Probabilidad y Estadística (Borradores, Curso 23)
Variables aleatorias (3)
Funciones

Sebastian Grynberg

11 de abril de 2011

... el único héroe válido es el héroe "en grupo",
nunca el héroe individual, el héroe solo.

(Héctor G. Oesterheld)

Índice

1. Funciones de variables aleatorias

Sea X una variable aleatoria definida sobre un espacio de probabilidad $(\Omega, \mathcal{A}, \mathbb{P})$. Sea $g : D \subseteq \mathbb{R} \to \mathbb{R}$ una función cuyo dominio D contiene al rango de X,

$$X(\Omega) := \{x(\omega) : \omega \in \Omega\}.$$

Entonces $Y = g(X)$ está bien definida y será una variable aleatoria si y sólo si

$$\{\omega \in \Omega : g(X) \leq y\} \in \mathcal{A} \qquad \text{para todo } y \in \mathbb{R}. \tag{1}$$

En otras palabras, si $g^{-1}((-\infty, y]) := \{x \in \mathbb{R} : g(x) \leq y\}$, el conjunto $\{X \in g^{-1}(-\infty, y]\}$ debe tener asignada probabilidad. Este es típicamente el caso. Por ejemplo, si X es discreta, cualquier función g cuyo dominio contenga al rango de X satisface (1). Si X no es discreta, cualquier función g seccionalmente continua cuyo dominio contenga al rango de X satisface (1).

1.1. Método básico: eventos equivalentes

Si queremos hallar la función de distribución de $Y = g(X)$ tenemos que calcular

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(g(X) \leq y) = \mathbb{P}(X \in g^{-1}(-\infty, y]). \tag{2}$$

Los siguientes ejemplos ilustran el *método básico* de cómo hacerlo.

Ejemplo 1.1 (Del péndulo a la distribución de Cauchy)**.** Sea Θ el ángulo de un péndulo medido desde la vertical cuyo extremo superior se encuentra sostenido del punto $(0,1)$. Sea $(X,0)$ el punto de intersección de la recta que contiene al péndulo y el eje x -ver la Figura 1-. Trigonometría mediante, sabemos que

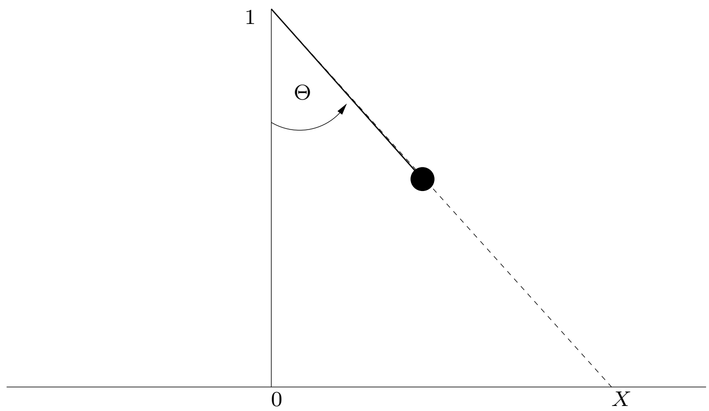

Figura 1: Pendulo.

$$X = \tan \Theta$$

Si el ángulo Θ es una variable aleatoria con distribución uniforme sobre el intervalo $(-\pi/2, \pi/2)$, cuál es la distribución de X?

Primero observamos que para cada $\theta \in (-\pi/2, \pi/2)$ tenemos que

$$\mathbb{P}(\Theta \leq \theta) = \frac{\theta - (-\pi/2)}{\pi/2 - (-\pi/2)} = \frac{\theta + \pi/2}{\pi} = \frac{1}{2} + \frac{\theta}{\pi}.$$

De esto se deduce que

$$\begin{aligned} \mathbb{P}(X \leq x) &= \mathbb{P}(\tan \Theta \leq x) \\ &= \mathbb{P}(\Theta \leq \arctan x) = \frac{1}{2} + \frac{1}{\pi} \arctan x, \end{aligned}$$

y derivando obtenemos que

$$f_X(x) = \frac{1}{\pi(1+x^2)}.$$

Teorema 1.2. *Sea X una variable aleatoria continua con función de distribución creciente. Entonces, $Y = F_X(X) \sim \mathcal{U}(0,1)$.*

