Cálculo de Probabilidades II — Febrero 2017 — Primera semana

Cuestión (2 puntos). Enunciar la desigualdad de Tchebychev y deducir que si la sucesión de variables aleatorias $\{X_n\}_{n\geq 1}$ converge a X en media de orden p, para $p > 0$, entonces también converge en probabilidad.

Ejercicio 1 (4 puntos). El punto P se toma con distribución uniforme en el interior de un triángulo rectángulo isósceles con cateto de longitud uno. Sea U la distancia de P al cateto más próximo y V la distancia de P a la hipotenusa.

(a) Calcular $\mathrm{P}\{U \geq u, V \geq v\}$ siendo $u \geq 0$, $v \geq 0$, y $2u + \sqrt{2}v \leq 1$.

(b) Del apartado anterior, deducir las funciones de densidad marginales de U y V, y calcular sus esperanzas. Determinar la función de densidad conjunta de (U, V).

Ejercicio 2 (6 puntos). La variable aleatoria (X, Y) tiene función de densidad

$$f(x, y) = 2(x + y) \quad \text{para } (x, y) \in C,$$

y $f(x, y) = 0$ si $(x, y) \notin C$, siendo $C = \{(x, y) \in \mathbb{R}^2 : 0 \leq y \leq x \leq 1\}$.

(a) Dado $(x, y) \in C$, calcular el valor $F(x, y)$ de la función de distribución de (X, Y) en el punto (x, y).

(b) Para cada $0 \leq x \leq 1$, calcular $\mathrm{E}[Y \mid X = x]$.

(c) Para cada $0 \leq x \leq 1$, calcular $\mathrm{E}[Y \mid X \leq x]$.

Nota máxima: 10 puntos.

Solución

Ejercicio 1. (a). Se supondrá que el triángulo del enunciado es el de vértices $(0,0)$, $(1,0)$, $(0,1)$. Se tiene que el vector aleatorio (U,V) toma valores en el conjunto

$$C = \{(u,v) \in \mathbb{R}^2 : u \geq 0,\ v \geq 0,\ 2u + \sqrt{2}v \leq 1\}.$$

En efecto, la variable V toma valores en el intervalo $[0, \sqrt{2}/2]$ (la máxima distancia a la hipotenusa se da cuando $P = (0,0)$) y, para cada valor $0 \leq v \leq \sqrt{2}/2$, la máxima distancia al cateto más próximo se da en $P = (\frac{1-\sqrt{2}v}{2}, \frac{1-\sqrt{2}v}{2})$.
Para (u,v) en el conjunto C, se da el suceso $\{U \geq u, V \geq v\}$ cuando P pertenece al triángulo de vértices (u,u), $(u, 1-u-\sqrt{2}v)$, $(1-u-\sqrt{2}v, u)$. Por tanto,

$$\mathrm{P}\{U \geq u, V \geq v\} = (1 - 2u - \sqrt{2}v)^2.$$

(b). Haciendo $v = 0$ en la expresión de (a), se obtiene $\mathrm{P}\{U \geq u\} = (1-2u)^2$, de donde la función de densidad de U es $f_U(u) = 4(1-2u)$ para $0 \leq u \leq 1/2$. Su esperanza es $\mathrm{E}[U] = 1/6$.
Análogamente, haciendo $u = 0$ en (a) se obtiene

$$\mathrm{P}\{V \geq v\} = (1 - \sqrt{2}v)^2 \quad \text{para } 0 \leq v \leq \sqrt{2}/2.$$

Su función de densidad es $f_V(v) = 2\sqrt{2} - 4v$ y su esperanza $\mathrm{E}[V] = \frac{\sqrt{2}}{6}$.
Finalmente, la función de distribución F del vector (U,V) es igual a

$$F(u,v) = 1 - \mathrm{P}\{U \geq u\} - \mathrm{P}\{V \geq v\} + P\{U \geq u, V \geq v\}.$$

Derivando respecto de u y v se llega a que la función de densidad de (U,V) es constante e igual a $4\sqrt{2}$ en C. Por tanto, (U,V) se distribuye uniformemente en C.
En este ejercicio, obsérvese que no es relevante que las desigualdades sean estrictas o no, puesto que se trata de distribuciones absolutamente continuas.

