EXAMEN ORDINARIO 2021 RESUELTO
EJERCICIO 1.

1) 1.25 puntos *Sean $(X_1,\ldots,X_{n_1})$, $(Y_1,\ldots,Y_{n_2})$ muestras aleatorias simples independientes de dos poblaciones, $\mathcal{N}(\mu_1,\sigma^2)$ y $\mathcal{N}(\mu_2,\sigma^2)$, respectivamente, con medias muestrales $\overline{X}$ e $\overline{Y}$.*

1a) **0.75 puntos** *Calcular el percentil r de la siguiente variable:*

$$Z=\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{\sqrt{\sum_{i=1}^{n_1}(X_i-\overline{X})^2+\sum_{i=1}^{n_2}(Y_i-\overline{Y})^2}}.$$

$$T=\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}=\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{\sum_{i=1}^{n_1}(X_i-\overline{X})^2+\sum_{i=1}^{n_2}(Y_i-\overline{Y})^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\to t(n_1+n_2-2)\ \ 0.375$$

$$T=\frac{\sqrt{n_1+n_2-2}}{\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}Z\ \ 0.125$$

$$r/100=P(Z<k)=P\left(T<\frac{\sqrt{n_1+n_2-2}}{\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}k\right)\Rightarrow\frac{\sqrt{n_1+n_2-2}}{\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}k=t_{n_1+n_2-2;1-r/100}\Rightarrow k=\frac{t_{n_1+n_2-2;1-r/100}}{\sqrt{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}.$$

0.25

1.1 $\mu_1-\mu_2=-1;\ n_1=10,\ n_2=8;\ r=99\to t_{16;0.01}=2.5835\Rightarrow k=\frac{2.5835}{4}\sqrt{\frac{1}{10}+\frac{1}{8}}=0.30637.$

1.2 $\mu_1-\mu_2=-2;\ n_1=6,\ n_2=5;\ r=90\to t_{9;0.1}=1.383\Rightarrow k=\frac{1.383}{3}\sqrt{\frac{1}{6}+\frac{1}{5}}=0.27915.$

1.3 $\mu_1-\mu_2=-2;\ n_1=7,\ n_2=11;\ r=90\to t_{16;0.1}=1.3368\Rightarrow k=\frac{1.3368}{4}\sqrt{\frac{1}{7}+\frac{1}{11}}=0.16158.$

1.4 $\mu_2-\mu_1=-1;\ n_1=7,\ n_2=4;\ r=99\to t_{9;0.01}=2.8214\Rightarrow k=\frac{2.8214}{3}\sqrt{\frac{1}{7}+\frac{1}{4}}=0.58947.$

1.5 $\mu_2-\mu_1=-1;\ n_1=8,\ n_2=10;\ r=95\to t_{16;0.05}=1.7459\Rightarrow k=\frac{1.7459}{4}\sqrt{\frac{1}{8}+\frac{1}{10}}=0.20704.$

1b) **0.5 puntos** *Calcular el estimador máximo verosímil de σ^2 basado en $(X_1,\ldots,X_{n_1},Y_1,\ldots,Y_{n_2})$.*

$$L_{x_i,y_j}(\sigma^2)=\frac{1}{\sigma^{n_1}(2\pi)^{n_1/2}}e^{-\frac{\sum_{i=1}^{n_1}(x_i-\mu_1)^2}{2\sigma^2}}\frac{1}{\sigma^{n_2}(2\pi)^{n_2/2}}e^{-\frac{\sum_{j=1}^{n_2}(y_j-\mu_2)^2}{2\sigma^2}}=\frac{1}{\sigma^{n_1+n_2}(2\pi)^{n_1+n_2/2}}e^{-\frac{\sum_{i=1}^{n_1}(x_i-\mu_1)^2+\sum_{j=1}^{n_2}(y_j-\mu_2)^2}{2\sigma^2}}$$

