cramér–Rao.

- b1) Sabiendo que $I_{X_1, \dots, X_n}(\theta) = \frac{2n}{\theta^2}$, calcular $E[X]$ y $\text{Var}[X]$.

- b2) Sabiendo que

$$\sum_{i=1}^n \frac{T(X_i)}{n}$$

es un estimador eficiente de $2/\theta$, calcular $T(x)$.

Ejercicio 3.

- a) Enunciar y demostrar el Teorema de Zehna definiendo previamente los siguientes conceptos: función de verosimilitud de un parámetro, función de verosimilitud de una función paramétrica y estimador máximo verosímil para funciones paramétricas.

- b) Calcular la función de verosimilitud de

$$\lambda = (\theta - 1)^2$$

asociada a una realización muestral cuyo máximo valor es 3 si

$$f_\theta(x) = e^{x-\theta}, \quad x \leq \theta, \theta > 0.$$

Ejercicio 4. Sea (X_1, \dots, X_n) una m.a.s. de X v.a. con función de densidad

$$f_\theta(x) = \frac{1}{\sqrt{8\theta}\sqrt{x-1}}, \quad 1 < x < 2\theta + 1. \quad (3)$$

Deducir el test más potente de tamaño arbitrario para contrastar

$$H_0 : \theta = \theta_0 \quad \text{frente a} \quad H_1 : \theta = \theta_1$$

siendo $\theta_1 < \theta_0$. Calcular la potencia.

Ejercicio 5.

a) Test de Kolmogorov–Smirnov.

- i) Plantear el problema de contraste.
- ii) Dar el valor del estadístico.
- iii) Enunciar el teorema que justifica su uso.
- iv) Ventajas frente al test χ^2 .

b) Se cuentan el número de tutorías a lo largo de un curso por 50 profesores. Se quiere contrastar si el número de tutorías por profesor se puede describir mediante una distribución de Poisson.

Número de tutorías	0	1	2	3	4	5
Número de profesores	2	5	10	14	12	7

Ejercicio 1. Sean (X_1, \dots, X_{n_1}) y (Y_1, \dots, Y_{n_2}) m.a.s. de X e Y , variables que siguen $N(\mu_1, 4)$ y $N(\mu_2, 5)$ respectivamente.

- a) Si $\mu_1 = 2$, $\mu_2 = 3$ y sean (X_1, \dots, X_8) y (Y_1, \dots, Y_{10}) dos muestras de tamaño 8 y 10 respectivamente con medias muestrales \bar{X}, \bar{Y} , calcular el percentil 99 de

$$V = \frac{\bar{X} - \bar{Y} + 1}{\sqrt{\frac{\sum_{i=1}^8 (X_i - \bar{X})^2}{4} + \frac{\sum_{i=1}^{10} (Y_i - \bar{Y})^2}{5}}}. \quad (1)$$

Del enunciado del problema sacamos que

$$X_i \rightsquigarrow N(\mu_1, 4), \quad i = 1, \dots, n_1$$

$$Y_j \rightsquigarrow N(\mu_2, 5) \quad j = 1, \dots, n_2$$

y en este apartado, $n_1 = 8$ y $n_2 = 10$. Buscamos obtener la distribución de V . Para ello, primero hallamos la distribución del numerador.

Como $\mu_1 = 2$ y $\mu_2 = 3$, entonces

$$\mathbb{E}[\bar{X} - \bar{Y} + 1] = \mathbb{E}[\bar{X}] - \mathbb{E}[\bar{Y}] + 1 = \mu_1 - \mu_2 + 1 = 2 - 3 + 1 = 0$$

Por teoría sabemos que

$$X \rightsquigarrow N(\mu_1, 4) \implies \bar{X} \rightsquigarrow N\left(\mu_1, \frac{4}{n_1}\right)$$

$$Y \rightsquigarrow N(\mu_2, 5) \implies \bar{Y} \rightsquigarrow N\left(\mu_2, \frac{5}{n_2}\right)$$

Así, asumiendo que (X_1, \dots, X_{n_1}) y (Y_1, \dots, Y_{n_2}) son independientes, entonces

$$\bar{X} - \bar{Y} \rightsquigarrow N\left(\mu_1 - \mu_2, \frac{4}{n_1} + \frac{5}{n_2}\right)$$

y se obtiene que

$$\text{Var}(\bar{X} - \bar{Y} + 1) = \text{Var}(\bar{X} - \bar{Y}) = \frac{4}{n_1} + \frac{5}{n_2} = \frac{4}{8} + \frac{5}{10} = \frac{1}{2} + \frac{1}{2} = 1$$

Deducimos entonces que $Z = \bar{X} - \bar{Y} + 1 \rightsquigarrow N(0, 1)$

Falta hallar la distribución del denominador. Usaremos que

$$\frac{(n_1 - 1)S_1^2}{\sigma_1^2} = \frac{\sum_{i=1}^{n_1} (X_i - \bar{X})^2}{\sigma_1^2} \rightsquigarrow \chi^2(n_1 - 1)$$

$$\frac{(n_2 - 1)S_2^2}{\sigma_2^2} = \frac{\sum_{j=1}^{n_2} (Y_j - \bar{Y})^2}{\sigma_2^2} \rightsquigarrow \chi^2(n_2 - 1)$$

y como estas últimas v.a. son independientes, por serlo las m.a.s. de X e Y , podemos aplicar la reproductividad de la distribución χ^2 :

$$W = \frac{\sum_{i=1}^{n_1} (X_i - \bar{X})^2}{\sigma_1^2} + \frac{\sum_{j=1}^{n_2} (Y_j - \bar{Y})^2}{\sigma_2^2} \rightsquigarrow \chi^2(n_1 - 1 + n_2 - 1) = \chi^2(n_1 + n_2 - 2) = \chi^2(16)$$

En estas condiciones puede aplicarse la construcción de la distribución T de Student

$$T = \frac{Z}{\sqrt{W/16}} \rightsquigarrow t(16)$$

y basta notar que

$$V = \frac{Z}{\sqrt{W}} = \frac{Z}{\sqrt{16 \cdot W/16}} = \frac{Z}{\sqrt{16} \sqrt{W/16}} = \frac{1}{4} \frac{Z}{\sqrt{W/16}} = \frac{1}{4} T$$

La tabla proporcionada para la T de Student cumple que

$$P(T \leq q_p) = p \iff P(T_n > t_{n,1-p}) = 1 - p$$

donde q_p es el percentil $p \in]0, 1[$, $t_{n,1-p}$ es el valor tabulado para la fila n y la columna $1 - p$, y $T_n \stackrel{not}{=} T \rightsquigarrow t(n)$. El percentil 99 de V será consecuentemente

$$q_{0,99}(V) = \frac{1}{4} t_{16;0,01} \approx \frac{1}{4} \cdot 2,5835 \approx 0,645875$$

- b) Calcular el intervalo de confianza para $\mu_1 - \mu_2$ de menor longitud esperada uniformemente a nivel de confianza $1 - \alpha$. ¿Cómo sería el intervalo si las varianzas fueran desconocidas pero iguales?

