Vectores aleatorios

PID 00253304

Ana Escudero
Alícia Miralles
Alícia Vila

Tiempo mínimo de dedicación recomendado: 3 horas

Universitat
Oberta
de Catalunya

Índice

Introducción

Hemos dedicado los módulos «Variables aleatorias» y «Funciones de variables aleatorias» al estudio de variables aleatorias simples o unidimensionales.

Ahora bien, a veces nos encontramos con fenómenos que están relacionados con más de una variable aleatoria a la vez. Por ejemplo, en un circuito en el que la resistencia, inductancia y capacidad estén modelizadas como variables aleatorias, tendremos que trabajar con tres variables aleatorias a la vez. En los sistemas de transmisión, a menudo tenemos una señal de entrada aleatoria. Dado que la variable de entrada es aleatoria, la señal de salida también lo es. En este caso, necesitaremos trabajar con dos variables aleatorias para poder encontrar la relación entre estas y poder caracterizar el sistema de transmisión. También, si en un circuito tomamos 5 medidas de un valor desconocido, como podría ser la intensidad de la corriente, el error en cada una de estas medidas podría ser modelizado por una variable aleatoria y entonces deberíamos trabajar con cinco variables aleatorias al mismo tiempo.

El tratamiento de vectores de variables aleatorias nos introducirá la necesidad de definir la probabilidad, distribución de probabilidad y función de densidad conjuntas. Imaginad que tenemos un vector aleatorio bidimensional, (X, Y), en el que las variables X e Y son, respectivamente, la altura y el peso de un estudiante. Podemos definir S_1 como el espacio muestral para la altura y S_2 como el espacio muestral para el peso. Cada una de estas variables tendrá su función de probabilidad. Sin embargo, si ahora definimos el espacio muestral $S = S_1 \times S_2$, el resultado de nuestro experimento nos dará una altura y un peso y podremos definir una probabilidad asociada a estos dos hechos.

En este módulo, utilizaremos los conceptos que hemos visto para el caso unidimensional en los módulos anteriores y los extrapolaremos al caso de los vectores aleatorios bidimensionales. En el apartado 1, veremos cómo se aplica este concepto a las variables aleatorias discretas. El apartado 2 será parecido al primero, pero consideraremos vectores de variables aleatorias continuas.

Objetivos

Los objetivos que hay que lograr en este módulo son los siguientes:

1. Entender el concepto y la utilidad de vector aleatorio y saber poner ejemplos.

2. Conocer los vectores de variable aleatoria discreta y continua.

3. Calcular las funciones de probabilidad conjunta y de probabilidad marginal.

4. Aplicar los conceptos de probabilidad condicionada e independencia a vectores aleatorios.

5. Relacionar las variables aleatorias de un vector mediante la covarianza y el coeficiente de correlación.

1. Vector aleatorio (X, Y) con X e Y variables aleatorias discretas

Empezamos el apartado definiendo qué entendemos por vector aleatorio bidimensional, en este caso, aplicado a variables aleatorias discretas.

Definición 1.1. Si X e Y son dos variables aleatorias discretas, se denomina **vector aleatorio discreto bidimensional** al vector (X, Y).

En general, dadas n variables aleatorias discretas, $X_1, X_2, \ldots, X_n$, hay que trabajar con el vector aleatorio discreto n-dimensional $(X_1, X_2, \ldots, X_n)$.

Terminología

A veces, (X, Y) se denomina *variable aleatoria bidimensional*, y $(X_1, X_2, \ldots, X_n)$, *variable aleatoria n-dimensional*.

1.1. Probabilidad conjunta. Probabilidad marginal

Al tratar con vectores aleatorios, aparecen dos conceptos nuevos que no habíamos tratado anteriormente: la **probabilidad conjunta** y la **probabilidad marginal**. A continuación, las definimos.

Definición 1.2. Sean X, Y dos variables aleatorias discretas en las que X toma los valores $\{a_1, a_2, \ldots, a_n\}$ e Y toma los valores $\{b_1, b_2, \ldots, b_m\}$. Para cada pareja de valores a_i, b_j, tenemos definida la **función de probabilidad conjunta**

$$P(X = a_i, Y = b_j) = P(\{X = a_i\} \cap \{Y = b_j\}) \tag{1}$$

También se utiliza la notación $P_{XY}(a_i, b_j) = P(X = a_i, Y = b_j)$.

Es decir, la **probabilidad conjunta** es la probabilidad de que la variable aleatoria X tome el valor a_i, y la variable aleatoria Y tome el valor b_j.

Para el caso particular en el que X toma los valores $\{a_1, a_2\}$ e Y, $\{b_1, b_2\}$, obtenemos la tabla de probabilidades conjuntas siguiente.

$Y \setminus X$	a_1	a_2	$P(Y=b_j)$
b_1	$P(X=a_1, Y=b_1)$	$P(X=a_2, Y=b_1)$	$P(Y=b_1)$
b_2	$P(X=a_1, Y=b_2)$	$P(X=a_2, Y=b_2)$	$P(Y=b_2)$
$P(X=a_i)$	$P(X=a_1)$	$P(X=a_2)$	1

Definimos a continuación el concepto de **probabilidad marginal**.

Definición 1.3. Para cada valor a_i, definimos la **función de probabilidad marginal** de la variable X,

$$P(X=a_i) = \sum_{j=1}^{m} P(X=a_i, Y=b_j), \tag{2}$$

es decir, la suma de las probabilidades conjuntas fijado un valor de X y para todos los valores de Y. De manera parecida, para cada b_j, definimos la **función de probabilidad marginal** de Y,

$$P(Y=b_j) = \sum_{i=1}^{n} P(X=a_i, Y=b_j). \tag{3}$$

En la última fila de la tabla anterior, obtenemos las probabilidades marginales de X. La casilla de esta fila donde aparece $P(X = a_1)$ nos da la probabilidad marginal de la variable X para el valor a_1, puesto que nos dice cuál es la probabilidad de obtener a_1 para todos los valores de Y. Es decir, $P(X=a_1) = P(X = a_1, Y = b_1) + P(X = a_1, Y = b_2)$. De manera análoga, en la última columna de la tabla anterior, obtenemos las probabilidades marginales de Y. Observad la última casilla de la tabla. Tiene un valor igual a 1 porque es la probabilidad de cualquier valor del espacio muestral.

Observación

Con las probabilidades marginales, trabajamos de la misma manera que con las probabilidades definidas para una sola variable. Podemos, por lo tanto, considerar los mismos parámetros que habíamos definido en el tema de variables aleatorias. En particular, σ_X y σ_Y son las desviaciones típicas de X y de Y, respectivamente.

Ejemplo 1.1

Un emisor manda un mensaje binario (formado con elementos de $\{0, 1\}$), de tamaño 2 y al azar. Por el canal de transmisión, se pueden producir errores. Sabemos que la probabilidad de que un bit llegue al receptor con error es $P(\text{error}) = 0,02$. Definimos las variables aleatorias de la manera siguiente: la variable aleatoria X cuenta el número de 0 que envía el emisor y la variable aleatoria Y cuenta el número de 0 que llegan al receptor. Calcularemos las probabilidades conjuntas y marginales. A partir de esto, calcularemos el valor medio o esperanza y la desviación típica. De este modo, podemos caracterizar las variables aleatorias X e Y y compararlas.

Si el mensaje para transmitir es de tamaño 2, tanto la variable X como la Y, que cuentan el número de ceros en el mensaje, pueden tomar los valores $\{0,1,2\}$. Puesto que X e Y pueden tomar 3 valores, las combinaciones posibles de las variables X e Y para poder calcular las probabilidades son $3 \cdot 3 = 9$. Así pues, podemos definir las probabilidades conjuntas siguientes:

- Se emite 11 y llega 11. $P(X=0,Y=0) = P(T=11)P(R=11\,|\,T=11) = \left(\frac{1}{2}\right)^2 \cdot 0{,}98^2 = 0{,}2401$. T es el mensaje transmitido y R representa el mensaje recibido.

- Se emite 11 y llega 01 o 10, es decir, tenemos error en uno de los dos bits. $P(X=0,Y=1) = P(T=11)P(R=01\,|\,T=11) + P(T=11)P(R=10\,|\,T=11) = \left(\frac{1}{2}\right)^2 \cdot 0{,}02 \cdot 0{,}98 + \left(\frac{1}{2}\right)^2 \cdot 0{,}98 \cdot 0{,}02 = \left(\frac{1}{2}\right)^2 \cdot 2 \cdot 0{,}98 \cdot 0{,}02 = 0{,}0098$.

- Se emite 11 y llega 00. $P(X=0,Y=2) = \left(\frac{1}{2}\right)^2 \cdot 0{,}02^2 = 0{,}0001$.

- Se emite 10 y llega 11 o se emite 01 y llega 11. $P(X=1,Y=0) = \left(\frac{1}{2}\right)^2 \cdot 0{,}98 \cdot 0{,}02 + \left(\frac{1}{2}\right)^2 \cdot 0{,}02 \cdot 0{,}98 = 0{,}0098$.

- Se emite 01 y llega 01, o se emite 01 y llega 10, o se emite 10 y llega 10, o se emite 10 y llega 01. $P(X=1,Y=1) = 2\left(\frac{1}{2}\right)^2 \cdot 0{,}98^2 + 2\left(\frac{1}{2}\right)^2 0{,}02^2 = 0{,}4804$.

- Se emite 01 o 10 y llega 00. $P(X=1,Y=2) = 2\left(\frac{1}{2}\right)^2 \cdot 0{,}02 \cdot 0{,}98 = 0{,}0098$.

- Se emite 00 y llega 11. $P(X=2,Y=0) = \left(\frac{1}{2}\right)^2 \cdot 0{,}02^2 = 0{,}0001$.

- Se emite 00 y llega 01 o 10. $P(X=2,Y=1) = 2\left(\frac{1}{2}\right)^2 \cdot 0{,}98 \cdot 0{,}02 = 0{,}0098$.

- Se emite 00 y llega 00. $P(X=2,Y=2) = \left(\frac{1}{2}\right)^2 \cdot 0{,}98^2 = 0{,}2401$.

