Probabilidad - 3er Curso (Doble Grado en Ingeniería Informática y Matemáticas)
Convocatoria ordinaria (11 de enero de 2023)

Apellidos, nombre:

---

1. **(1.5 puntos)** Justificar las siguientes relaciones:

   a) **(0.25 puntos)** Sean X_1 y X_2 variables aleatorias independientes e identicamente distribuidas según una ley Binomial, $B(3, \frac{1}{2})$. Justificar que $P[X_1 + X_2 = 8] = 0$.

   b) **(0.25 puntos)** Sean X_1, X_2 y X_3 variables aleatorias independientes e identicamente distribuidas según una ley de Poisson, $P(3)$. Justificar que $P[X_1 + X_2 + X_3 > 0] = \frac{e^9 - 1}{e^9}$.

   c) **(1 punto)** Sea $\{X_n\}_{n \in \mathbb{N}}$ una sucesión de variables aleatorias independientes e identicamente distribuidas según una ley de Bernoulli de parámetro p. Se considera $S_n = \sum_{i=1}^{n} X_i$, la variable aleatoria que define las sumas parciales de la sucesión. Justificar que $\dfrac{S_n}{n} \to^P p$.

2. **(1.5 puntos)** Sean $X_1, \ldots, X_n$ n variables aleatorias continuas, independientes e identicamente distribuidas con distribución uniforme en el intervalo $[0, 1]$. Deducir la expresión analítica de la función de densidad de la variable aleatoria $Z = max\{X_1, \ldots, X_n\}$.

3. **(5 puntos)** Dado el vector aleatorio (X, Y) con distribución uniforme en el recinto acotado limitado por el exterior de la parábola $y = x^2$, la recta de ecuación $2y + x = 1$ y la recta de ecuación $y = 0$:

   a) **(0.25 puntos)** Obtener la función de densidad conjunta.

   b) **(1.50 puntos)** Obtener la función de distribución de probabilidad conjunta.

   c) **(0.75 puntos)** Obtener las funciones de densidad condicionadas.

   d) **(0.50 puntos)** Obtener la probabilidad de que $X \geq \dfrac{1}{2}$.

   e) **(1.25 puntos)** Obtener la mejor aproximación mínimo cuadrática a la variable X conocidos los valores de la variable Y.

   f) **(0.50 puntos)** Obtener la mejor aproximación de la variable aleatoria Y sin observar la variable X y dar una medida del error cuadrático medio cometido en esta aproximación.

   g) **(0.25 puntos)** ¿Son X e Y variables aleatorias independientes? Justificar de forma muy breve la respuesta.

4. **(2 puntos)** Sea (X,Y) un vector aleatorio con distribución normal bidimensional. La moda de Y vale 4 y la $Var[Y/X = x_0] = \dfrac{Var(Y)}{2} \neq 0$. La curva de regresión de Y/X es $y = x + 5$ y el error cuadrático medio asociado a esta aproximación es 3.

   a) **(1.25 puntos)** Determinar los parámetros de la distribución de (X, Y).

   b) **(0.75 puntos)** Especificar la función generatriz de momentos de (X, Y).

Observaciones e indicaciones:

- En el **apartado 1.c** hay que demostrar el/los resultados empleados para justificar la relación pedida, salvo la Desigualdad de Chebychev.
- En el **apartado 3.b** se obtienen **hasta 1.25 puntos** si las integrales se dejan indicadas y **hasta 1.50 puntos** si se obtienen sus valores de forma explícita.

1. **(1.5 puntos)** Justificar las siguientes relaciones:

   a) **(0.25 puntos)** Sean X_1 y X_2 variables aleatorias independientes e identicamente distribuidas según una ley Binomial, $B(3, \frac{1}{2})$. Justificar que $P[X_1 + X_2 = 8] = 0$.

   b) **(0.25 puntos)** Sean X_1, X_2 y X_3 variables aleatorias independientes e identicamente distribuidas según una ley de Poisson, $P(3)$. Justificar que $P[X_1 + X_2 + X_3 > 0] = \frac{e^9 - 1}{e^9}$.

   c) **(1 punto)** Sea $\{X_n\}_{n \in \mathbb{N}}$ una sucesión de variables aleatorias independientes e identicamente distribuidas según una ley de Bernoulli de parámetro p. Se considera $S_n = \sum_{i=1}^{n} X_i$, la variable aleatoria que define las sumas parciales de la sucesión. Justificar que $\frac{S_n}{n} \to^P p$.

