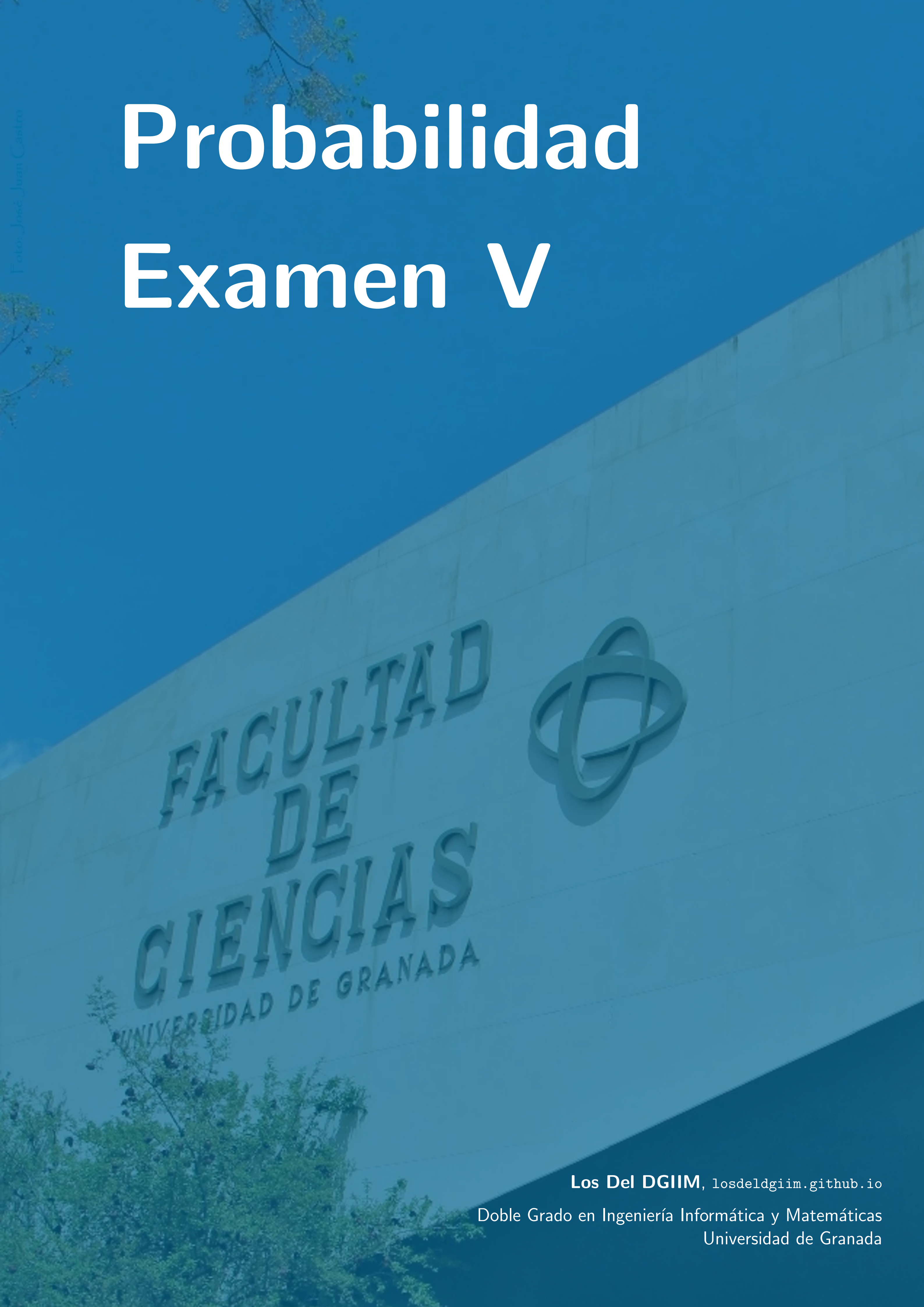
Probabilidad

Examen V



Los Del DGIIM, `losdeldgiim.github.io`

Doble Grado en Ingeniería Informática y Matemáticas
Universidad de Granada

Probabilidad
Examen V

Los Del DGIIM, `losdeldgiim.github.io`

Arturo Olivares Martos
José Juan Urrutia Milán

Granada, 2024-2025

Asignatura Probabilidad.

Curso Académico 2021-22.

Grado Doble Grado en Ingeniería Informática y Matemáticas.

Grupo Único.

Descripción Examen Ordinario

Fecha 21 de enero de 2022.

