Probabilidad
Examen VI

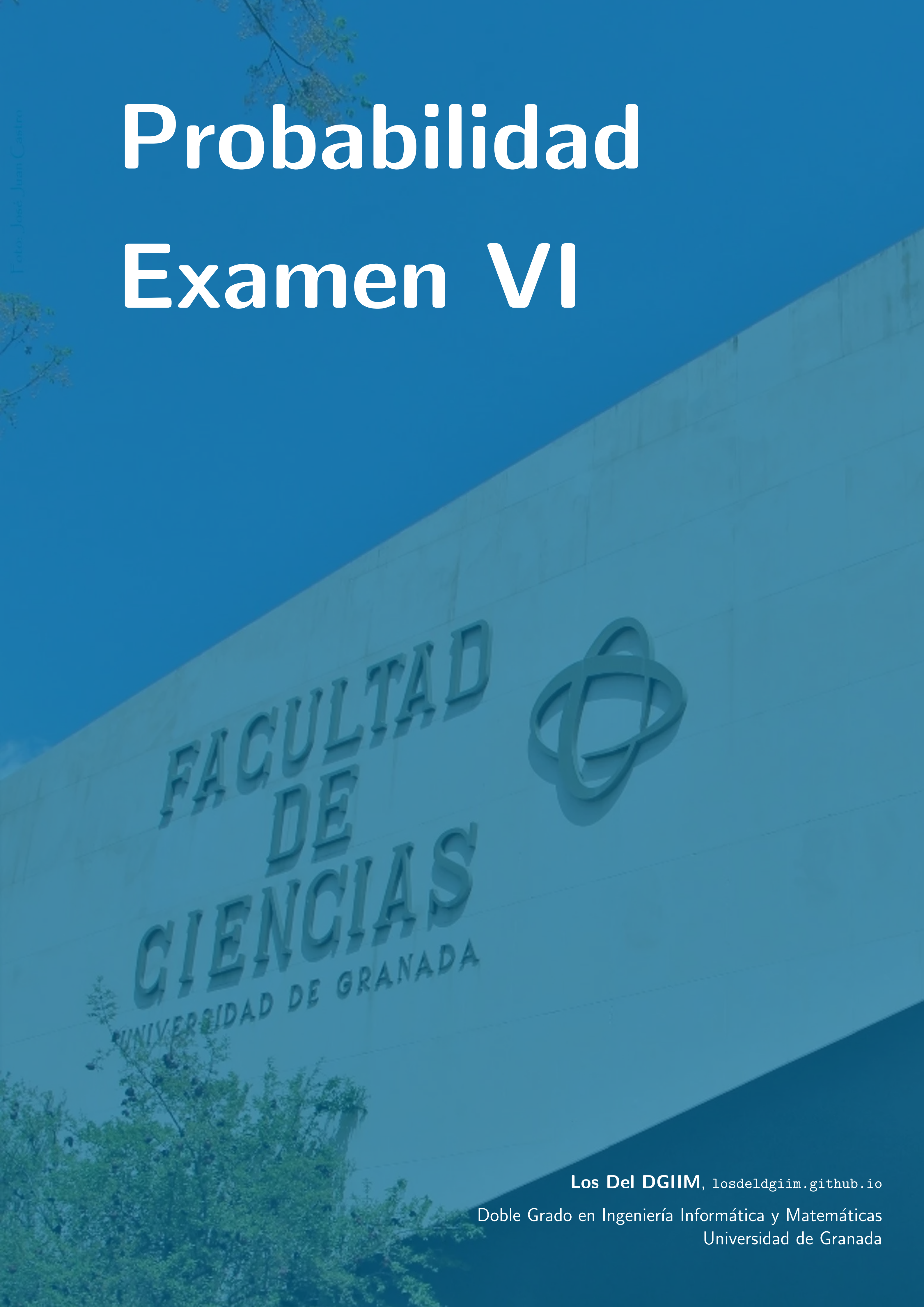


Foto: José Juan Castro

Los Del DGIIM, `losdeldgiim.github.io`

Doble Grado en Ingeniería Informática y Matemáticas
Universidad de Granada

Probabilidad
Examen VI

Los Del DGIIM, `losdeldgiim.github.io`

Arturo Olivares Martos
José Juan Urrutia Milán

Granada, 2024-2025

Asignatura Probabilidad.

