... señale claramente aquí qué ejercicios entrega

Entrego ejercicios 1 2 3 4

(Reservado para el corrector):

1	2	3	4	Nota
\	25	23	15	63

A Poff! Justo...

Por favor, resuelva cada ejercicio en hojas separadas. Numere todas las hojas y coloque en cada una su nombre y apellido. *Para aprobar es necesario tener al menos 60 puntos. Justifique todas sus respuestas.*

1. (25p.) Una fábrica produce bombones que se componen de dulce de leche y chocolate. La cantidad de chocolate en gramos en cada bombón es una variable aleatoria con distribución $N(\mu = 100, \sigma^2 = 9)$ mientras que la cantidad de dulce de leche en gramos es una variable con distribución continua desconocida con media 80 gramos y desvío estándar 4 gramos. Cabe aclarar que las cantidades de dulce de leche y la de chocolate en un bombón NO son independientes entre sí, pero se sabe que la covarianza entre ambos es de 5 gramos^2.

   (a) (12p.) La fábrica produce 100 bombones de forma independiente. Hallar una cota inferior no trivial (es decir, distinta de cero) para la probabilidad exacta de que el peso promedio de los 100 bombones esté entre los 175 y los 185 gramos.

   (b) (13p.) Ahora la fábrica produce una cantidad n de bombones, calcular n para que la probabilidad aproximada de que la suma de los pesos de los n bombones sea mayor a 10000 gramos sea al menos 0.99.

2. (35p.) Sea $X_1, \dots, X_n$ una muestra aleatoria con X_i una variable aleatoria con la siguiente densidad:

$$f_X(x) = \frac{3}{\theta^3} x^2 I_{(0,\theta]}(x), \quad \theta > 0$$

   (a) (4p.) Hallar un estimador de θ basado en el primer momento.

   (b) (4p.) Hallar un estimador de θ basado en el segundo momento.

   (c) (5p.) Hallar el estimador de θ de máxima verosimilitud.

   (d) (12p.) Analizar la consistencia de los tres estimadores.

   (e) (10p.) Hallar un intervalo de confianza de nivel exacto $1 - \alpha$ para θ basado en la muestra. Sugerencia: analizar la distribución de $T = \frac{\max_{1 \le i \le n} X_i}{\theta}$.

3. (25p.) Los canales de la ciudad de Hamburgo son conocidos por sospecharse de contar con un alt nivel de contaminación. Por un tema de salud pública, sería necesario tal vez llevar a cabo un costoso plan de concientización ciudadana sobre la manera de tratar los residuos contaminantes. El ayuntamiento está evaluando llevar a cabo el plan, pero lo hará solamente si tiene la certeza de que el porcentaje promedio de material nocivo para la salud humana (metales pesados, bacterias, etc) es mayor al 10%. Realiza una muestra en 25 puntos distintos de la ciudad y obtiene una contaminación promedio de 10,8%. Se sabe, por cuestiones de química de los contaminantes y particularidades de la zona, que es razonable suponer que la distribución del porcentaje de contaminación en un elemento de la muestra sigue una distribución normal de media μ desconocido (el valor de interés propiamente dicho) y varianza 4. Se supone que la muestra es independiente. El ayuntamiento desea que la probabilidad de llevar a cabo el plan de concientización cuando en realidad no hubiese sido estrictamente necesario sea de a lo sumo 0.04.

(a) (4p.) Definir las variables aleatorias involucradas en el problema y plantear las hipótesis a testear.

(b) (6p.) Consturir un test de nivel exacto 0,04 apropiado para el problema, especificando estadístico de test, distribución bajo el borde de la hipótesis nula y región de rechazo.

(c) (3p.) En base a la muestra observada, cuál es la decisión del ayuntamiento?

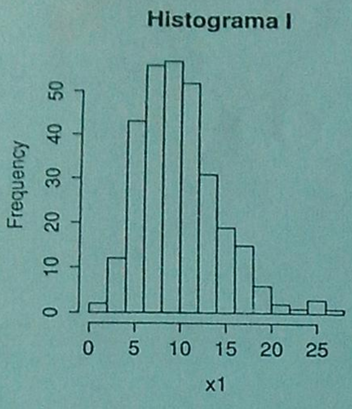
(d) (3p.) Calcule el p-valor.

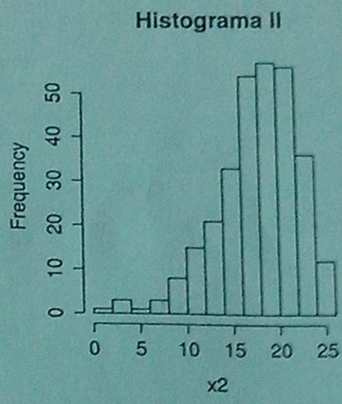
(e) (3p.) Supongamos que ahora se realiza un test de nivel 0,01. **Sin hacer ninguna cuenta**, rehaga el punto c).

