Probabilidad

Grado en Matemáticas

Curso 2018-2019

Hoja de Problemas 9

Fecha de Entrega: **Miércoles, 24 de Abril**

1. La siguiente tabla recoge la función de masa conjunta de las variables X e Y:

		X		
		-2	0	2
	0	0.2	0.2	0.2
Y	4	0.2	0	0.2

(a) ¿Son X e Y variables incorrelacionadas? Razonar la respuesta.

(b) ¿Son X e Y variables aleatorias independientes? Justificar la respuesta.

(c) Calcular $V(Y)$ y $V(Y|X=0)$.

(**Examen de Mayo de 2017**)

Solución: La siguiente tabla proporciona las probabilidades conjuntas de las variables X e Y, y las distribuciones marginales de cada variable:

		X			
		-2	0	2	
	0	0.2	0.2	0.2	0.6
Y	4	0.2	0	0.2	0.4
		0.4	0.2	0.4	1

(a) Las variables **incorrelacionadas** son aquellas cuya **covarianza es 0**. Por tanto, para responder a esta pregunta necesitamos calcular $Cov(X, Y)$.
Recordemos que una de las dos fórmulas de la covarianza (la que hace más sencillos cálculos) es

$$Cov(X,Y) = E(X \cdot Y) - E(X) \cdot E(Y).$$

En este caso tenemos que

$$\begin{aligned} E(X \cdot Y) &= (-2) \cdot 4 \cdot 0.2 + 2 \cdot 4 \cdot 0.2 = -8 \cdot 0.2 + 8 \cdot 0.2 = 0, \\ E(X) &= -2 \cdot 0.4 + 0 \cdot 0.4 + 2 \cdot 0.4 = 0, \\ E(Y) &= 0 \cdot 0.6 + 4 \cdot 0.4 = 1.6. \end{aligned}$$

Luego

$$Cov(X,Y) = 0 - 0 \cdot 1.6 = \mathbf{0},$$

y por tanto X e Y **sí son variables aleatorias incorrelacionadas**.

(b) La condición de independencia es más fuerte que la de incorrelación. Por tanto, aunque hayamos comprobado que X e Y están incorrelacionadas, esto no implica que sean independientes.

Las variables aleatorias X e Y serán independientes si y sólo si se verifica

$$P(X=k, Y=m) = P(X=k) \times P(Y=m)$$

para todo $k \in \{-2, 0, 2\}$ y todo $m \in \{0, 4\}$.

Puesto que (por ejemplo)

$$P(X=0, Y=0) = 0.2 \neq 0.12 = 0.2 \times 0.6 = P(X=0) \times P(Y=0),$$

las variables X e Y **no son independientes**, aunque sí estén incorrelacionadas.

(c) La función de masa marginal de la variable aleatoria Y es:

$$\begin{aligned} f_X(0) &= P(X=0) = 0.6, \\ f_X(4) &= P(X=4) = 0.4, \end{aligned}$$

o, escrito abreviadamente en forma matricial,

$$Y \equiv \begin{pmatrix} 0 & 4 \\ 0.6 & 0.4 \end{pmatrix}.$$

Como ya hemos visto,

$$E(Y) = 0 \cdot 0.6 + 4 \cdot 0.4 = 1.6$$

Además

$$E\left(Y^2\right) = 0^2 \cdot 0.6 + 4^2 \cdot 0.4 = 6.4$$

Luego

$$V(Y) = E\left(Y^2\right) - E(Y)^2 = 6.4 - 1.6^2 = \mathbf{3.84}$$

Por otra parte, la distribución de la variable Y condicionada por $X=0$ es

$$\begin{aligned} P(Y=-8|X=0) &= \frac{P(X=0, Y=0)}{P(X=0)} = \frac{f(0,2)}{f_X(0)} = \frac{0.2}{0.2} = 1, \\ P(Y=4|X=0) &= \frac{P(X=0, Y=4)}{P(X=0)} = \frac{f(0,4)}{f_X(0)} = \frac{0}{0.2} = 0, \end{aligned}$$

o

$$Y|X=0 \equiv \begin{pmatrix} 0 & 4 \\ 1 & 0 \end{pmatrix},$$

es decir

$$Y|X=0 \equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

Luego

$$\begin{aligned}
E(Y|X=0) &= 0 \times 1 = 0,\\
E(Y^2|X=0) &= 0^2 \times 1 = 0,\\
V(Y|X=0) &= E(Y^2|X=0) - [E(Y|X=0)]^2 = 0 - 0 = \mathbf{0}.
\end{aligned}$$