Observación

Observad que el mensaje recibido no es independiente del mensaje transmitido. Por este motivo, $P(R=11)$ sabiendo que hemos transmitido 11 es $P(R=11\,|\,T=11)$.

Calculamos ahora las probabilidades marginales. Es decir, fijamos el valor de una variable y sumamos para todos los valores de la otra variable:

- $P(X=0) = 0{,}25$, $P(X=1) = 0{,}5$, $P(X=2) = 0{,}25$,
- $P(Y=0) = 0{,}25$, $P(Y=1) = 0{,}5$, $P(Y=2) = 0{,}25$.

La tabla de probabilidades conjuntas es la siguiente.

$Y \setminus X$	0	1	2	$P(Y=b_j)$
0	0,2401	0,0098	0,0001	0,25
1	0,0098	0,4804	0,0098	0,5
2	0,0001	0,0098	0,2401	0,25
$P(X=a_i)$	0,25	0,5	0,25	1

Ahora calculamos algunos parámetros que nos dan información de cada una de las variables:

- La esperanza o valor medio de la variable X, que es la suma de los valores que puede tomar la variable aleatoria multiplicados por la probabilidad de aparecer, es decir:

$$\mathrm{E}(X) = 0 \cdot P(X=0) + 1 \cdot P(X=1) + 2 \cdot P(X=2) = 1.$$

Véase también

El valor medio se estudia en el módulo «Variables aleatorias» de esta asignatura.

- La varianza de la variable X, que se define como la esperanza del valor de la variable menos su esperanza y todo esto al cuadrado.

  Es decir: $\mathrm{Var}(X) = \mathrm{E}((X - \mathrm{E}(X))^2)$. Utilizando el teorema de la esperanza, $\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2$. Ahora ya podemos calcular la varianza:

$$\mathrm{E}(X^2) = 0^2 \cdot P(X=0) + 1^2 \cdot P(X=1) + 2^2 \cdot P(X=2) = 1,5.$$

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2 = 1,5 - 1^2 = 0{,}5.$$

- La desviación típica de la variable X que nos da información de la dispersión de los valores que toma X, en las mismas unidades: $\sigma_X = \sqrt{0{,}5} = 0{,}707$.

Podéis comprobar, haciendo los mismos cálculos, que para la variable Y obtenemos: $\mathrm{E}(Y) = 1$, $\mathrm{Var}(Y) = 0{,}5$ y $\sigma_Y = 0{,}707$.

1.2. Funciones de probabilidad condicionadas. Independencia de variables aleatorias

Ya hemos estudiado la noción de probabilidad condicionada. Vimos cómo podemos calcular la probabilidad de un suceso sabiendo que se había producido otro suceso. En aquel caso, nos referíamos a una sola variable aleatoria, X. Si el resultado de un experimento no nos daba ninguna pista sobre el resultado siguiente, hablábamos de sucesos independientes.

Véase también

Recordad la noción de probabilidad condicionada que vimos en el subapartado 2.3 del módulo «Introducción a la probabilidad».

La noción de probabilidad condicionada que veremos a continuación aquí es básicamente la misma, pero en este caso calcularemos la probabilidad de que la variable X tome un valor sabiendo cuál es el valor de la variable Y. Es decir, calcularemos la probabilidad de X condicionada a Y.

Definición 1.4. Función de probabilidad de X condicionada a Y.

La probabilidad de que la variable X tome el valor a_i sabiendo que Y toma el valor b_j (probabilidad de $X=a_i$ condicionada a $Y=b_j$) es:

$$P(X=a_i \mid Y=b_j) = \frac{P(X=a_i,\ Y=b_j)}{P(Y=b_j)}. \qquad (4)$$

Al trabajar con variables aleatorias bidimensionales, nos podríamos hacer la pregunta inversa: cuál es la probabilidad de obtener un cierto valor de Y sabiendo cuál es el valor que ha salido para la variable X. En este caso, hablaríamos de probabilidad condicionada de Y a X y se define de manera análoga a la anterior.

Definición 1.5. Función de probabilidad de Y condicionada a X.

La probabilidad de que la variable Y tome el valor b_j sabiendo que X toma el valor a_i (probabilidad de $Y = b_j$ condicionada a $X = a_i$) es:

$$P(Y = b_j \mid X = a_i) = \frac{P(X = a_i,\ Y = b_j)}{P(X = a_i)}. \qquad (5)$$

A veces, el resultado que conocemos (ya sea una realización de X o de Y) no nos da ninguna pista sobre la probabilidad de la otra variable, es decir, los resultados de las variables X e Y son independientes. Lo expresamos de la manera siguiente.

Definición 1.6. Las variables X e Y son **independientes** si, y solo si,

$$P(X = a_i,\ Y = b_j) = P(X = a_i)P(Y = b_j), \quad \forall i, j. \qquad (6)$$

Ejemplo 1.2

Con el mismo enunciado que en el ejemplo 1.1, nos hacemos las preguntas siguientes:

1) Sabiendo que el receptor ha recibido una palabra con un cero ($Y = 1$), ¿cuál es la probabilidad de que el emisor haya enviado la palabra 00?

$$P(X = 2 \mid Y = 1) = \frac{P(X = 2,\ Y = 1)}{P(Y = 1)} = \frac{0{,}0098}{0{,}5} = 0{,}0196.$$

2) ¿Las variables X e Y son independientes?
Dado que $P(X = 2, Y = 1) = 0{,}0098 \neq P(X = 2)P(Y = 1) = 0{,}125$, no son independientes. De hecho, hay una correlación alta entre el mensaje transmitido y el mensaje recibido, puesto que en el 96 % de los casos ($(1 - 0{,}02)^2 = 0{,}96$) esperamos recibir lo mismo que hemos transmitido.

1.3. Relación entre variables aleatorias discretas: covarianza y coeficiente de correlación

Hasta aquí hemos definido qué es un vector de variable aleatoria discreta, y nos hemos centrado en estudiar el caso de los vectores formados por dos variables aleatorias, es decir, los vectores aleatorios bidimensionales.

Hemos visto que parámetros como la esperanza, la varianza y la desviación estándar se pueden extender fácilmente al caso bidimensional partiendo de la definición que hemos visto para el caso unidimensional. Ahora bien, en el caso del vector aleatorio, nos podemos hacer una nueva pregunta que no nos habíamos planteado antes: ¿podemos medir cuál es la relación entre las variables aleatorias que forman el vector? Esta pregunta es la que intentaremos resolver en este subapartado.

Empezaremos definiendo tres parámetros que nos permiten caracterizar la relación entre dos variables aleatorias, X e Y.

Definición 1.7. Definimos la **esperanza del producto** como sigue:

$$\mathrm{E}(XY) = \sum_i \sum_j a_i b_j P(X = a_i,\ Y = b_j). \qquad (7)$$

Definición 1.8. La **covarianza** entre dos variables aleatorias X e Y se define de la manera siguiente:

$$\mathrm{Cov}(X,Y) = E[(X - \mathrm{E}(X))(Y - \mathrm{E}(Y))] \qquad (8)$$

$$= \sum_i \sum_j (a_i - \mathrm{E}(X))(b_j - \mathrm{E}(Y))P(X = a_i,\ Y = b_j).$$

Desarrollando la suma anterior, después de multiplicar los términos de los paréntesis llegamos a:

$$\mathrm{Cov}(X,Y) = \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y). \qquad (9)$$

Otro parámetro estadístico en lo referente a parejas de variables aleatorias es el coeficiente de correlación, que es una versión normalizada de la covarianza (o coeficiente de correlación lineal de Pearson).

Covarianza

Notad que después de simplificar, la covarianza es la esperanza del producto menos el producto de esperanzas de las dos variables aleatorias del vector.

Definición 1.9. El **coeficiente de correlación** entre las variables X e Y se define como la covarianza dividida entre las desviaciones estándar de las variables aleatorias:

$$\rho = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y}. \qquad (10)$$

Observación

La covarianza y el coeficiente de correlación son parámetros poblacionales que miden el nivel de relación lineal entre las variables.

La covarianza nos da información sobre la posible existencia de relación lineal entre X e Y, es decir, nos ayuda a saber si estas variables crecen de manera conjunta o no. Para dar una información normalizada, hay que dividir el valor de la covarianza por el producto de las desviaciones típicas de las dos variables. Esto es justamente lo que hace el coeficiente de correlación de Pearson.

Propiedades del coeficiente de correlación

1) El coeficiente de correlación siempre se encuentra en el margen siguiente:

$$-1 \leq \rho \leq 1. \tag{11}$$

2) Si ρ se encuentra cerca de 1 o -1, decimos que hay una fuerte correlación lineal entre X e Y.

3) Si, de media, Y aumenta cuando X aumenta: $\rho > 0$.

4) Si, de media, Y disminuye cuando X aumenta: $\rho < 0$.

Si ρ se encuentra cerca de 0, las variables presentan una correlación lineal débil o no hay correlación lineal.

Correlación cercana a 0

En caso de que el coeficiente de correlación sea cercano a 0, esto se puede deber al hecho de que no hay correlación de ningún tipo o bien a que la correlación que hay es no lineal (por ejemplo, cuadrática, cúbica, etc.).

Definición 1.10. Si $\rho = 0$, decimos que las variables son **linealmente incorrelacionadas**.

Relación entre independencia e incorrelación

Si X e Y son independientes, entonces $\text{Cov}(X, Y) = 0$ y $\rho = 0$. La implicación contraria, en general, no es cierta.

En efecto, suponemos que X e Y son independientes. El anterior resultado se obtiene teniendo en cuenta (6) y calculando con (7) y (9):

$$\begin{aligned} \text{E}(XY) &= \sum_i \sum_j a_i b_j P(X = a_i,\ Y = b_j) = \sum_i \sum_j a_i b_j P(X = a_i) P(Y = b_j) \\ &= \sum_i a_i P(X = a_i) \sum_j b_j P(Y = b_j) = \text{E}(X)\,\text{E}(Y). \end{aligned}$$

Entonces, $\text{Cov}(X, Y) = \text{E}(XY) - \text{E}(X)\,\text{E}(Y) = 0$.