Ya sabemos que

$$\bar{X} - \bar{Y} \rightsquigarrow N\left(\mu_1 - \mu_2, \frac{4}{n_1} + \frac{5}{n_2}\right)$$

Y como nos piden el intervalo de confianza (se asume que bilateral), entonces por el método del pivote visto en teoría, usando como pivote

$$T(X_1, \dots, X_{n_1}, Y_1, \dots, Y_{n_2}; \mu_1 - \mu_2) = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{4}{n_1} + \frac{5}{n_2}}} \rightsquigarrow N(0, 1)$$

este debe verificar, por definición de intervalo de confianza, lo siguiente

$$P_{\mu_1, \mu_2} \left(\lambda_1 < \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{4}{n_1} + \frac{5}{n_2}}} < \lambda_2 \right) \geq 1 - \alpha \quad \forall (\mu_1, \mu_2) \in \mathbb{R}^2$$

para ciertos $\lambda_1, \lambda_2 \in \mathbb{R}$. Buscamos aislar la diferencia de medias poblacionales en el centro de la cadena de desigualdades, luego

$$\begin{aligned} \lambda_1 &< \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{4}{n_1} + \frac{5}{n_2}}} < \lambda_2 \\ \iff \lambda_1 \sqrt{\frac{4}{n_1} + \frac{5}{n_2}} &< (\bar{X} - \bar{Y}) - (\mu_1 - \mu_2) < \lambda_2 \sqrt{\frac{4}{n_1} + \frac{5}{n_2}} \\ \iff \bar{Y} - \bar{X} + \lambda_1 \sqrt{\frac{4}{n_1} + \frac{5}{n_2}} &< -(\mu_1 - \mu_2) < \bar{Y} - \bar{X} + \lambda_2 \sqrt{\frac{4}{n_1} + \frac{5}{n_2}} \\ \iff \bar{X} - \bar{Y} - \lambda_2 \sqrt{\frac{4}{n_1} + \frac{5}{n_2}} &< \mu_1 - \mu_2 < \bar{X} - \bar{Y} - \lambda_1 \sqrt{\frac{4}{n_1} + \frac{5}{n_2}} \end{aligned}$$

La solución para obtener el intervalo de confianza de menor longitud esperada uniformemente a nivel de confianza $1 - \alpha$ se alcanza con $\lambda_1 = -z_{\alpha/2}$, $\lambda_2 = z_{\alpha/2}$, donde $P[Z > z_\alpha] > \alpha$, y el intervalo de confianza bilateral es

$$\left(\bar{X} - \bar{Y} - z_{\alpha/2} \sqrt{\frac{4}{n_1} + \frac{5}{n_2}}, \bar{X} - \bar{Y} + z_{\alpha/2} \sqrt{\frac{4}{n_1} + \frac{5}{n_2}} \right)$$

Si se pidiera cualquiera de los dos unilaterales, entonces

- a) Si $\lambda_1 = -\infty$, se sustituye el extremo superior del intervalo bilateral anterior por $+\infty$, y $z_{\alpha/2}$ por z_α .
- b) Si $\lambda_2 = +\infty$, se sustituye el extremo inferior del intervalo bilateral anterior por $-\infty$, y $z_{\alpha/2}$ por z_α .

Si las varianzas fueran desconocidas pero iguales, entonces el pivote sería

$$T(X_1, \dots, X_{n_1}, Y_1, \dots, Y_{n_2}; \mu_1 - \mu_2) = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \rightsquigarrow t(n_1 + n_2 - 2)$$

donde

$$S_p = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

y como la T de Student tiene las mismas propiedades de simetría respecto al origen que $N(0, 1)$, entonces el intervalo de confianza bilateral a nivel de confianza $1 - \alpha$ sería

$$\left(\bar{X} - \bar{Y} - t_{n_1+n_2-2;\alpha/2} S_p \sqrt{\frac{4}{n_1} + \frac{5}{n_2}}, \bar{X} - \bar{Y} + t_{n_1+n_2-2;\alpha/2} S_p \sqrt{\frac{4}{n_1} + \frac{5}{n_2}} \right)$$

donde $P[T_{n_1+n_2-2} > t_{n_1+n_2-2;\alpha}] = \alpha$ y nuevamente si se pidiera cualquiera de los dos unilaterales, entonces

- a) Si $\lambda_1 = -\infty$, se sustituye el extremo superior del intervalo bilateral anterior por $+\infty$, y $t_{n_1+n_2-2;\alpha/2}$ por $t_{n_1+n_2-2;\alpha}$.
- b) Si $\lambda_2 = +\infty$, se sustituye el extremo inferior del intervalo bilateral anterior por $-\infty$, y $t_{n_1+n_2-2;\alpha/2}$ por $t_{n_1+n_2-2;\alpha}$.

Ejercicio 2.

- a) Sea X una v.a. con función de densidad

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

Calcular el UMVUE y determinar para qué valores de n existe. ¿Es eficiente?

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 $T \stackrel{not}{\equiv} T(X_1, \dots, X_n)$) insesgada en $g(\theta) = \theta$ (como no se especifica de quién es el UMVUE, se asume que del parámetro) estimadora y con momento de segundo orden finito. Entonces $h(T)$ será el UMVUE.

Podría pensarse en un principio en buscar el estadístico suficiente y completo si la familia fuera de tipo exponencial, sin embargo, como $\mathcal{X} =]-\theta^2, +\infty[$, vemos que el conjunto de valores de la variable depende de θ , luego ya no tendríamos que la familia es de tipo exponencial.