Ejercicio 2. (a). El valor de la función de distribución $F(x,y)$ es la probabilidad del trapecio con vértices $(0,0)$, (y,y), (x,y), $(x,0)$ que se escribe como $0 \leq v \leq y$, $v \leq u \leq x$. Por tanto

$$F(x,y) = 2\int_0^y \int_v^x (u+v)\,du\,dv = x^2y + xy^2 - y^3$$

para $(x,y) \in C$.
(b). Para cada $0 \leq x \leq 1$ se tiene que $F(x,x) = x^3$ es la función de distribución marginal de X con densidad, por tanto, $f_X(x) = 3x^2$ en ese intervalo. La función de densidad de Y condicionada por $X = x$, con $0 < x \leq 1$, es

$$f(y|x) = \frac{2(x+y)}{3x^2} \quad \text{para } 0 \leq y \leq x$$

por lo que

$$\mathrm{E}[Y \mid X = x] = \int_0^x y f(y|x)\,dy = \frac{5x}{9}.$$

Esta expresión también es válida para $x = 0$.
(c). Fijado un valor $0 < x \leq 1$ y un valor $0 \leq y \leq x$ se tiene que

$$\mathrm{P}\{Y \leq y \mid X \leq x\} = \frac{F(x,y)}{F(x,x)} = \frac{x^2y + xy^2 - y^3}{x^3}.$$

Esta expresión, como función de y, es la función de distribución de Y condicionada por $X \leq x$. La función de densidad asociada es

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

Por tanto,

$$\mathrm{E}[Y \mid X \leq x] = \int_0^x y g(y|x) dy = \frac{5x}{12}.$$

La expresión es válida para $x = 0$.

Cálculo de Probabilidades II— Febrero 2017 — Segunda semana

Cuestión 1. (2 puntos) Caracterizar la convergencia en distribución de una sucesión $\{X_n\}$ de variables aleatorias en términos (a) de sus funciones de distribución, (b) en términos de las esperanzas $\mathrm{E}[g(X_n)]$ y (c) en términos de sus funciones características.

Ejercicio 1. (6 puntos) Un punto P del cuadrado $[0,1] \times [0,1]$ tiene coordenadas (X, Y). Otro punto Q se elige al azar en $[0,1] \times [0,1]$.

(a) Calcular la probabilidad de que la recta PQ tenga pendiente positiva.

(b) (X, Y) se elige con densidad $f(x, y) = kxy$ en la región $\{X < Y\}$ de $[0,1] \times [0,1]$ y Q al azar en $[0,1] \times [0,1]$, independientemente de P. Determinar la probabilidad de que la pendiente de la recta PQ sea positiva.

(c) P y Q se eligen al azar e independientemente en $[0,1] \times [0,1]$. Determinar la densidad de la pendiente de la recta PQ.

Ejercicio 2. (4 puntos)

(a) Si $\{X_k\}_{k \geq 1}$ es una sucesión de variables aleatorias incorreladas, con $\mathrm{E}[X_k] = \mu$ y $\sigma^2(X_k) \leq C$, probar que $S_n = 1/n \sum_{k=1}^n X_k$ converge en media cuadrática hacia μ.

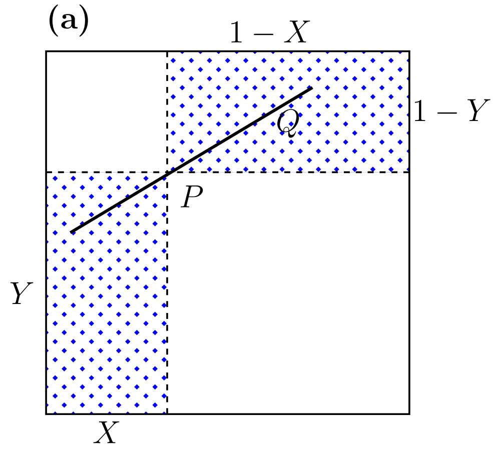
(b) Aplicarlo a establecer que si $\{X_k\}_{k \geq 1}$ son variables independientes con distribución común uniforme en $[-1, 1]$, entonces $Z_n = 1/n \sum_{k=1}^n X_k^2$ converge en media cuadrática y determinar el límite.