Notamos que, puesto que $\rho = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y}$, $\rho = 0 \Leftrightarrow C(X,Y) = 0$.

Veamos un ejemplo en el que se aplican todos estos conceptos.

Ejemplo 1.3

Observad la figura 1. En la parte izquierda, hemos representado las 10 parejas de valores (x_i, y_i) que toma un vector (X,Y). En este caso, se obtiene un coeficiente de correlación de 0,9. Puesto que se trata de un valor cercano a 1, las variables X e Y presentan una correlación lineal fuerte: a medida que aumenta el valor de una, también aumenta el valor de la otra. El hecho de que tengan una correlación lineal fuerte nos dice que los valores (x_i, y_i) se encuentran cerca de una recta. La recta que hemos representado la hemos obtenido por el método de los mínimos cuadrados. A pesar de que no es nuestro propósito obtenerla, la hemos representado para observar mejor la correlación lineal. En la figura de la derecha, hemos dado otro ejemplo en el que el comportamiento es completamente distinto. En este otro caso, se obtiene un coeficiente de correlación de 0,25, prácticamente no hay correlación.

Figura 1. El coeficiente de correlación lineal del gráfico de la izquierda es $0{,}9$, y el del gráfico de la derecha, $0{,}25$.

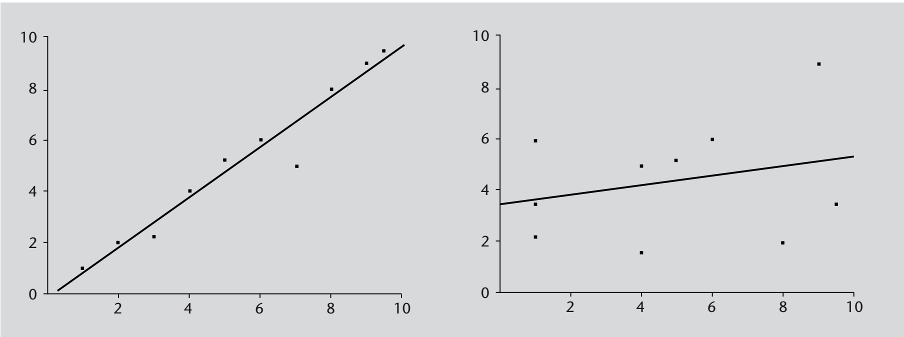


Figura 1

En la figura de la izquierda, las parejas de puntos están más correlacionadas que las de la figura de la derecha. Observad que la correlación en el primer caso es de 0,9, y en el segundo vale 0,25.

Ejemplo 1.4

Continuamos con el ejemplo 1.1, el del emisor binario, y calculamos ahora los parámetros definidos anteriormente. Calculamos, en primer lugar, la esperanza del producto:

$$\begin{aligned}
\mathrm{E}(XY) &= 0\cdot 0\cdot P(X=0,\ Y=0) + 0\cdot 1\cdot P(X=0,\ Y=1) + 0\cdot 2\cdot P(X=0,\ Y=2) + \\
&+ 1\cdot 0\cdot P(X=1,\ Y=0) + 1\cdot 1\cdot P(X=1,\ Y=1) + 1\cdot 2\cdot P(X=1,\ Y=2) + \\
&+ 2\cdot 0\cdot P(X=2,\ Y=0) + 2\cdot 1\cdot P(X=2,\ Y=1) + 2\cdot 2\cdot P(X=2,\ Y=2) = \\
&= 0{,}4804 + 2\cdot 0{,}0098 + 2\cdot 0{,}0098 + 4\cdot 0{,}2401 = 1,4.
\end{aligned}$$

Véase también

En el apartado 2 de este módulo, trataremos conceptos parecidos a los que hemos visto hasta ahora, pero aplicados a vectores aleatorios, en los que las variables aleatorias serán ahora continuas.

A continuación, y a partir de la esperanza del producto y del producto de esperanzas, como hemos visto en la definición 1.8, podemos calcular la covarianza:

$$\mathrm{Cov}(X,Y) = 1,4 - 1\cdot 1 = 0{,}4,$$

y el coeficiente de correlación:

$$\rho = \frac{0{,}4}{\sqrt{0{,}5}\,\sqrt{0{,}5}} = 0{,}8.$$

2. Vector aleatorio (X, Y) con X e Y variables aleatorias continuas

La estructura de este apartado es similar a la del anterior. En primer lugar, definiremos qué entendemos por vector aleatorio de variables continuas. A continuación, definiremos las funciones de distribución y de densidad conjuntas. También definiremos las funciones de densidad marginales. Observad que, en este caso, hablamos de funciones de densidad, puesto que trataremos variables aleatorias continuas. Como en el apartado anterior, veremos las probabilidades condicionadas y cuándo las variables del vector son independientes. Acabaremos el apartado evaluando si dos variables aleatorias varían de manera similar mediante la covarianza y el coeficiente de correlación.

Definición 2.1. Si X e Y son dos variables aleatorias continuas, se denomina **vector aleatorio continuo bidimensional** al vector (X, Y).

En general, dadas n variables aleatorias continuas, $X_1, X_2, \ldots, X_n$, hay que trabajar con el vector aleatorio continuo n-dimensional $(X_1, X_2, \ldots, X_n)$. En este apartado, nos centraremos en el caso $n = 2$.

Ya vimos en el tema de variables aleatorias que el tratamiento con variables continuas es muy diferente que con variables discretas. Empezamos definiendo las funciones de distribución conjunta y de densidad conjunta.

Véase también

En el módulo «Variables aleatorias», vimos las diferencias en el tratamiento de las variables aleatorias discretas y continuas. Estas diferencias son válidas también para el tratamiento de vectores aleatorios discretos y continuos.

2.1. Función de distribución conjunta. Función de densidad conjunta

Véase también

Recordad los conceptos de función de distribución y de densidad que vimos en el subapartado 3.1 del módulo «Variables aleatorias».

Definición 2.2. La **función de distribución conjunta**, F_{XY}, de las variables continuas X e Y es una aplicación de $\mathbb{R}^2$ a $[0, 1]$ definida por:

$$F_{XY}(x, y) = P(X \leq x,\ Y \leq y) \quad \forall (x, y) \in \mathbb{R}^2. \tag{12}$$

Si F_{XY} es continua y dos veces derivable, decimos que X e Y son **conjuntamente continuas**. La **función de densidad conjunta**, f_{XY}, es

$$f_{XY}(x, y) = \frac{\partial^2}{\partial x \partial y} F_{XY}(x, y). \tag{13}$$

En la anterior definición, F_{XY} y f_{XY} son funciones de dos variables, es decir, definidas en $\mathbb{R}^2$. En este caso, podemos derivar respecto a una de las variables, manteniendo la otra constante, y obtenemos las operaciones de derivada *parcial*. Dada una función $f(x,y)$, podemos derivarla respecto a x: $\frac{\partial f}{\partial x}$, o derivarla respecto a y: $\frac{\partial f}{\partial y}$. Para obtener la densidad conjunta f_{XY}, tenemos que hacer una derivada seguida de la otra a la función de distribución conjunta: $\frac{\partial^2}{\partial x \partial y} F_{XY}(x,y)$.

Propiedades de la función de densidad conjunta

1) La densidad conjunta es una función no negativa:

$$f_{XY}(x,y) \geq 0. \tag{14}$$

2) Para cada conjunto $A \subset \mathbb{R}^2$:

$$P((X,Y) \in A) = \int\int_A f_{XY}(x,y)\, dx\, dy. \tag{15}$$

3) Cálculo de la función de distribución a partir de la función de densidad:

$$F_{XY}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{XY}(u,v)\, du\, dv. \tag{16}$$

4) Normalización de la densidad:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x,y)\, dx\, dy = 1. \tag{17}$$

La propiedad (15) se podría considerar como la definición directa de la función de densidad. En este caso, (16) se deduce teniendo en cuenta que $F_{XY}(x,y)$ es la probabilidad de que $-\infty < X \leq x$ y $-\infty < Y \leq y$. Del mismo modo, (17) dice que la probabilidad de que X e Y tomen cualquier valor es 1. Notamos que haciendo $\frac{\partial^2}{\partial x \partial y} F_{XY}(x,y)$ a partir de (16) da efectivamente $f_{XY}(x,y)$.

Generalizando, lo que vimos en una dimensión, la probabilidad $P(A)$, en la que A es un subconjunto de $\mathbb{R}^2$, es determinada por el volumen por debajo de la función de densidad conjunta $f_{XY}(x,y)$ y que determina A (fórmula (15)).

Probabilidad $P(A)$

En el caso de variables aleatorias unidimensionales, encontrábamos la probabilidad mediante el área por debajo de la función de densidad. Observad que aquí estamos tratando con variables bidimensionales, y por este motivo hablamos de volumen por debajo de la función de densidad conjunta $f_{XY}(x,y)$.

Ejemplo 2.1

Un vector aleatorio (X, Y) tiene función de densidad:

$$f_{XY}(x,y) = \begin{cases} \dfrac{K}{(1+x+y)^3} & \text{si } x > 0, y > 0,\ x + y < 1, \\ 0 & \text{de otro modo.} \end{cases}$$

donde K es una constante.

1) ¿Qué valor tiene la constante K?
2) ¿Cuál es la probabilidad de que $Y > X$?
3) ¿Cuál es la probabilidad de que $X > 2Y$?
4) ¿Cuál es la probabilidad de que $X + Y > \frac{1}{2}$?

Notamos que la densidad está definida en un triángulo con vértice en los puntos $(0, 0)$, $(1, 0)$ y $(0, 1)$. Las integrales que calculamos se tienen que restringir a esta región. La podemos describir diciendo que x varía entre 0 y 1 y que, para cada x, y varía entre 0 y $1 - x$.

