Cálculo de Probabilidades I — Febrero 2014 — Segunda semana

Ejercicio 1. Una urna contiene tres bolas rojas y cinco bolas blancas. Se realiza una primera extracción de dos bolas de la urna (sin reemplazamiento) y, a continuación, se hace una segunda extracción de otras dos bolas (sin reemplazamiento). Sean X e Y el número de bolas rojas en las extracciones primera y segunda, respectivamente.

(a) Determinar la función de probabilidad conjunta de (X, Y).

(b) Hallar la función de probabilidad de X, su media y su varianza. Observar que X e Y tienen la misma distribución, pero no son independientes.

(c) Establecer la expresión

$$\mathrm{E}[Y|X] = 1 - \frac{1}{3}X$$

y deducir el valor del coeficiente de correlación entre X e Y.

Ejercicio 2.

(a) Sean X e Y dos variables aleatorias independientes con distribución geométrica de parámetros $0 < p < 1$ y $0 < q < 1$, respectivamente. Determinar la función de probabilidad de la variable aleatoria $Z = \min\{X, Y\}$.

Una urna contiene bolas rojas, azules y blancas en proporciones p_1, p_2 y p_3, respectivamente, siendo los p_i estrictamente positivos con $p_1 + p_2 + p_3 = 1$. Se extraen sucesivamente bolas de la urna con reemplazamiento. Sea T_1 el número de extracción en el que ha salido, por primera vez, una bola roja; y sea T_2 el número de extracción en el que ha salido, por primera vez, una bola azul.

(b) Hallar la función de probabilidad de $\min\{T_1, T_2\}$, que es el número de extracción en que ha salido, por primera vez, una bola roja o azul.

Solución

Ejercicio 1.
(a) Se observa que los posibles valores de (X, Y) son todos los pares de enteros (x, y) con $0 \le x, y \le 2$ salvo el $(2,2)$. Las probabilidades se calculan en función de las bolas de la primera extracción y, después, teniendo en cuenta las bolas que quedan en la urna, se analiza la segunda extracción.

- $P\{(X,Y) = (0,0)\} = \frac{\binom{3}{0}\binom{5}{2}}{\binom{8}{2}} \cdot \frac{\binom{3}{0}\binom{3}{2}}{\binom{6}{2}} = \frac{1}{14}$.
- $P\{(X,Y) = (1,0)\} = \frac{\binom{3}{1}\binom{5}{1}}{\binom{8}{2}} \cdot \frac{\binom{2}{0}\binom{4}{2}}{\binom{6}{2}} = \frac{3}{14} = P\{(X,Y) = (0,1)\}$.
- $P\{(X,Y) = (2,0)\} = \frac{\binom{3}{2}\binom{5}{0}}{\binom{8}{2}} \cdot \frac{\binom{1}{0}\binom{5}{2}}{\binom{6}{2}} = \frac{1}{14} = P\{(X,Y) = (0,2)\}$.
- $P\{(X,Y) = (1,1)\} = \frac{\binom{3}{1}\binom{5}{1}}{\binom{8}{2}} \cdot \frac{\binom{2}{1}\binom{4}{1}}{\binom{6}{2}} = \frac{2}{7}$.
- $P\{(X,Y) = (2,1)\} = \frac{\binom{3}{2}\binom{5}{0}}{\binom{8}{2}} \cdot \frac{\binom{1}{1}\binom{5}{1}}{\binom{6}{2}} = \frac{1}{28} = P\{(X,Y) = (1,2)\}$.

(b) Se deduce que

$$P\{X = 0\} = \frac{5}{14}, \quad P\{X = 1\} = \frac{15}{28}, \quad P\{X = 2\} = \frac{3}{28},$$

por lo que $E[X] = \frac{3}{4}$ y $\mathrm{Var}(X) = \frac{45}{112}$. Se tiene, efectivamente, que X e Y tienen la misma distribución. No son independientes porque

$$P\{X = 2, Y = 2\} = 0 \quad \text{mientras que} \quad P\{X = 2\} = P\{Y = 2\} > 0.$$

(c) Condicionando por $X = 0$ se tiene

$$P\{Y = 0|X = 0\} = \frac{1}{5}, \quad P\{Y = 1|X = 0\} = \frac{3}{5}, \quad P\{Y = 2|X = 0\} = \frac{1}{5},$$

por lo que $E[Y|X = 0] = 1$. Análogamente, se obtiene $E[Y|X = 1] = \frac{2}{3}$ y $E[Y|X = 2] = \frac{1}{3}$. Se deduce la expresión $E[Y|X] = 1 - \frac{1}{3}X$ para todos los valores que toma la variable aleatoria X.

La recta de regresión de Y sobre X tiene, pues, ecuación $y = 1 - \frac{1}{3}x$. Su pendiente es igual a

$$-\frac{1}{3} = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)} = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}}$$

donde se ha hecho uso de que $\mathrm{Var}(X) = \mathrm{Var}(Y)$. Esta última expresión se corresponde con el coeficiente de correlación; luego $\rho = -1/3$.

Ejercicio 2.
(a) La función de probabilidad de X es $\mathrm{P}\{X = n\} = (1-p)^{n-1}p$ para $n \geq 1$, y análogamente para Y. Se tiene pues que

$$\mathrm{P}\{X > n\} = (1-p)^n \quad \text{y} \quad \mathrm{P}\{Y > n\} = (1-q)^n \quad \text{para } n \geq 0.$$

Siendo X e Y independientes,

$$\begin{aligned} \mathrm{P}\{Z > n\} &= \mathrm{P}\{X > n, Y > n\} \\ &= \mathrm{P}\{X > n\} \cdot \mathrm{P}\{Y > n\} \\ &= \big((1-p)(1-q)\big)^n. \end{aligned}$$

Se deduce que Z tiene distribución geométrica de parámetro $1-(1-p)(1-q)$.
(b) Las variables aleatorias T_1 y T_2 tienen distribución geométrica de parámetros p_1 y p_2, respectivamente. Fijado $n \geq 0$, calculamos $\mathrm{P}\{\text{mín}\{T_1, T_2\} > n\}$. Se tiene la siguiente igualdad de sucesos.

$$\{\text{mín}\{T_1, T_2\} > n\} = \{T_1 > n\} \cap \{T_2 > n\} = \{\text{las primeras } n \text{ bolas son blancas}\}.$$

Por tanto,

$$\mathrm{P}\{\text{mín}\{T_1, T_2\} > n\} = p_3^n$$

y resulta que $\text{mín}\{T_1, T_2\}$ tiene distribución geométrica de parámetro $1 - p_3$.

Se observa que esta distribución no se corresponde con la del apartado (a), dado que $1-(1-p_1)(1-p_2) \neq 1-p_3$. Esto es porque las variables aleatorias T_1 y T_2 no son independientes. En efecto, si $T_1 = n$ entonces necesariamente $T_2 \neq n$, lo que implica que las variables no son independientes.