Nota máxima: 10 puntos.

Solución

Ejercicio 1

(a)



Fijado el punto P, la recta PQ tiene pendiente positiva cuando Q se encuentre en uno de los dos rectángulos señalados en la figura, delimitados por las paralelas a los ejes que pasan por P. Las áreas de esos rectángulos son XY y $(1-X)(1-Y)$ respectivamente. Luego, si R es la pendiente de la recta PQ, se cumple

$$\mathrm{P}\{R > 0 \mid X, Y\} = XY + (1-X)(1-Y)$$

condicionado por el hecho de que se conocen las coordenadas (X, Y) de P.

(b) La densidad $f(x, y)$ debe integrar 1 en la región $\{0 \leq x \leq y \leq 1\}$; es decir

$$1 = k \int_0^1 \int_x^1 xy \, dy \, dx = \frac{k}{2} \int_0^1 x(1-x^2) \, dx = k/8$$

de donde $k = 8$. Por tanto, la probabilidad de que la pendiente sea positiva resulta

$$\begin{aligned} \mathrm{P}\{R > 0\} &= 8 \int_0^1 \int_x^1 [xy + (1-x)(1-y)] \, xy \, dy \, dx \\ &= \frac{8}{3} \int_0^1 x^2(1-x^3) \, dx + \frac{4}{3} \int_0^1 x(1-x)[1 - 3x^2 + 2x^3] \, dx = \frac{5}{9}. \end{aligned}$$

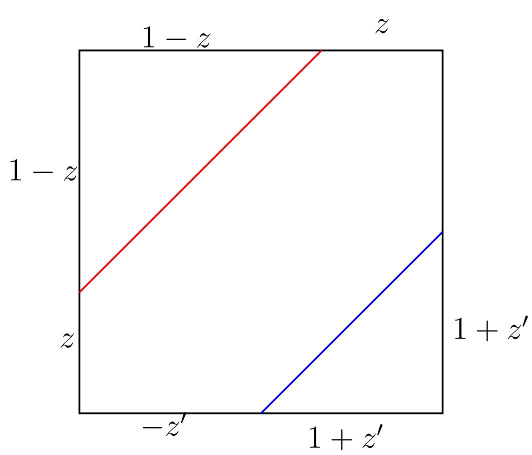
(c) Sean (U, V) las coordenadas de Q, de manera que la pendiente de la recta PQ es $R = (V-Y)/(U-X)$. Como X, Y, U, V son independientes y con distribución uniforme en $[0, 1]$, las variables $V - Y$ y $U - X$ son independientes y tienen la misma distribución.

Concretamente la distribución de $U - X$ es

$$\mathrm{P}\{U - X \leq z\} = \begin{cases} 1 - (1-z)^2/2 & \text{si } z \geq 0 \\ (1+z)^2/2 & \text{si } z < 0 \end{cases}$$

ya que, si $z > 0$, $\{U - X \leq z\}$ es la región por debajo de la línea roja; mientras que, si $z < 0$, se trata de la región por debajo de la línea azul. Por consiguiente, $U - X$ tiene densidad

$$g(z) = 1 - |z| \quad \text{para } z \in [-1, 1].$$



La independencia de $Z = U - X$ y $T = V - Y$ da como densidad conjunta de (Z, T)

$$g(z,t) = (1 - |z|)(1 - |t|) \quad \text{para } z, t \in [-1, 1].$$

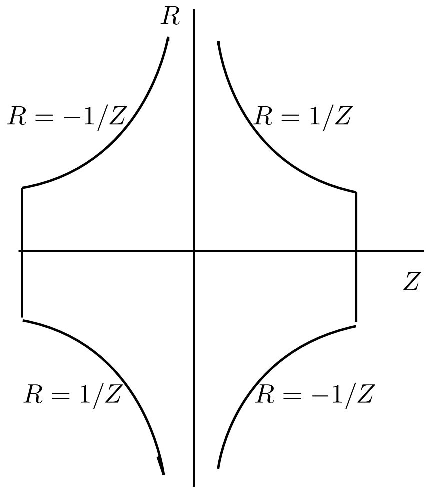
Ahora la densidad de $R = T/Z$ se obtiene mediante el cambio