1) K se determina imponiendo la condición de normalización (17):

$$1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)\,dx\,dy = \int_0^1 \left(\int_0^{1-x} \frac{K}{(1+x+y)^3}\,dy\right) dx$$

$$= K\int_0^1 \left[-\frac{1}{2(1+x+y)^2}\right]_{y=0}^{y=1-x} dx = K\int_0^1 \left(\frac{1}{2(1+x)^2} - \frac{1}{8}\right) dx$$

$$= K\left[-\frac{1}{2(1+x)} - \frac{x}{8}\right]_0^1 = \frac{K}{8},$$

de donde $K = 8$.

2) Aplicamos (15). Tenemos que integrar la densidad sobre la región intersección entre el triángulo donde está definido el vector (X, Y) y el semiplano $Y > X$. Resulta un triángulo con vértice $(0, 0)$, $(0, 1)$ y $(\frac{1}{2}, \frac{1}{2})$ (este último es la intersección entre $x + y = 1$ y $y = x$). Esto restringe y entre x y $1 - x$, siempre que x esté entre 0 y $\frac{1}{2}$:

$$P(Y > X) = \int_0^{\frac{1}{2}} \left(\int_x^{1-x} \frac{8}{(1+x+y)^3}\,dy\right) dx = \int_0^{\frac{1}{2}} \left[-\frac{4}{(1+x+y)^2}\right]_{y=x}^{y=1-x} dx$$

$$= \int_0^{\frac{1}{2}} \left(\frac{4}{(1+2x)^2} - 1\right) dx = \left[-\frac{2}{1+2x} - x\right]_0^{\frac{1}{2}} = \frac{1}{2},$$

3) El suceso $X > 2Y$ corresponde al triángulo de vértice $(0, 0)$, $(1, 0)$ y $(\frac{2}{3}, \frac{1}{3})$ (este último es la intersección entre $x + y = 1$ y $2y = x$). Ahora es más directo integrar primero x entre $2y$ y $1 - y$, y después integrar y de 0 a $\frac{1}{3}$:

$$P(X > 2Y) = \int_0^{\frac{1}{3}} \left(\int_{2y}^{1-y} \frac{8}{(1+x+y)^3}\,dx\right) dy = \int_0^{\frac{1}{3}} \left[-\frac{4}{(1+x+y)^2}\right]_{x=2y}^{x=1-y} dy$$

$$= \int_0^{\frac{1}{3}} \left(\frac{4}{(1+3y)^2} - 1\right) dy = \left[-\frac{4}{1+3y} - y\right]_0^{\frac{1}{3}} = \frac{1}{3}.$$

4) En este caso, es más simple la región complementaria $X + Y < \frac{1}{2}$, consistente en el triángulo de vértice $(0, 0)$, $(\frac{1}{2}, 0)$ y $(0, \frac{1}{2})$. Así:

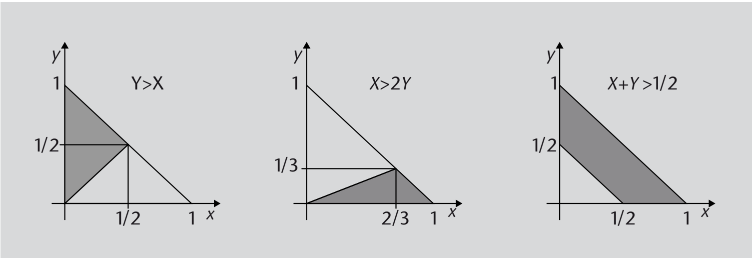
$$P\left(X + Y > \frac{1}{2}\right) = 1 - P\left(X + Y < \frac{1}{2}\right) = 1 - \int_0^{\frac{1}{2}} \left(\int_0^{\frac{1}{2}-x} \frac{8}{(1 + x + y)^3}\, dy\right) dx$$

$$= 1 - \int_0^{\frac{1}{2}} \left[-\frac{4}{(1 + x + y)^2}\right]_{y=0}^{y=\frac{1}{2}-x} dx = 1 - \int_0^{\frac{1}{2}} \left(\frac{4}{(1 + x)^2} - \frac{16}{9}\right) dx$$

$$= 1 - \left[-\frac{4}{1 + x} - \frac{16}{9}x\right]_0^{\frac{1}{2}} = \frac{5}{9}.$$

En la figura 2, se muestran los tres sucesos.

Figura 2. Sucesos del ejemplo 2.1



A continuación, definiremos un tipo particular de vector aleatorio bidimensional.

Definición 2.3. Distribución uniforme. Decimos que el vector aleatorio (X, Y) se distribuye **uniformemente** en la región $D \subset \mathbb{R}^2$ si la función de densidad conjunta es:

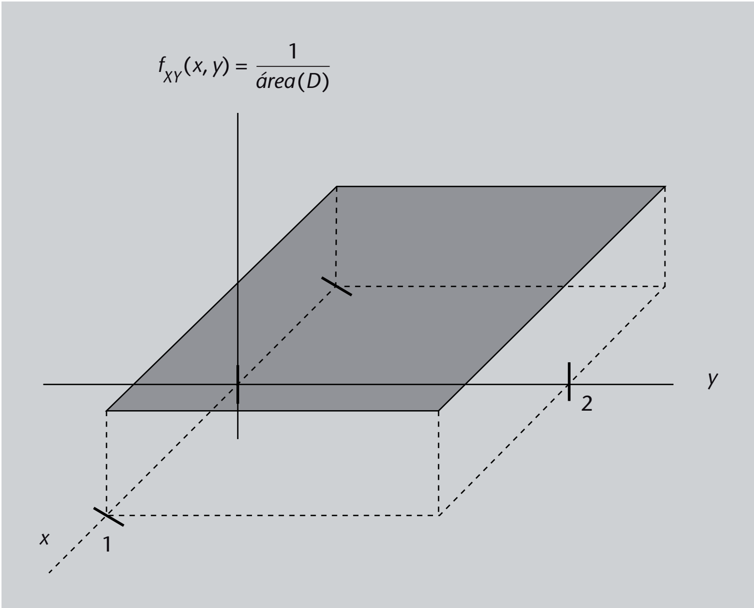
$$f_{XY}(x, y) = \begin{cases} \frac{1}{\text{área}(D)} & \text{si } (x, y) \in D \\ 0 & \text{de otro modo} \end{cases} \tag{18}$$

Es decir, para todos los puntos x e y que se encuentran dentro del dominio de definición de la variable aleatoria uniforme, el valor de $f_{XY}(x, y)$ es constante. En este caso, el cálculo de volúmenes por debajo de la función de densidad puede no requerir la utilización de las integrales dobles, como veremos en el ejemplo 2.4.

De manera intuitiva, podemos ver que si tenemos dos regiones dentro del área D con la misma área, las dos tienen la misma probabilidad.

En la figura 3, podéis ver un ejemplo de ello. En este caso, la región D es rectangular.

Figura 3. Representación gráfica de $f_{XY}(x, y)$ para el caso uniforme



2.2. Funciones de densidad marginales

En este subapartado, veremos cómo se pueden calcular las funciones de densidad marginales. Recordad que para el caso de las variables discretas, tal y como hemos visto en el subapartado 1.1, las funciones de probabilidad marginales las obteníamos para un cierto valor de una de las variables del vector y haciendo el sumatorio para todos los casos de la otra variable aleatoria. Aquí aplicaremos la misma idea, pero sustituyendo los sumatorios por integrales.

Definición 2.4. Funciones de densidad marginales

Densidad marginal de $\boldsymbol{X}$: fijado un valor de x, integramos para todos los valores posibles de y:

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy. \tag{19}$$

Densidad marginal de $\boldsymbol{Y}$: fijado un valor de y, integramos para todos los valores posibles de x:

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\, dx. \tag{20}$$

A partir de las expresiones anteriores, podemos encontrar los parámetros que caracterizan a cada una de las variables: la esperanza, el momento de orden 2 y la varianza para cada una de las variables del vector.

La definición de estos parámetros es la misma que habíamos hecho para el caso unidimensional.* En este caso, sin embargo, utilizamos las funciones de densidad marginales:

*Véase el subapartado 3.3 del módulo «Variables aleatorias».

- Esperanza de X y esperanza de Y:

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} x\, f_X(x)\, dx, \qquad \mathrm{E}(Y) = \int_{-\infty}^{\infty} y\, f_Y(y)\, dy.$$

- Momentos de orden 2:

$$\mathrm{E}(X^2) = \int_{-\infty}^{\infty} x^2\, f_X(x)\, dx, \qquad \mathrm{E}(Y^2) = \int_{-\infty}^{\infty} y^2\, f_Y(y)\, dy.$$

- Varianza de X y varianza de Y:

$$\sigma_X^2 = \mathrm{E}(X^2) - \mathrm{E}(X)^2, \qquad \sigma_Y^2 = \mathrm{E}(Y^2) - \mathrm{E}(Y)^2.$$

2.3. Funciones de densidad condicionadas. Variables independientes

En este subapartado, veremos cuándo las dos variables aleatorias de nuestro vector son **independientes** o, al contrario, cuándo el resultado de una nos da alguna pista sobre la otra. En este segundo caso, hablamos de **densidad condicionada**.

Definición 2.5. Las variables continuas X e Y son **independientes** si y solo si:

$$f_{XY}(x,y) = f_X(x) f_Y(y), \qquad \forall\, x,\, y. \tag{21}$$

Es decir, X e Y son independientes si la función de densidad conjunta es igual al producto de las funciones de densidad marginales, y viceversa.

Definición 2.6. Funciones de densidad condicionadas

Se define la **función de densidad de X condicionada a $Y = y$** como:

$$f(x\,|\,y) = \frac{f_{XY}(x,y)}{f_Y(y)}. \tag{22}$$

De manera análoga, la **función de densidad de Y condicionada a $X = x$** como:

$$f(y\,|\,x) = \frac{f_{XY}(x,y)}{f_X(x)}. \tag{23}$$

Observación

Observad que si las variables son independientes, las densidades condicionadas: $f(x\,|\,y) = \frac{f_X(x)f_Y(y)}{f_Y(y)} = f_X(x)$, $f(y\,|\,x) = f_Y(y)$.

Ahora veremos algunos ejemplos para aclarar todos estos conceptos.

