Vectores Aleatorios

Pablo L. De Nápoli

Departamento de Matemática
Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires

Probabilidades y Estadística para Matemática
Segundo cuatrimestre de 2021

Parte I

Vectores Aleatorios

Vectores aleatorios

Las ideas anteriores sobre variables aleatorias continuas, pueden generalizarse para considerar vectores aleatorios.

Definición

Sea $(\Omega, \mathcal{E}, P)$ un espacio de probabilidad. Un vector aleatorio n-diemensional es una función $X : \Omega \to \mathbb{R}^n$ con la propiedad de que si $I = (a_1, b_1] \times (a_2, b_2] \times \ldots \times (a_2, b_n]$ es un intervalo de $\mathbb{R}^n$ entonces $X^{-1}(I) = \{\omega \in \Omega : X(\omega) \in I\} \in \mathcal{E}$, es decir está definida la probabilidad $P\{X \in I\}$ de que X pertenezca a I.

Obsevación: Dar un vector aleatorio n-dimensional es equivalente a dar n variables aleatorias $X_1, X_2, \ldots, X_n$.

Recordamos un ejemplo de vector aleatorio discreto

En La distribución multinomial consideramos experimentos con muchos varios posibles, en lugar de un experimento con sólo dos resultados. Consideramos un experimento con N resultados posibles, y supongamos que la probabilidad de que ocurra el i-ésimo resultado en una realización del experimento es p_i, de modo que:

$$\sum_{i=1}^{N} p_i = 1$$

Supongamos que repetimos el experimento n veces en condiciones independientes, y llamemos X_i a la cantidad de veces que ocurre el i-ésimo resultado, de modo que:

$$X_1 + X_2 + \ldots + X_N = n$$

Entonces, la distribución de probabilidades conjunta de las X_i viene dada por:

$$P\{X_1 = k_1, X_2 = k_2, \ldots, X_N = k_N\} = \frac{n!}{k_1! k_2 \ldots k_N!} \, p_1^{k_1} p_2^{k_2} \ldots p_N^{k_N}$$

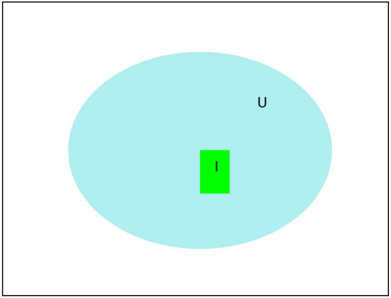
si $k_1 + k_2 + \ldots + k_N = N$ (y cero en caso contrario). Notamos que $X = (X_1, X_2, \ldots, X_N)$ es un vector aleatorio N-dimensional.

Distribución uniforme en un abierto

Consideramos un conjunto abierto $U \subset \mathbb{R}^n$ con medida $m(U)$ (longitud, área, volumen) finita y positiva. Podemos definir la distribución uniforme en U de la siguiente manera: decimos que X tiene distribución uniforme en U si X pertenece a U con probabilidad 1, y si para cualquier intervalo $I \subset U$, se cumple que:

$$P\{X \in I\} = \frac{m(I)}{m(U)} \quad \forall\, I \subset U$$

Esta noción generaliza la que vimos antes de distribución uniforme en un intervalo $U = (a, b)$ de la recta.



Una observación sobre esta definición

Si usamos la integral de Riemann, la noción de medida asociada es la medida de Peano-Jordan que podríamos definir así para $U \subset \mathbb{R}^n$ acotado:

$$m(U) = \int_R I_U(x)\, dx \quad I_U(x) = \begin{cases} 1 & \text{si} \quad x \in U \\ 0 & \text{si} \quad x \notin U \end{cases}$$

donde R es un rectángulo (intervalo) cualquiera que contenga a U, siempre que I_U sea integrable (en el sentido de Riemann) en U.
Pero esta noción no no está definida para cualquier abierto (¡hay ejemplos de abiertos cuyo borde no tiene medida cero!) y no es una medida σ-aditiva.