Esto es lo que se conoce como una **variable degenerada** (en este caso en 0), ya que no es una verdadera variable aleatoria, sino una constante.

2. Martín es un antiguo profesor e investigador de la URJC, y doctor en Estadística, que ahora se dedica a vender pañuelos en un semáforo. Dada su experiencia científica previa, ha comprobado, después de varios meses trabajando en la calle, que sólo le compran clines uno de cada 20 conductores de los que paran en los semáforos.

   Está lloviendo y hace frío, pero Martín necesita vender un paquete más para poder pagar al menos el recibo del agua.

   ¿Cuántos coches tienen que pasar como mínimo por ese semáforo para que la probabilidad de que Martín venda al menos un paquete de pañuelos sea al menos 0.8?

   Solución: Para cada coche que pasa por el semáforo, la probabilidad de que le compre un paquete de clines a Martín es

$$p = \frac{1}{20} = 0.05$$

   Definimos la variable aleatoria

$$C_n = \text{número de paquetes de pañuelos vendidos tras pasar } n \text{ coches},$$

   cuya distribución es binomial, con n repeticiones y probabilidad de éxito $p = 0.05$:

$$C_n \sim Bin(n, 0.05)$$

   Tenemos que determinar el mínimo valor de n para el cual se verifica

$$P(C_n \geq 1) \geq 0.8$$

   Para cada $n \in \mathbb{N}^+$ se tiene que

$$P(C_n \geq 1) = 1 - P(C_n = 0) = 1 - (1 - 0.05)^n = 1 - 0.95^n,$$

   luego necesitamos que se verifique

$$1 - 0.95^n \geq 0.8,$$

   o lo que es equivalente

$$0.95^n \leq 1 - 0.8 = 0.2$$

   Para despejar n de esta desigualdad tomamos logaritmos (recordemos que el logaritmo de cualquier número del intervalo (0,1) es negativo):

$$0.95^n \leq 0.2 \iff \ln(0.95^n) \leq \ln(0.2) \iff n\ln(0.95) \leq \ln(0.2) \iff n \geq \frac{\ln(0.2)}{\ln(0.95)} = \frac{-1.6094}{-0.0513} = 31.38$$

   Por lo tanto, para que la probabilidad de que Martín venda al menos un paquete de pañuelos sea al menos 0.8, tendrán que pasar **al menos 32 coches** por el semáforo.

   Obervación: Otra forma de razonar este problema es a partir de una distribución geométrica.

3. Una urna contiene 11 bolas: 5 negras, 2 blancas y el resto rojas. De esta urna **se van a extraer dos bolas al azar y sin remplazamiento**.

   Un tahur, que está considerando realizar alguna apuesta sobre esta extracción, está interesando en conocer la cantidad de bolas negras y blancas que pueden salir en este sorteo.

   Sea N el número de bolas negras y y B el número de bolas blancas seleccionadas en esta extracción.

   Determinar la función de masa conjunta del vector (N, B), y razonar si se trata o no de dos variables aleatorias **incorrelacionadas** y si son variables **independientes**.