A)

$X_1, X_2 \leadsto B(n=3, p=1/2)$ y son ind $\Longrightarrow$ Por reproductividad

$X_1 + X_2 \leadsto B(n=3+3=6, p=1/2)$,

$P[X_1 + X_2 = x] = \binom{6}{x} \left(\frac{1}{2}\right)^x \cdot \left(\frac{1}{2}\right)^{6-x}$ $x = 0, 1 \ldots 6$ $\Longrightarrow$

$P[X_1 + X_2 = 8] = 0$ Suceso Imposible!!!

B)

De la misma forma, por reproductividad, $X_1 + X_2 + X_3 \leadsto P(\lambda = 3+3+3 = 9)$

$P[X_1 + X_2 + X_3 > 0] = 1 - P[X_1 + X_2 + X_3 = 0] = 1 - \frac{\lambda^0}{0!} e^{-\lambda} =$
$1 - e^{-9} = \frac{e^9 - 1}{e^9}$

2. **(1.5 puntos)** Sean $X_1,\ldots,X_n$ n variables aleatorias continuas, independientes e identicamente distribuidas con distribución uniforme en el intervalo $[0,1]$. Deducir la expresión analítica de la función de densidad de la variable aleatoria $Z = max\{X_1,\ldots,X_n\}$.

$$\left.\begin{array}{ll} X_1,\ldots X_n \text{ id. distr.} \Rightarrow & F_{X_i}(x) = F(x) \quad \forall i = 1 \ldots n \\ X_1 \ldots X_n \text{ ind} \Rightarrow & F_{(X_1\ldots X_n)}(x_1\ldots x_n) = \prod_{i=1}^{n} F_{X_i}(x_i) \end{array}\right\} \Rightarrow F_{(X_1\ldots X_n)}(x_1\ldots x_n) = \prod_{i=1}^{n} F(x_i)$$

$$F_{\text{máx}(X_1\ldots X_n)}(y) = P[\text{máx}(X_1\ldots X_n) \leq y] = P[X_1 \leq y \ldots X_n \leq y] = \prod_{i=1}^{n} F_{X_i}(y) = F(y)^n$$

$$f_{\text{máx}(X_1\ldots X_n)}(x) = \frac{dF_{\text{máx}(X_1\ldots X_n)}}{dx}(x) = n\,F(x)^{n-1}\,F'(x) = n\,f(x)\,F(x)^{n-1}, \text{ donde}$$

f es la fdd de X_i $\forall i = 1 \ldots n$

sabemos $f(x) = \begin{cases} \frac{1}{1-0} = 1 & x \in [0,1] \\ 0 & \text{en otro caso} \end{cases}$

$$F(x) = \int_0^x 1 \cdot dy = x \quad \forall x \in [0,1]$$

$$f_{\text{máx}(X_1\ldots X_n)}(x) = \begin{cases} n x^{n-1} & \forall x \in [0,1] \\ 0 & \text{en otro caso} \end{cases}$$

3. **(5 puntos)** Dado el vector aleatorio (X,Y) con distribución uniforme en el recinto acotado limitado por el exterior de la parábola $y = x^2$, la recta de ecuación $2y + x = 1$ y la recta de ecuación $y = 0$:

a) **(0.25 puntos)** Obtener la función de densidad conjunta.

b) **(1.50 puntos)** Obtener la función de distribución de probabilidad conjunta.