En los cursos de real, se ve una noción más general, la medida de Lebesgue que es σ-aditiva y está definida sobre usa σ-álgebra. Con esta noción más general, U puede ser cualquier abierto, e incluso´ U e I podrían ser conjuntos medibles Lebesgue cualesquiera, siempre que $m(U) > 0$.

Densidad conjunta de un vector aleatorio

Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función integrable tal que $0 \leq f(x) \leq 1$, y

$$\int_{\mathbb{R}^n} f(x)\, dx = 1$$

Decimos que el vector X se distribuye según la densidad conjunta $f(x)$ si para cualquier intervalo $I \subset \mathbb{R}^n$, tenemos que:

$$P\{X \in I\} = \int_I f(x)\, dx$$

Más generalmente, podríamos reemplazar I por cualquier conjunto medible.

La distribución uniforme en U corresponde a la elección:

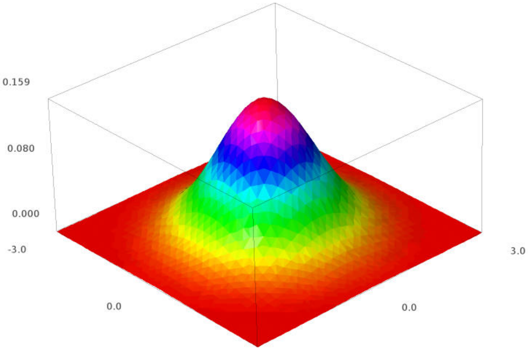
$$f(x) = \frac{1}{m(U)}\, I_U(x)$$

Distribución normal bivariada

Uuna posible generalización de la distribución normal a dos dimensiones (normal bi-variada), se obtiene especificando que el vector (X, Y) se distribuye según la densidad conjunta:

$$f(x,y) = \frac{1}{2\pi} e^{-(x^2+y^2)/2}$$

Veremos más adelante que esta densidad corresponde al caso especial de dos variables aleatorias independientes con esperanza 0 y variancia 1.



Distribución normal bivariada (2)

Más generalmente, decimos que el vector aleatorio n-dimensional X tiene distribución normal multivariada si se distribuye según una densidad de la forma:

$$f(x) = ce^{-q(x-\mu)}$$

donde $q(x) = x^t Ax$ es una forma cuadrática definida positiva (asociada a una matriz simétrica definida positiva $A \in \mathbb{R}^{n\times n}$), y c es una constante elegida de modo que la integral de f sobre todo $\mathbb{R}^n$ dé 1. Más adelante volveremos sobre este concepto.

Funciones de distribución

La noción de función de distribución puede generalizarse a vectores aleatorios.

Definición

Si $X : \Omega \to \mathbb{R}^n$ *es un vector aleatorio, su* **función de distribución conjunta** *es la función* $F : \mathbb{R}^n \to \mathbb{R}$ *dada por:*

$$F(x_1, x_2, \ldots, x_n) = P\{X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n\}$$

Por ejemplo, si X es un vector aleatorio que se distribuye según la densidad conjunta $f(x)$, entonces su función de distribución conjunta es:

$$F(x_1, x_2, \ldots, x_n) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_n} f(\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n) \, d\tilde{x}_1 \, d\tilde{x}_2 \, \ldots \, d\tilde{x}_n$$

Funciones de distribución (2)