Demostración. El análisis se reduce a examinar el comportamiento de la función de distribución de Y sobre el intervalo $(0,1)$. Para cada $y \in (0,1)$ vale que

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(F_X(X) \leq y) = \mathbb{P}(X \leq F_X^{-1}(y)) = F_X(F_X^{-1}(y)) = y.$$

$\square$

Corolario 1.3. *Sea X una variable aleatoria continua con función de distribución creciente. Sea Y una variable aleatoria cualquiera. Entonces X puede transformarse en una copia de Y haciendo lo siguiente: $\hat{Y} = F_Y^{-1}(F_X(X))$, donde F_Y^{-1} es la inversa generalizada de Y.*

Ejemplo 1.4. Construir una moneda equilibrada X usando una variable aleatoria T con distribución exponencial de intensidad 1.

$$\hat{X} = \mathbf{1}\left\{\frac{1}{2} < 1 - e^{-T} < 1\right\}.$$

$\square$

El siguiente ejemplo puede considerarse un prototipo que ilustra cómo tratar con las funciones de variables aleatorias cuando no son inyectivas.

Ejemplo 1.5 (Prototipo)**.** Sea X una variable aleatoria cualquiera y sea $Y = X^2$. Queremos determinar la distribución de Y.

1. Cálculo explícito de la función de distribución. La función de distribución de Y se calcula observando que $g(x) = x^2$ y utilizando la fórmula (2)

$$F_Y(y) = \mathbb{P}\left(X \in g^{-1}((-\infty, y])\right).$$

En este caso, el conjunto $g^{-1}((-\infty, y])$ adopta la forma

$$\begin{aligned} g^{-1}((-\infty, y]) &= \left\{x \in \mathbb{R} : x^2 \leq y\right\} \\ &= \begin{cases} [-\sqrt{y}, \sqrt{y}] & \text{si } y \geq 0, \\ \emptyset & \text{si } y < 0. \end{cases} \end{aligned}$$

Por lo tanto,

$$F_Y(y) = \mathbb{P}\left(-\sqrt{y} \leq X \leq \sqrt{y}\,\right) \mathbf{1}\{y \geq 0\} = \left(F_X(\sqrt{y}) - F_X(-\sqrt{y}-)\right) \mathbf{1}\{y \geq 0\}. \tag{3}$$

En particular, si X es continua, $\mathbb{P}(X = x) = 0$ para todo $x \in \mathbb{R}$, y la identidad (3) adopta la forma

$$F_Y(y) = \left(F_X(\sqrt{y}) - F_X(-\sqrt{y})\right) \mathbf{1}\{y > 0\}. \tag{4}$$

2. Cálculo explícito de la densidad de probabilidades. Si X es absolutamente continua con densidad de probabilidades $f_X(x)$, la densidad de probabilidades de $Y = X^2$ se obtiene derivando la función de distribución $F_Y(y)$. De la identidad (4) se deduce que:

$$\begin{aligned} f_Y(y) &= \frac{d}{dy} F_Y(y) \\ &= \left(f_X(\sqrt{y}) \frac{1}{2\sqrt{y}} - f_X(-\sqrt{y}) \frac{1}{-2\sqrt{y}}\right) \mathbf{1}\{y > 0\} \\ &= \frac{1}{2\sqrt{y}} \left(f_X(\sqrt{y}) + f_X(-\sqrt{y})\right) \mathbf{1}\{y > 0\}. \end{aligned} \tag{5}$$

$\square$

Ejemplo 1.6 (De continua a discreta)**.** Sea $U \sim \mathcal{U}(0,1]$ Hacemos $Y = [10\,U]$, donde $[x]$ representa la parte entera de $x \in \mathbb{R}$. Queremos determinar la función de probabilidad de Y.