Curso Académico 2019-20.

Grado Grado en Matemáticas y Doble Grado en Física y Matemáticas.

Grupo Único.

Descripción Examen Ordinario

Fecha 23 de junio de 2020.

TEORÍA (2.5 puntos)

Ejercicio 1. Obtener la expresión analítica para la función de densidad del máximo y del mínimo de n variables aleatorias independientes e identicamente distribuidas con función de distribución marginal F.

Sean $X_1, \dots, X_n$ variables aleatorias independientes e idénticamente distribuidas con función de distribución $F_{X_i} = F$ para todo $i \in \{1, \dots, n\}$. Consideramos el vector aleatorio $X = (X_1, \dots, X_n)$.

Respecto a la distribución del máximo, tenemos que:

$$F_{\text{máx}(X)}(x) = P[X_1 \leqslant x, \dots, X_n \leqslant x] = \prod_{i=1}^{n} P[X_i \leqslant x] = F(x)^n \qquad \forall x \in \mathbb{R}.$$

Por tanto, su función de densidad es:

$$f_{\text{máx}(X)}(x) = \frac{d}{dx} F_{\text{máx}(X)}(x) = nF(x)^{n-1} f(x) \qquad \forall x \in \mathbb{R}.$$

donde f es la función de densidad común a todas las variables aleatorias, que además cumple $f = F'$.

Respecto a la distribución del mínimo, tenemos que:

$$F_{\text{mín}(X)}(x) = 1 - P[X_1 > x, \dots, X_n > x] = 1 - \prod_{i=1}^{n} P[X_i > x] = 1 - (1 - F(x))^n \qquad \forall x \in \mathbb{R}.$$

Por tanto, su función de densidad es:

$$f_{\text{mín}(X)}(x) = \frac{d}{dx} F_{\text{mín}(X)}(x) = n(1 - F(x))^{n-1} f(x) \qquad \forall x \in \mathbb{R}.$$

Ejercicio 2. Dada una sucesión $\{X_n\}_{n\in\mathbb{N}}$ de variables aleatorias independientes e idénticamente distribuidas tales que existen $E[X_n] = \mu$ y $E[X_n^2] < \infty$, probar que:

$$\frac{S_n}{n} \xrightarrow{P} \mu$$

Este resultado está probado como corolario directo de la Ley de Kolmogorov.

PROBLEMAS (7.5 puntos)

Ejercicio 3. Una persona recibe un ticket que puede acumular para canjear por un premio si al elegir al azar dos números cualesquiera, X entre 0 y 1 e Y entre 1 y 3, su suma $X + Y < 2$. Este procedimiento se puede repetir de forma indefinida e independiente. Calcular el número de elecciones necesarias para poder obtener un premio que requiere de al menos 10 tickets acumulados con probabilidad superior a 0,9495.

Observación. Aunque el enunciado no lo especifica, entendemos que X e Y son variables aleatorias continuas (podrían ser discretas).

Tenemos que:

$$X \sim \mathcal{U}([0,1]) \quad Y \sim \mathcal{U}([1,3])$$

Por tanto, sus funciones de densidad son:

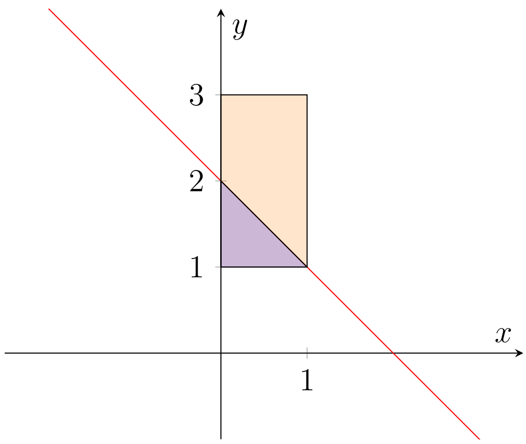
$$f_X(x) = \begin{cases} 1 & \text{si } x \in [0,1] \\ 0 & \text{en otro caso} \end{cases}$$

$$f_Y(y) = \begin{cases} 1/2 & \text{si } y \in [1,3] \\ 0 & \text{en otro caso} \end{cases}$$