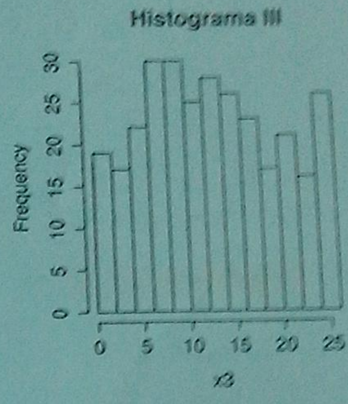
(f) (6p.) Supongamos que ahora se muestrean en n puntos y se considera el test original de nivel 0,04. Calcular un n de forma tal que la probabilidad de decidir llevar a cabo el plan de concientización cuando μ vale 11 sea al menos 0.90.

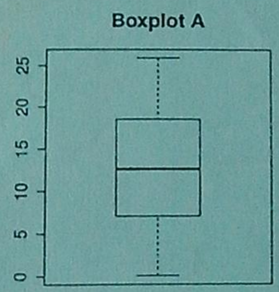
15/15

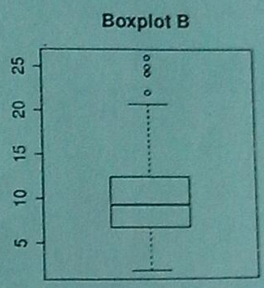
4. (15p.) Los vectores $\mathbf{x1}$, $\mathbf{x2}$ y $x3$ son muestras aleatorias de 300 observaciones cada una provenientes de las distribuciones continuas F_1, F_2 y F_3 respectivamente. A continuación se muestran sus histogramas y sus boxplots.

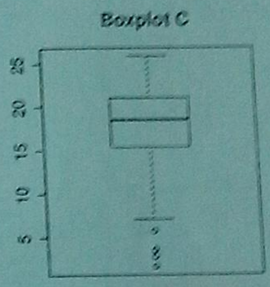












Llamamos $\tilde{\mu}_1$, $\tilde{\mu}_2$ y $\tilde{\mu}_3$ a las medianas de F_1, F_2 y F_3

(a) (2p.) Indique a qué histograma corresponde cada boxplot.

A con III, B con I y C con II.

(b) (2.5p.) La distribución cuya muestra generó el boxplot C, ¿Qué arriesgarías, tiene mayor la media o la mediana? Responder mirando sólo el boxplot

Tiene mayor la mediana que la media.

(c) (3p.) Ordene de menor a mayor, según parecen indicar los histogramas, los números $F_1(10)$, $F_2(10)$ y $F_3(10)$.

$F_2(10)$, $F_3(10)$, $F_1(10)$

En los siguientes items indique verdadero o falso.

(d) (1.5p.) Se puede afirmar casi con seguridad que F_2 es una distribución normal. F

(e) (1.5p.) Mirando los boxplot, podemos afirmar que casi con seguridad la distribución que generó el boxplot A tiene mayor distancia intercuartil que las otras. V

(f) (1.5p.) Mirando los histogramas, podemos afirmar que casi con seguridad: $F_2(\tilde{\mu}_1) > \frac{1}{2}$ F

(g) (1.5p.) Podemos afirmar sin temor a la menor duda que: $F_2(\tilde{\mu}_2) = \frac{1}{2}$ V

(h) (1.5p.) Mirando los histogramas, podemos afirmar que casi con seguridad $F_1(\tilde{\mu}_2) > \frac{1}{2}$ V

2

a) Estimador de θ basado en el primer momento

$$E(X) = \int_{-\infty}^{\infty} x \, \frac{3}{\theta^3} x^2 \, I_{(0,\theta]}(x)\,dx = \frac{3}{\theta^3}\int_0^{\theta} x^3\,dx =$$

4/4

$$= \frac{3}{\theta^3}\left(\frac{x^4}{4}\right)\Bigg|_0^{\theta} = \frac{3}{\theta^3}\left(\frac{\theta^4}{4}\right) = \frac{3}{4}\theta \quad \rightarrow \boxed{\hat{\theta}_{M1} = \overline{X}\,\frac{4}{3}}$$

$\overline{X} \xrightarrow{P} E(X)$ por LGN

b) Estimador de θ basado en el segundo momento

$$E(X^2) = \int_{-\infty}^{\infty} x^2 \cdot \frac{3}{\theta^3} x^2 \, I_{(0,\theta]}(x)\,dx =$$