   (**Examen de Mayo de 2018**)

   Solución: La siguiente tabla proporciona la función de masa conjunta del vector (N, B) y las distribuciones marginales de las dos variables:

		N			
	f(x,y)	**0**	**1**	**2**	Total
B	**0**	$\frac{12}{110}$	$\frac{40}{110}$	$\frac{20}{110}$	$\frac{72}{110}$
B	**1**	$\frac{16}{110}$	$\frac{20}{110}$	**0**	$\frac{36}{110}$
B	**2**	$\frac{2}{110}$	**0**	**0**	$\frac{2}{110}$
	Total	$\frac{30}{110}$	$\frac{60}{110}$	$\frac{20}{110}$	**1**

   Para determinar si las variables N y B son o no incorrelacionadas hay que determinar si su covarianza es 0. En este caso se tiene que

$$E(N \times B) = \frac{20}{110} = \frac{2}{11},$$
$$E(N) = \frac{100}{110} = \frac{10}{11},$$
$$E(B) = \frac{40}{110} = \frac{4}{11},$$

   luego

$$Cov(N, B) = E(N \times B) - E(N) \times E(B) = \frac{2}{11} - \frac{10}{11} \times \frac{4}{11} = \mathbf{-0.1488}$$

   Puesto que $Cov(N, B) \neq 0$ **las variables N y B no están incorrelacionadas**.

   Esto implica que N **y** B **no pueden ser independientes**.

4. Enunciar la desigualdad de Chebyshev, y demostrarla para el caso de una variable aleatoria continua.

   (**Examen de Mayo de 2017**)

   Solución: La desigualdad de Chebyshev establece que dada una variable aleatoria X con esperanza $E(X) = \mu$ y varianza $V(X) = \sigma^2 < \infty$, entonces, para cualquier $\epsilon > 0$ se verifica

$$P(\,|X - E(X)| \geq \epsilon) \leq \frac{V(X)}{\epsilon^2},$$

   es decir,

$$P(\,|X - \mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2}.$$

Demostración para el caso continuo: La varianza de la variable aleatoria continua X, con función de densidad f_X, responde a la fórmula

$$\sigma^2 = V(X) = \int_{x \in S_X} (X - \mu)^2 \cdot f_X(x)\, dx$$

que puede descomponerse en dos sumandos:

$$\sigma^2 = \int_{\{x \in S_X : |x-\mu| < \epsilon\}} (X - \mu)^2 \cdot f_X(x)\, dx + \int_{\{x \in S_X : |x-\mu| \geq \epsilon\}} (X - \mu)^2 \cdot f_X(x)\, dx$$

Puesto que

$$\int_{\{x \in S_X : |x-\mu| < \epsilon\}} (X - \mu)^2 \cdot f_X(x)\, dx \geq 0$$

se tiene que

$$\begin{aligned}
\sigma^2 &\geq \int_{\{x \in S_X : |x-\mu| \geq \epsilon\}} (X - \mu)^2 \cdot f_X(x)\, dx \\
&\geq \int_{\{x \in S_X : |x-\mu| \geq \epsilon\}} \epsilon^2 f_X(x)\, dx \\
&= \epsilon^2 \int_{\{x \in S_X : |x-\mu| \geq \epsilon\}} f_X(x)\, dx \\
&= \epsilon^2 \cdot P(\,|X - \mu| \geq \epsilon),
\end{aligned}$$

y despejando de esta desigualdad se obtiene

$$P(\,|X - \mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2}.$$

c.q.d.

5. La siguiente tabla recoge la función de masa conjunta de las variables X e Y:

		Y		
	$f(x,y)$	-2	0	2
X	-2	0.07	0.08	0
	-1	0.06	0.14	0.10
	0	0.17	0.11	0.02
	2	0.05	0.12	0.08

(a) Hallar $P(X = Y)$.
(b) Determinar si las variables X e Y son independientes.
(c) Calcular $P(X \times Y \geq 0)$
(d) Hallar $P(|X| = 2)$.
(e) Calcular $V(Y)$ y $V(Y|X = -2)$.
(f) Hallar $Cov(X, Y)$.

