Cálculo de Probabilidades II— Febrero 2015 — Primera semana

Ejercicio 1. El vector aleatorio (X, Y) tiene función de densidad

$$f(x, y) = \frac{12}{7} x(x + y) \quad \text{para } 0 \leq x \leq 1 \text{ y } 0 \leq y \leq 1.$$

(a) Calcular la función de densidad de Y condicionada por $X = x$, para $0 \leq x \leq 1$, y calcular $\mathrm{E}[Y|X = x]$.

(b) Determinar la función de distribución F del vector aleatorio (X, Y).

(c) Fijado un valor $0 \leq x \leq 1$, calcular $\mathrm{E}[Y|X \leq x]$.

Ejercicio 2. En un juego de dardos, un lanzador se sitúa a una distancia L de una diana rectangular de altura d colgada en una pared. El dardo es lanzado desde la altura media de la diana, cuya distancia al suelo es mayor que $\sqrt{3}L/3$, mientras que $d/2 < \sqrt{3}L/3$, de modo que la diana no llega al suelo. Debido al mal pulso del jugador, la trayectoria del dardo (que se supone recta) hace con la horizontal un ángulo aleatorio A, cuya función de densidad es

$$f(\alpha) = \cos \alpha \quad \text{para } -\tfrac{\pi}{6} \leq \alpha \leq \tfrac{\pi}{6}.$$

(a) Efectuado el lanzamiento, el dardo queda a una distancia D del centro de la línea media de la diana. Hallar la función de distribución de la variable aleatoria D y calcular la probabilidad de que el dardo haya quedado fuera de la diana.

(b) Se utiliza un sistema de puntuación continua, de tal forma que la puntuación del lanzamiento es igual a $10(1 - \frac{2D}{d})$ cuando el dardo entra en la diana, y cero si el dardo queda fuera de la diana. Determinar la puntuación esperada de un lanzamiento del jugador.

Solución

Ejercicio 1.
(a) La función de densidad marginal de X es $f_X(x) = \frac{6}{7}x(2x+1)$ para $0 \leq x \leq 1$. Por tanto, la función de densidad de Y condicionada por $X = x$, para $0 < x \leq 1$ es

$$f_{Y|X}(y|x) = \frac{2(x+y)}{2x+1} \quad \text{para } 0 \leq y \leq 1.$$

Aunque la función de densidad de X tome el valor $f_X(0) = 0$, y en principio no se pueda decir que $f_{Y|X}(y|0) = f(0,y)/f_X(0)$, utilizando directamente la definición de la pág. 162 puede verse que la expresión de la densidad condicionada es también válida para $x = 0$. Por tanto, la esperanza condicionada pedida es

$$\mathrm{E}[Y|X = x] = \int_0^1 y\frac{2(x+y)}{2x+1}dy = \frac{3x+2}{3(2x+1)} \quad \text{para } 0 \leq x \leq 1.$$

(b) La función de distribución F en el punto (x, y), con $0 \leq x, y \leq 1$ vale

$$F(x,y) = \int_0^x \int_0^y f(u,v)dvdu = \frac{4}{7}x^3y + \frac{3}{7}x^2y^2.$$

Por tanto

$$F(x,y) = \begin{cases} 1 & \text{si } x \geq 1 \text{ e } y \geq 1; \\ \frac{4}{7}x^3 + \frac{3}{7}x^2 & \text{si } 0 \leq x \leq 1 \text{ e } y \geq 1; \\ \frac{4}{7}y + \frac{3}{7}y^2 & \text{si } x \geq 1 \text{ y } 0 \leq y \leq 1; \\ \frac{4}{7}x^3y + \frac{3}{7}x^2y^2 & \text{si } 0 \leq x, y \leq 1; \\ 0 & \text{si } x \leq 0 \text{ o } y \leq 0. \end{cases}$$