Ejemplo 2.2

Una señal de comunicación empieza en el instante X y acaba en el instante Y, dados por el vector aleatorio (X,Y) con función de densidad:

$$f_{XY}(x,y) = \begin{cases} e^{-y} & \text{si } 0 < x < y, \\ 0 & \text{de otro modo.} \end{cases}$$

1) Calculamos las probabilidades de los siguientes sucesos:
$A =$ «La duración de la señal es inferior a 2».
$B =$ «En $t{=}2$ la señal es activa».
$C =$ «En $t{=}2$ la señal ya ha empezado y en $t = 1$ todavía no ha acabado».
Notamos que A equivale a $Y - X < 2$, B equivale a $X < 2, Y > 2$ y C equivale a $X{<}2, Y{>}1$.
Para A, la región es $0 < x < \infty$ y $x < y < x + 2$:

$$P(A) = \int_0^\infty \left(\int_x^{x+2} e^{-y} dy \right) dx = \int_0^\infty (e^{-x}-e^{-x-2}) dx = (1-e^{-2}) \int_0^\infty e^{-x} dx = 1-e^{-2}.$$

Para B, la región es $0 < x < 2$ y $2 < y < \infty$:

$$P(B) = \int_0^2 \left(\int_2^\infty e^{-y} dy \right) dx = \int_0^2 e^{-2} dx = 2e^{-2}.$$

Para C, la región tiene dos partes. Si $0 < x < 1$, entonces $1 < y < \infty$. Si $1 < x < 2$, entonces $x < y < \infty$:

$$P(C) = \int_0^1 \left(\int_1^\infty e^{-y} dy \right) dx + \int_1^2 \left(\int_x^\infty e^{-y} dy \right) dx$$

$$= \int_0^1 e^{-1} dx + \int_1^2 e^{-x} dx = e^{-1} + (e^{-1} - e^{-2}) = 2e^{-1} - e^{-2}.$$

2) Calculamos las funciones de densidad condicionadas $f(x\,|\,y)$ y $f(x\,|\,y)$.
Necesitamos primero las funciones de densidad marginales:
Si $x > 0$

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy = \int_{x}^{\infty} e^{-y}\,dy = [-e^{-y}]_x^{\infty} = e^{-x}.$$

Si $y > 0$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx = \int_{0}^{y} e^{-y}\,dx = e^{-y}[x]_0^{y} = ye^{-y}.$$

Así: $f_X(x) = \begin{cases} e^{-x} & \text{si } x > 0, \\ 0 & \text{de otro modo,} \end{cases}$ $\qquad f_Y(y) = \begin{cases} ye^{-y} & \text{si } y > 0, \\ 0 & \text{de otro modo.} \end{cases}$

Ahora calculamos: Dado $y > 0$, si $0 < x < y$:

$$f(x\,|\,y) = \frac{f_{XY}(x,y)}{f_Y(y)} = \frac{e^{-y}}{ye^{-y}} = \frac{1}{y}.$$

Dado $x > 0$, si $x < y < \infty$:

$$f(y\,|\,x) = \frac{f_{XY}(x,y)}{f_X(x)} = \frac{e^{-y}}{e^{-x}} = e^{-(y-x)}.$$

Así: $f(x\,|\,y) = \begin{cases} \frac{1}{y} & \text{si } 0 < x < y, \\ 0 & \text{de otro modo,} \end{cases}$ $\qquad f(y\,|\,x) = \begin{cases} e^{-(y-x)} & \text{si } x < y < \infty, \\ 0 & \text{de otro modo.} \end{cases}$

3) Si la señal acaba en el instante $t = 4$, ¿cuál es la probabilidad de que en el instante $t = 1$ todavía no hubiera empezado?
Tenemos como condición $Y = 4$ y como suceso $X > 1$, así que utilizaremos la densidad de X condicionada a Y:

$$f(x\,|\,Y = 4) = \frac{1}{4}, \qquad 0 < x < 4.$$

La probabilidad pedida es:

$$P(X > 1\,|\,Y = 4) = \int_1^4 \frac{1}{4}\,dx = \frac{3}{4}.$$

Ejemplo 2.3

La señal de entrada, X (voltios), en un canal de comunicaciones, se encuentra distribuida uniformemente en el intervalo $[-2, 2]$. La señal de salida, Y (voltios), es la suma de la señal de entrada más un ruido que se encuentra uniformemente distribuido en el intervalo $[-3, 3]$. Calculamos las probabilidades condicionadas $P(Y \le 0\,|\,X = 1)$, $P(Y \le y\,|\,X = 1)$ y $P(Y \le y\,|\,X = x)$. Es decir, ¿cuál es la probabilidad de que la señal de salida sea menor o igual que cero sabiendo que la señal de entrada es igual a 1 V? ¿Cuál es la probabilidad de cualquier valor de la señal de salida y, sabiendo que la señal de entrada es 1 V? Y finalmente, de manera más genérica, ¿cuál es esta probabilidad para cualquier valor x de la señal de entrada?

Si la señal de entrada es x, entonces dado un valor cualquiera de x, la señal de salida, y, podrá tomar un valor dentro del margen $[-3, 3]$ centrado en el valor de x concreto. Es decir, la variable Y se distribuye uniformemente en $[x-3, x+3]$. Por ejemplo, si la señal de entrada es de un voltio, $X = 1$, entonces Y se distribuye uniformemente en $[-2, 4]$.

Calculamos ahora la probabilidad condicionada: $P(Y \leq 0 \,|\, X = 1) = \frac{2}{6}$. ¿Cómo hemos hecho este cálculo? Observad que $Y \leq 0$ para el intervalo $[-2, 0]$ respecto al intervalo total $[-2, 4]$.

Ahora responderemos la segunda cuestión: $P(Y \leq y \,|\, X = 1) = \frac{y+2}{6}$, para $-2 < y < 4$. Observad que, en este caso, hemos tomado todo el intervalo de valores posibles para la variable Y sabiendo que X toma un valor igual a 1.

El caso más general, $P(Y \leq y \,|\, X = x) = \frac{y-x+3}{6}$, para $x-3 < y < x+3$. Observad que aquí hemos expresado la variable Y en función de los valores posibles de X.

2.4. Relación entre variables aleatorias continuas: covarianza y coeficiente de correlación

Definimos los mismos parámetros que hemos visto para vectores discretos como en el caso de las variables aleatorias unidimensionales. Ahora los sumatorios se harán integrales.

Definición 2.7.

Esperanza del producto:

$$\mathrm{E}(XY) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_{XY}(x,\ y)\, dx\, dy. \tag{24}$$

Covarianza:

$$\mathrm{Cov}(X, Y) = \mathrm{E}[(X - \mathrm{E}(X))(Y - \mathrm{E}(Y))] \tag{25}$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mathrm{E}(X))(y - \mathrm{E}(Y)) f_{XY}(x,\ y)\, dx\, dy.$$

Al igual que en el caso discreto, se obtiene la propiedad:

$$\mathrm{Cov}(X, Y) = \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y). \tag{26}$$

Coeficiente de correlación:

$$\rho = \frac{\mathrm{Cov}(X, Y)}{\sigma_X \sigma_Y}. \tag{27}$$

Como en el caso de variables aleatorias discretas, el parámetro covarianza da información sobre la relación lineal entre X e Y, es decir, si las variables crecen de manera conjunta o no. Para dar una información más significativa hay que normalizar la covarianza, y por eso definimos el coeficiente de correlación.

Observación

Observad las propiedades de la covarianza y del coeficiente de correlación para los casos de las variables discretas y continuas, y observad que son idénticas.

Propiedades de la covarianza y del coeficiente de correlación

1) $-1 \leq \rho \leq 1$.

2) Si ρ se encuentra cerca de 1 o -1, decimos que hay una fuerte correlación lineal entre X e Y.

3) Si X e Y aumentan o disminuyen conjuntamente, $\rho > 0$.

4) Si una de las variables aumenta al disminuir la otra (o al revés), $\rho < 0$.

5) Si ρ se encuentra cerca de 0, las variables presentan una correlación lineal débil o no hay correlación lineal. En el caso particular de $\rho = 0$, decimos que las variables son linealmente incorrelacionadas (podría haber otro tipo de tipo de correlación no lineal).

6) Si X e Y son independientes, entonces $\mathrm{Cov}(X, Y) = 0$ y $\rho = 0$. La implicación contraria, en general, no es cierta.

Ejemplo 2.4

(X, Y) es un vector aleatorio bidimensional uniforme en la región limitada por el triángulo, T, de lados sobre las rectas $y = 0$, $x = 3$ e $y = x$. Lo podéis ver en la figura 4.

Figura 4. Densidad uniforme en el dominio T

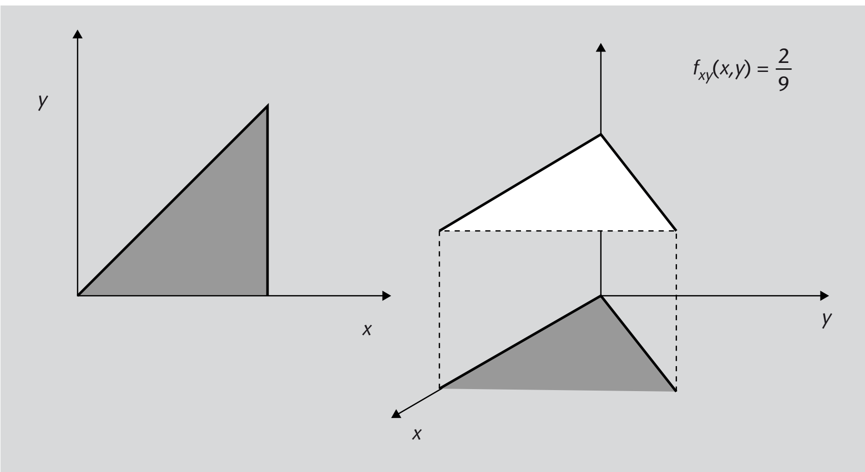


Figura 4

Observad cómo determinamos el dominio donde están definidas las variables x e y y la densidad uniforme dentro de este dominio.