$$\ln L_{x_i,y_j}(\sigma^2)=-\frac{n_1+n_2}{2}\ln(2\pi)-\frac{n_1+n_2}{2}\ln\sigma^2-\frac{\sum_{i=1}^{n_1}(x_i-\mu_1)^2+\sum_{j=1}^{n_2}(y_j-\mu_2)^2}{2\sigma^2}$$

$$\frac{d\ln L_{x_i,y_j}(\sigma^2)}{d\sigma^2}=-\frac{n_1+n_2}{2\sigma^2}+\frac{\sum_{i=1}^{n_1}(x_i-\mu_1)^2+\sum_{j=1}^{n_2}(y_j-\mu_2)^2}{2\sigma^4}=0\Rightarrow\hat{\sigma}^2=\frac{\sum_{i=1}^{n_1}(x_i-\mu_1)^2+\sum_{j=1}^{n_2}(y_j-\mu_2)^2}{n_1+n_2}.$$

EJERCICIO 2.

2.1) **2.5 puntos** *Sea $(X_1,\dots,X_n)$ una muestra aleatoria simple de una variable con función de densidad:*

$$f_\theta(x)=\frac{2x}{(\theta+1)^2},\quad 0<x<\theta+1.\quad (a=2;\ k=-1)$$

a) Encontrar, si existe, el UMVUE para θ.

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

$$\Theta=(-1,+\infty),\quad \chi^n=\mathbb{R}^{+n}$$

- *Estadístico suficiente:* $T=\max X_i$. 0.25

$$f_\theta^n(x_1,\dots,x_n)=\frac{2^n x_1\cdots x_n}{(\theta+1)^{2n}}I_{(0,+\infty)}(\theta+1-\max x_i),\quad (x_1,\dots,x_n)\in\mathbb{R}^{+n}.$$

- *Distribución de* $T=\max X_i$: 0.25

$$F_\theta(x)=\int_0^x\frac{2t}{(\theta+1)^2}dt=\frac{t^2}{(\theta+1)^2}\Big|_0^x=\boxed{\frac{x^2}{(\theta+1)^2},\quad 0<x<\theta+1.}\Rightarrow F_\theta^T(t)=[F_\theta(t)]^n=\frac{t^{2n}}{(\theta+1)^{2n}},\quad 0<t<\theta+1.$$

$$f_\theta^T(t)=\frac{2n\,t^{2n-1}}{(\theta+1)^{2n}},\quad \boxed{0<t<\theta+1.}$$

- *Completitud de* $T=\max X_i$: 0.25

$$E_\theta[g(T)]=\int_0^{\theta+1}g(t)\frac{2nt^{2n-1}}{(\theta+1)^{2n}}dt=0,\ \forall\theta>-1\Rightarrow\boxed{\int_0^{\theta+1}g(t)t^{2n-1}dt}=0,\ \forall\theta>-1\Rightarrow g(\theta+1)(\theta+1)^{2n-1}=0,\ \forall\theta>-1.$$

$$\Rightarrow g(t)t^{2n-1}=0,\ \forall t>0\Rightarrow g(t)=0,\ \forall t>0\Rightarrow P_\theta(g(T)=0)\geq P_\theta(T>0)=1,\ \forall\theta>-1$$

- *Cálculo del UMVUE:*

$$E_\theta[h(T)]=\int_0^{\theta+1}h(t)\frac{2nt^{2n-1}}{(\theta+1)^{2n}}dt=\theta,\ \forall\theta>-1\Rightarrow\int_0^{\theta+1}h(t)2nt^{2n-1}dt=\theta(\theta+1)^{2n},\ \forall\theta>-1.$$

Puesto que $\lim_{\theta\to-1}\theta(\theta+1)^{2n}=0,\ \forall n\in\mathbb{N}$, tiene sentido trabajar con n arbitrario:

- $h(\theta+1)2n(\theta+1)^{2n-1}=(\theta(\theta+1)^{2n})'=(\theta+1)^{2n}+2n\theta(\theta+1)^{2n-1}\Rightarrow h(\theta+1)=\frac{\theta+1}{2n}+\theta,\ \forall\theta>-1\ (\theta+1>0).$

  $\Rightarrow h(T)=\frac{T}{2n}+T-1,\ (T>0).$ 0.75
- $T>0\Rightarrow h(T)>-1$ es estimador de $\theta>-1$. 0.125
- T está acotado para todo $\theta\Rightarrow h(T)$ está acotado para todo $\theta\Rightarrow$ es de segundo orden. 0.125

$$\text{UMVUE para }\theta\ \to\ \frac{T}{2n}+T-1.$$

- *Función de verosimilitud y estimador máximo verosímil:*

$$L_{x_1,\dots,x_n}(\theta)=\begin{cases}0, & \theta<\max x_i-1\\ \frac{2^n x_1\cdots x_n}{(\theta+1)^{2n}}, & \theta>\max x_i-1\end{cases}\quad 0.25$$

$L_{x_1,\dots,x_n}(\theta)$ decrece con θ para $\theta>\max x_i-1\Rightarrow\widehat{\theta}(X_1,\dots,X_n)=\max X_i-1$. 0.25

$\widehat{\theta}(X_1,\dots,X_n)$ no es insesgado ya que es función del suficiente y completo y, de ser insesgado, coincidiría con el UMVUE. 0.25

EJERCICIO 3

3.1) 2.25 puntos $\boxed{k=1, \quad \sigma^2 = 2}$

3a) **0.375** *Sea X una variable continua con distribución en una familia regular en el sentido de Fréchet-Cramér-Rao, $X \to \{P_\theta;\ \theta \in \Theta\}$. Probar que si $\int_{\mathcal{X}} \frac{\partial^2 f_\theta(x)}{\partial \theta^2} dx = 0$, la función de información verifica $I_X(\theta) = -E_\theta\left[\frac{\partial^2 \ln f_\theta(X)}{\partial \theta^2}\right]$.*

- $$\frac{\partial \ln f_\theta(X)}{\partial \theta} = \frac{\frac{\partial f_\theta(X)}{\partial \theta}}{f_\theta(X)} \Rightarrow \frac{\partial^2 \ln f_\theta(X)}{\partial \theta^2} = \frac{\frac{\partial^2 f_\theta(X)}{\partial \theta^2} f_\theta(X) - \left(\frac{\partial f_\theta(X)}{\partial \theta}\right)^2}{(f_\theta(X))^2} = \frac{\frac{\partial^2 f_\theta(X)}{\partial \theta^2}}{f_\theta(X)} - \frac{\left(\frac{\partial f_\theta(X)}{\partial \theta}\right)^2}{(f_\theta(X))^2} =$$
$$= \frac{\frac{\partial^2 f_\theta(X)}{\partial \theta^2}}{f_\theta(X)} - \left(\frac{\partial \ln f_\theta(X)}{\partial \theta}\right)^2.$$
- $$E_\theta\left[\frac{\partial^2 \ln f_\theta(X)}{\partial \theta^2}\right] = \int_{\mathcal{X}} \frac{\partial^2 f_\theta(x)}{\partial \theta^2} dx - I_X(\theta) = -I_X(\theta).$$

3b) **1.875** *Sea $(X_1, \ldots, X_n)$ una muestra aleatoria simple de una variable con función de densidad:*

$$f_\theta(x) = \frac{1}{2x\sqrt{\pi}} \exp\left\{-\frac{(\ln x - \theta)^2}{4}\right\}, \quad x > 0.$$