(c) Para cada $0 \leq x \leq 1$, se tiene que

$$\mathrm{P}\{Y \leq y | X \leq x\} = \frac{F(x,y)}{F(x,1)} = \frac{4xy+3y^2}{4x+3} \quad \text{para } 0 \leq y \leq 1,$$

por lo que la función de densidad de Y condicionada por $X \leq x$ es

$$g(y|x) = \frac{2(2x+3y)}{4x+3} \quad \text{para } 0 \leq y \leq 1.$$

Resulta entonces que

$$\mathrm{E}[Y|X \leq x] = \int_0^1 yg(y|x)dy = \frac{2(x+1)}{4x+3} \quad \text{para } 0 \leq x \leq 1.$$

(Para el caso particular $x = 0$, en el que $F(0,1) = 0$ se hace el mismo análisis que anteriormente.) Se observa que $\mathrm{E}[Y|X \leq x] \neq \mathrm{E}[Y|X = x]$ salvo —como era de esperar— cuando es $x = 0$.

Ejercicio 2.

(a) La variable aleatoria D toma valores en el intervalo $[0, \sqrt{3}L/3]$ y, en particular, el dardo siempre llega a la pared y no da en el suelo. La función de distribución de la variable A es $F(\alpha) = \frac{1}{2} + \operatorname{sen}\alpha$ para $-\pi/6 \leq \alpha \leq \pi/6$. Se cumple la siguiente igualdad de sucesos

$$\{D \leq z\} = \big\{|A| \leq \arctan(z/L)\big\} \quad \text{para cada } 0 \leq z \leq \sqrt{3}L/3.$$

Por tanto,

$$\mathrm{P}\{D \leq z\} = 2 \operatorname{sen}\arctan(z/L) = \frac{2z}{\sqrt{L^2 + z^2}} \quad \text{para } 0 \leq z \leq \sqrt{3}L/3.$$

Su función de densidad es

$$g(z) = \frac{2L^2}{(L^2 + z^2)^{3/2}} \quad \text{para } 0 \leq z \leq \sqrt{3}L/3.$$

Hay probabilidad positiva de que el dardo no dé en la diana, puesto que $d/2 < \sqrt{3}L/3$. La probabilidad de que el dardo no dé en la diana es

$$\mathrm{P}\{D > d/2\} = 1 - \frac{2d}{\sqrt{4L^2 + d^2}}.$$

(b) La puntuación esperada del jugador es

$$\int_0^{d/2} 10\Big(1 - \frac{2z}{d}\Big) f(z) dz,$$

que es igual a

$$10\mathrm{P}\{D \leq d/2\} - \frac{40L^2}{d} \int_0^{d/2} \frac{z}{(L^2 + z^2)^{3/2}} dz.$$

La primitiva de la función dentro de la integral es $-\frac{1}{\sqrt{L^2+z^2}}$, por lo que operando resulta que la puntuación esperada es

$$20 \frac{\sqrt{d^2 + 4L^2} - 2L}{d}$$

Esta cantidad se puede expresar en función de la razón entre la distancia a la diana y la mitad de la longitud de la diana $2L/d = r$ y su valor es

$$20(\sqrt{r^2 + 1} - r).$$

Cálculo de Probabilidades II— Febrero 2015 — Segunda semana

Ejercicio 1. Para cada $n \geq 1$, la variable aleatoria X_n tiene función de densidad

$$f_n(x) = ne^{-x}\big(1 - e^{-x}\big)^{n-1} \quad \text{para } x \geq 0.$$

Determinar el límite en distribución de las siguientes sucesiones de variables aleatorias:

(a) $\big\{e^{-X_n}\big\}_{n\geq 1}$.

(b) $\big\{n\big(1 - e^{-X_n}\big)\big\}_{n\geq 1}$.

(c) $\big\{ne^{-X_n}\big\}_{n\geq 1}$.

Ejercicio 2. Sea F la función de distribución de la variable aleatoria real X, y sea G la función de distribución de la variable aleatoria $F(X)$.

(a) Probar que $G(y) \leq y$ para todo $0 \leq y \leq 1$.