PARTE 1 (2.5 puntos)

Ejercicio 1 (0.25 puntos)**.** Sean X_1, X_2, X_3 variables aleatorias independientes e idénticamente distribuidas según una ley Binomial, $B(3, 1/2)$. Justificar que:

$$P[X_1 + X_2 + X_3 = 8] = \frac{9}{2^9}$$

Tenemos que:

$$X_1, X_2, X_3 \sim B(3, 1/2)$$

Por la reproductividad de la distribución binomial, como son independientes, tenemos que:

$$X_1 + X_2 + X_3 \sim B(9, 1/2)$$

Por tanto, usando la función masa de probabilidad de la distribución binomial, tenemos que:

$$P[X_1 + X_2 + X_3 = 8] = \binom{9}{8}\left(\frac{1}{2}\right)^8\left(\frac{1}{2}\right)^1 = \frac{9}{2^9}$$

Ejercicio 2 (0.25 puntos)**.** Sean X_1 y X_2 variables aleatorias independientes e idénticamente distribuidas según una ley de Poisson, $\mathcal{P}(3)$. Justificar que:

$$P[X_1 + X_2 > 0] = \frac{e^6 - 1}{e^6}$$

Tenemos que:

$$X_1, X_2 \sim \mathcal{P}(3)$$

Por la reproductividad de la distribución de Poisson, por ser independientes tenemos que:

$$X_1 + X_2 \sim \mathcal{P}(6)$$

Por tanto, tenemos que:

$$P[X_1 + X_2 > 0] = 1 - P[X_1 + X_2 = 0] = 1 - \frac{e^{-6}6^0}{0!} = 1 - e^{-6} = \frac{e^6 - 1}{e^6}$$

Ejercicio 3. Para predecir los valores de una variable aleatoria X a partir de los de otra variable aleatoria Y se considera un modelo lineal:

1. **(0.50 puntos)** Obtener de forma razonada los coeficientes del modelo lineal considerado.

   Se busca aproximar X como $\widehat{X} = aY + b$. Para ello, se minimiza el error cuadrático medio:

$$\begin{aligned}
\text{E.C.M.}(X \mid Y) &= E[(X - \widehat{X})^2] = E\left[(X - aY - b)^2\right] = \\
&= E\left[X^2 - 2aXY - 2bX + a^2Y^2 + 2abY + b^2\right] = \\
&= E[X^2] - 2aE[XY] - 2bE[X] + a^2E[Y^2] + 2abE[Y] + b^2
\end{aligned}$$

   Para ello, se busca minizar la siguiente función:

$$L(a, b) = E.C.M.(X \mid Y) = E[X^2] - 2aE[XY] - 2bE[X] + a^2E[Y^2] + 2abE[Y] + b^2$$

Se tiene demostrado en Teoría que llegamos a la siguiente expresión:

$$\begin{cases} a = \dfrac{\mathrm{Cov}[X,Y]}{\mathrm{Var}[Y]} \\ b = E[X] - aE[Y] \end{cases}$$

2. **(0.75 puntos)** Si $x - y = 1$ y $2y - 3x = -1$ son las dos rectas de regresión para el vector (X, Y), se pide: identificar la recta de regresión del apartado anterior; obtener una medida de la bondad del ajuste y calcular la esperanza del vector (X, Y).

   Suponemos que las rectas de regresión de X sobre Y y de Y sobre X son $x - y = 1$ y $2y - 3x = -1$, respectivamente. Por tanto, tenemos que:

$$\begin{aligned} x = y + 1 &= \frac{\mathrm{Cov}[X,Y]}{\mathrm{Var}[Y]} \cdot y + E[X] - \frac{\mathrm{Cov}[X,Y]}{\mathrm{Var}[Y]} \cdot E[Y] \\ y = \frac{3}{2}x - \frac{1}{2} &= \frac{\mathrm{Cov}[X,Y]}{\mathrm{Var}[X]} \cdot x + E[Y] - \frac{\mathrm{Cov}[X,Y]}{\mathrm{Var}[X]} \cdot E[X] \end{aligned}$$

   Identificando términos, obtenemos que:

$$\frac{\mathrm{Cov}[X,Y]}{\mathrm{Var}[X]} \cdot \frac{\mathrm{Cov}[X,Y]}{\mathrm{Var}[Y]} = \rho^2_{X,Y} = 1 \cdot \frac{3}{2} = \frac{3}{2} > 1$$

   Por tanto, llegamos a una contradicción, por lo que la suposición es incorrecta. La recta de regresión de Y sobre X es $y = x - 1$ y la de X sobre Y es $x = {}^2/_3 y + {}^1/_3$.

   La proporción de varianza de cada variable que queda explicada por el modelo de regresión lineal es el coeficiente de determinación, que en este caso es:

$$\rho^2_{X,Y} = \frac{2}{3} \cdot 1 = \frac{2}{3} \approx 0{,}667\,\%$$

   Por último, por identificación de términos, tenemos el siguiente sistema:

$$\left\{ \begin{array}{rcl} E[X] - E[Y] & = & 1 \\ E[Y] - \dfrac{3}{2}E[X] & = & -\dfrac{1}{2} \end{array} \right\} \Longrightarrow \begin{cases} E[X] = -1 \\ E[Y] = -2 \end{cases}$$

   Por tanto, tenemos que:

$$E[(X,Y)] = \begin{pmatrix} -1 & -2 \end{pmatrix}$$

Ejercicio 4 (0.75 puntos)**.** Las componentes de un vector aleatorio continuo son variables aleatorias continuas. Sin embargo, en general, un conjunto de variables aleatorias continuas no da lugar a un vector aleatorio continuo. Justificar que este recíproco sí es cierto si consideramos un conjunto X_1, X_2, $\dots$, X_n de variables aleatorias continuas independientes.