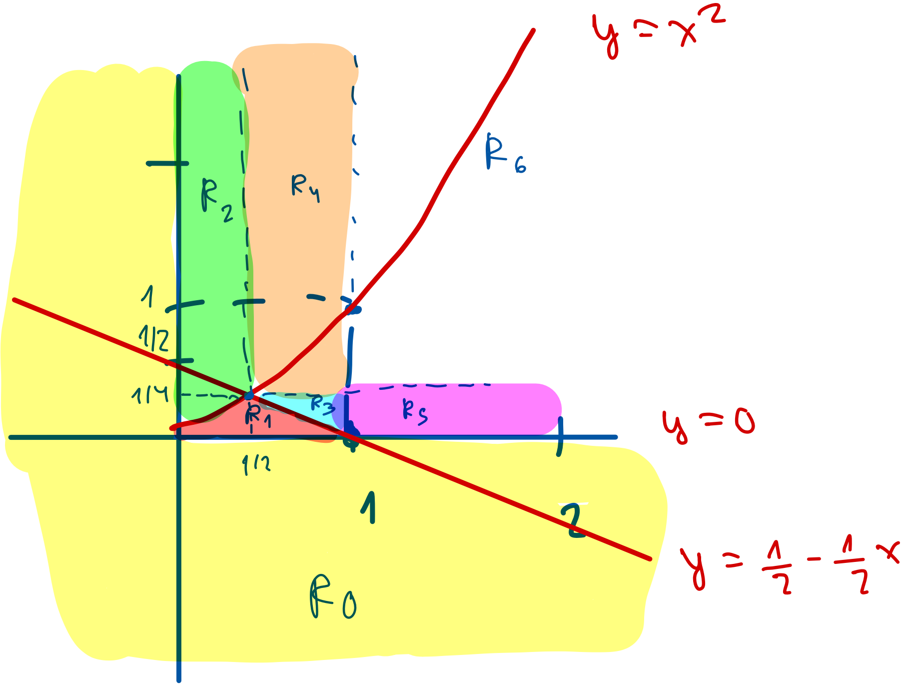
c) **(0.75 puntos)** Obtener las funciones de densidad condicionadas.

d) **(0.50 puntos)** Obtener la probabilidad de que $X \geq \frac{1}{2}$.

e) **(1.25 puntos)** Obtener la mejor aproximación mínimo cuadrática a la variable X conocidos los valores de la variable Y.

f) **(0.50 puntos)** Obtener la mejor aproximación de la variable aleatoria Y sin observar la variable X y dar una medida del error cuadrático medio cometido en esta aproximación.

g) **(0.25 puntos)** ¿Son X e Y variables aleatorias independientes? Justificar de forma muy breve la respuesta.



$$x^2 = \frac{1}{2} - \frac{1}{2}x \Rightarrow 2x^2 + x - 1 = 0 \Rightarrow x = \frac{-1 \pm \sqrt{1+8}}{4} = \frac{-1 \pm 3}{4} = \begin{cases} 1/2 \\ -1 \end{cases}$$

A)

$$\int_0^{1/2}\int_0^{x^2} k\,dy\,dx + \int_{1/2}^{1}\int_0^{\frac{1-x}{2}} k\,dy\,dx = k\left(\int_0^{1/2} x^2\,dx + \int_{1/2}^{1} \frac{1-x}{2}\,dx\right) = k\left(\frac{1}{24} + \frac{1}{16}\right) = 1$$

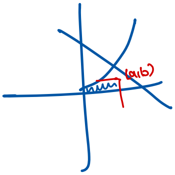
$$\Rightarrow k = 48/5 \Rightarrow f(x,y) = \begin{cases} 48/5 & (x,y) \in R_1 \\ 0 & \text{en otro caso} \end{cases}$$

B)

$R_0 = \{(x,y) \in \mathbb{R}^2 / x<0 \vee y<0\}$ $\quad F(a,b) = 0$

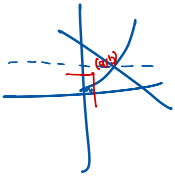
$R_7 = \{(x,y) \in \mathbb{R}^2 / 1<x \wedge \frac{1}{4} < y\}$ $\quad F(a,b) = 1$

$R_1 = \{(x,y) \in \mathbb{R}^2 / 0<x<\frac{1}{2}, y<x^2 \wedge \frac{1}{2}<x<1, y< \frac{1-x}{2}$



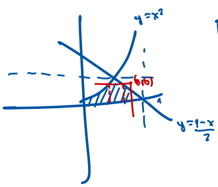
$$F(a,b) = \int_0^b \int_{\sqrt{y}}^a k \, dx \, dy$$