$$\begin{cases} Z = Z \\ R = T/Z \end{cases} \quad \text{cuyo inverso} \quad \begin{cases} Z = Z \\ T = ZR \end{cases}$$

tiene jacobiano de valor absoluto $|Z|$, de modo que la densidad de (Z, R) es

$$h(z,r) = (1 - |z|)(1 - |zr|)|z|$$

en la región transformada que se muestra en la figura (los lados $Z = \pm 1$ se conservan, mientras que $T = 1$ y $T = -1$ se transforman en las hipérbolas $R = 1/Z$ y $R = -1/Z$).



Por último, hay que integrar en z para obtener la marginal de R. Teniendo en cuenta la simetría de $h(z, r)$ se obtiene:
Para $r \in [0, 1]$

$$h(r) = 2\int_0^1 (1 - z)(1 - zr)z\,dz = \frac{2 - r}{6}.$$

Para $r > 1$

$$h(r) = 2\int_0^{1/r} (1 - z)(1 - zr)z\,dz = \frac{2r - 1}{6r^3}.$$

Mientras que si r es negativo

$$h(r) = \frac{2 + r}{6} \quad \text{si } r > -1 \quad \text{y} \quad h(r) = \frac{2r + 1}{6r^3} \quad \text{si } r < -1.$$

Ejercicio 2
(a) Como la varianza de una suma de variables incorreladas es la suma de las varianzas de los sumandos, será

$$\mathrm{E}[(S_n/n - \mu)^2] = \sigma^2(S_n/n) = \frac{1}{n^2}\sum_{k=1}^{n} \sigma^2(X_k) \leq \frac{C}{n} \longrightarrow 0$$

con lo cual $S_n/n \to \mu$ en $\mathcal{L}^2$.

(b) En el caso de que $\{X_k\}$ sean variables independientes, también $\{X_k^2\}$ son independientes y, por consiguiente, incorreladas. Además

$$\mathrm{E}[X_k^2] = \frac{1}{2}\int_{-1}^{1} x^2\,dx = \frac{1}{3} \quad \text{y} \quad \sigma^2(X_k^2) = \frac{1}{2}\int_{-1}^{1} x^4\,dx - \Big(\frac{1}{3}\Big)^2 = \frac{4}{45}.$$

De acuerdo con (a), se cumple $Z_n/n \longrightarrow 1/3$ en $\mathcal{L}^2$.

Nota: La interpretación de la afirmación $Z_n/n \to 1/3$ en probabilidad es que, en dimensión n muy grande, la transformación

$$(x_1, x_2, \ldots, x_n) \mapsto (x_1, x_2, \ldots, x_{n-1}, (x_1^2 + x_2^2 + \cdots + x_n^2)^{1/2})$$

transforma cada punto del cubo $[-1, 1]^n$ en un punto que está en la capa de grosor 2ε alrededor de la semiesfera $(x_1^2 + x_2^2 + \cdots + x_n^2)^{1/2} = \sqrt{n/3}$, con probabilidad muy próxima a 1 (dependiente de ε y n). Si ε es muy pequeño, la mayor parte del volumen de $[-1, 1]^n$ acaba casi sobre la superficie de la semiesfera.

Cálculo de Probabilidades II — Septiembre 2017

Cuestión (2 puntos). Definir la convolución de dos funciones de distribución F y G. Indicar qué ocurre cuando una de ellas es absolutamente continua. Relacionar la convolución con el Cálculo de Probabilidades.

Ejercicio (10 puntos). Se elige al azar un punto (X, Y) uniformemente en el cuadrado $[0, 1]^2$ y se considera el triángulo T de vértices $(0, 0)$, $(X, 1)$ y $(1, Y)$.

(a) Determinar la distribución del área Z del triángulo T y su moda.

(b) Calcular la media y la desviación típica de Z.

(c) Hallar la probabilidad de que pertenezca a T un punto Q elegido al azar e independientemente en el triángulo de vértices $(0, 0)$, $(1, 0)$ y $(0, 1)$.