La noción de función de distribución resulta más complicada que en el caso de variables aleatorias unidimensionales. En el caso unidimensional, la probabilidad de que la variable X tome un valor en el intervalo $(a,b]$ viene dada, en términos de la función de distribución F_X, por:

$$P\{X \in (a,b]\} = P\{X \leq b\} - P\{X \leq a\} = F_X(b) - F_X(a)$$

En cambio si (X,Y) es un vector aleatorio bidimensional con función de distribución conjunta F, y $R = (a,b] \times (c,d]$ es un rectángulo (semiabierto) en $\mathbb{R}^2$, la probabilidad de que (X,Y) tome un valor en R es (por la fórmula de inclusiones y exclusiones):

$$\begin{aligned} P\{(X,Y) \in R\} = {} & P\{X \leq b, Y \leq d\} - P\{X \leq a, Y \leq d\} \\ & - P\{X \leq b, Y \leq c\} + P\{X \leq a, Y \leq c\} \end{aligned}$$

$$P\{(X,Y) \in R\} = F(b,d) - F(a,d) - F(b,c) + F(a,c) := \Delta F(R)$$

Esta cantidad es necesariamente no negativa, esta es la generalización bidimensional del hecho de que en el caso unidimensional la función de distribución es creciente.

Funciones de distribución (3)

Una fórmula análoga (¡pero más complicada!) es cierta para vectores aleatorios en más dimensiones. Por ello, la noción de función de distribución no resultará tan útil como lo era en el caso unidimensional (y con frecuencia resulta más cómodo pensar directamente en términos de probabilidades asignadas a rectángulos, o subconjuntos más generales de $\mathbb{R}^n$).

Distribuciones de probabilidad marginales

Consideramos para simplificar la notación, un vector aleatorio bidimensional (X, Y), donde X e Y serán variables aleatorias unidimensionales.
Investiguemos qué relación existe entre la función de distribución conjunta F del vector (X, Y) y las funciones de distribución F_X y F_Y de cada variable por separado.
Notemos que:

$$F_X(x) = P\{X \leq x\} = P\{X \leq x, Y \leq +\infty\} = F(x, +\infty) = \lim_{y \to +\infty} F(x, y)$$

Similarmente,

$$F_Y(y) = \lim_{x \to +\infty} F(x, y)$$

F_X y F_Y se conocen como las funciones de distribución marginales del vector aleatorio (X, Y).

Densidades de probabilidad marginales

Consideremos ahora el caso particular, en que el vector aleatorio (X, Y) se distribuye según la densidad conjunta $f(x, y)$, su función de distribución será entonces:

$$F(x_0, y_0) = P\{X \leq x_0, Y \leq y_0\} = \int_{-\infty}^{x_0} \int_{-\infty}^{y_0} f(x, y)\, dx\, dy$$

y en consecuencia sus funciones de distribución marginales vendrán dadas por:

$$F_X(x_0) = P\{X \leq x_0\} = \int_{-\infty}^{x_0} \int_{-\infty}^{\infty} f(x, y)\, dx\, dy$$

$$F_Y(y_0) = P\{Y \leq y_0\} = \int_{-\infty}^{+\infty} \int_{-\infty}^{y_0} f(x, y)\, dx\, dx$$

Densidades de probabilidad marginales (2)

Utilizando el teorema de Fubini, podemos escribir F_X como una integral reiterada:

$$F_X(x_0) = \int_{-\infty}^{x_0} \left(\int_{-\infty}^{\infty} f(x,y)\, dy \right) dx$$

Esta igualdad significa que el vector aleatorio X se distribuye según la densidad:

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

Similarmente, el vector aleatorio Y se distribuye según la densidad:

$$f_Y(x) = \int_{-\infty}^{\infty} f(x,y)\, dy$$

f_X y f_Y se conocen como las densidades marginales de probabilidad del vector aleatorio (X, Y).

Un ejemplo

Antes consideramos un vector aleatorio (X, Y) que se distribuía según la densidad conjunta

$$f(x,y) = \frac{1}{2\pi} e^{-(x^2+y^2)/2} = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \cdot \frac{1}{\sqrt{2\pi}} e^{-y^2/2}$$

Entonces

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f(x,y)\, dy \\ &= \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-y^2/2}\, dy = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \end{aligned}$$

Luego $X \sim N(0,1)$. Similarmente, por simetría,

$$f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}$$

y también $Y \sim N(0,1)$.