Buscamos calcular en primer lugar $P[X+Y<2]$. Para ello, por ser X e Y independientes, tenemos que:

$$f_{X,Y}(x,y) = f_X(x)f_Y(y) = \begin{cases} 1/2 & \text{si } (x,y) \in [0,1] \times [1,3] \\ 0 & \text{en otro caso} \end{cases}$$

Veamos el conjunto gráficamente, junto con la recta $x+y=2$:



Por tanto, tenemos que:

$$P[X+Y<2] = \int_0^1 \int_1^{2-x} \frac{1}{2} dydx = \frac{1}{2}\int_0^1 (2-x-1)dx = \frac{1}{2}\int_0^1 (1-x)dx = \frac{1}{2}\left[x - \frac{x^2}{2}\right]_0^1 =$$
$$= \frac{1}{2}\left(1-\frac{1}{2}\right) = \frac{1}{4}$$

Por tanto, la probabilidad de obtener un ticket es $p = 1/4$. Sea ahora Z_i la variable aleatoria que toma el valor 1 si se obtiene un ticket en la i-ésima elección y 0 en otro caso. Entonces, $Z_i \sim \mathcal{B}(1,p)$, con $p = 1/4$.

Consideramos ahora $S_n = \sum\limits_{i=1}^{n} Z_i$, que por la propiedad reproductiva de la distribución binomial, sabemos que $S_n \sim \mathcal{B}(n,p)$. De hecho, S_n es el número de tickets obtenidos en las n elecciones. Por tanto, buscamos calcular n tal que:

$$P[S_n \geqslant 10] > 0{,}9495$$

Por ser $S_n \sim \mathcal{B}(n,p)$, tenemos que:

$$E[S_n] = np \qquad \mathrm{Var}[S_n] = np(1-p)$$

Por tanto, buscamos n tal que:

$$\begin{aligned} P[S_n \geqslant 10] > 0{,}9495 &\Longleftrightarrow P\left[\frac{S_n - E[S_n]}{\sqrt{\mathrm{Var}[S_n]}} \geqslant \frac{10-np}{\sqrt{np(1-p)}}\right] > 0{,}9495 \\ &\Longleftrightarrow P\left[\frac{S_n - E[S_n]}{\sqrt{\mathrm{Var}[S_n]}} \geqslant \frac{10-{}^{n}/_{4}}{\sqrt{n}\cdot {}^{\sqrt{3}}/_{4}}\right] > 0{,}9495 \end{aligned}$$

Por el Teorema de Lévy, tenemos que:

$$\left\{\frac{S_n - E[S_n]}{\sqrt{\mathrm{Var}[S_n]}}\right\} \xrightarrow{L} \mathcal{N}(0,1)$$

Por tanto, sea $Z \sim \mathcal{N}(0,1)$, buscamos n tal que:

$$P[Z \geqslant \frac{40-n}{\sqrt{n}\cdot\sqrt{3}}] > 0{,}9495 \Longleftrightarrow P[Z \leqslant -\frac{40-n}{\sqrt{n}\cdot\sqrt{3}}] > 0{,}9495$$

Consultando la tabla de la distribución normal estándar, obtenemos que:

$$-\frac{40-n}{\sqrt{n}\cdot\sqrt{3}} > 1{,}64 \Longleftrightarrow 40-n < -1{,}64\sqrt{n}\cdot\sqrt{3} \Longleftrightarrow n - 2{,}84\sqrt{n} - 40 > 0$$

Resolvemos la igualdad:

$$\sqrt{n} = \frac{2{,}84 \pm \sqrt{2{,}84^2 + 4\cdot 40}}{2} = \frac{2{,}84 \pm 12{,}96}{2} = 7{,}9 \Longrightarrow n = 62{,}41$$

Evaluando, vemos que como solución sirve cualquier $n > 62{,}41$. Como se entiende que se pide el número entero más pequeño, la respuesta es $n = 63$.