Definimos la siguiente función continua en $\mathbb{R}^n$:

$$f_{(X_1,X_2,\ldots,X_n)}: \begin{array}{rcl} \mathbb{R} & \longrightarrow & \mathbb{R} \\ (x_1, x_2, \ldots, x_n) & \longmapsto & \prod\limits_{i=1}^{n} f_{X_i}(x_i) \end{array}$$

y probaremos que es la función de densidad de $(X_1, X_2, \ldots, X_n)$. En primer lugar, está bien definida por estarlo cada una de las funciones de densidad de las variables aleatorias. Veamos que:

$$F_{(X_1,X_2,\ldots,X_n)}(x_1, x_2, \ldots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} \prod_{i=1}^{n} f_{X_i}(t_i)\, dt_1 \cdots dt_n$$

Para ello, como son independientes, tenemos que:

$$P[X_1 \leqslant x_1, X_2 \leqslant x_2, \ldots, X_n \leqslant x_n] = \prod_{i=1}^{n} P[X_i \leqslant x_i]$$

Por ser f_{X_i} la función de densidad de X_i, tenemos que:

$$F_{(X_1,X_2,\ldots,X_n)}(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} \int_{-\infty}^{x_i} f_{X_i}(t_i)\, dt_i$$

Por tanto, tenemos que la función que hemos definido efectivamente es la función de densidad de $(X_1, X_2, \ldots, X_n)$.

Aunque no sería necesario, veamos que cumple las condiciones de toda función de densidad. En primer lugar, es no negativa, ya que cada término es mayor o igual que 0. Veamos ahora que integra 1:

$$\begin{aligned} \int_{\mathbb{R}^n} f_{(X_1,X_2,\ldots,X_n)}(x_1, x_2, \ldots, x_n)\, dx_1 \cdots dx_n &= \int_{\mathbb{R}^n} \prod_{i=1}^{n} f_{X_i}(x_i)\, dx_1 \cdots dx_n = \\ &= \prod_{i=1}^{n} \int_{\mathbb{R}} f_{X_i}(x_i)\, dx_i = \prod_{i=1}^{n} 1 = 1 \end{aligned}$$

PARTE 2 (7.5 puntos)

Ejercicio 1 (5 puntos)**.** Dado el vector aleatorio continuo (X, Y) distribuido uniformemente en el recinto

$$C = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1 \ \wedge \ x, y < 0\}$$

Observación. A tener en cuenta:

- En el **apartado 1.2** se obtiene **hasta 1 punto** si las integrales se dejan indicadas y **hasta 1.5 puntos** si se obtienen sus primitivas de forma explícita.
- Si necesitara obtener la primitiva de la función $f(x) = \sqrt{1 - x^2}$, realizar el cambio de variable unidimensional $x = \text{sen}(t)$.

- $\operatorname{arc\,sen}(0) = 0$, $\operatorname{arc\,sen}(-1) = -\frac{\pi}{2}$, $\operatorname{arc\,sen}\left(-\frac{1}{\sqrt{2}}\right) = -\frac{\pi}{4}$.
- $\cos^2(x) = \frac{1+\cos(2x)}{2}$, $\operatorname{sen}^2(x) = \frac{1-\cos(2x)}{2}$.

1. **(0.25 puntos)** Obtener la función de densidad conjunta.
2. **(1.50 puntos)** Obtener la función de distribuición de probabilidad conjunta.
3. **(0.75 puntos)** Obtener las funciones de densidad condicionadas.
4. **(0.50 puntos)** Obtener la probabilidad de que $X - Y > 0$.
5. **(1.50 puntos)** Obtener la mejor aproximaión mínimo cuadrática a la variable aleatoria Y conocidos los valores de la variable X y el error cuadrático medio de esta aproximación.
6. **(0.50 puntos)** Obtener una media de la bondad del ajuste del apartado anterior.

Ejercicio 2 (2.5 puntos)**.** Dado el vector aleatorio (X, Y) con función generatriz de momentos

$$M_{(X,Y)}(t_1, t_2) = exp\left(\frac{2t_1 + 4t_1^2 + 9t_2^2 + 6t_1t_2}{2}\right)$$

1. **(0.75 puntos)** Obtener la razón de correlación y el coeficiente de correlación lineal de las variables (X, Y).
2. **(0.75 puntos)** Indicar las distribuciones de las variables aleatorias $Y/X = 1$ y $X/Y = 0$.
3. **(1 punto)** Obtener la distribución de probabilidad del vector aleatorio $(2X, Y-X)$. Justificar que las variabes aleatorias $2X$ y $Y - X$ tienen cierta asociación lineal en sentido negativo.