1) Encontramos la función de densidad conjunta, funciones de densidad marginales y el valor esperado de cada una de las variables. ¿Las variables X e Y son independientes?

Dado que la distribución es uniforme,

$$f_{XY}(x,y) = \begin{cases} \frac{1}{\text{área triángulo}} = \frac{2}{9} & \text{si } (x,y) \in T, \\ 0 & \text{de otro modo.} \end{cases}$$

Área del triángulo

El área del triángulo es el producto de la base y la altura, dividido entre dos. En este caso: $\frac{3 \cdot 3}{2}$.

Calculamos la densidad marginal de X:

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy = \begin{cases} \int_0^x \frac{2}{9}\,dy = \frac{2}{9}x & \text{si } x \in [0,3], \\ 0 & \text{de otro modo.} \end{cases}$$

Observad los límites de integración. Si tomamos un valor cualquiera de x (lo podéis comprobar sobre la figura 4, haciendo una recta vertical sobre el triángulo), la y se encuentra entre cero y la recta $x = y$.

Calculamos ahora la densidad marginal de Y:

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dx = \begin{cases} \int_y^3 \frac{2}{9}\,dx = \frac{2}{9}(3-y) & \text{si } y \in [0,3], \\ 0 & \text{de otro modo.} \end{cases}$$

En este caso, fijamos un valor cualquiera de y trazando una recta horizontal sobre el triángulo. Observad que x se encuentra entre las rectas $x = y$ y $x = 3$, que son los valores que nos definen los límites de integración.

Límites de integración

Observad que ahora tomamos como límites de integración todo el intervalo de variación de las variables aleatorias $[0,3]$, a diferencia de cuando hemos calculado las densidades marginales para cada una de las variables.

Valor esperado de X:

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} x\, f_X(x)\,dx = \frac{2}{9}\int_0^3 x^2\,dx = \frac{2}{9} \cdot \frac{3^3}{3} = 2.$$

Valor esperado de Y:

$$\mathrm{E}(Y) = \int_{-\infty}^{\infty} y\, f_Y(y)\,dy = \frac{2}{9}\int_0^3 (3-y)y\,dy = \frac{2}{9}\left(\frac{3^3}{2} - \frac{3^3}{3}\right) = 1.$$

Las variables no son independientes porque: $f_{XY}(x,y) \neq f_X(x) f_Y(y)$.

2) Encontramos el coeficiente de correlación.

Momento de orden 2 de X:

$$\mathrm{E}(X^2) = \int_{-\infty}^{\infty} x^2\, f_X(x)\,dx = \frac{2}{9}.\int_0^3 x^3\,dx = \frac{2}{9} \cdot \frac{3^4}{4} = \frac{9}{2}.$$

Momento de orden 2 de Y:

$$\mathrm{E}(Y^2) = \int_{-\infty}^{\infty} y^2\, f_Y(y)\,dy = \frac{2}{9}\int_0^3 (3-y)y^2\,dy = \frac{2}{9} \cdot \left(3^3 - \frac{3^4}{4}\right) = \frac{3}{2}.$$

Varianza de X y σ_X:

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2 = \frac{9}{2} - 4 = \frac{1}{2}, \qquad \sigma_X = \sqrt{\frac{1}{2}}.$$

Varianza de Y y σ_Y:

$$\mathrm{Var}(Y) = \mathrm{E}(Y^2) - \mathrm{E}(Y)^2 = \frac{3}{2} - 1 = \frac{1}{2}, \qquad \sigma_Y = \sqrt{\frac{1}{2}}.$$

Esperanza del producto:

$$\mathrm{E}(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, y\, f_{XY}(x,y)\, dx\, dy$$

$$= \frac{2}{9}\int_{x=0}^{3}\int_{y=0}^{x} x\, y\, dy\, dx = \frac{2}{9}\int_{x=0}^{3} \frac{x^3}{2}\, dx = \frac{2}{9}\cdot\frac{3^4}{8} = \frac{9}{4}.$$

Covarianza y coeficiente de correlación:

$$\mathrm{Cov}(X,Y) = \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y) = \frac{9}{4} - 2\cdot 1 = \frac{1}{4}, \qquad \rho = \frac{\frac{1}{4}}{\sqrt{\frac{1}{2}}\sqrt{\frac{1}{2}}} = \frac{1}{2}.$$

3) Encontramos las probabilidades: $P(X < \frac{1}{2})$, $P(Y < \frac{1}{2})$ y $P(XY < \frac{1}{4})$:

$$P\left(X < \frac{1}{2}\right) = \int_0^{\frac{1}{2}} \frac{2}{9}\, x\, dx = \frac{1}{36}.$$

Representamos el volumen que determina la zona del triángulo donde $x < \frac{1}{2}$ (figura 5):

$$P\left(Y < \frac{1}{2}\right) = \int_0^{\frac{1}{2}} \frac{2}{9}\,(3-y)\, dy = \frac{11}{36}.$$

Representamos el volumen que determina la zona del triángulo donde $y < \frac{1}{2}$ (figura 5):

$$P\left(XY < \frac{1}{4}\right) = \int_{x=0}^{\frac{1}{2}}\int_{y=0}^{x} \frac{2}{9}\, dy\, dx + \int_{\frac{1}{2}}^{3}\int_{y=0}^{\frac{1}{4x}} \frac{2}{9}\, dy\, dx$$

$$= \frac{2}{9}\int_0^{\frac{1}{2}} x\, dx + \frac{2}{9}\int_{\frac{1}{2}}^{3} \frac{1}{4x}\, dx = \frac{1}{36} + \frac{\ln 3 + \ln 2}{18} = 0{,}127.$$

En la figura 5, mostramos la zona del triángulo donde $xy < \frac{1}{4}$ y el volumen que genera.

Figura 5. Función de densidad uniforme en el triángulo T

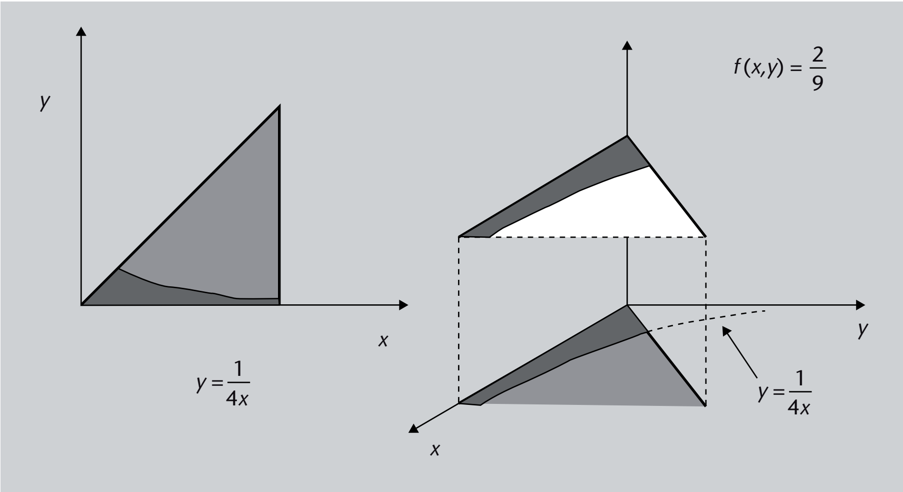


Figura 5

La zona del triángulo donde definimos las variables aleatorias genera un volumen determinado.

Resumen

En este módulo, hemos visto los vectores aleatorios. Estos pueden ser discretos (apartado 1 del módulo) o continuos (apartado 2 del módulo).

Dadas X e Y dos variables aleatorias discretas, podemos definir un vector aleatorio discreto, (X, Y). De este vector aleatorio hemos definido la **probabilidad conjunta** y la **probabilidad marginal**:

- Probabilidades conjuntas: $P(X=a_i, Y=b_j) = P(\{X=a_i\} \cap \{Y=b_j\})$.
- Probabilidad marginal de X (de manera análoga, podríamos definir la de Y):
  $P(X=a_i) = \sum_{j=1}^{m} P(X=a_i, Y=b_j)$.

También hemos visto el concepto de **probabilidad condicionada**:

$$P(X=a_i \,|\, Y=b_j) = \frac{P(X=a_i,\ Y=b_j)}{P(Y=b_j)}.$$

Y hemos definido el concepto de **independencia**. Hemos visto que X e Y son independientes si y solo si: $P(X=a_i,\ Y=b_j) = P(X=a_i)P(Y=b_j)$ para todo i, j.

Para finalizar, hemos aprendido a medir el grado de similitud de las variables aleatorias que forman el vector mediante los conceptos siguientes:

- **Esperanza del producto**: $\mathrm{E}(XY) = \sum_i \sum_j a_i b_j P(X=a_i,\ Y=b_j)$.
- **Covarianza**: $\mathrm{Cov}(X,Y) = \mathrm{E}[(X-\mathrm{E}(X))(Y-\mathrm{E}(Y))] = \mathrm{E}(XY)-\mathrm{E}(X)\,\mathrm{E}(Y)$.
- **Coeficiente de correlación lineal de Pearson**: $\rho = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y}$.

El apartado 2 de este módulo lo hemos dedicado a los vectores aleatorios bidimensionales continuos. Concretamente, si X e Y son variables aleatorias continuas, podemos definir el vector aleatorio continuo (X, Y), y hemos visto los conceptos asociados siguientes:

- **Función de distribución conjunta**: $F_{XY}(x,y) = P(X \leq x,\ Y \leq y) \quad \forall (x,y) \in \mathbb{R}^2$.
- **Función de densidad conjunta**: si F_{XY} es dos veces derivable, la f_{XY}, es $f_{XY}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{XY}(x,y)$.
- **Densidad marginal de $\boldsymbol{X}$**: $f_X(x) = \int_{-\infty}^{\infty} f(x,y)\, dy$. (De manera análoga, podemos definir la densidad marginal de Y.)