En primer lugar observamos que la variable aleatoria Y es el primer dígito del desarrollo decimal de un número elegido al azar sobre el intervalo $(0,1)$. Los posibles valores de Y son $0,1,\dots,9$. Para cada $y \in \{0,1,\dots,9\}$ vale que

$$\mathbb{P}(Y = y) = \mathbb{P}\left(\frac{y}{10} < U \leq \frac{y+1}{10}\right) = \frac{1}{10}.$$

□

1.2. Variables discretas

Sea X una variable aleatoria discreta a valores $x_1, x_2, \dots$. De la relación $Y = g(X)$ se deduce que los posibles valores de Y son $y_1 = g(x_1), y_2 = g(x_2), \dots$. Si la función de probabilidad de X está dada por

$$p_X(x_i) = p_i, \qquad i = 1,2,\dots$$

La función de probabilidad de Y se determina por

$$p_Y(y_i) = \mathbb{P}(Y = y_i) = \mathbb{P}(X \in g^{-1}(y_i)) = \sum_{x \in g^{-1}(y_i)} p_x$$

Ejercicios adicionales

1. Sea X una variable aleatoria discreta con función de probabilidad dada por

$$p_X(x) = \frac{1}{8}\left(4\mathbf{1}\{x=-1\} + 2\mathbf{1}\{x=0\} + \mathbf{1}\{x \in \{1,2\}\}\right)$$

Hallar la función de probabilidad de Y para

(**a**) $Y = 2X + 1$,

(**b**) $Y = 2X^2 + 1$.

1.3. Funciones a trozos: dividir y conquistar

Sea X una variable y sea $A_1, \dots, A_k$ una partición de $\mathbb{R}$ tal que $\mathbb{P}(X \in A_i) > 0$ para todo i. Consideramos una función a trozos definida por

$$g(x) = \sum_{i=1}^{k} g_i(x)\mathbf{1}\{x \in A_i\}, \tag{6}$$

donde, para cada $i = 1, \dots, k$, $g_i : \mathbb{R} \to \mathbb{R}$, es una función tal que $g_i(X)$ es una variable aleatoria. Si se quiere hallar la distribución de

$$Y = g(X) = \sum_{i=1}^{k} g_i(X)\mathbf{1}\{X \in A_i\}$$

se puede hacer lo siguiente: considerar las variables truncadas $X_i = X|X \in A_i$, hallar las distribuciones de las variables $Y_i = g_i(X_i)$ y luego ponderarlas con los pesos $\mathbb{P}(X \in A_i)$. Esto es,

$$F_Y(y) = \sum_{i=1}^{k} F_{Y_i}(y)\mathbb{P}(X \in A_i). \tag{7}$$

En efecto, por una parte tenemos que

$$\begin{aligned} F_Y(y) &= \mathbb{P}\left(\sum_{j=1}^{k} g_j(X)\mathbf{1}\{X \in A_j\} \le y\right) \\ &= \sum_{i=1}^{k} \mathbb{P}\left(\sum_{j=1}^{k} g_j(X)\mathbf{1}\{X \in A_j\} \le y,\, X \in A_i\right) \\ &= \sum_{i=1}^{k} \mathbb{P}\left(g_i(X) \le y,\, X \in A_i\right) \\ &= \sum_{i=1}^{k} \mathbb{P}\left(X \in g_i^{-1}(-\infty, y] \cap A_i\right). \end{aligned} \tag{8}$$

Por otra parte,

$$F_{Y_i}(y) = \mathbb{P}(g_i(X_i) \le y) = \mathbb{P}(X_i \in g^{-1}(-\infty, y]) = \frac{\mathbb{P}(X \in g^{-1}(-\infty, y] \cap A_i)}{\mathbb{P}(X \in A_i)}.$$

Equivalentemente,

$$P(X \in g^{-1}(-\infty, y] \cap A_i) = F_{Y_i}(y)\mathbb{P}(X \in A_i). \tag{9}$$

Combinando (8) y (9) se obtiene (7). $\square$

1.4. Funciones inyectivas suaves

Teorema 1.7 (Cambio de variables)**.** Sea X una variable aleatoria absolutamente continua con densidad de probabilidades $f_X(x)$. Sea $Y = g(X)$, donde g es una función monótona con derivada no nula. Entonces Y es absolutamente continua y admite una densidad de probabilidades de la forma

$$f_Y(y) = \left.\frac{f_X(x)}{|g'(x)|}\right|_{x=g^{-1}(y)}. \tag{10}$$

Demostración.