La suficiencia se obtendrá entonces por medio del Teorema de Factorización de Neyman-Fisher. Calculamos la función conjunta

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

Ahora, vemos que $x > \theta^2 \iff I_{\mathbb{R}^+}(x - \theta^2) = 1$, de donde se deduce que

$$\prod_{i=1}^n f_\theta(x_i) = \prod_{i=1}^n \frac{\theta}{2}x_i^{-3/2} \neq 0 \iff I_{\mathbb{R}^+}(x_{(1)} - \theta^2) = 1$$

luego

$$\begin{aligned} f_\theta^n(X_1, \dots, X_n) &= \prod_{i=1}^n \frac{\theta}{2}x_i^{-3/2} I_{\mathbb{R}^+}(x_{(1)} - \theta^2) = \left(\frac{\theta}{2}\right)^n \prod_{i=1}^n x_i^{-3/2} I_{\mathbb{R}^+}(x_{(1)} - \theta^2) = \\ &= \left(\frac{\theta}{2}\right)^n \prod_{i=1}^n x_i^{-3/2} I_{\mathbb{R}^+}(x_{(1)} - \theta^2) \end{aligned}$$

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

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

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 :

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

donde en $(*)$ se ha usado que, para que $f_\theta(x)$ sea función de densidad, debe ser no negativa e integrar a 1, lo cual implica necesariamente que $\theta > 0$.

La función de densidad del estadístico será entonces

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

Ahora, consideramos h una función medible verificando

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

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

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

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

$$-H(\theta^2) = 0 \iff h(\theta^2)\theta^{2-(n/2+1)} = 0 \stackrel{\theta > 0}{\iff} h(\theta^2) = 0$$

Por tanto, $\mathbb{R}^+ \subseteq \{t : h(t) = 0\}$, y consecuentemente

$$1 \geq P[h(T) = 0] \geq P[T \in \mathbb{R}^+] = 1 \implies P[h(T) = 0] = 1$$

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 θ y de segundo orden finito. Sea h (independiente de la anterior) función medible tal que

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

Supongamos que $h(t) = ct^\alpha$, con $\alpha, c \in \mathbb{R}$ a priori. Entonces

$$\int_{\theta^2}^{+\infty} ct^{\alpha-(n/2+1)}dt = c \int_{\theta^2}^{+\infty} t^{\alpha-(n/2+1)}dt$$

y esta última integral sabemos que converge si y solo si $\alpha - (n/2 + 1) < -1$, lo cual equivale a que

$$\alpha - (n/2 + 1) < -1 \iff \alpha - \frac{n}{2} < 0 \iff \alpha < \frac{n}{2}$$

En tal caso, obtenemos el resultado de la integral impropia

$$\begin{aligned}\int_{\theta^2}^{+\infty} t^{\alpha-(n/2+1)}dt &= \lim_{m \rightarrow +\infty} \int_{\theta^2}^m t^{\alpha-(n/2+1)}dt = \lim_{m \rightarrow +\infty} \frac{1}{\alpha - n/2} [t^{\alpha-n/2}]_{\theta^2}^m = \\ \lim_{m \rightarrow +\infty} \frac{1}{\alpha - n/2} (m^{\alpha-n/2} - \theta^{2(\alpha-n/2)}) &= \frac{1}{\alpha - n/2} \cdot \left(\lim_{m \rightarrow +\infty} \frac{1}{m^{n/2-\alpha}} - \theta^{2(\alpha-n/2)} \right) = \\ \frac{\theta^{2(\alpha-n/2)}}{n/2 - \alpha}\end{aligned}$$

Sustituyendo

$$c \int_{\theta^2}^{+\infty} t^{\alpha-(n/2+1)}dt = \frac{2}{n\theta^{n-1}} \iff c \cdot \frac{\theta^{2(\alpha-n/2)}}{n/2 - \alpha} = \frac{2}{n\theta^{n-1}} = \frac{2}{n} \theta^{1-n}$$

y por comparación entre exponentes de θ de las expresiones de los dos miembros de la igualdad

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

Debe considerarse a partir de ahora que $n \geq 2$ (para que la integral converja), y usando ese valor de α para determinar la constante e igualando constantes nuevamente obtenemos el valor de c

$$c \cdot \frac{1}{n/2 - \alpha} = \frac{2}{n} \iff c \cdot \frac{2}{n-1} = c \cdot \frac{1}{n/2 - 1/2} = \frac{2}{n} \iff c = \frac{n-1}{n}$$

Y hemos llegado a que

$$h(t) = ct^\alpha = \frac{n-1}{n}t^{1/2} = \frac{n-1}{n}\sqrt{t} \quad t > \theta^2$$

Por construcción $h(T)$ es insesgada en θ , y además

$$h(t) = \frac{n-1}{n}\sqrt{t} > 0$$

puesto que $n \geq 2$ y $t > \theta^2 > 0$ luego $\text{Im}(h) \subseteq \mathbb{R}^+ = \Theta$, de tal manera que $h(T)$ también es estimador. Queda comprobar que tiene momento de segundo orden finito.

Ello se cumplirá en caso de que

$$E_\theta[h(T)^2] < +\infty \iff \left(\frac{n-1}{n}\right)^2 E_\theta[T] < +\infty \iff E_\theta[T] < +\infty$$

Calculamos $E_\theta[T]$

$$E_\theta[T] \stackrel{\text{def}}{=} \int_{\theta^2}^{+\infty} tf_T(t)dt = \frac{n\theta^n}{2} \int_{\theta^2}^{+\infty} t^{-n/2} dt$$

Y como ya se ha visto antes, dicha integral converge si y solo si

$$-n/2 < -1 \iff -n < -2 \iff n > 2 \iff n \geq 3$$

En tal caso, la integral resulta ser

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

y juntando todo

$$\begin{aligned} E_\theta[(h(T))^2] &= \left(\frac{n-1}{n}\right)^2 E_\theta[T] = \left(\frac{n-1}{n}\right)^2 \frac{n\theta^n}{2} \frac{\theta^{2(1-n/2)}}{n/2 - 1} = \\ &= \frac{(n-1)^2}{n^2} \frac{\cancel{n}\theta^{\cancel{n}}}{\cancel{2}} \frac{\cancel{2}}{n-2} \frac{\theta^2}{\cancel{\theta^{\cancel{n}}}} = \frac{(n-1)^2}{n(n-2)} \theta^2 < +\infty \end{aligned}$$

imponiendo que $n \geq 3$ (para $n = 2$ existe el estimador insesgado, pero no tiene momento de segundo orden finito). Por lo tanto, por el Teorema de Lehmann-Scheffé, podemos concluir que $E[h(T)/T] = h(T)$ es el UMVUE para θ , y existe siempre que $n \geq 3$.