- **Densidad de X condicionada a $Y = y$**: $f(x \mid y) = \frac{f_{XY}(x,y)}{f_Y(y)}$. (De manera análoga, podemos definir la densidad de Y condicionada a $X = x$.)
- **Esperanza del producto**: $\mathrm{E}(XY) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_{XY}(x,y)\, dx\, dy$.
- **Covarianza**: $\mathrm{Cov}(X,Y) = \mathrm{E}[(X - \mathrm{E}(X))(Y - \mathrm{E}(Y))] = \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y)$.
- **Coeficiente de correlación**: $\rho = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y}$.

Actividades

1. Disponemos de un canal de comunicaciones que caracterizamos con las variables aleatorias X y Y, la velocidad de transmisión del canal y la proporción de errores que introduce el canal, respectivamente. La función de densidad conjunta del vector aleatorio (X, Y) está expresada en función de una constante k:

$$f_{XY}(x,y) = \begin{cases} k(x+y) & \text{si } \ 0 < x < 2, \ 0 < y < 2 \\ 0 & \text{de otro modo} \end{cases}$$

a) Encontrad el valor de k.
Pista: de manera análoga a como sucedía en el caso de funciones de densidad de una variable aleatoria, el volumen total determinado por la función de densidad conjunta $\mathbb{R}^2$ tiene que valer 1 -fórmula (17).

b) Encontrad las funciones de densidad marginales de X e Y.

c) Determinad si X e Y son independientes.

2. Para el vector aleatorio (X, Y) del problema anterior:

a) Encontrad las funciones de densidad condicional $f(x \,|\, y)$ y $f(y \,|\, x)$.

b) Encontrad $P\left(0 < Y < \frac{1}{2} \,|\, X = 1\right)$.

3. Dada una señal acústica que representamos mediante la variable aleatoria X, la introducimos en un amplificador que genera una nueva variable aleatoria $Y = aX + b$ (a y b constantes, a distinta de cero). Se pide:

a) Encontrad la covarianza de X e Y y expresadla en términos de σ_X^2, la varianza de X.
Pista: recordad que la esperanza es un operador lineal.

b) Utilizando el hecho de que $\mathrm{Var}(Y) = \mathrm{Var}(aX + b) = a^2\,\mathrm{Var}(X)$, encontrad* el coeficiente de correlación de X e Y.

* Atención: durante los cálculos, recordad que una desviación típica siempre es positiva, mientras que la constante a podría ser positiva o negativa.

4. La función de densidad conjunta de un vector aleatorio (X, Y) es:

$$f_{XY}(x,y) = \begin{cases} 6y & \text{si } \ 0 < y < x < 1 \\ 0 & \text{de otro modo} \end{cases}$$

a) Encontrad la función de densidad marginal $f_X(x)$.

b) Encontrad la función de densidad condicional $f(y \,|\, x)$.

c) Calculad el valor esperado condicional correspondiente, es decir, $\mathrm{E}(Y \,|\, X = x)$.

5. La función de probabilidad conjunta de dos señales digitales que se representan mediante un vector aleatorio (X, Y) es:

$$P_{XY}(x,y) = \begin{cases} 0{,}45 & \text{si } \ x = 0, \ y = 0 \\ 0{,}1 & \text{si } \ x = 1, \ y = 0 \\ 0{,}05 & \text{si } \ x = 0, \ y = 1 \\ 0{,}4 & \text{si } \ x = 1, \ y = 1 \\ 0 & \text{de otro modo} \end{cases}$$

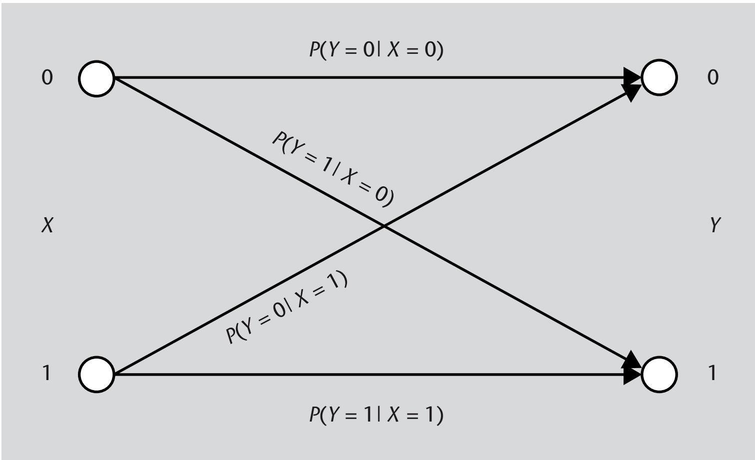
a) Calculad las funciones de probabilidad marginal de X y de Y.

b) Encontrad el valor medio y la varianza de X. Lo mismo para Y.

c) Encontrad la covarianza y el coeficiente de correlación de X e Y.

6. Probad que no pueden existir dos variables aleatorias X e Y para las cuales $\mathrm{E}(X) = 3$, $\mathrm{E}(Y) = 2$, $\mathrm{E}(X^2) = 10$, $\mathrm{E}(Y^2) = 29$ y $\mathrm{E}(XY) = 0$
Pista: haced la prueba por reducción al absurdo, es decir, suponed que es cierto lo que dice el enunciado y tratad de buscar una contradicción.

7. Considerad el canal de comunicación que se muestra abajo. Sea (X,Y) un vector aleatorio, en el que X es la entrada del canal e Y es la salida. Sabemos que $P(X=0) = 0{,}5$, $P(Y=1\,|\,X=0) = 0{,}1$, y que $P(Y=0\,|\,X=1) = 0{,}2$. Se pide:



a) Encontrad la función de probabilidad conjunta $P_{XY}(x,y)$.

b) Encontrad las funciones de probabilidad marginales $P_X(x)$ y $P_Y(y)$.

c) ¿Son X e Y independientes?

8. La función de distribución conjunta de un vector aleatorio (X,Y) está determinada por:

$$F_{XY}(x,y) = \begin{cases} \frac{y+e^{-x(y+1)}}{y+1} - e^{-x} & \text{si } x,\, y > 0 \\ 0 & \text{de otro modo} \end{cases}$$

Encontrad la función de densidad conjunta, $f_{XY}(x,y)$, y utilizad un software (por ejemplo, Wiris) para representarla gráficamente.

9. La función de densidad conjunta de un vector aleatorio (X,Y) está determinada por:

$$f_{XY}(x,y) = \begin{cases} 2 & \text{si } 0 < y < x < 1 \\ 0 & \text{de otro modo} \end{cases}$$

a) Encontrad las funciones de densidad marginales de X e Y.

b) Encontrad las funciones de densidad condicional $f(x\,|\,y)$ y $f(y\,|\,x)$.

10. En un servidor que procesa peticiones de cliente, definimos dos tiempos de espera: X_1, que es el tiempo de espera de la petición a la cola del servidor, y X_2, que es el tiempo que el servidor emplea en procesar la petición.

Las variables aleatorias X_1 y X_2 tienen la función de densidad conjunta siguiente:

$$f_{X_1X_2}(x_1, x_2) = \begin{cases} e^{-(x_1+x_2)} & \text{si } 0 < x_1 < \infty, 0 < x_2 < \infty \\ 0 & \text{de otro modo} \end{cases}$$

Calculad:

a) Probabilidad de que la petición esté más de una hora en el sistema (tiempo a la cola más tiempo de procesamiento).

b) Las densidades marginales de X_1 y X_2.

c) ¿Son independientes X_1 y X_2?

d) La probabilidad de que la petición pase más de una hora esperando en la cola.

Solucionario

1.

a)

$$1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)dxdy = k\int_0^2\int_0^2 (x+y)dxdy$$

$$= k\int_0^2 \left[\frac{x^2}{2} + xy\right]_{x=0}^{x=2} dy = k\int_0^2 (2+2y)dy = 8k \;\Rightarrow\; k = \frac{1}{8}.$$

b) La función de densidad marginal de X es:

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)dy = \frac{1}{8}\int_0^2 (x+y)dy = \frac{1}{8}\left[xy + \frac{y^2}{2}\right]_{y=0}^{y=2} = \begin{cases} \frac{1}{4}(x+1) & 0 < x < 2 \\ 0 & \text{de otro modo} \end{cases}$$

De manera análoga, la función de densidad marginal de Y es:

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)dy = \frac{1}{8}\int_0^2 (x+y)dy = \frac{1}{8}\left[xy + \frac{y^2}{2}\right]_{y=0}^{y=2} = \begin{cases} \frac{1}{4}(y+1) & 0 < y < 2 \\ 0 & \text{de otro modo} \end{cases}$$

c) X e Y no son independientes, puesto que: $f_{XY}(x,y) \neq f_X(x)f_Y(y)$ $(\frac{1}{8}(x+y) \neq \frac{1}{4}(x+1)\cdot\frac{1}{4}(y+1))$.

2.

a)

$$f(y\,|\,x) = \frac{\frac{1}{8}(x+y)}{\frac{1}{4}(x+1)} = \frac{x+y}{2(x+1)}, \qquad 0 < y < 2,\; 0 < x < 2.$$

$$f(x\,|\,y) = \frac{\frac{1}{8}(x+y)}{\frac{1}{4}(y+1)} = \frac{x+y}{2(y+1)}, \qquad 0 < x < 2,\; 0 < y < 2.$$

b) La probabilidad pedida es:

$$P\left(0 < Y < \frac{1}{2}\,|\,X = 1\right) = \int_0^{1/2} f(y\,|\,x=1)dy = \int_0^{1/2} \frac{1+y}{4}dy = \frac{5}{32}.$$

3.

a) Notad que:

$$\mathrm{E}(XY) = E\left[X(aX+b)\right] = a\,\mathrm{E}(X^2) + b\,\mathrm{E}(X).$$

$$\mathrm{E}(Y) = \mathrm{E}(aX+b) = a\,\mathrm{E}(X) + b.$$

Por lo tanto,

$$\mathrm{Cov}(X,Y) = \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y) = a\,\mathrm{E}(X^2) + b\,\mathrm{E}(X) - \mathrm{E}(X)(a\,\mathrm{E}(X) + b)$$

$$= a(\mathrm{E}(X^2) - \mathrm{E}(X)^2) = a\,\mathrm{Var}(X) = a\sigma_X^2.$$

b) Notad que:* $\sigma_Y^2 = a^2\sigma_X^2 \;\Rightarrow\; \sigma_Y = |a|\,\sigma_X.$

*Véase el ejemplo 3.2 del módulo «Funciones de variables aleatorias».