Sabiendo que $E_\theta[\ln X] = \theta$, $Var_\theta[\ln X] = 2$ y que la familia de distribuciones es regular, se pide:

b1) **1.125** *Calcular la función de información asociada a la muestra y encontrar la clase de funciones paramétricas que admiten estimador eficiente y los correspondientes estimadores.*

$$f_\theta(x) = \exp\left\{-\ln(2x\sqrt{\pi}) - \frac{(\ln x - \theta)^2}{4}\right\}, \quad x > 0 \Rightarrow \frac{\partial f_\theta(x)}{\partial \theta} = \frac{1}{2}[\ln x - \theta] f_\theta(x) \ \Rightarrow \ \frac{\partial \ln f_\theta(x)}{\partial \theta} = \frac{1}{2}[\ln x - \theta]$$

- $I_X(\theta) = \frac{1}{4} E_\theta\left[(\ln X - \theta)^2\right] = \frac{Var_\theta[\ln X]}{4} = \frac{1}{2} \Rightarrow I_{(X_1,\ldots,X_n)}(\theta) = nI_X(\theta) = \frac{n}{2}, \quad \forall \theta \in \mathbb{R}.$ *0.375*
- $\frac{\partial \ln f_\theta^n(x_1,\ldots,x_n)}{\partial \theta} = \sum_{i=1}^n \frac{\partial \ln f_\theta(x_i)}{\partial \theta} = \frac{1}{2}\left[\sum_{i=1}^n \ln x_i - n\theta\right], \quad x_1, \ldots, x_n > 0$ *0.25*
- $a(\theta) = \frac{1}{2}, \quad T(X_1,\ldots,X_n) = \sum_{i=1}^n \ln X_i, \quad g(\theta) = n\theta, \quad g'(\theta) = n \longrightarrow I_{(X_1,\ldots,X_n)}(\theta) = \frac{n}{2} = a(\theta)g'(\theta)$ *0.25*
- $T = \sum_{i=1}^n \ln X_i \in \mathbb{R}$ *estimador de* $g(\theta) = n\theta \in \mathbb{R}$. *0.125*

$$\forall a \neq 0, b \in \mathbb{R}, \ a\sum_{i=1}^n \ln X_i + b \text{ estimador eficiente de } an\theta + b. \ 0.125$$

b2) **0.5** *Calcular $E_\theta[U \ln X_i]$, siendo $U \equiv U(X_1, \ldots, X_n)$ un estimador insesgado de $1/\theta$ y regular.*

U *insesgado en* $1/\theta$ *es regular* $\Leftrightarrow -\frac{1}{\theta^2} = E_\theta\left[U \frac{\partial \ln f_\theta^n(X_1,\ldots,X_n)}{\partial \theta}\right] = \sum_{i=1}^n E_\theta\left[U \frac{\partial \ln f_\theta(X_i)}{\partial \theta}\right] = nE_\theta\left[U \frac{\partial \ln f_\theta(X_i)}{\partial \theta}\right]$

$E_\theta\left[U \frac{\partial \ln f_\theta^n(X_i)}{\partial \theta}\right] = \frac{1}{2} E_\theta[U(\ln X_i - \theta)] = \frac{1}{2} E_\theta[U \ln X_i] - \frac{1}{2} = -\frac{1}{n\theta^2} \Rightarrow E_\theta[U \ln X_i] = 1 - \frac{2}{n\theta^2}, \ i = 1, \ldots, n.$

b3) **0.25** *Calcular la cota para la varianza de estimadores regulares, insesgados en $3\theta^2 + 2$, y justificar si se alcanza o no dicha cota.*

$$\frac{(6\theta)^2}{I_{(X_1,\ldots,X_n)}(\theta)} = \frac{72\theta^2}{n} \quad \textit{NO SE ALCANZA}$$

EJERCICIO 4.

