

Probabilidad

Examen V

FACULTAD
DE
CIENCIAS
UNIVERSIDAD DE GRANADA



Los Del DGIIM, losdeldgiim.github.io

Doble Grado en Ingeniería Informática y Matemáticas
Universidad de Granada



Esta obra está bajo una Licencia Creative Commons Atribución-NoComercial-SinDerivadas 4.0 Internacional (CC BY-NC-ND 4.0).

Eres libre de compartir y redistribuir el contenido de esta obra en cualquier medio o formato, siempre y cuando des el crédito adecuado a los autores originales y no persigas fines comerciales.

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 $\hat{X} = aY + b$. Para ello, se minimiza el error cuadrático medio:

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

Para ello, se busca minizar la siguiente función:

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

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

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

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:

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

Identificando términos, obtenemos que:

$$\frac{\text{Cov}[X, Y]}{\text{Var}[X]} \cdot \frac{\text{Cov}[X, Y]}{\text{Var}[Y]} = \rho_{X,Y}^2 = 1 \cdot \frac{3}{2} = \frac{3}{2} > 1$$

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 $x = \frac{2}{3}y + \frac{1}{3}$.

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:

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

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

$$\begin{cases} E[X] - E[Y] = 1 \\ E[Y] - \frac{3}{2}E[X] = -\frac{1}{2} \end{cases} \implies \begin{cases} E[X] = -1 \\ E[Y] = -2 \end{cases}$$

Por tanto, tenemos que:

$$E[(X, Y)] = (-1 \quad -2)$$

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 :

$$\begin{aligned} f_{(X_1, X_2, \dots, X_n)} : \quad \mathbb{R} &\longrightarrow \mathbb{R} \\ (x_1, x_2, \dots, x_n) &\longmapsto \prod_{i=1}^n f_{X_i}(x_i) \end{aligned}$$

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

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

Para ello, como son independientes, tenemos que:

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

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

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

Por tanto, tenemos que la función que hemos definido efectivamente es la función de densidad de (X_1, X_2, \dots, 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:

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

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 $x = \sin(t)$.

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

1. **(0.25 puntos)** Obtener la función de densidad conjunta.

En primer lugar, y debido a que lo usaremos con mucha frecuencia, resolveremos la siguiente integral de forma genérica. Dados $a, b \in [0, 1]$, $a < b$, resolveremos la siguiente integral:

$$\int_a^b \sqrt{1-x^2} dx$$

Haciendo el cambio de variable $x = \sen(t)$, tenemos que:

$$\begin{aligned} \int_a^b \sqrt{1-x^2} dx &= \int_{\arcsen(a)}^{\arcsen(b)} \cos(t) \cos(t) dt = \int_{\arcsen(a)}^{\arcsen(b)} \cos^2(t) dt = \quad (1) \\ &= \int_{\arcsen(a)}^{\arcsen(b)} \frac{1+\cos(2t)}{2} dt = \frac{1}{2} \int_{\arcsen(a)}^{\arcsen(b)} 1+\cos(2t) dt = \\ &= \frac{1}{2} \left[t + \frac{\sen(2t)}{2} \right]_{\arcsen(a)}^{\arcsen(b)} = \\ &= \frac{1}{2} \left[\arcsen(b) - \arcsen(a) + \frac{\sen(2 \arcsen(b))}{2} - \frac{\sen(2 \arcsen(a))}{2} \right] = \\ &= \frac{1}{2} \left[\arcsen(b) - \arcsen(a) + \frac{2b\sqrt{1-b^2}}{2} - \frac{2a\sqrt{1-a^2}}{2} \right] \stackrel{(*)}{=} \\ &\stackrel{(*)}{=} \frac{1}{2} \left[\arcsen(b) - \arcsen(a) + b\sqrt{1-b^2} - a\sqrt{1-a^2} \right] \end{aligned}$$

donde en (*) hemos empleado, en primer lugar, el seno del ángulo doble, por lo que:

$$\sen(2 \arcsen(x)) = 2 \sen(\arcsen(x)) \cos(\arcsen(x)) = 2x\sqrt{1-x^2}$$

y ahí hemos empleado el valor de $\cos(\arcsen(x))$. Como $\arcsen(x) \in [-\pi/2, \pi/2]$, su coseno es positivo. Por tanto:

$$\cos(\arcsen(x)) = \sqrt{1-\sen^2(\arcsen(x))} = \sqrt{1-x^2}$$

Una vez hecha esta integral (que usaremos en varios casos), procedemos con el ejercicio. Veamos en primer lugar la forma de C , que se muestra en la Figura 1.