Respecto a la eficiencia, sabemos por un corolario visto en teoría que solo existen estimadores eficientes para familias de tipo exponencial (además de regulares en el sentido de Fréchet-Cramér-Rao), pero al principio se ha comprobado que la familia del problema no podía serlo, dado que el espacio muestral \mathcal{X} dependía del parámetro θ , por lo que, como no existen estimadores insesgados eficientes de θ , en particular el UMVUE obtenido no es eficiente.

b) Sea

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

en una familia regular según Fréchet–Cramér–Rao.

b1) Sabiendo que $I_{X_1, \dots, X_n}(\theta) = \frac{2n}{\theta^2}$, calcular $E[X]$ y $\text{Var}[X]$.

Por ser regular según FCR, sabemos que se verifica que

$$E_\theta \left[\frac{\partial \ln f_\theta(x)}{\partial \theta} \right] = 0$$

Así

$$\ln f_\theta(x) = \ln \theta + \ln(T(x)) - \theta x \implies \frac{\partial \ln f_\theta(x)}{\partial \theta} = \frac{1}{\theta} - x$$

La esperanza de X puede obtenerse ya

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

Como

$$\text{Var}_\theta \left[\frac{\partial \ln f_\theta(x)}{\partial \theta} \right] = I_X(\theta)$$

usando la aditividad de la función de información de Fisher respecto a la v.a. y respecto a la m.a.s. de la v.a., deducimos que

$$I_{X_1, \dots, X_n}(\theta) = nI_X(\theta) \iff \frac{2n}{\theta^2} = nI_X(\theta) \iff I_X(\theta) = \frac{2}{\theta^2}$$

de donde

$$\text{Var}_\theta \left[\frac{1}{\theta} - X \right] = \text{Var}_\theta(X) = I_X(\theta) = \frac{2}{\theta^2}$$

b2) Sabiendo que

$$\sum_{i=1}^n \frac{T(X_i)}{n}$$

es un estimador eficiente de $2/\theta$, calcular $T(x)$.

Sea $g(\theta) = 2/\theta$. Denotamos $\bar{T} \stackrel{\text{not}}{=} \sum_{i=1}^n T(X_i)/n$. Como \bar{T} es eficiente, en particular, por definición, alcanza la cota de FCR, es decir, existe $a(\theta) \neq 0$ tal que

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

Obtenemos la función conjunta

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

ahora aplicamos logaritmos y calculamos su parcial respecto de θ

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

$$\frac{\partial \ln f_\theta^n(X_1, \dots, X_n)}{\partial \theta} = \frac{n}{\theta} - \sum_{i=1}^n X_i$$

Por otro lado,

$$a(\theta)(\bar{T} - g(\theta)) = a(\theta) \left(\sum_{i=1}^n \frac{T(X_i)}{n} - \frac{2}{\theta} \right)$$

Igualando

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

El lado izquierdo es una función lineal, que depende de la muestra solo a través de $\sum_{i=1}^n X_i$. Por la igualdad de la cota de FCR, el lado derecho debe hacerlo de la misma forma, luego deben existir constantes, $\alpha, \beta \in \mathbb{R}$ tales que

$$\sum_{i=1}^n T(X_i) = \sum_{i=1}^n (\alpha \cdot X_i + \beta)$$

de donde se deduce que $T(x) = \alpha x + \beta$. Usando que \bar{T} es eficiente para $g(\theta)$, en particular también es insesgado en $g(\theta)$, es decir, verifica

$$\mathbb{E}_\theta[\bar{T}] = g(\theta) = \frac{2}{\theta}$$

y por linealidad de la esperanza, y por ser X_i pertenecientes a una m.a.s. de X v.a. $i = 1, \dots, n$, entonces

$$\mathbb{E}_\theta[\bar{T}] = \mathbb{E}_\theta \left[\frac{1}{n} \sum_{i=1}^n T(X_i) \right] = \frac{1}{n} \sum_{i=1}^n \mathbb{E}_\theta[T(X_i)] \stackrel{i.d.}{=} \frac{1}{n} \cdot n \mathbb{E}_\theta[T(X)] = \mathbb{E}_\theta[T(X)]$$

Podemos entonces hallar $\mathbb{E}_\theta[T]$

$$\frac{2}{\theta} = \mathbb{E}_\theta[\bar{T}] = \mathbb{E}_\theta[T(X)] = \alpha \mathbb{E}_\theta[X] + \beta = \alpha \frac{1}{\theta} + \beta = \frac{\alpha}{\theta} + \beta$$

Necesariamente, $\alpha = 2$ y $\beta = 0$, luego

$$T(x) = 2x$$

Ejercicio 3.

- a) Enunciar y demostrar el Teorema de Zehna definiendo previamente los siguientes conceptos: función de verosimilitud de un parámetro, función de verosimilitud de una función paramétrica y estimador máximo verosímil para funciones paramétricas.

Definición 0.1 (Función de Verosimilitud de un Parámetro). Sea X una v.a. con distribución en una familia paramétrica de distribuciones, $\{F_\theta, \theta \in \Theta\}$. Sea $f_\theta(x)$ la f.m.p. (caso discreto) ó la f.d.d. (caso continuo) de X . Se considera X_1, \dots, X_n una m.a.s. de X y sea $f_\theta^n(x_1, \dots, x_n)$ su f.m.p. ó f.d.d. (respectivamente) conjunta con $\theta \in \Theta$. Para cada x_1, \dots, x_n realización muestral, se define la *función de verosimilitud* asociada a dichos valores de la muestra como una función de θ de la siguiente forma:

$$\begin{aligned} L_{x_1, \dots, x_n} : \Theta &\longrightarrow \mathbb{R}^+ \cup \{0\} \\ \theta &\longmapsto L_{x_1, \dots, x_n}(\theta) = f_\theta^n(x_1, \dots, x_n) \end{aligned}$$