(b) Sean D_F y D_G los puntos de discontinuidad de F y G, respectivamente. Probar que: dado $0 \leq y \leq 1$,

$$y \in D_G \iff \{x \in I\!R : F(x) = y\} \cap D_F \neq \emptyset.$$

(c) Probar que si F es continua entonces G es la distribución uniforme en el intervalo $[0, 1]$ (se puede hacer este apartado sin haber resuelto el (b)).

Solución

Ejercicio 1. La función de distribución de X_n es

$$F_n(x) = (1 - e^{-x})^n \quad \text{para } x \geq 0.$$

(a) Sea $Y_n = e^{-X_n}$. La variable Y_n toma valores en el intervalo $(0, 1]$. Sea $0 < y \leq 1$. Se tiene que

$$\mathrm{P}\{Y_n \leq y\} = \mathrm{P}\{e^{-X_n} \leq y\} = \mathrm{P}\{X_n \geq -\log y\} = 1 - F_n(-\log y) = 1 - (1 - y)^n.$$

Su límite cuando $n \to \infty$ es 1. Por tanto, $Y_n \xrightarrow{d} 0$.

(b) Sea $Z_n = n(1 - e^{-X_n})$, que toma valores en $[0, n)$. Fijado $z \geq 0$ se toma n tal que $n > z$. Se cumple

$$\mathrm{P}\{Z_n \leq z\} = \mathrm{P}\Big\{X_n \leq -\log\Big(1 - \frac{z}{n}\Big)\Big\} = \Big(\frac{z}{n}\Big)^n.$$

Su límite cuando $n \to \infty$ es 0. Por tanto, Z_n no converge en ley a a ninguna distribución.

(c) Sea $W_n = ne^{-X_n}$, que toma valores en $(0, n]$. Fijado $z \geq 0$ se toma n tal que $n > z$. Se cumple

$$\mathrm{P}\{W_n \leq z\} = \mathrm{P}\Big\{X_n \geq -\log\frac{z}{n}\Big\} = 1 - \Big(1 - \frac{z}{n}\Big)^n.$$

Su límite cuando $n \to \infty$ es $1 - e^{-z}$. Por tanto, W_n converge en ley a una distribución exponencial de parámetro 1.

Ejercicio 2.
(a) Fijado $0 \leq y \leq 1$, se define el conjunto

$$C = \{x \in \mathbb{R} : F(x) \leq y\}.$$

Se tiene que $G(y) = \mathrm{P}\{X \in C\}$.

Si $y = 1$ entonces $C = \mathbb{R}$ y se cumple que $G(1) = 1$. Supóngase ahora que $0 \leq y < 1$. Si $y = 0$ y $F(x) > 0$ para todo $x \in \mathbb{R}$ entonces C es vacío y $G(0) = 0$. En caso contrario, C es un intervalo no vacío que es necesariamente de la forma $(-\infty, x_0]$ o $(-\infty, x_0)$ para algún valor $x_0 \in \mathbb{R}$.

- Si $C = (-\infty, x_0]$ entonces $G(y) = \mathrm{P}\{X \in C\} = F(x_0) \leq y$ (porque el propio punto x_0 pertenece a C).
- Si $C = (-\infty, x_0)$ entonces $G(y) = \mathrm{P}\{X \in C\} = F(x_0^-) = \lim_{x \uparrow x_0} F(x)$. Puesto que los puntos $x < x_0$ pertenecen a C verifican $F(x) \leq y$; de ahí que también sea $G(y) \leq y$.

En conclusión, para todo $0 \le y \le 1$ se tiene $G(y) \le y$ con, además, $G(1) = 1$.

Dadas dos funciones de distribución F_1 y F_2 se dice que $F_1 \preceq F_2$ (leído F_2 domina a F_1) cuando $F_1(x) \ge F_2(x)$ para todo $x \in I\!R$. Se interpreta como que la distribución F_2 toma los valores "grandes" con mayor probabilidad que F_1, puesto que es $1 - F_1(x) \le 1 - F_2(x)$, que son las probabilidades de los intervalos $(x, +\infty)$ bajo las medidas de probabilidad asociadas a F_1 y F_2, respectivamente.