Solución: La siguiente tabla proporciona las probabilidades conjuntas de las variables X e Y y las distribuciones marginales de las dos variables:

		Y			
	$f(x,y)$	-2	0	2	
	-2	0.07	0.08	0	0.15
X	-1	0.06	0.14	0.10	0.30
	0	0.17	0.11	0.02	0.30
	2	0.05	0.12	0.08	0.25
		0.35	0.45	0.20	1

(a)

$$\begin{aligned} P(X=Y) &= P(X=-2, Y=-2) + P(X=0, Y=0) + P(X=2, Y=2) \\ &= 0.07 + 0.11 + 0.08 \\ &= \mathbf{0.26}. \end{aligned}$$

(b) Las variables aleatorias X e Y serán independientes si y sólo si se verifica

$$P(X=k, Y=m) = P(X=k) \times P(Y=m)$$

para todo $k \in \{-2, -2, 0, 2\}$ y todo $m \in \{-2, 0, 2\}$.
Puesto que (por ejemplo)

$$P(X=-2, Y=2) = 0 \neq 0.03 = 0.15 \times 0.20 = P(X=-2) \times P(Y=2),$$

las variables X e Y **no son independientes**.

(c)

$$\begin{aligned} P(X \times Y \geq 0) &= 1 - P(X \times Y < 0) \\ &= 1 - P(X=-2, Y=2) - P(X=-1, Y=2) - P(X=2, Y=-2) \\ &= 1 - 0 - 0.10 - 0.05 \\ &= \mathbf{0.85} \end{aligned}$$

(d)

$$P(|X|=2) = P(X=-2) + P(X=2) = 0.15 + 0.25 = \mathbf{0.4}$$

(e) La función de masa **marginal** de la variable aleatoria Y es

$$P(Y=-2) = 0.35, \quad P(Y=0) = 0.45, \quad P(Y=2) = 0.20,$$

o escrito abreviadamente

$$Y \equiv \left(\begin{array}{ccc} -2 & 0 & 2 \\ 0.35 & 0.45 & 0.20 \end{array} \right),$$

y, en consecuencia,

$$\begin{aligned} E(Y) &= -2 \times 0.35 + 0 \times 0.45 + 2 \times 0.2 = -0.3, \\ E(Y^2) &= 4 \times 0.35 + 0 \times 0.45 + 4 \times 0.2 = 2.2. \end{aligned}$$

Luego

$$V(Y) = E(Y) - E(Y^2) = 2.2 - (-0.3)^2 = \mathbf{2.11}.$$

Por otra parte, la función de masa de la variable aleatoria Y **condicionada por el suceso** $X = -2$ es

$$P(Y=-2|X=-2) = \frac{7}{15}, \quad P(Y=0|X=-2) = \frac{8}{15}, \;\; P(Y=2|X=-2) = 0,$$

es decir,

$$Y|X=-2 \equiv \left(\begin{array}{cc} -2 & 0 \\ \\ \frac{7}{15} & \frac{8}{15} \end{array} \right),$$

y por tanto,

$$\begin{aligned} E(Y|X=-2) &= -2 \times \frac{7}{15} + 0 \times \frac{8}{15} = -\frac{14}{15}, \\ E(Y^2|X=-2) &= 4 \times \frac{7}{15} + 0 \times \frac{8}{15} = \frac{28}{15}, \end{aligned}$$

y en consecuencia,

$$V(Y|X=-2) = E(Y^2|X=-2) - [E(Y|X=-2)]^2 = \frac{28}{15} - \left(-\frac{14}{15}\right)^2 = \mathbf{0.9956}.$$

(f)

$$\begin{aligned} E(X) &= -2 \times 0.15 - 1 \times 0.3 + 0 \times 0.3 + 2 \times 0.25 = -0.1, \\ E(Y) &= -0.3, \\ E(X \times Y) &= (-2) \times (-2) \times 0.07 + (-1) \times (-2) \times 0.06 + \ldots + 2 \times 2 \times 0.08 = 0.32. \end{aligned}$$

Luego

$$Cov(X,Y) = E(X \times Y) - E(X) \times E(Y) = 0.32 - (-0.1) \times (-0.3) = \mathbf{0.29}.$$