Por tanto, debe realizar al menos 63 elecciones para obtener un premio con probabilidad superior a 0,9495.

Ejercicio 4. De un vector aleatorio bidimensional (X,Y) se sabe que sus rectas de regresión son $5y - x + 1 = 0$ y $2x - 5y + 2 = 0$:

1. Identificar la recta de regresión de Y sobre X.

   Supongamos que la recta de regresión de Y sobre X es $y = {}^{x}/_{5} - {}^{1}/_{5}$, luego la recta de regresión de X sobre Y es $x = {}^{5y}/_{2} - 1$. Por tanto, como el coeficiente de determinación es el producto de las pendientes de las rectas de regresión, tenemos que:

   $$\rho^2_{X,Y} = \frac{1}{5}\cdot\frac{5}{2} = \frac{1}{2} < 1$$

   Por tanto, el resultado es coherente y la recta de regresión de Y sobre X es:

   $$y = \frac{x}{5} - \frac{1}{5}$$

   Si hubiésemos hecho la suposición contraria, habríamos obtenido un coeficiente de determinación mayor que 1, lo cual no tiene sentido.

2. Obtener una medida de la proporción de varianza de cada variable que queda explicada por el modelo de regresión lineal.

   Esta medida es el coeficiente de determinación, que en este caso es $\rho^2_{X,Y} = {}^1\!/\!_2$. Por tanto, el 50 % de la varianza de Y queda explicada por el modelo de regresión lineal.

3. Calcular la esperanza del vector bidimensional.

   Sabemos que este es el punto de corte de las rectas de regresión. Por tanto, resolviendo el sistema de ecuaciones:

$$\begin{aligned} 5y - x + 1 &= 0 \\ 2x - 5y + 2 &= 0 \end{aligned}$$

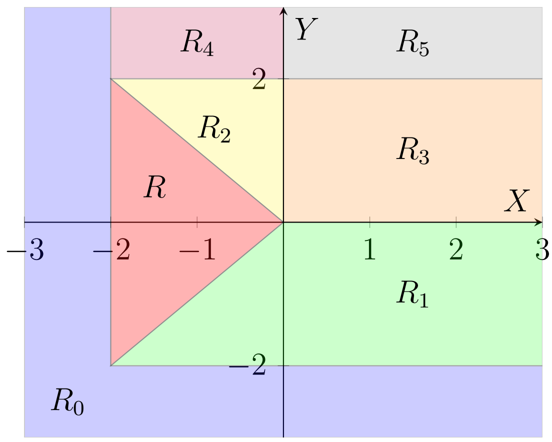
   obtenemos que $x = -3$ e $y = {}^{-4}\!/\!_5$. Por tanto, tenemos que:

$$E[(X,Y)] = (-3, -{}^4\!/\!_5)$$

Ejercicio 5. Dado el vector bidimensional (X,Y) distribuido uniformemente en el recinto limitado $R = \{(x,y) \in \mathbb{R}^2 \mid -2 \leqslant x \leqslant y \leqslant -x\}$:

1. Obtener su función de densidad conjunta.

   Representamos en primer lugar dicho conjunto:



   Tenemos que, para $x \in [-2,0]$ y $y \in [x,-x]$, la función de densidad conjunta es:

$$f_{(X,Y)}(x,y) = k, \qquad k \in \mathbb{R}^+$$

   Para que sea una función de densidad, hemos de tener que:

$$1 = \int_{\mathbb{R}^2} f = \int_R f = \int_R k$$

   Tenemos dos opciones:

Integración Normal:

$$1 = \int_R k = \int_{-2}^{0} \int_{x}^{-x} k \, dy \, dx = \int_{-2}^{0} k(-2x) \, dx =$$
$$= -2k \left[\frac{x^2}{2} \right]_{-2}^{0} = -2k \left[0 - 2 \right] = 4k \Longrightarrow k = \frac{1}{4}$$

Razonando según la Forma de R:

$$1 = \int_R k = k\lambda(R) = k \cdot \frac{4 \cdot 2}{2} = 4k \Longrightarrow k = \frac{1}{4}$$

En cualquier caso, para el valor de $k = 1/4$, la función de densidad conjunta es integrable, no negativa e integra 1.