1. La función g *es creciente*: $g(x_1) \leq g(x_2)$ para $x_1 \leq x_2$. En tal caso la función inversa g^{-1} también es creciente. En consecuencia,

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \leq y) = \mathbb{P}(g(X) \leq y) \\ &= \mathbb{P}\left(X \leq g^{-1}(y)\right) = F_X\left(g^{-1}(y)\right). \end{aligned} \tag{11}$$

La función $F_Y(y)$ es derivable, debido a que es una composición de funciones derivables. Derivando con respecto a y y usando la regla de la cadena se obtiene

$$\frac{d}{dy}F_Y(y) = \frac{d}{dy}F_X(g^{-1}(y)) = \frac{f_X(g^{-1}(y))}{g'(g^{-1}(y))}.$$

2. La función g *es decreciente*: $g(x_1) \geq g(x_2)$ para $x_1 \leq x_2$. En este caso la función inversa g^{-1} también es decreciente. En consecuencia,

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \leq y) = \mathbb{P}(g(X) \leq y) \\ &= \mathbb{P}\left(X \geq g^{-1}(y)\right) = 1 - F_X\left(g^{-1}(y)-\right). \end{aligned} \tag{12}$$

Derivando con respecto a y se obtiene

$$\frac{d}{dy}F_Y(y) = \frac{d}{dy}\left(1 - F_X(g^{-1}(y))\right) = -\frac{f_X(g^{-1}(y))}{g'(g^{-1}(y))}.$$

□

Corolario 1.8 (Cambio lineal)**.** Dados $a > 0$ y $b \in \mathbb{R}$, la densidad de probabilidades de $Y = aX + b$ adopta la forma

$$f_Y(y) = \frac{1}{a}f_X\left(\frac{y-b}{a}\right). \tag{13}$$

En palabras, desde el punto de vista de la densidad de probabilidades, el cambio lineal $y = ax + b$ efectúa una *traslación en* b seguida de un *cambio de escala de* 1 *en* a sobre la densidad original. Cuando el parámetro a se achica, los valores de Y tienden a estar más concentrados (alrededor del valor medio) y cuando a se agranda, tienden a dispersarse. □

Ejemplo 1.9 (Variables exponenciales)**.** Se dice que la variable aleatoria Y tiene *distribución exponencial de intensidad* $\lambda > 0$, y se denota $Y \sim \text{Exp}(\lambda)$, si $Y = \frac{1}{\lambda}X$, donde X es una variable aleatoria absolutamente continua que admite una densidad de probabilidades de la forma $f_X(x) = e^{-x}\mathbf{1}\{x \geq 0\}$. De (13) se deduce que Y admite una densidad de probabilidades de la forma

$$f_Y(y) = \lambda e^{-\lambda y}\mathbf{1}\{y \geq 0\}.$$

□

Ejemplo 1.10 (Variables Normales)**.** Sean $\mu \in \mathbb{R}$ y $\sigma > 0$. Se dice que la variable aleatoria Y tiene distribución *normal de parámetros* μ*,* σ^2, y se denota $Y \sim \mathcal{N}(\mu, \sigma^2)$, si $Y = \sigma X + \mu$, donde X es una variable aleatoria absolutamente continua con densidad de probabilidades $\varphi(x)$ definida por:

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}. \tag{14}$$

De (13) se deduce que Y admite una densidad de probabilidades de la forma

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right).$$

□

Nota Bene. Las fórmulas (11) y (12) permiten calcular explícitamente la función de distribución, F_Y, para transformaciones monótonas (continuas) $Y = g(X)$, independientemente de la clase de variable que sea X. ¿Qué hacer cuando la transformación g es suave pero no es inyectiva?