Definición 0.2 (Función de Verosimilitud de una Función Paramétrica). Sea $g : \Theta \rightarrow \Lambda$ una función paramétrica. En el contexto de la Definición 0.1, para cada x_1, \dots, x_n realización muestral, se define la *función de verosimilitud* de $\lambda = g(\theta)$ asociada a dicha realización como:

$$\begin{aligned} M_{x_1, \dots, x_n} : \Lambda &\longrightarrow \mathbb{R}^+ \cup \{0\} \\ \lambda &\longmapsto M_{x_1, \dots, x_n}(\lambda) = \sup_{\theta \in g^{-1}(\lambda)} L_{x_1, \dots, x_n}(\theta) \end{aligned}$$

Definición 0.3 (Estimador Máximo Verosímil para Funciones Paramétricas). Un estimador $\hat{\lambda}(X_1, \dots, X_n)$ de λ es *estimador de máxima verosimilitud* (EMV) de λ si:

$$\forall (x_1, \dots, x_n) \in \mathcal{X}^n, \quad M_{x_1, \dots, x_n}(\hat{\lambda}(x_1, \dots, x_n)) = \max_{\lambda \in \Lambda} M_{x_1, \dots, x_n}(\lambda)$$

Teorema 0.1 (de Invarianza de Zehna). *Sea X una v.a. con distribución en una familia paramétrica de distribuciones, $\{F_\theta, \theta \in \Theta\}$. Sea X_1, \dots, X_n una m.a.s. de X . Sea g una función medible. Si $\hat{\theta}(X_1, \dots, X_n)$ es EMV de θ , entonces $g(\hat{\theta}(X_1, \dots, X_n))$ es EMV de $g(\theta)$.*

Demostración. Sea $\lambda = g(\theta)$ y, fijada una realización muestral, $(x_1, \dots, x_n) \in \mathcal{X}^n$, notemos $\hat{\lambda} \equiv g(\hat{\theta}(x_1, \dots, x_n))$ (de esta manera, $\hat{\theta}(x_1, \dots, x_n) \in g^{-1}(\hat{\lambda})$).

Obtenemos la función de verosimilitud de la función paramétrica usando la Definición 0.2:

$$M_{x_1, \dots, x_n}(\hat{\lambda}) \stackrel{\text{def}}{=} \sup_{\theta \in g^{-1}(\hat{\lambda})} L_{x_1, \dots, x_n}(\theta) \stackrel{\hat{\theta} \in g^{-1}(\hat{\lambda})}{=} L_{x_1, \dots, x_n}(\hat{\theta}(x_1, \dots, x_n)) \quad (*)$$

donde en la última igualdad se ha usado que $\hat{\theta}(X_1, \dots, X_n)$ es EMV de θ , luego maximiza la función de verosimilitud de la Definición 0.1. Ahora, vemos que

$$\forall \lambda \in \Lambda, \quad M_{x_1, \dots, x_n}(\lambda) = \sup_{\theta \in g^{-1}(\lambda)} L_{x_1, \dots, x_n}(\theta) \leq \sup_{\theta \in \Theta} L_{x_1, \dots, x_n}(\theta) = L_{x_1, \dots, x_n}(\hat{\theta}(x_1, \dots, x_n)) \quad (**)$$

la última igualdad nuevamente por ser $\hat{\theta}(X_1, \dots, X_n)$ es EMV de θ . Deducimos entonces que

$$\forall \lambda \in \Lambda, \quad M_{x_1, \dots, x_n}(\lambda) \stackrel{(**)}{\leqslant} L_{x_1, \dots, x_n}(\hat{\theta}(x_1, \dots, x_n)) \stackrel{(*)}{=} M_{x_1, \dots, x_n}(\hat{\lambda})$$

Es decir, $\hat{\lambda} = g(\hat{\theta}(x_1, \dots, x_n))$ maximiza M_{x_1, \dots, x_n} , para cada $(x_1, \dots, x_n) \in \mathcal{X}^n$. Por la Definición 0.3, $g(\hat{\theta}(X_1, \dots, X_n))$ es el EMV de $\lambda = g(\theta)$. \square

b) Calcular la función de verosimilitud de

$$\lambda = (\theta - 1)^2$$

asociada a una realización muestral cuyo máximo valor es 3 si

$$f_\theta(x) = e^{x-\theta}, \quad x \leqslant \theta, \theta > 0.$$

Vemos que $\mathcal{X} =]-\infty, \theta]$. Sea $(x_1, \dots, x_n) \in \mathcal{X}^n$ tal que $x_{(n)} = 3$. Calculamos la función de densidad conjunta, asumiendo a partir de ahora que $\theta > 0$ (en otro caso, $f_\theta(x) = 0$).

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

Ahora, vemos que $x \leqslant \theta \iff I_{\mathbb{R}_0^-}(x - \theta) = 1$ de donde se deduce que

$$\prod_{i=1}^n f_\theta(x_i) = \prod_{i=1}^n e^{x_i - \theta} \neq 0 \iff I_{\mathbb{R}_0^-}(x_i - \theta) = 1 \quad \forall i = 1, \dots, n$$

y a su vez

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

luego

$$f_\theta^n(x_1, \dots, x_n) = \prod_{i=1}^n e^{x_i - \theta} I_{\mathbb{R}_0^-}(x_{(n)} - \theta) = e^{\left(\sum_{i=1}^n x_i - n\theta\right)} I_{\mathbb{R}_0^-}(x_{(n)} - \theta)$$

Por la Definición 0.1

$$L_{x_1, \dots, x_n}(\theta) = f_\theta^n(x_1, \dots, x_n) = e^{\left(\sum_{i=1}^n x_i - n\theta\right)} I_{\mathbb{R}_0^-}(x_{(n)} - \theta) \quad \forall \theta \in \Theta = \mathbb{R}^+$$

El enunciado nos dice que $\lambda = g(\theta) = (\theta - 1)^2$. Por la Definición 0.2

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

Resolvemos $(\theta - 1)^2 = \lambda$ para expresar θ explícitamente en función de λ . Vemos que $\lambda \geqslant 0$, y $(\theta - 1)^2 = \lambda \iff \theta - 1 = \pm\sqrt{\lambda} \iff \theta = 1 \pm \sqrt{\lambda}$.