La función de densidad conjunta es constante, por lo que:

$$f(x, y) = \begin{cases} k, & (x, y) \in C \\ 0, & \text{en otro caso.} \end{cases}$$

Para que f sea una función de densidad, tenemos que:

$$1 = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy$$

Hay dos opciones:

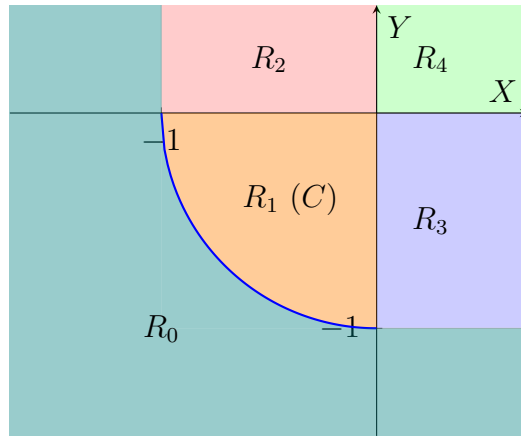


Figura 1: Recinto C .

Integrando de la forma usual: Es necesario que:

$$1 = \int_{-1}^0 \int_{-\sqrt{1-x^2}}^0 k \, dy \, dx = k \int_{-1}^0 \sqrt{1-x^2} \, dx \stackrel{(1)}{=} \frac{k}{2} \left[0 + \frac{\pi}{2} \right] = \frac{k\pi}{4} \implies k = \frac{4}{\pi}.$$

Razonando la forma de C : Sabemos que C es un cuarto de círculo de radio 1, por lo que su área es $\pi/4$. Por tanto, tenemos que:

$$1 = \int_C f(x, y) = k \int_C 1 = k \cdot \lambda(C) = k \cdot \frac{\pi}{4} \implies k = \frac{4}{\pi}.$$

2. **(1.50 puntos)** Obtener la función de distribución de probabilidad conjunta.

Distinguiamos en función de los valores de (x, y) :

- Si $(x, y) \in R_0$: $(x < -1 \text{ o } y < -1 \text{ o } -1 < x, y < 0, x^2 + y^2 \geq 1)$

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

- Si $(x, y) \in R_1$: $(-1 < x, y < 0 \text{ y } x^2 + y^2 < 1)$

$$\begin{aligned} F_{(X,Y)}(x, y) &= \int_{-\sqrt{1-y^2}}^x \int_{-\sqrt{1-u^2}}^y \frac{4}{\pi} \, dv \, du = \frac{4}{\pi} \int_{-\sqrt{1-y^2}}^x y + \sqrt{1-u^2} \, du \stackrel{(1)}{=} \\ &\stackrel{(1)}{=} \frac{4}{\pi} \left[[yu]_{-\sqrt{1-y^2}}^x + \frac{1}{2} \left[\arcsen(x) - \arcsen(-\sqrt{1-y^2}) + \right. \right. \\ &\quad \left. \left. + x\sqrt{1-x^2} + \sqrt{1-y^2}\sqrt{1-1+y^2} \right] \right] = \\ &= \frac{4}{\pi} \left[yx + y\sqrt{1-y^2} + \frac{1}{2} \left[\arcsen(x) - \arcsen(-\sqrt{1-y^2}) + \right. \right. \\ &\quad \left. \left. + x\sqrt{1-x^2} - y\sqrt{1-y^2} \right] \right] = \\ &= \frac{4}{\pi} \left[yx + \frac{y\sqrt{1-y^2}}{2} + \frac{1}{2} \left[\arcsen(x) - \arcsen(-\sqrt{1-y^2}) + x\sqrt{1-x^2} \right] \right] = \\ &= \frac{2}{\pi} \left[2xy + y\sqrt{1-y^2} + \arcsen(x) + \arcsen(\sqrt{1-y^2}) + x\sqrt{1-x^2} \right] \end{aligned}$$