Parte II

Esperanzas de Funciones de Vectores Aleatorios- Covarianza

Esperanza de funciones de vectores aleatorios.

Sea (X, Y) un vector aleatorio bidimensional, y $\varphi : \mathbb{R}^2 \to \mathbb{R}$ una función continua. La fórmula para la esperanza de una función de una variable aleatoria puede generalizarse para funciones de vectores aleatorios:

$$E[\varphi(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \varphi(x, y) \, dF(x, y)$$

donde $\varphi : \mathbb{R}^2 \to \mathbb{R}$ y la integral que aparece en el segundo miembro es una integral doble de Riemann-Stieltjes.

Integral doble de Riemann-Stieltjes

Para definir este concepto puede procederse como en análisis II, considerando primero la integral

$$\int_a^b \int_c^d \varphi(x, y)\, dF(x, y)$$

en un rectángulo $R = (a, b] \times (c, d]$ de $\mathbb{R}^2$. Consideramos una partición π del rectángulo R en rectángulos más pequeños $R_{ij} = (x_i, x_{i+1}] \times (y_j, y_{j+1}]$, definida por una partición π_x del intervalo $(a, b]$:

$$a = x_0 < x_1 < \ldots < x_M = b$$

y otra partición π_y del intervalo $(c, d]$:

$$a = y_0 < y_1 < \ldots < y_N = b$$

Integral doble de Riemann-Stieltjes (2)

Elegimos puntos intermedios $\xi_i \in (x_i, x_{i+1}]$ y $\eta_j \in (y_j, y_{j+1}]$, y consideramos sumas de Riemann-Stieltjes dobles:

$$S_\pi(\varphi, F) = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \varphi(\xi_i, \eta_j) \Delta F(R_{ij})$$

siendo

$$\Delta F(R_{ij}) = F(x_{i+1}, y_{j+1}) - F(x_i, y_{j+1}) - F(x_{i+1}, y_j) + F(x_i, y_j)$$

que de acuerdo a la fórmula que vimos antes representa la probabilidad de que el vector (X, Y) tome un valor en el rectángulo R_{ij}.

Integral doble de Riemann-Stieltjes (3)

Definamos la norma $|\pi|$ de la partición π como el máximo de las normas de las particiones π_x y π_y. Entonces si, cuando la norma de la partición π tiende a cero, las sumas

$$S_\pi(\varphi, F) = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \varphi(\xi_i, \eta_j) \Delta F(R_{ij})$$

convergen a un número I, diremos que la integral

$$\int_a^b \int_c^d \varphi(x, y) \, dF(x, y)$$

existe, y que toma el valor I. Análogamente a lo que sucede en el caso unidimensional, podemos demostrar que esto sucede si F es la función de distribución de un vector aleatorio, y φ es continua.

Integral doble de Riemann-Stieltjes (4)

La intergral impropia, sobre todo el plano, que aparece en la fórmula

$$E[\varphi(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \varphi(x, y)\, dF(x, y)$$

puede definirse como el límite de integrales sobre rectángulos:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \varphi(x, y)\, dF(x, y) = \lim_{\substack{a,c \to -\infty \\ b,d \to +\infty}} \int_{a}^{b} \int_{c}^{d} \varphi(x, y)\, dF(x, y)$$

Esperanza de vectores aleatorios

Para justificar intuitivamente la fórmula para la esperanza de $\varphi(X, Y)$, podemos proceder como en el caso discreto. Suponemos que $(X.Y)$ está concentrado en un rectángulo $R = (a, b] \times (c, d]$. Dada una partición de R, definimos las variables aleatorias discretas X_π e Y_π que aproximan a X e Y por:

$$X_\pi = \xi_i \text{ si } X \in (x_i, x_{i+1}]$$

$$Y_\pi = \eta_j \text{ si } Y \in (y_j, y_{j+1}]$$

y observando que:

$$E[\varphi(X_\pi, Y_\pi)] = S_\pi(\varphi, F)$$

Por lo que cuando la norma de la partición π tiende a cero, obtenemos formalmente la fórmula

$$E[\varphi(X, Y)] = \int\int_R \varphi(x, y)\, dF(x, y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \varphi(x, y)\, dF(x, y)$$