Ejemplo 1.11. Sea $X \sim \mathcal{N}(0, 1)$. Según la fórmula (5) la densidad de probabilidades de $Y = X^2$ es

$$f_Y(y) \quad = \quad \frac{1}{2\sqrt{y}} \left(\varphi(\sqrt{y}) + \varphi(-\sqrt{y})\right) \mathbf{1}\{y > 0\},$$

donde $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$. Por lo tanto,

$$f_Y(y) \quad = \quad \frac{1}{\sqrt{2\pi}} y^{-1/2} e^{-y/2} \mathbf{1}\{y > 0\}.$$

En otras palabras, si $X \sim \mathcal{N}(0, 1)$, entonces $X^2 \sim \Gamma(1/2, 1/2)$. □

El Teorema 1.7 puede generalizarse del siguiente modo

Teorema 1.12 (Cambio de variables II)**.** Sea X una variable aleatoria absolutamente continua con densidad de probabilidades $f_X(x)$. Sea $Y = g(X)$, donde g es una función derivable con derivada no nula (salvo en finitos puntos). Si para cada $y \in \mathbb{R}$, el conjunto $g^{-1}(y) = \{x \in \mathbb{R} : g(x) = y\}$ es discreto, entonces Y es absolutamente continua y admite una función densidad de probabilidades de la forma

$$f_Y(y) = \sum_{x \in g^{-1}(y)} \frac{f_X(x)}{|g'(x)|}. \tag{15}$$

Se sobreentiende que si $g^{-1}(y) = \emptyset$, $f_Y(y) = 0$.

Ejercicios adicionales

2. [James p.98] Si X tiene densidad $f_X(x)$, cuál es la densidad de $Y = \cos X$?

2. Funciones de vectores aleatorios

2.1. Método básico: eventos equivalentes

Sea $\mathbf{X} = (X_1, \dots, X_n)$ un vector aleatorio definido sobre un espacio de probabilidad $(\Omega, \mathcal{A}, \mathbb{P})$ y sea $g : \mathbb{R}^n \to \mathbb{R}$ una función cualquiera. Entonces, $Z := g(\mathbf{X})$ será una variable aleatoria si y solo si

$$\{\omega \in \Omega : g(\mathbf{X}(\omega)) \leq z\} \in \mathcal{A} \qquad \text{para todo } z \in \mathbb{R}.$$

La función de distribución de Z, $F_Z(z)$, se puede calcular mediante la función de distribución de $\mathbf{X}$ de la siguiente manera:

$$F_Z(z) = \mathbb{P}(Z \leq z) = \mathbb{P}(g(\mathbf{X}) \leq z) = \mathbb{P}(\mathbf{X} \in \mathcal{B}_z), \tag{16}$$

donde $\mathcal{B}_z := g^{-1}((-\infty, z]) = \{\mathbf{x} \in \mathbb{R}^n : g(\mathbf{x}) \leq z\}$.

Caso bidimensional continuo. Sea (X, Y) un vector aleatorio con densidad conjunta $f_{X,Y}(x, y)$. Cualquier función continua a valores reales $g : \mathbb{R}^2 \to \mathbb{R}$ define una nueva variable aleatoria $Z := g(X, Y)$. La función de distribución de Z, $F_Z(z) = \mathbb{P}(Z \leq z)$, se puede obtener a partir de la densidad conjunta de X e Y de la siguiente forma:

1. Para cada $z \in \mathbb{R}$ se determina el conjunto $\mathcal{B}_z \subset \mathbb{R}^2$ de todos los puntos (x, y) tales que $g(x, y) \leq z$.

2. Integrando la densidad conjunta $f_{X,Y}(x, y)$ sobre el conjunto $\mathcal{B}_z$ se obtiene la función de distribución de Z:

$$F_Z(z) = \iint_{\mathcal{B}_z} f_{X,Y}(x, y)dxdy. \tag{17}$$

3. Si se quiere obtener la densidad de Z basta derivar, con respecto a z, la función de distribución $F_Z(z)$. $\square$

Ejemplo 2.1. Sean X e Y dos variables aleatorias independientes cada una con distribución uniforme sobre el intervalo $[-1, 1]$. Se quiere hallar la función de distribución y la densidad de $Z = |X - Y|$.

La función de distribución de la variable $Z = |X - Y|$ se puede obtener observando la Figura 2.