- Si $(x, y) \in R_2$: $(-1 < x < 0 \text{ y } 0 < y)$

$$\begin{aligned} F_{(X,Y)}(x, y) &= \int_{-1}^x \int_{-\sqrt{1-u^2}}^0 \frac{4}{\pi} dv du = \frac{4}{\pi} \int_{-1}^x \sqrt{1-u^2} du \stackrel{(1)}{=} \\ &\stackrel{(1)}{=} \frac{2}{\pi} \left[\arcsen(x) + \frac{\pi}{2} + x\sqrt{1-x^2} \right] \end{aligned}$$

- Si $(x, y) \in R_3$: $(0 < x \text{ y } -1 < y < 0)$

$$\begin{aligned} F_{(X,Y)}(x, y) &= \int_{-\sqrt{1-y^2}}^0 \int_{-\sqrt{1-u^2}}^y \frac{4}{\pi} dv du = \frac{4}{\pi} \int_{-\sqrt{1-y^2}}^0 y + \sqrt{1-u^2} du \stackrel{(1)}{=} \\ &\stackrel{(1)}{=} \frac{4}{\pi} \left[[yu]_{-\sqrt{1-y^2}}^0 + \frac{1}{2} \left[\arcsen(0) - \arcsen\left(-\sqrt{1-y^2}\right) + \right. \right. \\ &\quad \left. \left. + 0\sqrt{1-0^2} + \sqrt{1-y^2}\sqrt{1-1+y^2} \right] \right] = \\ &= \frac{4}{\pi} \left[y\sqrt{1-y^2} + \frac{1}{2} \left[\arcsen\left(\sqrt{1-y^2}\right) - y\sqrt{1-y^2} \right] \right] = \\ &= \frac{2}{\pi} \left[y\sqrt{1-y^2} + \arcsen\left(\sqrt{1-y^2}\right) \right] \end{aligned}$$

- Si $(x, y) \in R_4$: $(0 < x \text{ y } 0 < y)$

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

Por tanto, la función de distribución conjunta es:

$$F_{(X,Y)}(x, y) = \begin{cases} 0, & (x, y) \in R_0 \\ \frac{2}{\pi} \left[2xy + y\sqrt{1-y^2} + \arcsen(x) + \arcsen\left(\sqrt{1-y^2}\right) + x\sqrt{1-x^2} \right], & (x, y) \in R_1 \\ \frac{2}{\pi} \left[\arcsen(x) + \frac{\pi}{2} + x\sqrt{1-x^2} \right], & (x, y) \in R_2 \\ \frac{2}{\pi} \left[y\sqrt{1-y^2} + \arcsen\left(\sqrt{1-y^2}\right) \right], & (x, y) \in R_3 \\ 1, & (x, y) \in R_4 \end{cases}$$

3. (0.75 puntos) Obtener las funciones de densidad condicionadas.

Para ello, calculamos en primer lugar las marginales. Tenemos que:

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{+\infty} f_{X,Y}(x, y) dy = \int_{-\sqrt{1-x^2}}^0 \frac{4}{\pi} dy = \frac{4}{\pi} \cdot \sqrt{1-x^2} \quad \forall x \in [-1, 0] \\ f_Y(y) &= \int_{-\infty}^{+\infty} f_{X,Y}(x, y) dx = \int_{-\sqrt{1-y^2}}^0 \frac{4}{\pi} dx = \frac{4}{\pi} \cdot \sqrt{1-y^2} \quad \forall y \in [-1, 0] \end{aligned}$$

Por tanto, dado $y^* \in [-1, 0]$, tenemos que:

$$f_{X|Y=y^*}(x) = \frac{f_{X,Y}(x, y^*)}{f_Y(y^*)} = \frac{\frac{4}{\pi}}{\frac{4}{\pi} \cdot \sqrt{1-y^{*2}}} = \frac{1}{\sqrt{1-y^{*2}}} \quad \forall x \in [-\sqrt{1-y^{*2}}, 0]$$

De forma análoga, dado $x^* \in [-1, 0]$, tenemos que:

$$f_{Y|X=x^*}(y) = \frac{f_{X,Y}(x^*, y)}{f_X(x^*)} = \frac{\frac{4}{\pi}}{\frac{4}{\pi} \cdot \sqrt{1-x^{*2}}} = \frac{1}{\sqrt{1-x^{*2}}} \quad \forall y \in [-\sqrt{1-x^{*2}}, 0]$$