(d) Si se repite $n = 2$ veces independientemente la elección del punto (X, Y), hallar la probabilidad de que el punto Q pertenezca a la intersección de los dos triángulos T_1 y T_2 obtenidos.

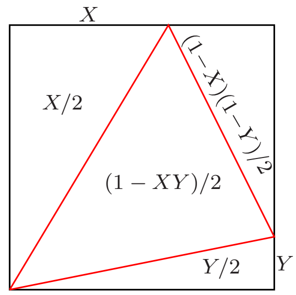
(e) Se repite $n = 100$ veces independientemente la elección del punto (X, Y) para formar los triángulos correspondientes: $T_1, T_2, \dots, T_n$. Después se elige al azar un punto P en el cuadrado $[0, 1]^2$ y se calcula K = *número esperado de triángulos* T_i *a los que pertenece* P (con los triángulos T_i fijos). Indicar, mediante la aproximación oportuna, la probabilidad de que sea $36 \leq K \leq 39$.

Nota máxima: 10 puntos

Solución

Ejercicio 1.

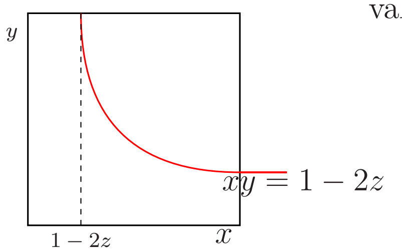
(a) El área Z se expresa



$$Z = 1 - \frac{X}{2} - \frac{Y}{2} - \frac{(1-X)(1-Y)}{2} = \frac{1-XY}{2}.$$

(También es el módulo del producto vectorial de $(1, Y, 0)$ y $(X, 1, 0)$).

$Z \leq z$ equivale a $XY \geq 1 - 2z$ y el máximo valor de Z es $1/2$ (cuando X o Y son cero).



Luego, para $z \in [0, 1/2]$, la función de distribución de Z vale (con el convenio "$0 \log 0 = 0$")

$$\begin{aligned} F(z) = \mathrm{P}\{XY \geq 1 - 2z\} &= \int_{1-2z}^{1} \int_{(1-2z)/x}^{1} dy\, dx \\ &= \int_{1-2z}^{1} \left(1 - \frac{1-2z}{x}\right) dx \\ &= 2z + (1-2z)\log(1-2z), \end{aligned}$$

La función de densidad correspondiente resulta

$$f(z) = -2\log(1-2z) \quad \text{para } z \in [0, 1/2).$$

Se trata de una función creciente en $[0, 1/2)$, pues su derivada $4/(1-2z)$ es positiva. Por tanto, la moda es $m = 1/2$.

(b) Mejor que calcular $\int_0^{1/2} z f(z)\, dz$ es observar que

$$\mathrm{E}[Z] = \frac{1 - \mathrm{E}[XY]}{2} = \frac{1}{2} - \frac{1}{2} \int_0^1 x\, dx \int_0^1 y\, dy = \frac{3}{8}.$$

Análogamente, como $\mathrm{E}[X^2 Y^2] = \int_0^1 x^2\, dx \int_0^1 y^2\, dy = 1/9$, es

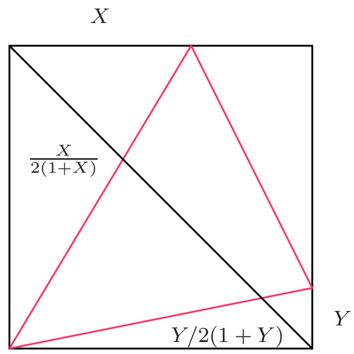
$$\mathrm{E}[Z^2] = \frac{1}{4}\mathrm{E}[(1-XY)^2] = \frac{1}{4}\mathrm{E}[1 - 2XY - X^2Y^2] = \frac{1 - 1/2 + 1/9}{4} = \frac{11}{72}.$$

Con lo cual $\sigma_Z^2 = 11/72 - (3/8)^2 = 7/(8^2\, 3^2)$ y $\sigma_Z = \sqrt{7}/24$.