El caso cuando hay una densidad conjunta

El caso que más nos va a interesar, es cuando el vector aleatorio (X, Y) se distribuye según una densidad conjunta $f(x, y)$. En este caso, como ocurría en el caso unidimensional, la esperanza de $\varphi(X, Y)$ puede calcularse mediante una integral de Riemann ordinaria, en lugar de una integral de Riemann-Stieltjes:

$$E[\varphi(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \varphi(x, y)\, f(x, y)\, dx\, dy$$

Covarianza

Un caso importante de aplicación de las fórmulas anteriores es cuando queremos calcular la covarianza de dos variables aleatorias en el caso continuo. Recordamos que por definición:

$$\text{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$$

siendo $\mu_X = E[X]$, $\mu_Y = E[Y]$. Entonces tomando $\varphi(x, y) = (x - \mu_X)(y - \mu_Y)$ en las fórmulas anteriores, tenemos que:

$$\text{Cov}(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y)\, dF(x, y)$$

en el caso general, y

$$\text{Cov}(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y)\, f(x, y)\, dx\, dy$$

si el vector (X, Y) admite una densidad conjunta.

El ejemplo otra vez

Volvamos a considerar el ejemplo de un vector aleatorio (X, Y) que se distribuía según la densidad conjunta

$$f(x,y) = \frac{1}{2\pi} e^{-(x^2+y^2)/2} = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \cdot \frac{1}{\sqrt{2\pi}} e^{-y^2/2}$$

Ya vimos que $X, Y \sim N(0,1)$ por lo que $\mu_X = \mu_Y = 0$. Calculemos

$$\begin{aligned}
\mathrm{Cov}(X, Y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy\ f(x,y)\ dx\ dy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x \cdot y \cdot \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \cdot \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dx\ dy \\
&= \left(\int_{-\infty}^{\infty} x \cdot \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\ dx \right) \cdot \left(\int_{-\infty}^{\infty} y \cdot \frac{1}{\sqrt{2\pi}} e^{-y^2/2}\ dy \right) \\
&= E[X] \cdot E[Y] = 0
\end{aligned}$$

Esperanza de una suma

Una de las propiedade más básicas de la esperanza es su linealidad. Sin embargo, es difícil justificar su validez en general partiendo de la definición que dimos con la integral de Stieltjes, ya que la función de distribución F_X no depende linealmente de la variable X. Utilizando la fómula que vimos recién, podríamos sin embargo dar una justificación de que $E[X+Y] = E[X] + E[Y]$ para el caso en que X e Y admiten una densidad conjunta continua y esperanza finita

$$\begin{aligned}
E[X+Y] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x+y)\, f(x,y)\, dx\, dy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x\, f(x,y)\, dx\, dy + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y\, f(x,y)\, dx\, dy \\
&= \int_{\infty}^{\infty} x \cdot f_X(x)\, dx + \int_{\infty}^{\infty} y \cdot f_Y(y)\, dy \\
&= E[X] + E[Y]
\end{aligned}$$

Esta cuenta imita la que hicimos antes en el caso discreto.