4.1) 2.75 puntos $\boxed{a=2, \quad k=2}$

Sea $(X_1,\dots X_n)$ una muestra aleatoria simple de una variable X con funciń de densidad:

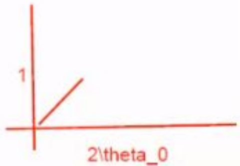
$$f_\theta(x) = \frac{x}{2\theta^2}, \quad 0 < x < 2\theta.$$

a) *Deducir el test de razón de verosimilitudes de tamaño α para contrastar $H_0 : \theta = \theta_0$ frente a $H_1 : \theta < \theta_0$, detallando el espacio paramétrico, el espacio muestral, y representando gráficamente el estadístico de contraste. ¿Qué tamaños se alcanzan? Justificar la respuesta.*

$$\begin{matrix} H_0 : \theta = \theta_0 \\ H_1 : \theta < \theta_0 \end{matrix} \longrightarrow \Theta = (0, \theta_0], \quad \chi^n = \{(x_1,\dots,x_n) \in \mathbb{R}^{+n}, \ \max x_i < 2\theta_0\} = (0,\ 2\theta_0)^n. \quad 0.25$$

$$(x_1,\dots,x_n) \in \chi^n \ \to \ L_{x_1,\dots,x_n}(\theta) = \begin{cases} 0, & 0 < \theta \le \max x_i/2 \\ \dfrac{x_1\cdots x_n}{2^n\theta^{2n}}, & \max x_i/2 < \theta \le \theta_0, \end{cases} \to \widehat{\theta}(x_1,\dots,x_n) = \max x_i/2. \ 0.5$$

- $\lambda(x_1,\dots,x_n) = \dfrac{L_{x_1,\dots,x_n}(\theta_0)}{\sup\limits_{\theta\le\theta_0} L_{x_1,\dots,x_n}(\theta)} = \dfrac{(\max x_i/2)^{2n}}{\theta_0^{2n}} = \left(\dfrac{\max x_i}{2\theta_0}\right)^{2n}, \quad 0 < \max x_i < 2\theta_0.$ 0.375



- $\lambda(x_1,\dots,x_n)$ crece estrictamente de 0 a 1 con $\max x_i$. 0.125
- $\varphi(X_1,\dots,X_n) = \begin{cases} 1 & \text{si } \max X_i < c' \quad (0 < c' \le 2\theta_0) \\ 0 & \text{si } \max X_i \ge c' \end{cases}$ 0.125+0.125
- $\alpha = P_{\theta_0}(\max X_i < c') = \left(\displaystyle\int_0^{c'} \frac{x}{2\theta_0^2}dx\right)^n = \left(\frac{c'^2}{4\theta_0^2}\right)^n \Rightarrow c' = 2\theta_0\alpha^{1/2n}.$ 0.25
- $c' = 2\theta_0\alpha^{1/2n} \le 2\theta_0, \quad \forall\alpha\in[0,1]$, por lo que se alcanzan todos los tamaños. 0.25

$$\boxed{\textit{TRV de tamaño } \alpha \in [0,1] \ \to \ \varphi(X_1,\dots,X_n) = \begin{cases} 1 & \text{si } \max X_i < 2\theta_0\alpha^{1/2n} \\ 0 & \text{si } \max X_i > 2\theta_0\alpha^{1/2n}. \end{cases}}$$

b) *Aplicando la definición de intervalo de confianza, deducir el nivel de confianza del intervalo obtenido a partir del test del apartado a).*

- Región de aceptación $\to 2\theta_0 < \max X_i\alpha^{-1/2n} \ \Rightarrow \ (0,\ \max X_i\alpha^{-1/2n}/2)$ es IC para θ al nivel $1-\alpha$. 0.25
- Deducción del nivel de confianza: 0.5

$$F_\theta(x) = P_\theta(X \le x) = \frac{x^2}{(2\theta)^2}, \quad 0 < x < 2\theta, \quad \forall\theta > 0.$$

$$P_\theta\big((0,\ \max X_i\alpha^{-1/2n}/2) \ni \theta\big) = P_\theta\big(\max X_i > 2\theta\alpha^{1/n}\big) = 1 - \big(F_\theta(2\theta\alpha^{1/2n})\big)^n = 1 - (\alpha^{1/n})^n = 1 - \alpha, \quad \forall\theta > 0.$$

EJERCICIO 5.