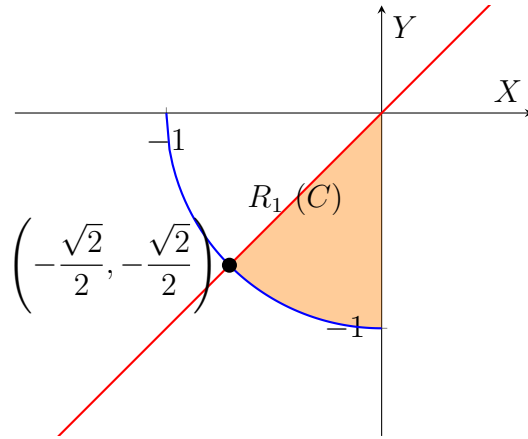


Figura 2: Conjunto en el que se intersecan C y $X - Y > 0$.

4. **(0.50 puntos)** Obtener la probabilidad de que $X - Y > 0$.

Veamos gráficamente este conjunto en la Figura 2, en el que el punto de corte es:

$$-\sqrt{1-x^2} = x \implies 1-x^2 = x^2 \implies x = -\frac{\sqrt{2}}{2} \implies \left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right)$$

Por tanto, tenemos que:

$$\begin{aligned} P[X - Y > 0] &= \int_{-\sqrt{2}/2}^0 \int_{-\sqrt{1-x^2}}^x \frac{4}{\pi} dy dx = \frac{4}{\pi} \int_{-\sqrt{2}/2}^0 x + \sqrt{1-x^2} dx = \\ &= \frac{4}{\pi} \left[\left[\frac{x^2}{2} \right]_{-\sqrt{2}/2}^0 + \frac{1}{2} \left[\arcsin(0) - \arcsin\left(-\frac{\sqrt{2}}{2}\right) + 0\sqrt{1-0^2} + \frac{\sqrt{2}}{2}\sqrt{1-2/4} \right] \right] = \\ &= \frac{4}{\pi} \left[-\frac{1}{4} + \frac{1}{2} \left[\arcsin\left(\frac{\sqrt{2}}{2}\right) + \frac{\sqrt{2}}{2}\sqrt{1-2/4} \right] \right] = \frac{4}{\pi} \left[-\frac{1}{4} + \frac{1}{2} \left[\frac{\pi}{4} + \frac{1}{2} \right] \right] = \\ &= -\frac{1}{\pi} + \frac{1}{2} + \frac{1}{\pi} = \frac{1}{2} \end{aligned}$$

5. **(1.50 puntos)** Obtener la mejor aproximación mínimo cuadrática a la variable aleatoria Y conocidos los valores de la variable X y el error cuadrático medio de esta aproximación.

La mejor aproximación por mínimos cuadrados es la curva de regresión de Y sobre X :

$$\begin{aligned} E[Y | X = x] &= \int_{-\infty}^{+\infty} y \cdot f_{Y|X=x}(y) dy = \int_{-\sqrt{1-x^2}}^0 y \cdot \frac{1}{\sqrt{1-x^2}} dy = \frac{1}{\sqrt{1-x^2}} \left[\frac{y^2}{2} \right]_{-\sqrt{1-x^2}}^0 = \\ &= \frac{1}{\sqrt{1-x^2}} \left[0 - \frac{1-x^2}{2} \right] = \frac{1}{\sqrt{1-x^2}} \left[-\frac{1-x^2}{2} \right] = -\frac{\sqrt{1-x^2}}{2} \quad \forall x \in [-1, 0] \end{aligned}$$

Por tanto, la mejor aproximación por mínimos cuadrados es:

$$E[Y | X] = -\frac{\sqrt{1-X^2}}{2}$$

El error cuadrático medio de esta aproximación es:

$$\text{E.C.M.}(E[Y | X]) = E[\text{Var}[Y | X]] = E[Y^2] - E[E^2[Y | X]]$$

Tenemos que:

$$E[Y^2] = \int_{-\infty}^{+\infty} y^2 \cdot f_Y(y) dy = \int_{-1}^0 y^2 \cdot \frac{4}{\pi} \cdot \sqrt{1-y^2} dy = \frac{4}{\pi} \int_{-1}^0 y^2 \cdot \sqrt{1-y^2} dy$$