2. Obtener su función de distribución conjunta. Distinguimos casos:

   - Si $x < -2$ o $y < -2$ (Zona R_0):

   $$F_{(X,Y)}(x, y) \int_{-\infty}^{x} \int_{-\infty}^{y} f_{(X,Y)}(u, v) \, dv \, du = 0$$

   - Si $x \in [-2, 0]$ y $y \in [x, -x]$ (Zona R):

   $$F_{(X,Y)}(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{(X,Y)}(u, v) \, dv \, du = \int_{-2}^{x} \int_{u}^{y} \frac{1}{4} \, dv \, du =$$
   $$= \int_{-2}^{x} \frac{1}{4}(y - u) \, du = \frac{1}{4} \left[yu - \frac{u^2}{2} \right]_{-2}^{x} = \frac{1}{4} \left[yx - \frac{x^2}{2} + 2y + 2 \right]$$

   - Si $y < 0$ y $x > y$ (Zona R_1):

   $$F_{(X,Y)}(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{(X,Y)}(u, v) \, dv \, du = \int_{-2}^{y} \int_{u}^{y} \frac{1}{4} \, dv \, du =$$
   $$= \int_{-2}^{y} \frac{1}{4}(y - u) \, du = \frac{1}{4} \left[yu - \frac{u^2}{2} \right]_{-2}^{y} = \frac{1}{4} \left[y^2 - \frac{y^2}{2} + 2y + 2 \right] =$$
   $$= \frac{1}{4} \left[\frac{y^2}{2} + 2y + 2 \right]$$

   - Si $x \in [-2, 0]$ y $y \in [-x, 2]$ (Zona R_2):

   $$F_{(X,Y)}(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{(X,Y)}(u, v) \, dv \, du = \int_{-2}^{-y} \int_{u}^{y} \frac{1}{4} \, dv \, du + \int_{-y}^{x} \int_{u}^{-u} \frac{1}{4} \, dv \, du =$$
   $$= \int_{-2}^{-y} \frac{1}{4}(y - u) \, du + \int_{-y}^{x} \frac{1}{4}(-2u) \, du = \frac{1}{4} \left[yu - \frac{u^2}{2} \right]_{-2}^{-y} + \frac{1}{4} \left[-u^2 \right]_{-y}^{x} =$$
   $$= \frac{1}{4} \left[\cancel{y(-y)} - \frac{(-y)^2}{2} + 2y + 2 \right] + \frac{1}{4} \left[-x^2 + \cancel{y^2} \right] =$$
   $$= \frac{1}{4} \left[2y + 2 - \frac{y^2}{2} - x^2 \right]$$

- Si $y \in [0,2]$ y $x \geqslant 0$ (Zona R_3):

$$F_{(X,Y)}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{(X,Y)}(u,v)\,dv\,du = \int_{-2}^{-y}\int_{u}^{y}\frac{1}{4}\,dv\,du + \int_{-y}^{0}\int_{u}^{-u}\frac{1}{4}\,dv\,du =$$
$$= \int_{-2}^{-y}\frac{1}{4}(y-u)\,du + \int_{-y}^{0}\frac{1}{4}(-2u)\,du = \frac{1}{4}\left[yu - \frac{u^2}{2}\right]_{-2}^{-y} + \frac{1}{4}\left[-u^2\right]_{-y}^{0} =$$
$$= \frac{1}{4}\left[\cancel{y(-y)} - \frac{(-y)^2}{2} + 2y + 2\right] + \frac{1}{4}\left[0 - \cancel{y^2}\right] =$$
$$= \frac{1}{4}\left[2y + 2 - \frac{y^2}{2}\right]$$

- Si $x \in [-2,0]$ y $y \geqslant 2$ (Zona R_4):

$$F_{(X,Y)}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{(X,Y)}(u,v)\,dv\,du = \int_{-2}^{x}\int_{u}^{-u}\frac{1}{4}\,dv\,du =$$
$$= \int_{-2}^{x}\frac{1}{4}(-2u)\,du = \frac{1}{4}\left[-u^2\right]_{-2}^{x} = \frac{1}{4}\left[-x^2+4\right]$$