En este ejercicio se ha probado que la distribución de $F(X)$ domina a la distribución uniforme en el intervalo $[0, 1]$ para cualquier F.

(b) Supongamos que $0 < y \le 1$ es un punto de discontinuidad de G (nótese que $0 \notin D_G$). Esto significa que hay probabilidad positiva de que sea $F(X) = y$. Sea

$$K = \{x \in I\!R : F(x) = y\}, \quad \text{por lo que } \mathrm{P}\{X \in K\} > 0.$$

El conjunto K es un intervalo de la forma $[x_0, y_0)$ para algunos $x_0 < y_0 \le +\infty$, o bien de la forma $[x_0, y_0]$ para $x_0 \le y_0 < \infty$. En cualquier caso se tiene que $\mathrm{P}\{X \in K\} = F(x_0) - F(x_0^-)$, por lo que x_0 es un punto de discontinuidad de F. Se ha probado pues que $\{x \in I\!R : F(x) = y\} \cap D_F \ne \emptyset$.

Recíprocamente, sea x un punto de discontinuidad de F con $F(x) = y$. Puesto que hay probabilidad positiva de que sea $X = x$, también habrá probabilidad positiva de que sea $F(X) = F(x) = y$, luego es $y \in D_G$.

Como corolario se obtiene que G es continua si y solamente si F es continua.

(c) Ya se sabe que $G(1) = 1$ y $G(0) = 0$. Dado $0 < y < 1$, utilizando la notación del apartado (a), se tiene que el conjunto $C = \{x \in I\!R : F(x) \le y\}$ es un intervalo cerrado (por ser F continua) de la forma $(-\infty, x_0]$. Además, se tiene que $F(x_0) = y$, pues si fuese $F(x_0) < y$ existiría $\epsilon > 0$ con $F(x_0 + \epsilon) < y$, por lo que sería $x_0 + \epsilon \in C$. Por tanto, $G(y) = \mathrm{P}\{X \in C\} = F(x_0) = y$.

Se ha establecido que si F es continua entonces G es la distribución uniforme en el intervalo $[0, 1]$.

Cálculo de Probabilidades 2 — Septiembre 2015

Cuestión 1 (1 punto). Dar la definición de función de distribución sobre $\mathbb{R}$.

Cuestión 2 (1 punto). Dar la definición de matriz de covarianzas de una variable aleatoria k-dimensional y enunciar sus principales propiedades.

Ejercicio 1 (4 puntos). Sean X e Y dos variables aleatorias tales que X tiene distribución geométrica de parámetro p tomando valores en $1, 2, 3, \ldots$, es decir,

$$\mathrm{P}\{X = k\} = (1-p)^{k-1}p \quad \text{para } k \geq 1$$

para un cierto valor $0 < p < 1$, y que, condicionada por $X = k$, la variable Y tiene distribución exponencial de parámetro k.

(a) Calcular la función de densidad de Y y hallar la función de probabilidad de X condicionada por $Y = y$, siendo $y \geq 0$.

(b) Determinar la función de distribución de Y condicionada por $X > k$ para cada $k \geq 0$.

Ejercicio 2 (4 puntos). Las variables aleatorias $\{X_n\}_{n\geq 1}$ son independientes y tienen todas función de densidad

$$f(x) = \frac{2}{x^3} \quad \text{si } x \geq 1$$

y $f(x) = 0$ si $x < 1$. Hallar el límite en distribución de las siguientes sucesiones de variables aleatorias cuando $n \to \infty$:

(a) $n \cdot \big(\min\{X_1, \ldots, X_n\} - 1\big)$;

(b) $\max\{X_1, \ldots, X_n\}/\sqrt{n}$.