Aplicamos ahora el cambio de variable $y = \sin(t)$, por lo que:

$$\begin{aligned} E[Y^2] &= \frac{4}{\pi} \int_{-\pi/2}^0 \sin^2(t) \cdot \sqrt{1-\sin^2(t)} \cdot \cos(t) dt = \frac{4}{\pi} \int_{-\pi/2}^0 \sin^2(t) \cdot \cos^2(t) dt = \\ &= \frac{4}{\pi} \int_{-\pi/2}^0 \frac{1-\cos(2t)}{2} \cdot \frac{1+\cos(2t)}{2} dt = \frac{1}{\pi} \int_{-\pi/2}^0 1 - \cos^2(2t) dt = \\ &= \frac{1}{\pi} \int_{-\pi/2}^0 1 - \frac{1+\cos(4t)}{2} dt = \frac{1}{\pi} \left[t - \frac{t}{2} - \frac{\sin(4t)}{8} \right]_{-\pi/2}^0 = \\ &= \frac{1}{\pi} \left[\frac{t}{2} - \frac{\sin(4t)}{8} \right]_{-\pi/2}^0 = \frac{1}{\pi} \left[\frac{\pi}{4} \right] = \frac{1}{4} \end{aligned}$$

Por otro lado, tenemos que:

$$\begin{aligned} E[E^2[Y | X]] &= E \left[\left(-\frac{\sqrt{1-X^2}}{2} \right)^2 \right] = E \left[\frac{1-X^2}{4} \right] = \frac{1}{4} E[1-X^2] = \\ &= \frac{1}{4} [1 - E[X^2]] \stackrel{(*)}{=} \frac{1}{4} \left[1 - \frac{1}{4} \right] = \frac{3}{16} \end{aligned}$$

donde en (*) hemos empleado que, como las funciones de densidad de X y Y son iguales, $E[X^2] = E[Y^2]$. Por tanto, tenemos que:

$$\text{E.C.M.}(E[Y | X]) = E[Y^2] - E[E^2[Y | X]] = \frac{1}{4} - \frac{3}{16} = \frac{4-3}{16} = \frac{1}{16}$$

6. **(0.50 puntos)** Obtener una media de la bondad del ajuste del apartado anterior.

La media de la bondad del ajuste es el valor de $\eta_{Y|X}^2$:

$$\eta_{Y|X}^2 = \frac{\text{Var}[E[Y | X]]}{\text{Var}[Y]} = 1 - \frac{\text{E.C.M.}(E[Y | X])}{\text{Var}[Y]}$$

Tenemos que:

$$\begin{aligned} E[Y] &= \int_{-\infty}^{+\infty} y \cdot f_Y(y) dy = \int_{-1}^0 y \cdot \frac{4}{\pi} \cdot \sqrt{1-y^2} dy = \frac{4}{\pi} \int_{-1}^0 y \cdot \sqrt{1-y^2} dy = \\ &= \frac{4}{\pi} \left[-\frac{1}{2} \cdot \frac{(1-y^2)^{3/2}}{3/2} \right]_{-1}^0 = \frac{4}{\pi} \left[-\frac{1}{3} \right] = -\frac{4}{3\pi} \\ E[Y^2] &= \frac{1}{4} \end{aligned}$$

$$\text{Var}[Y] = E[Y^2] - E^2[Y] = \frac{1}{4} - \left(-\frac{4}{3\pi} \right)^2 = \frac{1}{4} - \frac{16}{9\pi^2} = \frac{9\pi^2 - 64}{36\pi^2}$$

Por tanto, tenemos que:

$$\begin{aligned}\eta_{Y|X}^2 &= 1 - \frac{\text{E.C.M.}(E[Y | X])}{\text{Var}[Y]} = 1 - \frac{\frac{1}{16}}{\frac{9\pi^2 - 64}{36\pi^2}} = 1 - \frac{9\pi^2}{4(9\pi^2 - 64)} = \frac{36\pi^2 - 256 - 9\pi^2}{4(9\pi^2 - 64)} = \\ &= \frac{27\pi^2 - 256}{4(9\pi^2 - 64)} \approx 0,10553\end{aligned}$$

Por tanto, vemos que el ajuste no es adecuado.