Parte III

Independencia

Independencia

Nuestro siguiente objetivo será extender a variables no discretas la noción de independencia:

Definición

Dos variables aleatorias X e Y se dicen independientes, cuando para todo $a < b$ y todo $c < d$ los eventos $\{X \in (a, b]\}$ e $\{Y \in (c, d]\}$ son independientes. Es decir (en virtud de la definición de eventos independientes), si vale que:

$$P\{a < X \leq b, c < Y \leq d\} = P\{a < X \leq b\} \cdot P\{c < Y \leq d\}$$

Lema

Supongamos que el vector (X, Y) admite una densidad conjunta continua $f(x, y)$. Entonces las variables X e Y son independientes, si y sólo si f se factoriza en la forma:

$$f(x, y) = f_X(x) f_Y(y)$$

siendo f_X y f_Y las densidades marginales de probabilidad.

Continuación del ejemplo

Volvamos a considerar el ejemplo de un vector aleatorio (X, Y) que se distribuía según la densidad conjunta

$$f(x,y) = \frac{1}{2\pi} e^{-(x^2+y^2)/2}$$

Como

$$f(x,y) = f_X(x) \cdot f_Y(y)$$

donde

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}$$

vemos que esta densidad describe dos variables con distribución normal estándar $N(0,1)$ independientes.

Un lema: Teorema de diferenciación para integrales

Lema

Supongamos que f *es continua en* (x_0, y_0) *y consideramos la integral*

$$I_{hk} = \frac{1}{hk} \int \int_{R_{hk}} f(x, y)\, dx\, dy$$

siendo R_{hk} *el rectángulo*

$$R_{hk} = (x_0, x_0 + h] \times (y_0, y0 + k]$$

donde $h, k > 0$ *entonces*

$$I_{hk} \to f(x_0, y_0) \; cuando \; (h, k) \to 0$$

Prueba del lema

Como f es continua en (x_0, y_0), dado $\varepsilon > 0$ podemos elegir $\delta > 0$ tal que

$$|f(x,y) - f(x_0,y_0)| < \varepsilon$$

si $\|(x - x_0, y - y_0)\|_\infty = \max(x - x_0, y - y_0) < \delta$. Entonces

$$\begin{aligned}
|I_{hk} - f(x_0,y_0)| &= \left| \frac{1}{hk} \int\int_{R_{hk}} f(x,y)\, dx\, dy - f(x_0,y_0) \right| \\
&= \left| \frac{1}{hk} \int\int_{R_{hk}} f(x,y)\, dx\, dy - \frac{1}{hk} \int\int_{R_{hk}} f(x_0,y_0)\, dx\, dy \right| \\
&\leq \frac{1}{hk} \int\int_{R_{hk}} |f(x,y) - f(x_0,y_0)|\, dx\, dy \\
&\leq \frac{1}{hk} \int\int_{R_{hk}} \varepsilon\, dx\, dy = \varepsilon
\end{aligned}$$

si $\|(h,k)\|_\infty = \max(|h|,|k|) < \delta$.

Demostración

Supongamos primero que X e Y son independientes, y que el vector (X, Y) se distribuye según la densidad conjunta $f(x, y)$. Entonces X se distribuye según la densidad marginal f_X y similarmente Y se distribuye según la densidad marginal que definimos antes.
Entonces dado $(x_0, y_0) \in \mathbb{R}^2$ y $h, k > 0$, tenemos que:

$$P\{x_0 < X \leq x_0 + h, y_0 < Y \leq y_0 + k\} = \int_{x_0}^{x_0+h} \int_{y_0}^{y_0+k} f(x, y)\, dx\, dy \tag{1}$$

$$P\{x_0 < X \leq x_0 + h\} = \int_{x_0}^{x_0+h} f_X(x)\, dx \tag{2}$$

$$P\{y_0 < Y \leq y_0 + k\} = \int_{y_0}^{y_0+k} f_X(y)\, dy \tag{3}$$

En virtud de la definición de variables aleatorias independientes, vemos que:

$$\frac{P\{x_0 < X \leq x_0 + h, y_0 < Y \leq y_0 + k\}}{hk}$$

$$= \frac{P\{x_0 < X \leq x_0 + h\}}{h} \frac{P\{x_0 < X \leq x_0 + k\}}{k} \tag{4}$$