5) **1.25 puntos** *Para contrastar si la opinión sobre la cobertura de red móvil es la misma en los pueblos de una comarca, se ha efectuado una encuesta en cada uno de ellos. La siguiente tabla muestra las respuestas a la pregunta: ¿Hay buena cobertura de red móvil en su pueblo?*

	Sí	No	No sabe	No contesta
Pueblo 1	n_{11}	n_{12}	n_{13}	n_{14}
Pueblo 2	n_{21}	n_{22}	n_{23}	n_{24}
Pueblo 3	n_{31}	n_{32}	n_{33}	n_{34}

a) *Describir las variables usadas en este problema y especificar formalmente la hipótesis a contrastar en términos de la distribución dichas variables.*

$$X_1, X_2, X_3 : \text{Opinión sobre la cobertura en los pueblos } 1,\ 2,\ 3.$$

$$A_j \equiv \text{Sí, NO NS, NC} \to H_0 : \ P(X_1 = A_j) = P(X_2 = A_j) = P(X_3 = A_j),\ j = 1, \ldots, 4. \quad 0.25$$

b) *Especificar qué parámetros deben estimarse para calcular el estadístico de contraste, cuáles son los estimadores usados y justificar por qué se usan estos estimadores.*

Hay que estimar las frecuencias esperadas bajo H_0. Bajo H_0 todas las variables tienen la misma distribución y, por tanto, hay que estimar $p_j^0 = P_{H_0}(X = A_j),\ j = 1, \ldots, 4$, siendo X una variable con dicha distribución común. Estas probabilidades se estiman por máxima verosimilitud, basándose en el número total de datos en A_j, $N_{\cdot j} = \sum_{i=1}^{3} N_{ij}$. Ya que $N_{\cdot j} \to B(n, p_j^0)$, siendo n el total de datos observados, el estimador de p_j^0 es $\frac{N_{\cdot j}}{n}$, con . 0.25

c.1) *A la vista de los datos, ¿qué conclusión se obtiene sobre el problema planteado al nivel de significación 0.05? (trabajar con cuatro cifras decimales).* 0.75

	Sí	No	No sabe	No contesta	n_i
Pueblo 1	7	15	1	2	25
Pueblo 2	12	6	3	4	25
Pueblo 3	10	7	7	1	25
$n_{\cdot j}$	29	28	11	7	75

$\frac{n_i \cdot n_{\cdot j}}{n}$	Sí	No	No sabe	No contesta
Pueblo 1	9.6667	9.3333	3.6667	2.3333
Pueblo 2	9.6667	9.3333	3.6667	2.3333
Pueblo 3	9.6667	9.3333	3.6667	2.3333

Se agrupan la dos últimas categorías

	Sí	No	No sabe/No contesta	n_i
Pueblo 1	7	15	3	25
Pueblo 2	12	6	7	25
Pueblo 3	10	7	8	25
$n_{\cdot j}$	29	28	18	75

$e_{ij} = \frac{n_i \cdot n_{\cdot j}}{n}$	Sí	No	No sabe/No contesta
Pueblo 1	9.6667	9.3333	6
Pueblo 2	9.6667	9.3333	6
Pueblo 3	9.6667	9.3333	6

$\frac{(n_{ij} - e_{ij})^2}{e_{ij}}$	Sí	No	No sabe/No contesta
Pueblo 1	0.7356	3.4405	1.5
Pueblo 2	0.5632	1.1905	0.1667
Pueblo 3	0.0115	0.5833	0.6667

$\chi^2_{exp} = 8.8580$ 0.5

R.C. $\chi^2_{exp} > \chi^2_{4,0.05} = 9.4877$. No se rechaza H_0, los datos no muestran evidencias para rechazar que la opinión sobre la cobertura de red móvil es la misma en los tres pueblos estudiados de la comarca. 0.25