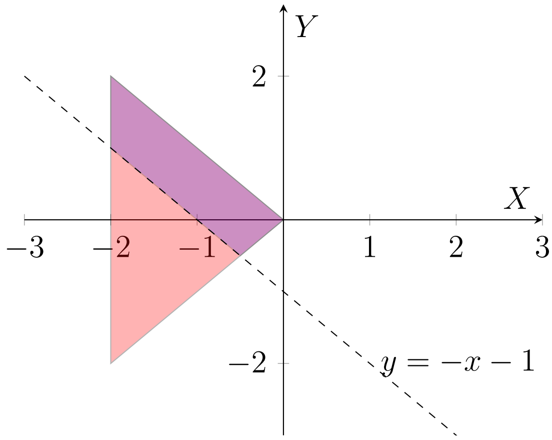
- Si $x \geqslant 0, y \geqslant 2$ (Zona R_5):

$$F_{(X,Y)}(x,y) = 1$$

Por tanto, la función de distribución conjunta es:

$$F_{(X,Y)}(x,y) = \begin{cases} 0 & \text{si } x < -2 \text{ o } y < -2 \\ \frac{1}{4}\left[yx - \frac{x^2}{2} + 2y + 2\right] & \text{si } x \in [-2,0] \text{ y } y \in [x,-x] \\ \frac{1}{4}\left[\frac{y^2}{2} + 2y + 2\right] & \text{si } y < 0 \text{ y } x > y \\ \frac{1}{4}\left[2y + 2 - \frac{y^2}{2} - x^2\right] & \text{si } x \in [-2,0] \text{ y } y \in [-x,2] \\ \frac{1}{4}\left[2y + 2 - \frac{y^2}{2}\right] & \text{si } y \in [0,2] \text{ y } x \geqslant 0 \\ \frac{1}{4}\left[-x^2+4\right] & \text{si } x \in [-2,0] \text{ y } y \geqslant 2 \\ 1 & \text{si } x \geqslant 0, y \geqslant 2 \end{cases}$$

3. Obtener la probabilidad de que $X+Y+1 \geqslant 0$. Veamos qué conjunto representa $X+Y+1 \geqslant 0$:



Tenemos por tanto que:

$$\begin{aligned}
P[X+Y+1 \geqslant 0] &= \int_{-2}^{-1/2} \int_{-x-1}^{-x} \frac{1}{4}\, dy\, dx + \int_{-1/2}^{0} \int_{x}^{-x} \frac{1}{4}\, dy\, dx = \\
&= \int_{-2}^{-1/2} \frac{1}{4}(-x+1+x)\, dx + \int_{-1/2}^{0} \frac{1}{4}(-2x)\, dx = \\
&= \int_{-2}^{-1/2} \frac{1}{4}\, dx + \int_{-1/2}^{0} -\frac{1}{2} x\, dx = \\
&= \frac{1}{4}\left[x\right]_{-2}^{-1/2} - \frac{1}{2}\left[\frac{x^2}{2}\right]_{-1/2}^{0} = \frac{1}{4}\left[\frac{-1}{2}+2\right] - \frac{1}{2}\left[0-\frac{1}{8}\right] = \frac{7}{16}
\end{aligned}$$

Ejercicio 6. Dado el vector bidimensional (X, Y) con la siguiente función masa de probabilidad conjunta:

$X \backslash Y$	0	1	2
1	$1/4$	0	0
2	0	$1/4$	0
3	$1/4$	0	$1/4$

1. Obtener la mejor aproximación minimo-cuadrática a la variable Y conocidos valores de la variable X, así como calcular una medida de la bondad del ajuste.