Demostración (2)

De la expresión (2) cuando $h \to 0$, deducimos que:

$$\frac{P\{x_0 < X \leq x_0 + h\}}{h} = \frac{F_X(x_0 + h) - F_X(x_0)}{h} \to f_X(x_0)$$

por el teorema fundamental del cálculo (siendo f_X continua en x_0).
Similarmente, cuando $k \to 0$, (3) y el teorema fundamental del cálculo nos dicen que:

$$\frac{P\{y_0 < Y \leq y_0 + k\}}{k} = \frac{F_Y(y_0 + k) - F_Y(y_0)}{k} \to f_Y(y_0)$$

Demostración (3)

Por otro lado,

$$P\{x_0 < X \leq x_0 + h, y_0 < Y \leq y_0 + k\} = \int\int_{R_{hk}} f(x,y)\, dx\, dy$$

siendo

$$R_{hk} = (x_0, x_0 + h] \times (x_0, x_0 + k]$$

Por el teorema de diferenciación para integrales (lema **??**), deducimos que:

$$\frac{P\{x_0 < X \leq x_0 + h, y_0 < Y \leq y_0 + k\}}{hk} \to f(x_0, y_0)$$

cuando $h, k \to 0$, siempre que f sea continua en el punto (x_0, y_0).
En consecuencia, cuando $h, k \to 0$, a partir de la relación (4), obtenemos que:

$$f(x_0, y_0) = f_X(x_0) f_Y(y_0) \tag{5}$$

Esto prueba una de las implicaciones del teorema

Demostración (4)

Para probar la afirmación recíproca, supongamos que la densidad conjunta f puede expresarse en la forma:

$$f(x,y) = f_X(x) f_Y(y)$$

siendo f_X y f_Y dos densidades de probabilidad. Notemos que entonces, f_X y f_Y deben ser entonces necesariamente las densidades marginales que definimos antes, como se deduce integrando respecto de x y de y.

Entonces, en virtud del teorema de Fubini,

$$P\{a < X \leq b, c < Y \leq d\} = \int_a^b \int_c^d f(x,y)\, dx\, dy =$$

$$= \left(\int_a^b f_X(x)\, dx \right) \left(\int_c^d f_Y(y)\, dy \right) = P\{a < X \leq b\} \cdot P\{c < Y \leq d\}$$

por lo que se deduce que X e Y son variables aleatorias independientes.

Esperanza del producto de variables independientes

Notemos, que el significado de esta demostración, es que la relación (5), es una "expresión infinitesimal" de la definición de independencia.
Como corolario obtenemos:

Corolario

Si X e Y son variables aleatorias independientes con esperanza finita, que se distribuyen según una densidad conjunta continua $f(x, y)$ entonces XY tiene esperanza finita y se tiene que

$$E[XY] = E[X] \cdot E[Y]$$

Nota: Esta fórmula es válida para variables aleatorias en general, pero la demostración no es fácil sin herramientas de análisis real.

Demostración

Nuevamente usamos la fórmula para la esperanza de un vecotr aleatorio:

$$
\begin{aligned}
E[XY] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (xy)\, f(x,y)\, dx\, dy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (xy)\, f_X(x) f_y(Y)\, dx\, dy \\
&= \left(\int_{-\infty}^{\infty} x f_X(x) \right) \cdot \left(\int_{-\infty}^{\infty} y f_X(y) \right) \\
&= E[X] \cdot E[Y]
\end{aligned}
$$

Para justificar rigurosamente este cálculo, hay que hacerlo primero con $|xy|$ en lugar de xy, lo que conduce a $E(|XY|) = E(|X|)E(|Y|)$, con lo que se establece que la integral doble es absolutamente convergente y se justifica la aplicación del teorema de Fubini.

Notamos que esta cuenta generaliza la que hicimos antes en el ejemplo de la normal bidimensional.