4/4

$$= \frac{3}{\theta^3}\int_0^{\theta} x^4\,dx = \frac{3}{\theta^3}\left(\frac{x^5}{5}\right)\Bigg|_0^{\theta} = \frac{3}{\cancel{\theta^3}}\,\frac{\theta^{\cancel{5}\,2}}{5} = \frac{3}{5}\theta^2 \quad \boxed{\hat{\theta}_{M2} = \sqrt{\frac{5}{3}\,\overline{X^2}}}$$

$\overline{X^2} \xrightarrow{P} E(X^2)$ por LGN

c) Estimador de θ de máxima verosimilitud

5/5

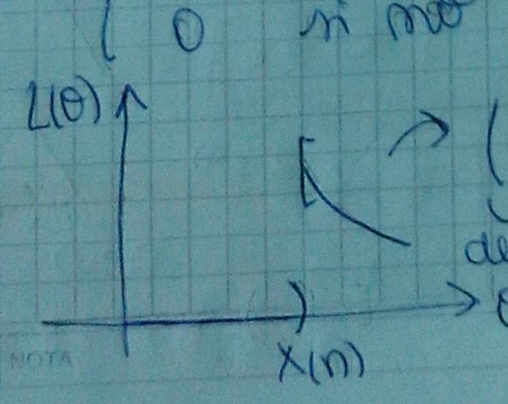
$$f_X(\vec{X}) = \prod_{i=1}^{n} f_X(x_i) = L(\theta)$$

Busco el argumento que maximiza $L(\theta)$:

$$L(\theta) = \prod_{i=1}^{n} \frac{3}{\theta^3} x_i^2 \, I_{(0,\theta]}(x_i) = \left(\frac{3}{\theta^3}\right)^n \cdot \prod_{i=1}^{n} x_i^2 \cdot I_{(0,\theta]}(x_i)$$

$$\prod_{i=1}^{n} I_{(0,\theta]}(x_i) = \begin{cases} 1 & \text{si } 0 \le x_i \le \theta \ \forall i \\ 0 & \text{si no} \end{cases} =$$

$$= \begin{cases} 1 & \text{si } X_{(n)} \le \theta \\ 0 & \text{si no} \end{cases}$$



$$\underbrace{\left(\frac{3}{\theta^3}\right)^n}_{\text{decreciente}} \cdot \underbrace{\prod_{i=1}^{n} x_i^2}_{\text{constante mayor a } 0} \rightarrow \boxed{\hat{\theta}_{MV} = X_{(n)}}$$

a) Consistencia de los tres estimadores: 12/12

Al fin alguien se acordó que vale xq g es continua!! Gracias!!

* $\hat{\theta}_{MO_1}$

$\overline{X} \xrightarrow{P} E(X)$ por L.G.N.

Tomo $g(x) = \frac{4}{3}x$ (es continua $\forall \mathbb{R}$)

$\Rightarrow \frac{4}{3}\overline{X} \xrightarrow{P} \frac{4}{3}E(X) \Leftrightarrow \hat{\theta}_{MO_1} \xrightarrow{P} \cancel{\frac{4}{3}}\cancel{\frac{3}{4}}\theta$

$\therefore$ como $\hat{\theta}_{MO_1} \xrightarrow{P} \theta$, $\boxed{\hat{\theta}_{MO_1} \text{ es consistente}}$

* $\hat{\theta}_{MO_2}$:

Ahora tomo $g(x) = \sqrt{\frac{5}{3}x}$ (continua $\forall \mathbb{R}_{>0}$, y $\overline{X^2} > 0$ y $E(X^2) > 0$)

$\overline{X} \xrightarrow{P} E(X) \Rightarrow \overline{X^2} \xrightarrow{P} E(X^2) \Rightarrow g(\overline{X^2}) \xrightarrow{P} g(E(X^2))$

$\Leftrightarrow \hat{\theta}_{MO_2} \xrightarrow{P} \theta$

$\therefore$ $\boxed{\hat{\theta}_{MO_2} \text{ es consistente}}$

* $\hat{\theta}_{MV}$:

porque son disjuntos

$P(|\hat{\theta}_{MV} - \theta| > \varepsilon) = \underbrace{P(X_{(n)} - \theta > \varepsilon)}_{1} + \underbrace{P(X_{(n)} - \theta < -\varepsilon)}_{2} = 3$

$1 = P(X_{(n)} > \theta + \varepsilon) = 0$ porque X se encuentra entre 0 y θ (por $I_{(0,\theta]}(x)$ de $f_X(x)$)

$2 = P(X_{(n)} < \theta - \varepsilon) = P\left(\bigcap_{i=1}^{n} (X_i < \theta - \varepsilon)\right) = [P(X_1 < \theta - \varepsilon)]^n =$ porque son i.i.d.