(c) Fijado el punto (X, Y) la probabilidad de que Q pertenezca a T es

$$\mathrm{P}\{Q \in T | X, Y\} = \frac{1}{2} - \frac{X}{2(1+X)} - \frac{Y}{2(1+Y)}$$

como corresponde a restar al área $1/2$ del triángulo en que se elige Q el área de los triángulos que quedan fuera de T. (Los vértices de dichos triángulos que no están sobre los ejes son intersección de la recta $x + y = 1$ con las rectas $x = Xy$ e $y = Yx$ respectivamente).



Como X e Y tienen distribución uniforme en $[0, 1]$, será

$$\begin{aligned}
\mathrm{P}\{Q \in T\} &= \int_0^1 \int_0^1 \left(\frac{1}{2} - \frac{x}{2(1+x)} - \frac{y}{2(1+y)}\right) dy\, dx \\
&= \int_0^1 \left(\frac{1}{2} - \frac{x}{2(1+x)} - \frac{1}{2}[1 - \log(1+y)]_0^1\right) dx \\
&= \frac{1}{2}\log 2 - \frac{1}{2}[1 - \log(1+x)]_0^1 = \log 2 - \frac{1}{2} \simeq 0'193.
\end{aligned}$$

(d) Si $(X_1, Y_1), (X_2, Y_2)$ son los dos puntos elegidos, la intersección de los triángulos obtenidos, $T_1 \cap T_2$, es el triángulo T que correspondería a $(X = \max\{X_1, X_2\}, Y = \max\{Y_1, Y_2\})$. La probabilidad $\mathrm{P}\{Q \in T | X, Y\}$ no ha variado, pero ahora X e Y tienen distribución $\mathrm{P}\{X \leq x\} = \mathrm{P}\{X_1 \leq x, X_2 \leq x\} = x^2$ para $x \in [0, 1]$, de densidad $f(x) = 2x$ en $[0, 1]$. Por consiguiente

$$\begin{aligned}
\mathrm{P}\{Q \in T\} &= \int_0^1 \int_0^1 \left(\frac{1}{2} - \frac{x}{2(1+x)} - \frac{y}{2(1+y)}\right) 2y\, dy\, 2x\, dx \\
&= \int_0^1 \left(\frac{1}{2} - \frac{x}{2(1+x)} - \int_0^1 \left(y - 1 + \frac{1}{1+y}\right) dy\ 2x\, dx\right. \\
&= \int_0^1 \left(1 - \log 2 - \frac{x}{2(1+x)}\right) 2x\, dx = \frac{3}{2} - 2\log 2 \simeq 0'114.
\end{aligned}$$

(Para cualquier $n > 2$ fijo, el cálculo es el mismo sustituyendo la densidad $2x$ por nx^{n-1} en $[0, 1]$. Para calcular la probabilidad de que Q pertenezca a la unión de los triángulos T_i, hay que sustituir X e Y por $\min\{X_i\}$ y $\min\{Y_i\}$ respectivamente, cuyas densidades son ambas $n(1-x)^{n-1}$ en $[0, 1]$).

(e) Si $U_i = \begin{cases} 1 & \text{si } P \in T_i \\ 0 & \text{si } P \notin T_i \end{cases}$, será

$$K = \mathrm{E}\Big[\sum_{i=1}^n U_i \,|T_1, \ldots, T_n\Big] = \sum_{i=1}^n \mathrm{E}[U_i | T_i] = \sum_{i=1}^n Z_i$$

siendo Z_i el área del triángulo T_i cuya distribución y características se han analizado en los apartados (a) y (b). En virtud del Teorema Central del Límite,

$$\frac{K_n - 3n/8}{\sqrt{7n}/24} \xrightarrow{d} \mathcal{N}(0,1) \quad \text{y, en particular con } n = 100, \quad \frac{K - 37'5}{5\sqrt{7}/12} \approx V$$

donde V tiene distribución $\mathcal{N}(0,1)$. Por consiguiente

$$\mathrm{P}\{36 \leq K \leq 39\} = \mathrm{P}\{-1'36 \leq V \leq 1'36\} = \phi(1'36) - \phi(-1'36) = 2\phi(1'36) - 1$$

siendo ϕ la función de distribución $\mathcal{N}(0,1)$. (El valor aproximado de la probabilidad es $0'8264$).