$R_2 = \{(x,y) \in \mathbb{R}^2 / 0<x<\frac{1}{2}, 0<y<1/4, x^2<y\}$



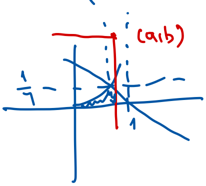
$$F(a,b) = \int_0^a \int_0^{x^2} k \, dy \, dx$$

$R_3 = \{(x,y) \in \mathbb{R}^2 / 0<x<\frac{1}{2}, 0<y<1/4, x^2<y\}$



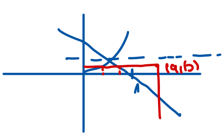
$$F(a,b) = \int_0^{\sqrt{b}} \int_0^{x^2} k \, dy \, dx + \int_{\sqrt{b}}^{1-2b} \int_0^b k \, dy \, dx + \int_{1-2b}^a \int_0^{\frac{1-x}{2}} k \, dy \, dx$$

$R_4 = \{(x,y) \in \mathbb{R}^2 / \frac{1}{2}<x<1, \frac{1}{4}<y\}$



$$F(a,b) = \int_0^{1/2} \int_0^{x^2} k \, dy \, dx + \int_{1/2}^a \int_0^{\frac{1-x}{2}} k \, dy \, dx$$

$R_5 = \{(x,y) \in \mathbb{R}^2 / 1<x, 0<y<1/4$



$$F(a,b) = \int_0^{\sqrt{b}} \int_0^{x^2} k \, dy \, dx + \int_{\sqrt{b}}^{1-2b} \int_0^b k \, dy \, dx + \int_{1-2b}^1 \int_0^{\frac{1-x}{2}} k \, dy \, dx$$

C)

$$f_X(x) = \begin{cases} \int_0^{x^2} k\,dy = kx^2 & 0<x<\frac{1}{2} \\ \int_0^{\frac{1-x}{2}} k\,dy = \frac{k}{2}(1-x) & \frac{1}{2}<x<1 \end{cases}$$

$$f_Y(y) = \int_{\sqrt{y}}^{1-2y} k\,dx = k(1-2y-\sqrt{y}) \qquad 0<y<\frac{1}{4}$$

$$f_{Y|X=x}(y) = \frac{f(x,y)}{f_X(x)} = \begin{cases} \frac{1}{x^2} & \left(0<x<\frac{1}{2}\right)\ 0<y<x^2 \\ \frac{1-x}{2} & \left(\frac{1}{2}<x<1\right)\ 0<y<\frac{1-x}{2} \end{cases}$$

$$f_{X|Y=y}(x) = \frac{f(x,y)}{f_Y(y)} = \frac{1}{1-2y-\sqrt{y}} \qquad \left(0<y<\frac{1}{4}\right)\ 0<y<x^2 \vee y<\frac{1-x}{2}$$

D)

$$P\left[X \geq \frac{1}{2}\right] = 1-P\left[X \leq \frac{1}{2}\right] = 1-\int_0^{1/2} kx^2\,dx = \frac{3}{5}$$

E)

$$E[X/Y=y] = \int_{\text{supp}\,f_{X/Y=y}} x\, f_{X/Y=y}(x)\,dx = \int_{\sqrt{y}}^{1-2y} x\,\frac{1}{1-2y-\sqrt{y}}\,dx = \frac{(1-2y)^2-(\sqrt{y})^2}{2(1-2y-\sqrt{y})} =$$

$$\frac{(1-2y-\sqrt{y})(1-2y+\sqrt{y})}{2(1-2y-\sqrt{y})} = \frac{1-2y+\sqrt{y}}{2} \Rightarrow$$

$$E[X/Y] = \frac{1-2Y+\sqrt{Y}}{2}$$

F)

La mejor aproximación de Y sin observar X es

$$E[Y] = \int_{\mathbb{R}} y\, f_Y(y)\, dy = \int_0^{1/4} yk(1-2y-\sqrt{y})\, dy = \frac{5}{48}\left[\frac{y^2}{2} - \frac{2y^3}{3} - \frac{y^{5/2}}{5/2}\right]_0^{1/4} =$$

$$\frac{2}{25} = 0'08$$

Sin Observar X, $ECM = Var[Y] = E[Y^2] - E[Y]^2 = 0'0098 - 0'08^2 = 0'0034$

$$E[Y^2] = \int_{\mathbb{R}} y^2 f_Y(y)\, dy = \frac{5}{48}\left[\frac{y^3}{3} - \frac{y^4}{2} - \frac{y^{7/2}}{7/2}\right]_0^{1/4} = \frac{49}{5000} = 0'0098$$

G)

No son independientes pues las condicionadas no coinciden con las marginales respectivas de cada variable.