   Completamos en primer lugar la función masa de probabilidad conjunta calculando las marginales:

$X \backslash Y$	0	1	2	
1	$1/4$	0	0	$1/4$
2	0	$1/4$	0	$1/4$
3	$1/4$	0	$1/4$	$1/2$
	$1/2$	$1/4$	$1/4$	

   Mostramos ahora en la siguiente tabla la distribución condicionada a X, es decir, $P[Y|X = x]$:

$X\backslash Y$	0	1	2
1	1	0	0
2	0	1	0
3	$1/2$	0	$1/2$

Por tanto, tenemos que:

$$\begin{aligned} E[Y|X=1] &= 0\cdot 1 + 1\cdot 0 + 2\cdot 0 = 0 \\ E[Y|X=2] &= 0\cdot 0 + 1\cdot 1 + 2\cdot 0 = 1 \\ E[Y|X=3] &= 0\cdot 1/2 + 1\cdot 0 + 2\cdot 1/2 = 1 \end{aligned}$$

Por tanto, la mejor aproximación minimo-cuadrática a la variable Y conocidos valores de la variable X es:

$$\hat{Y} = \begin{cases} 0 & \text{si } X = 1 \\ 1 & \text{si } X = 2,3 \end{cases}$$

Calculamos ahora una medida de la bondad del ajuste. Calculamos previamente:

$$\begin{aligned} E[Y] &= 0\cdot 1/2 + 1\cdot 1/4 + 2\cdot 1/4 = 3/4 \\ E[Y^2] &= 0^2\cdot 1/2 + 1^2\cdot 1/4 + 2^2\cdot 1/4 = 5/4 \\ \text{Var}[Y] &= E[Y^2] - E[Y]^2 = 5/4 - 9/16 = 11/16 \\ E[(E[Y|X])^2] &= \sum_{x\in\{1,2,3\}} E[Y|X=x]^2 P[X=x] = 0^2\cdot 1/4 + 1^2\cdot 1/4 + 1^2\cdot 1/2 = 3/4 \\ E[E[Y|X]] &= E[Y] = 3/4 \\ \text{Var}[E[Y|X]] &= E[(E[Y|X])^2] - E[E[Y|X]]^2 = 3/4 - 9/16 = 3/16 \end{aligned}$$

Por tanto, la medida de la bondad del ajuste es:

$$\eta^2_{Y/X} = \frac{\text{Var}[E[Y|X]]}{\text{Var}[Y]} = \frac{3/16}{11/16} = \frac{3}{11}$$

2. Obtener la aproximación lineal minimo-cuadrática de la variable Y para $X = 1$.

   Podríamos pensar que se trata de la mejor aproximación minimo-cuadrática de la variable Y para $X=1$, que sería $E[Y\mid X=1]=0$. Sin embargo, esta no es la aproximación lineal. La aproximación lineal minimo-cuadrática de la variable Y sobre X es:

$$Y - E[Y] = \frac{\text{Cov}[X,Y]}{\text{Var}[X]}(X - E[X])$$

   Calculamos por tanto dichos valores:

$$\begin{aligned} E[X] &= 1\cdot 1/4 + 2\cdot 1/4 + 3\cdot 1/2 = 9/4 \\ E[X^2] &= 1^2\cdot 1/4 + 2^2\cdot 1/4 + 3^2\cdot 1/2 = 23/4 \\ \text{Var}[X] &= E[X^2] - E[X]^2 = 23/4 - 81/16 = 11/16 \\ E[XY] &= 2\cdot 1\cdot 1/4 + 3\cdot 2\cdot 1/4 = 2 \\ \text{Cov}[X,Y] &= E[XY] - E[X]E[Y] = 2 - 9/4\cdot 3/4 = 5/16 \end{aligned}$$

Por tanto, la aproximación lineal minimo-cuadrática de la variable Y es:

$$Y - \frac{3}{4} = \frac{5/16}{11/16}\left(X - \frac{9}{4}\right) \Longleftrightarrow Y = \frac{5}{11}X - \frac{3}{11}$$

Por tanto, evaluando en $X = 1$, obtenemos que la aproximación lineal minimo-cuadrática de la variable Y para $X = 1$ es

$$Y = \frac{2}{11}$$

3. Obtener el error cuadrático medio para la aproximación del primer apartado.

   Tenemos que:

   $$\text{E.C.M.}(E[Y \mid X]) = E[\text{Var}[Y \mid X]] = \text{Var}[Y] - \text{Var}[E[Y \mid X]] = \frac{11}{16} - \frac{3}{16} = \frac{8}{16} = \frac{1}{2}$$