A priori habría dos candidatos para cada λ . Sin embargo, por restricciones del problema, $3 = x_{(n)} \leq \theta$, lo que implica que

$$\theta = 1 + \sqrt{\lambda} \geq 3 \iff \sqrt{\lambda} \geq 2 \iff \lambda \geq 4$$

$$\theta = 1 - \sqrt{\lambda} \geq 3 \iff -\sqrt{\lambda} \geq 2 \iff \sqrt{\lambda} \leq -2$$

La última opción no puede darse por ser $\lambda \geq 0$, por tanto, nos quedamos con la primera. Así, si $\lambda \geq 4$ (en otro caso, $M_{x_1, \dots, x_n}(\lambda) = 0$)

$$M_{x_1, \dots, x_n}(\lambda) = L_{x_1, \dots, x_n}(1 + \sqrt{\lambda}) = e^{\left(\sum_{i=1}^n x_i - n(1 + \sqrt{\lambda})\right)}$$

Aunque no se pide, como $e^{\sum_{i=1}^n x_i}$ es fijo, y $1 + \sqrt{\lambda}$ es creciente como función de λ y $e^{-n(1+\sqrt{\lambda})}$ es decreciente como función de λ , y $M_{x_1, \dots, x_n}(\lambda) \neq 0 \iff \lambda \in [4, +\infty[$, entonces el máximo se alcanza en el extremo inferior del intervalo, es decir $\hat{\lambda}(x_1, \dots, x_n) = 4$. Esto puede comprobarse también con el Teorema 0.1 pues $L_{x_1, \dots, x_n}(\theta)$ es decreciente en $[x_{(n)}, +\infty[= [3, +\infty[$. Por el mismo razonamiento, $\hat{\theta}(x_1, \dots, x_n) = 3$, luego $\hat{\lambda} = (\hat{\theta} - 1)^2 = (3 - 1)^2 = 4$.

Ejercicio 4. Sea (X_1, \dots, X_n) una m.a.s. de X v.a. con función de densidad

$$f_\theta(x) = \frac{1}{\sqrt{8\theta}\sqrt{x-1}}, \quad 1 < x < 2\theta + 1. \quad (3)$$

Deducir el test más potente de tamaño arbitrario para contrastar

$$H_0 : \theta = \theta_0 \quad \text{frente a} \quad H_1 : \theta = \theta_1$$

siendo $\theta_1 < \theta_0$. Calcular la potencia.

Tenemos un contraste de hipótesis simple frente a hipótesis simple, por lo que sabemos por el Lema de Neyman-Pearson que el Test de Neyman-Pearson será el más potente, para cualquier tamaño $\alpha \in]0, 1]$. Si el tamaño fuera $\alpha = 0$, sabemos entonces que el test más potente es

$$\varphi_0(X_1, \dots, X_n) = \begin{cases} 1 & \text{si } f_0^n(X_1, \dots, X_n) = 0 \\ 0 & \text{si } f_0^n(X_1, \dots, X_n) > 0 \end{cases}$$

denotando por $f_0^n \stackrel{\text{not}}{\equiv} f_{\theta_0}^n$ a la función conjunta bajo la hipótesis nula, y lo mismo con f_1^n y la función conjunta bajo la hipótesis alternativa. Como el tamaño es 0, entonces $E_{\theta_0}[\varphi] = 0 \implies \varphi = 0$ bajo H_0 (no rechaza nunca). Como $\theta_1 < \theta_0$, no existe ninguna región en la que $f_{\theta_0} = 0$ y $f_{\theta_1} > 0$, luego el test más potente es $\varphi \equiv 0$, y entonces su potencia en θ_1 es $\beta_\varphi(\theta_1) = E_{\theta_1}[\varphi] = E_{\theta_1}[0] = 0$.

Para dar el test de hipótesis, obtenemos las funciones de densidad conjuntas bajo la hipótesis nula y bajo la hipótesis alternativa.

A partir de ahora asumimos que $1 < x_i \quad \forall i = 1, \dots, n$. De otra manera, $f_\theta^n(x_1, \dots, x_n) = 0$. La función de densidad conjunta es

$$f_\theta^n(x_1, \dots, x_n) \stackrel{\text{indep.}}{=} \prod_{i=1}^n \frac{1}{\sqrt{8\theta}\sqrt{x_i-1}}$$

Y vemos que

$$x < 2\theta + 1 \iff I_{\mathbb{R}^-}(x - (2\theta + 1)) = 1$$

luego

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

Por tanto

$$f_\theta^n(x_1, \dots, x_n) = \prod_{i=1}^n \frac{1}{\sqrt{8\theta}\sqrt{x_i-1}} I_{\mathbb{R}^-}(x_{(n)} - (2\theta + 1)) =$$

$$\frac{1}{(\sqrt{8\theta})^n} \prod_{i=1}^n \frac{1}{\sqrt{x_i-1}} I_{\mathbb{R}^-}(x_{(n)} - (2\theta + 1)) = (8\theta)^{-n/2} \prod_{i=1}^n (x_i - 1)^{-1/2} I_{\mathbb{R}^-}(x_{(n)} - (2\theta + 1))$$

Se tiene entonces que

$$f_0^n(X_1, \dots, X_n) = (8\theta_0)^{-n/2} \prod_{i=1}^n (X_i - 1)^{-1/2} I_{\mathbb{R}^-}(X_{(n)} - (2\theta_0 + 1))$$

$$f_1^n(X_1, \dots, X_n) = (8\theta_1)^{-n/2} \prod_{i=1}^n (X_i - 1)^{-1/2} I_{\mathbb{R}^-}(X_{(n)} - (2\theta_1 + 1))$$

Consideramos el cociente

$$\frac{f_1^n(X_1, \dots, X_n)}{f_0^n(X_1, \dots, X_n)} = \frac{(8\theta_1)^{-n/2} \prod_{i=1}^n (X_i - 1)^{-1/2} I_{\mathbb{R}^-}(X_{(n)} - (2\theta_1 + 1))}{(8\theta_0)^{-n/2} \prod_{i=1}^n (X_i - 1)^{-1/2} I_{\mathbb{R}^-}(X_{(n)} - (2\theta_0 + 1))} = \\ \left(\frac{\theta_0}{\theta_1}\right)^{n/2} \frac{I_{\mathbb{R}^-}(X_{(n)} - (2\theta_1 + 1))}{I_{\mathbb{R}^-}(X_{(n)} - (2\theta_0 + 1))}$$

y como $\theta_1 < \theta_0$ por el enunciado, entonces $2\theta_1 + 1 < 2\theta_0 + 1$, y distinguimos como sigue:

1. Si $X_{(n)} \leq 2\theta_1 + 1$, ambas funciones indicadoras valen 1

$$\frac{f_1^n(X_1, \dots, X_n)}{f_0^n(X_1, \dots, X_n)} = \left(\frac{\theta_0}{\theta_1}\right)^{n/2}$$

2. Si $2\theta_1 + 1 < X_{(n)} \leq 2\theta_0 + 1$, la función indicadora del numerador valdría 0 y la del denominador 1

$$\frac{f_1^n(X_1, \dots, X_n)}{f_0^n(X_1, \dots, X_n)} = 0$$

3. Si $X_{(n)} > 2\theta_0 + 1$, entonces $f_0^n(X_1, \dots, X_n) = 0$, luego el suceso tendrá probabilidad 0 bajo la hipótesis nula, y no influiría en el tamaño del test.

Tenemos hasta ahora que

$$\frac{f_1^n(X_1, \dots, X_n)}{f_0^n(X_1, \dots, X_n)} = \begin{cases} k & \text{si } X_{(n)} \leq 2\theta_1 + 1 \\ 0 & \text{si } 2\theta_1 + 1 < X_{(n)} \leq 2\theta_0 + 1 \end{cases} \quad k = \left(\frac{\theta_0}{\theta_1}\right)^{n/2}$$

El cociente solo toma dos valores, y, en particular, nunca ocurre que

$$f_1^n(X_1, \dots, X_n) > k f_0^n(X_1, \dots, X_n)$$

de tal manera que el test con tamaño α más potente nunca rechazaría (a priori) con probabilidad 1 la hipótesis nula. La región en la que se puede producir el rechazo de H_0 es

$$\mathcal{C} = \{(x_1, \dots, x_n) \in \mathcal{X}^n : x_{(n)} \leq 2\theta_1 + 1\}$$

Dicho rechazo, por la forma del test, solo puede darse cuando $f_1^n(X_1, \dots, X_n) = k f_0^n(X_1, \dots, X_n)$, y el Test de Neyman-Pearson de tamaño α es

$$\varphi(X_1, \dots, X_n) = \begin{cases} \gamma(X_1, \dots, X_n) & \text{si } X_{(n)} \leq 2\theta_1 + 1 \\ 0 & \text{si } X_{(n)} > 2\theta_1 + 1 \end{cases}$$

donde $\gamma(X_1, \dots, X_n) = \gamma \in [0, 1]$ sabemos por teoría que es constante. Para determinar γ , se impone el tamaño α , es decir

$$\alpha = P_{\theta_0}[(X_1, \dots, X_n) \in \mathcal{C}] = \gamma P_{\theta_0}(X_{(n)} \leq 2\theta_1 + 1)$$

Hallamos $F_{\theta_0}(t)$

$$F_{\theta_0}(t) = \int_1^t \frac{1}{\sqrt{8\theta_0}\sqrt{x-1}} dx = \frac{1}{\sqrt{8\theta_0}} \int_1^t (x-1)^{-1/2} dx = \frac{1}{\sqrt{8\theta_0}} \left[\frac{(x-1)^{1/2}}{1/2} \right]_1^t = \frac{2\sqrt{t-1}}{\sqrt{8\theta_0}} = \frac{2\sqrt{t-1}}{2\sqrt{2\theta_0}} = \sqrt{\frac{t-1}{2\theta_0}}$$

Sabemos que la función de distribución del máximo verifica

$$P_{\theta_0}(X_{(n)} \leq 2\theta_1 + 1) = (F_{\theta_0}(2\theta_1 + 1))^n = \left(\sqrt{\frac{2\theta_1}{2\theta_0}} \right)^n = \left(\frac{\theta_1}{\theta_0} \right)^{n/2}$$

y juntando todo

$$\alpha = \gamma P_{\theta_0}(X_{(n)} \leq 2\theta_1 + 1) = \gamma \left(\frac{\theta_1}{\theta_0} \right)^{n/2}$$

de donde

$$\gamma = \frac{\alpha}{p_0}, \quad p_0 = \left(\frac{\theta_1}{\theta_0} \right)^{n/2}$$

Como $\gamma \in [0, 1] \iff 0 \leq \alpha/p_0 \leq 1$, pero $\alpha \in [0, 1]$ y $p_0 > 0$, se tiene que

$$\gamma \in [0, 1] \iff \alpha/p_0 \leq 1 \iff \alpha \leq p_0$$

Así, si $\alpha \leq p_0$, la potencia en θ_1 es

$$\beta_\varphi(\theta_1) = \gamma \cdot P_{\theta_1}[(X_1, \dots, X_n) \in \mathcal{C}]$$

Como bajo la hipótesis alternativa H_1 , siempre se cumple que $X_{(n)} \leq 2\theta_1 + 1$, entonces $\beta_\varphi(\theta_1) = \gamma$.

Si, por el contrario, $\alpha > p_0$, hay que rechazar también fuera de \mathcal{C} para conseguir el tamaño α , por lo tanto, el Test de Neyman-Pearson sería

$$\varphi(X_1, \dots, X_n) = \begin{cases} 1 & \text{si } X_{(n)} \leq 2\theta_1 + 1 \\ \gamma(X_1, \dots, X_n) \equiv \gamma & \text{si } 2\theta_1 + 1 < X_{(n)} \leq 2\theta_0 + 1 \\ 0 & \text{si } X_{(n)} > 2\theta_0 + 1 \end{cases}$$

Imponemos el tamaño

$$\alpha = p_0 + \gamma(1 - p_0) \iff \gamma = \frac{\alpha - p_0}{1 - p_0}$$

Y en este caso, bajo H_1 , $0 \leq P_{\theta_1}[2\theta_1 + 1 < X_{(n)} \leq 2\theta_0 + 1] \leq P_{\theta_1}[2\theta_1 + 1 < X_{(n)}] = 0$, donde la primera desigualdad se ha usado que

$$\{(x_1, \dots, x_n) \in \mathcal{X}^n : 2\theta_1 + 1 < X_{(n)} \leq 2\theta_0 + 1\} \subseteq \{(x_1, \dots, x_n) \in \mathcal{X}^n : 2\theta_1 + 1 < X_{(n)}\}$$

y luego que $P_{\theta_1}[X_{(n)} \leq 2\theta_1 + 1] = 1 \iff P_{\theta_1}[2\theta_1 + 1 < X_{(n)}] = 0$. Finalmente, la potencia en este caso es

$$\beta_\varphi(\theta_1) = E_{\theta_1}[\varphi(X_1, \dots, X_n)] = 1 \cdot P_{\theta_1}[X_{(n)} \leq 2\theta_1 + 1] +$$

$$\gamma \cdot P_{\theta_1}[2\theta_1 + 1 < X_{(n)} \leq 2\theta_0 + 1] + 0 \cdot P_{\theta_1}[X_{(n)} > 2\theta_0 + 1] = 1 \cdot 1 + \gamma \cdot 0 + 0 \cdot 0 = 1$$

Ejercicio 5.

a) Test de Kolmogorov–Smirnov.

i) Plantear el problema de contraste.