4. **(2 puntos)** Sea (X,Y) un vector aleatorio con distribución normal bidimensional. La moda de Y vale 4 y la $Var[Y/X = x_0] = \frac{Var(Y)}{2} \neq 0$. La curva de regresión de Y/X es $y = x + 5$ y el error cuadrático medio asociado a esta aproximación es 3.

   a) **(1.25 puntos)** Determinar los parámetros de la distribución de (X,Y).

   b) **(0.75 puntos)** Especificar la función generatriz de momentos de (X,Y).

A)

$$(X,Y) \rightsquigarrow N_2\left(\mu = (\mu_X, \mu_Y),\ \Sigma = \begin{pmatrix} \sigma_X^2 & \sigma_X\sigma_Y\rho \\ \sigma_X\sigma_Y\rho & \sigma_Y^2 \end{pmatrix}\right)$$

$$- \ Mo[Y] = E[Y] = 4$$

$$- \ Y/X{=}x_0 \rightsquigarrow N\left(\mu_Y + \frac{\sigma_Y}{\sigma_X}\rho(x_0 - \mu_X),\ \sigma_Y^2(1-\rho^2)\right)$$

$$Var[Y/X = x_0] = \sigma_Y^2(1-\rho^2) = \frac{\sigma_Y^2}{2} \Rightarrow \sigma_Y^2\left(\frac{1}{2} - \rho^2\right) = 0 \overset{\sigma_Y^2 \neq 0}{\Longrightarrow} \rho^2 = \frac{1}{2}$$

$$\Rightarrow \rho = \frac{1}{+\sqrt{2}} > 0$$

$$- \ ECM(rr(Y/X)) = \sigma_Y^2(1-\rho^2) = \frac{1}{2}\sigma_Y^2 = 3 \Rightarrow \sigma_Y^2 = 6$$

$$- \ y = x + 5 \text{ es } rr\ Y/X \Rightarrow$$

$$y - \mu_Y = \frac{\sigma_Y}{\sigma_X}\rho(x - \mu_X) \Rightarrow y = \mu_Y - \mu_X\frac{\sigma_Y}{\sigma_X}\rho + \frac{\sigma_Y}{\sigma_X}\rho x =$$

$$y = 4 - \mu_X\underbrace{\frac{\sigma_Y}{\sigma_X}\cdot\frac{1}{\sqrt{2}}}_{1} + \underbrace{\frac{\sigma_Y}{\sigma_X}\frac{1}{\sqrt{2}}}_{1}x = 5 + x \Rightarrow \begin{cases} \sigma_Y = \sigma_X\sqrt{2} \\ 4 - \mu_X = 5 \Rightarrow \mu_X = -1 \end{cases}$$

$$\sigma_Y = \sigma_X\sqrt{2} \Rightarrow \sigma_Y^2 = 2\sigma_X^2 \Rightarrow \sigma_X^2 = \frac{6}{2} = 3$$

$$\mu = (-1, 4), \quad \Sigma = \begin{pmatrix} 3 & 3 \\ 3 & 6 \end{pmatrix}$$

B)

$$M_{(X,Y)}(t_1,t_2) = \exp\left(t_1\mu_X + t_2\mu_Y + \frac{t_1^2\sigma_X^2 + t_2^2\sigma_Y^2 + 2t_1t_2\sigma_X\sigma_Y\rho}{2}\right) =$$

$$e^{-t_1 + 4t_2 + \frac{3t_1^2 + 6t_2^2 + 6t_1t_2}{2}} \quad \forall (t_1,t_2) \in \mathbb{R}^2$$