$= \left[\int_{-\infty}^{\theta-\varepsilon} \frac{3}{\theta^3} x^2 \cdot I_{(0,\theta]}(x)\,dx\right]^n = \left[\frac{3}{\theta^3}\int_0^{\theta-\varepsilon} x^2\,dx\right]^n =$

$= \left[\frac{3}{\theta^3}\left(\frac{x^3}{3}\right)\Big|_0^{\theta-\varepsilon}\right]^n = \left[\frac{\cancel{3}}{\theta^3}\frac{(\theta-\varepsilon)^3}{\cancel{3}}\right]^n = \left[\frac{(\theta-\varepsilon)^3}{\theta^3}\right]^n \xrightarrow[n\to\infty]{} 0$

< 1 ($\theta > 0$ y $\varepsilon > 0$)

$\Rightarrow 3 \xrightarrow[n\to\infty]{} 0 \Rightarrow \boxed{\hat{\theta}_{MV} \text{ es consistente}}$

HOJA Nº 1 | 1

FECHA

(8) Plan si % promedio > 10%

m.a. 25 puntos con cont promedio = 10,8%

P(plan cuando no es necesario) ≤ 0,04

P(rechazar H_0 | H_0 verdadera)

a) X_i = porcentaje de contaminación en el i-ésimo punto de la ciudad.

$X_i \sim N(\mu, 4)$

$X_1, \dots, X_{25}$ i.i.d.

$\boxed{H_0: \mu \leq 10}$

$\boxed{H_1: \mu > 10}$

4/4

b) Estadístico del test: $\boxed{\frac{(\bar{X} - 10)\sqrt{25}}{2}}$

4/6

Bajo la hipótesis de H_0: $\boxed{\frac{(\bar{X} - 10)\sqrt{25}}{2} \xrightarrow{\mathcal{D}} Z \sim N(0,1)}$ por T.C.L.

Ojo que $Z = \frac{(\bar{X} - 10)}{2} 5 \sim N(0,1)$ no necesito el TCL. El test te lo piden exacto!

Región de rechazo:

$P_{H_0}(Z > z_\alpha) = 0,04 = 1 - P_{H_0}(Z \leq z_\alpha)$

$\Leftrightarrow P_{H_0}(Z \leq z_\alpha) = 0,96$

$\Rightarrow z_\alpha = \Phi(0,96) = 1,76$

$\Rightarrow$ R.R.: $\boxed{Z > 1,76}$

c) $z_{OBS} = \frac{(10,8 - 10)\sqrt{25}}{2} = 2$ y $z_{OBS} > 1,76$

3/3

∘8 La decisión del ayuntamiento debería ser llevar a cabo el plan pues existen evidencias suficientes para rechazar H_0

NOTA

d) $P(Z > Z_{OBS}) = P(Z > 2) = 1 - P(Z \leq 2) = 1 - 0,9772 =$
$= \boxed{0,0228}$

3/3

e) Con un test de nivel 0,01 no existen evidencias suficientes para rechazar H_0 y, por lo tanto, la decisión del ayuntamiento debería ser no llevar a cabo el plan.

3/3

f) n puntos
test de nivel 0,04
$n \,/\, P(\text{rechazar } H_0 \mid \mu = 11) > 0,90$

6/6

$P(\text{rechazar } H_0 \mid \mu=11) = P_{\mu=11}(Z > 1,76) =$

$= P\left(\frac{\bar{x} - 11}{2}\sqrt{n} + \frac{11-10}{2}\sqrt{n} > 1,76 \right) =$

$\xrightarrow{D} Z \sim N(0,1)$ bajo $\mu = 11$

$= P\left(Z > 1,76 - \frac{1}{2}\sqrt{n} \right) = 1 - P\left(Z \leq 1,76 - \frac{1}{2}\sqrt{n} \right) \Rightarrow$

$\Rightarrow 1 - \Phi\left(1,76 - \frac{1}{2}\sqrt{n} \right) > 0,90$

$\Updownarrow$

$-\Phi\left(1,76 - \frac{1}{2}\sqrt{n}\right) > 0,90 - 1 = -0,1$

$\Updownarrow$

$\Phi\left(1,76 - \frac{1}{2}\sqrt{n}\right) < 0,1$

Perfecto!!

$\Updownarrow$

$1,76 - \frac{1}{2}\sqrt{n} < \Phi^{-1}(0,1) = -\Phi^{-1}(0,9) = -1,29$

Despejo n:

$n > 37,21 \Rightarrow \boxed{n > 38}$

NOTA