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

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

ii) Dar el valor del estadístico.

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

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

iii) Enunciar el teorema que justifica su uso.

El test se basa en el teorema de Glivenko-Cantelli:

Teorema 0.2 (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 .*

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

iv) Ventajas frente al test χ^2 .

Se asume que nos referimos al test χ^2 de Pearson.

- 1) El test de Kolmogorov–Smirnov no necesita hacer particiones de los datos, mientras que el test χ^2 de Pearson sí.
- 2) Para v.a. continuas, es más apropiado usar el test de Kolmogorov–Smirnov que el test χ^2 de Pearson.

- 3) 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.
- b) Se cuentan el número de tutorías a lo largo de un curso por 50 profesores. Se quiere contrastar si el número de tutorías por profesor se puede describir mediante una distribución de Poisson.

Número de tutorías	0	1	2	3	4	5
Número de profesores	2	5	10	14	12	7

En este caso, como la variable aleatoria es discreta, y tenemos frecuencias, es más apropiado usar el test χ^2 de Pearson. Sea

$$X \equiv \text{“Número de tutorías por un profesor”}$$

El contraste a resolver es

$$\begin{cases} H_0 : X \rightsquigarrow \mathcal{P}(\lambda) \\ H_1 : X \not\rightsquigarrow \mathcal{P}(\lambda) \end{cases} \quad \lambda > 0$$

Vemos que la hipótesis nula es compuesta, luego primeramente debemos estimar el valor del parámetro λ . Por el método de los momentos, un estimador del parámetro de la distribución de Poisson λ es la media muestral. Así

$$\hat{\lambda} = \bar{X} = \frac{0 \cdot 2 + 1 \cdot 5 + 2 \cdot 10 + 3 \cdot 14 + 4 \cdot 12 + 5 \cdot 7}{50} = \frac{150}{50} = 3$$

El constrauste adaptado sería

$$\begin{cases} H_0 : X \rightsquigarrow \mathcal{P}(3) \\ H_1 : X \not\rightsquigarrow \mathcal{P}(3) \end{cases}$$

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

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

las probabilidades teóricas bajo H_0 con el parámetro λ estimado por $\hat{\lambda}$.

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 $q = 1$ el número de parámetros estimados.

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

$$E_i^* = n\hat{p}_i \geq 5 \quad \forall i = 1, \dots, k, \quad n = 50, \quad \hat{p}_i = P[X = i] = e^{-3} \frac{3^i}{i!}$$

Primero vamos tanteando.

$$E_0^* = 50 \cdot \hat{p}_0 = 50 \cdot e^{-3} \frac{3^0}{0!} \approx 2,49$$

$$E_1^* = 50 \cdot e^{-3} \frac{3^1}{1!} \approx 7,47$$

$$E_2^* = 50 \cdot e^{-3} \frac{3^2}{2!} \approx 11,2$$

$$E_3^* = 50 \cdot e^{-3} \frac{3^3}{3!} \approx 11,2$$

$$E_4^* = 50 \cdot e^{-3} \frac{3^4}{4!} \approx 8,4$$

$$E_5^* = 50 \cdot e^{-3} \frac{3^5}{5!} \approx 5,05$$

$$E_6^* = 50 \cdot e^{-3} \frac{3^6}{6!} \approx 2,52$$

Vemos que $E_i^* \geq 5 \quad \forall i = 1, \dots, 5$. Sin embargo, $E_0^*, E_6^* < 5$, y $x_{(k)} = 5$, por lo que una partición sería la siguiente

$$A_1 = \{0, 1\}, \quad A_i = \{i\}, \quad A_5 = \{\geq 5\}, \quad i = 2, 3, 4$$

En este caso, $k = 5$ (número de clases tras agrupar), luego $\chi^2(N_1, \dots, N_k) \rightsquigarrow \chi^2(3)$. Las frecuencias observadas, denotadas por $O_i, i = 1, \dots, k$, son

$$O_1 = N_0 + N_1 = 2 + 5 = 7, \quad O_2 = 10, \quad O_3 = 14, \quad O_4 = 12, \quad O_5 = 7$$

Las frecuencias esperadas son

$$E_1 = E_0^* + E_1^* \approx 2,49 + 7,47 = 9,96, \quad E_2 = E_2^* \approx 11,2 \quad E_3 = E_3^* \approx 11,2 \quad E_4 = E_4^* \approx 8,4$$

$$E_5 = 50 - \sum_{i=1}^4 E_i \approx 50 - (9,96 + 11,2 + 11,2 + 8,4) = 9,24$$

Obtenemos

$$\begin{aligned} \chi^2_{exp} &= \sum_{i=1}^5 \frac{(O_i - E_i)^2}{E_i} = \frac{(7 - 9,96)^2}{9,96} + \frac{(10 - 11,2)^2}{11,2} + \frac{(14 - 11,2)^2}{11,2} + \frac{(12 - 8,4)^2}{8,4} + \\ &\quad \frac{(7 - 9,24)^2}{9,24} \approx 3,79 \end{aligned}$$

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

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

y χ^2_{exp} el valor del estadístico obtenido con la muestra observada. Usando que $k - q - 1 = 3$, obtenemos

$$p\text{-valor} \approx P[\chi^2(3) \geq 3,79] \approx 0,3$$

Como el p-valor es grande (respecto a los niveles habituales de significación), no hay evidencia para rechazar H_0 , luego puede suponerse que $X \rightsquigarrow \mathcal{P}(3)$.