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) .

La función generatriz de momentos de una normal bivalente es:

$$M_{(X,Y)}(t_1, t_2) = \exp\left(\mu_1 t_1 + \mu_2 t_2 + \frac{\sigma_1^2 t_1^2 + \sigma_2^2 t_2^2 + 2\rho\sigma_1\sigma_2 t_1 t_2}{2}\right)$$

Vemos por tanto que (X, Y) sigue una distribución normal bivalente con parámetros:

$$\begin{cases} \mu_1 = 1 \\ \mu_2 = 0 \\ \sigma_1^2 = 4 \\ \sigma_2^2 = 9 \\ \rho\sigma_1\sigma_2 = 3 \end{cases}$$

Por tanto, tenemos que $(X, Y) \sim \mathcal{N}_2(\mu, \Sigma)$, con:

$$\mu = \begin{pmatrix} 1 & 0 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 4 & 3 \\ 3 & 9 \end{pmatrix}$$

Por tanto, el coeficiente de correlación lineal es:

$$\rho_{X,Y} = \rho = \frac{3}{\sigma_1\sigma_2} = \frac{3}{2 \cdot 3} = \frac{1}{2}$$

Por otro lado, sabemos que las curvas de regresión coinciden con las rectas de regresión, luego la razón de correlación es:

$$\eta_{Y|X}^2 = \eta_{X|Y}^2 = \rho_{X,Y}^2 = \frac{1}{4}$$

2. **(0.75 puntos)** Indicar las distribuciones de las variables aleatorias $Y | X = 1$ y $X | Y = 0$.

Sabemos que, dados $x^*, y^* \in \mathbb{R}$, las distribuciones condicionadas son:

$$Y \mid X = x^* \sim \mathcal{N}\left(\mu_2 + \rho \frac{\sigma_2}{\sigma_1}(x^* - \mu_1), \sigma_2^2(1 - \rho^2)\right)$$

$$X \mid Y = y^* \sim \mathcal{N}\left(\mu_1 + \rho \frac{\sigma_1}{\sigma_2}(y^* - \mu_2), \sigma_1^2(1 - \rho^2)\right)$$

Por tanto, tenemos que:

$$Y \mid X = 1 \sim \mathcal{N}\left(0 + \frac{1}{2} \cdot \frac{3}{2}(1 - 1), 9\left(1 - \frac{1}{4}\right)\right) = \mathcal{N}(0, 27/4)$$

$$X \mid Y = 0 \sim \mathcal{N}\left(1 + \frac{1}{2} \cdot \frac{2}{3}(0 - 0), 4\left(1 - \frac{1}{4}\right)\right) = \mathcal{N}(1, 3)$$

3. **(1 punto)** Obtener la distribución de probabilidad del vector aleatorio $(2X, Y - X)$. Justificar que las variables aleatorias $2X$ y $Y - X$ tienen cierta asociación lineal en sentido negativo.

Tenemos que:

$$(2X, Y - X) = (X \ Y) A, \quad A = \begin{pmatrix} 2 & -1 \\ 0 & 1 \end{pmatrix}$$

Por tanto, notando $X' = 2X$ y $Y' = Y - X$, tenemos que $(X', Y') \sim \mathcal{N}(\mu A, A^t \Sigma A)$, donde:

$$\mu A = (1 \ 0) \begin{pmatrix} 2 & -1 \\ 0 & 1 \end{pmatrix} = (2 \ -1)$$

$$A^t \Sigma A = \begin{pmatrix} 2 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 4 & 3 \\ 3 & 9 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 8 & 6 \\ -1 & 6 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 16 & -2 \\ -2 & 7 \end{pmatrix}$$

Por tanto, tenemos que:

$$\rho_{X', Y'} = \frac{-2}{\sqrt{16} \sqrt{7}} \approx -0,18$$

$$\rho_{X', Y'}^2 = \frac{2^2}{16 \cdot 7} = \frac{1}{28} = 0,03$$

Por tanto, como $\rho_{X', Y'}^2$ es muy cercano a 0, no tienen apenas asociación. No obstante, de tenerla, como $\rho_{X', Y'}$ es negativo, es en sentido negativo.