El coeficiente de correlación de X e Y es:

$$\rho = \frac{\mathrm{Cov}(X,Y)}{\sigma_X\sigma_Y} = \frac{a\sigma_X^2}{\sigma_X\cdot|a|\,\sigma_X} = \frac{a}{|a|} = \begin{cases} 1 & a>0 \\ -1 & a<0 \end{cases}$$

4.

a)

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)dy = \int_0^x 6y dy = 3x^2.$$

$$f_X(x) = \begin{cases} 3x^2 & 0<x<1 \\ 0 & \text{de otro modo} \end{cases}$$

b)

$$f(y\,|\,x) = \frac{f_{XY}(x,y)}{f_X(x)} = \begin{cases} 2\frac{y}{x^2} & 0<y<x<1 \\ 0 & \text{de otro modo} \end{cases}$$

c) El valor pedido es:

$$\mathrm{E}(Y\mid X=x) = \int_{-\infty}^{\infty} y f(y\,|\,x)dy = \frac{1}{x^2}\int_0^x 2y^2 dy = \frac{2}{3}x.$$

5.

a) Probabilidad marginal de X:

$$P(X=0) = P_{XY}(0,0) + P_{XY}(0,1) = 0{,}5, \qquad P(X=1) = P_{XY}(1,0) + P_{XY}(1,1) = 0{,}5.$$

Probabilidad marginal de Y:

$$P(Y=0) = P_{XY}(0,0) + P_{XY}(1,0) = 0{,}55, \qquad P(Y=1) = P_{XY}(0,1) + P_{XY}(1,1) = 0{,}45.$$

b)

$$\mathrm{E}(X) = \sum_i x_i P(X=x_i) = 0 \cdot 0{,}5 + 1 \cdot 0{,}5 = 0{,}5.$$

$$\mathrm{E}(X^2) = \sum_i x_i^2 P(X=x_i) = 0^2 \cdot 0{,}5 + 1^2 \cdot 0{,}5 = 0{,}5.$$

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2 = 0{,}5 - 0{,}5^2 = 0{,}25.$$

Análogamente, se obtiene:

$$\mathrm{E}(Y) = 0{,}45, \qquad \mathrm{Var}(Y) = 0{,}2475.$$

c)

$$\mathrm{E}(XY) = \sum_i \sum_j x_i y_j P_{XY}(x_i, y_j) = 0 \cdot 0 \cdot 0{,}45 + 0 \cdot 1 \cdot 0{,}05 + 1 \cdot 0 \cdot 0{,}1 + 1 \cdot 1 \cdot 0{,}4 = 0{,}4.$$

$$\mathrm{Cov}(X,Y) = \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y) = 0{,}4 - 0{,}5 \cdot 0{,}45 = 0{,}175.$$

$$\rho = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{0{,}175}{\sqrt{0{,}25 \cdot 0{,}2475}} = 0{,}704.$$

6. Si suponemos que se satisfacen todas las condiciones del enunciado, entonces:

$$\mathrm{Cov}(X,Y) = \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y) = 0 - 3 \cdot 2 = -6$$

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2 = 10 - 9 = 1 \implies \sigma_X = 1.$$

$$\mathrm{Var}(Y) = \mathrm{E}(Y^2) - \mathrm{E}(Y)^2 = 29 - 4 = 25 \implies \sigma_Y = 5.$$

$$\rho = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{-6}{1 \cdot 5} = -\frac{6}{5} < -1.$$

Contradicción, puesto que tiene que ser $-1 \leq \rho \leq 1$.

Por lo tanto, no se pueden dar a la vez todas las condiciones del enunciado.

7.

a) Notad que $P(X=1) = 0{,}5$, $P(Y=0 \mid X=0) = 0{,}9$ y . $P(Y=1 \mid X=1) = 0{,}8$

Por lo tanto:

$P_{XY}(0,0) = P(X=0, Y=0) = P(Y=0 \mid X=0)P(X=0) = 0{,}45.$

$P_{XY}(0,1) = P(X=0, Y=1) = P(Y=1 \mid X=0)P(X=0) = 0{,}05.$

$P_{XY}(1,0) = P(X=1, Y=0) = P(Y=0 \mid X=1)P(X=1) = 0{,}10.$

$P_{XY}(1,1) = P(X=1, Y=1) = P(Y=1 \mid X=1)P(X=1) = 0{,}40.$

b) Las funciones de probabilidad marginales son:

Marginal de X:

$P_X(0) = P(X=0) = P_{XY}(0,0) + P_{XY}(0,1) = 0{,}5.$

$P_X(1) = P(X=1) = P_{XY}(1,0) + P_{XY}(1,1) = 0{,}5.$

Marginal de Y:

$P_Y(0) = P(Y=0) = P_{XY}(0,0) + P_{XY}(1,0) = 0{,}55.$

$P_Y(1) = P(Y=1) = P_{XY}(0,1) + P_{XY}(1,1) = 0{,}45.$

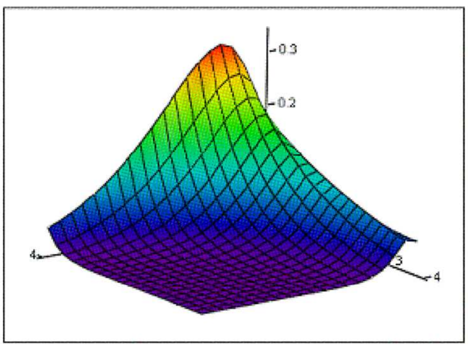
c) X e Y no son independientes, puesto que $P_{XY}(0,0) = 0{,}45 \neq P_X(0)P_Y(0) = 0{,}275.$

8. Para x, $y > 0$, se tiene que:

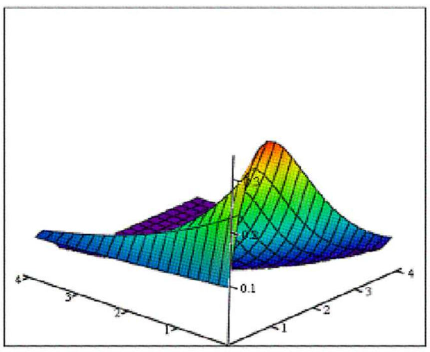
$$f_{XY}(x,y) = \frac{\partial^2 F_{XY}}{\partial x \partial y}(x,y) = \frac{\partial}{\partial y}\left(\frac{\partial F_{XY}}{\partial x}(x,y)\right) = \frac{\partial}{\partial y}(e^{-x} - e^{-x(y+1)}) = xe^{-x(y+1)}.$$

Entonces: $f_{XY}(x,y) = \begin{cases} xe^{-x(y+1)} & x,\ y > 0 \\ 0 & \text{de otro modo} \end{cases}$

A continuación, se muestra la gráfica de la función de densidad conjunta desde diferentes perspectivas:



s



s

9.

a) Si $0 < x < 1$ entonces $f_X(x) = \int_0^x 2dy = 2x$. Por lo tanto: $f_X(x) = \begin{cases} 2x & 0 < x < 1 \\ 0 & \text{de otro modo} \end{cases}$

Si $0 < y < 1$ entonces $f_Y(y) = \int_y^1 2dx = 2(1-y)$. Por lo tanto: $f_Y(y) = \begin{cases} 2(1-y) & 0 < y < 1 \\ 0 & \text{de otro modo} \end{cases}$

b) Si $0 < y < x < 1$ entonces $f(y\,|\,x) = \frac{f_{XY}(x,y)}{f_X(x)} = \frac{2}{2x} = \frac{1}{x}$. Por lo tanto: $f(y\,|\,x) = \begin{cases} \frac{1}{x} & 0 < y < x < 1 \\ 0 & \text{de otro modo} \end{cases}$

Si $0 < y < x < 1$ entonces $f(x\,|\,y) = \frac{f_{XY}(x,y)}{f_Y(y)} = \frac{1}{1-y}$. Por lo tanto: $f(x\,|\,y) = \begin{cases} \frac{1}{1-y} & 0 < y < x < 1 \\ 0 & \text{de otro modo} \end{cases}$

10.

a) Calculamos la probabilidad de que el tiempo de espera total no supere 1 hora, es decir, $X_1 + X_2 > 1$.

$$P(X_1 + X_2 > 1) = 1 - P(X_1 + X_2 < 1) = 1 - \int_0^1 \left(\int_0^{1-x_1} e^{-(x_1+x_2)} dx_2 \right) dx_1$$

$$1 - \int_0^1 e^{-x_1}(1 - e^{-(1-x_1)}) dx_1 = 1 - \int_0^1 (e^{-x_1} - e^{-1}) dx_1$$

$$1 - [-e^{-x_1} - e^{-1}x_1]_0^1 = 2e^{-1}.$$

b) Las funciones de densidad marginales son las siguientes:

$$f_{X_1}(x_1) = \int_0^\infty e^{-(x_1+x_2)} dx_2 = e^{-x_1} \int_0^\infty e^{-x_2} dx_2 = e^{-x_1}, \quad x_1 > 0.$$

Análogamente, para x_2:

$$f_{X_2}(x_2) = \int_0^\infty e^{-(x_1+x_2)} dx_1 = e^{-x_2} \int_0^\infty e^{-x_1} dx_2 = e^{-x_2}, \quad x_2 > 0.$$

c) X_1 y X_2 son independientes puesto que, para $x_1 > 0, x_2 > 0$:

$$f_{X_1X_2}(x_1, x_2) = e^{-(x_1+x_2)} = f_{X_1}(x_1) f_{X_2}(x_2) = e^{-x_1} e^{-x_2}.$$

d) La probabilidad de que el tiempo de espera en la cola sea mayor que una hora es:

$$P(X_1 > 1) = \int_1^\infty e^{-x_1} dx_1 = e^{-1}.$$