Solución

Ejercicio 1. (a). La función de densidad de Y viene dada por

$$f_Y(y) = \sum_{k=1}^{\infty} (1-p)^{k-1} p \cdot k e^{-ky} \quad \text{para } y \geq 0,$$

que es la suma de las probabilidades marginales de X multiplicadas por la densidad de Y condicionada por $X = k$. Se obtiene

$$f_Y(y) = pe^{-y}\big(1-(1-p)e^{-y}\big)^{-2} \quad \text{para } y \geq 0.$$

La función de probabilidad de X condicionada por $Y = y$, para $y \geq 0$, es

$$\frac{(1-p)^{k-1}p \cdot ke^{-ky}}{f_Y(y)} = k\big(1-(1-p)e^{-y}\big)^2\big((1-p)e^{-y}\big)^{k-1} \quad \text{para } k \geq 1.$$

Por tanto, la distribución de $X - 1$ condicionada por $Y = y$ es binomial negativa de parámetros 2 y $1-(1-p)e^{-y}$.

(b). Se calcula

$$\begin{aligned} \mathrm{P}\{Y > y \mid X > k\} &= \frac{\mathrm{P}\{Y > y, X > k\}}{\mathrm{P}\{X > k\}} \\ &= \frac{\sum_{r=k+1}^{\infty} \mathrm{P}\{Y > y, X = r\}}{(1-p)^k}. \end{aligned}$$

Puesto que la distribución de Y condicionada por $X = r$ es exponencial de parámetro r, se tiene

$$\mathrm{P}\{Y > y, X = r\} = \mathrm{P}\{Y > y \mid X = r\}\mathrm{P}\{X = r\} = e^{-ry} \cdot (1-p)^{r-1}p,$$

luego

$$\sum_{r=k+1}^{\infty} \mathrm{P}\{Y > y, X = r\} = \frac{e^{-y(k+1)}(1-p)^k p}{1-(1-p)e^{-y}}.$$

Se deduce finalmente que

$$\mathrm{P}\{Y > y \mid X > k\} = \frac{pe^{-y(k+1)}}{1-(1-p)e^{-y}},$$

deduciéndose la expresión de la función de distribución.

Ejercicio 2. La función de distribución F asociada a f es

$$F(x) = 1 - \frac{1}{x^2} \quad \text{para } x \geq 1$$

y $F(x) = 0$ para $x < 1$.
(a). La variable aleatoria $Y_n = n \cdot \big(\text{mín}\{X_1, \ldots, X_n\} - 1\big)$ toma valores en $[0, \infty)$. Dado $y \geq 0$ se tiene

$$\begin{aligned} \mathrm{P}\{Y_n \leq y\} &= 1 - \mathrm{P}\Big\{\text{mín}\{X_1, \ldots, X_n\} > 1 + \frac{y}{n}\Big\} \\ &= 1 - \big(1 - F(1 + y/n)\big)^n \\ &= 1 - \Big(1 + \frac{y}{n}\Big)^{-2n}, \end{aligned}$$

cuyo límite cuando $n \to \infty$ es $1 - e^{-2y}$. Por tanto, $\{Y_n\}_{n \geq 1}$ converge en distribución a una ley exponencial de parámetro 2.

(b). Dado $z > 0$ se toma n lo bastante grande para que sea $z\sqrt{n} \geq 1$. Siendo

$$Z_n = \text{máx}\{X_1, \ldots, X_n\}/\sqrt{n},$$

se tiene

$$\begin{aligned} \mathrm{P}\{Z_n \leq z\} &= \mathrm{P}\Big\{\text{máx}\{X_1, \ldots, X_n\} \leq z\sqrt{n}\Big\} \\ &= \big(F(z\sqrt{n})\big)^n \\ &= \Big(1 - \frac{1}{nz^2}\Big)^n, \end{aligned}$$

cuyo límite cuando $n \to \infty$ es e^{-1/z^2}. Cuando $z \leq 0$, es claro que $\mathrm{P}\{Z_n \leq z\} = 0$. Por tanto, $\{Z_n\}_{n \geq 1}$ converge en ley a la distribución dada por

$$G(z) = \begin{cases} e^{-1/z^2} & \text{si } z > 0, \\ 0 & \text{si } z \leq 0. \end{cases}$$