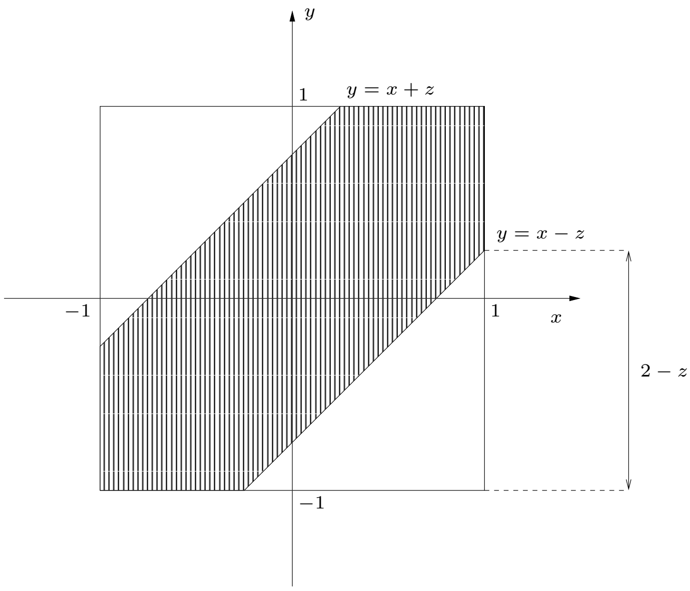

Figura 2: La región sombreada representa los puntos del cuadrado $[-1, 1] \times [-1, 1]$ tales que $|x - y| \leq z$, $0 \leq z \leq 2$ y su área es $4 - (2 - z)^2 = 4z - z^2$.

Debido a que las variables aleatorias X e Y son independientes y uniformemente distribuidas obre el intervalo $[-1, 1]$, tenemos que

$$\mathbb{P}((X, Y) \in \mathcal{B}) = \frac{\text{área}(\mathcal{B})}{4},$$

cualquiera sea la región $\mathcal{B}$ contenida en el cuadrado $[-1, 1] \times [-1, 1]$ para la que tenga sentido la noción de área. En consecuencia,

$$\mathbb{P}(|X - Y| \leq z) = \frac{4z - z^2}{4}$$

para todo $z \in [0, 2]$. Derivando esta última expresión respecto de z se obtiene la densidad de $Z = |X - Y|$:

$$f_Z(z) = \left(\frac{2 - z}{2}\right) \mathbf{1}\{z \in (0, 2)\}.$$

$\square$

Caso bidimensional discreto. Sea (X, Y) un vector aleatorio discreto sobre un espacio de probabilidad $(\Omega, \mathcal{A}, \mathbb{P})$, con función de probabilidad conjunta $p_{X,Y}(x, y)$. Sea $g : \mathbb{R}^2 \to \mathbb{R}$ una función cualquiera, $Z := g(X, Y)$ es una nueva variable aleatoria, cuya función de probabilidad, $p_Z(z)$, se obtiene de la siguiente manera:

$$p_Z(z) = \mathbb{P}(Z = z) = \mathbb{P}(g(X, Y) = z) = \sum_{(x,y) \in \mathcal{B}_z} p_{X,Y}(x, y), \tag{18}$$

donde $\mathcal{B}_z = \{(x, y) \in X(\Omega) \times Y(\Omega) : g(x, y) = z\}$. $\square$

2.1.1. Suma de variables

Ejemplo 2.2 (Suma)**.** Sean X, Y dos variables aleatorias con densidad conjunta $f_{X,Y}(x,y)$ y sea $Z = X + Y$. Para cada $z \in \mathbb{R}$, $\mathcal{B}_z = \{(x,y) \in \mathbb{R}^2 : y \leq z - x\}$. Usando la fórmula (17) se obtiene la función de distribución de Z

$$F_Z(z) = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{z-x} f_{X,Y}(x,y)dy \right) dx. \tag{19}$$

La densidad de Z se obtiene derivando respecto de z la función de distribución $F_Z(z)$

$$f_Z(z) = \frac{d}{dz} F_Z(z) = \int_{-\infty}^{\infty} f_{X,Y}(x, z-x)dx. \tag{20}$$

Ejemplo 2.3 (Suma de variables independientes)**.** Sean X, Y dos variables aleatorias continuas e independientes con densidad conjunta $f_{X,Y}(x,y) = f_X(x)f_Y(y)$. Según la fórmula (20) que aparece en el Ejemplo 2.2, la densidad de probabilidades de la suma $Z = X + Y$ es

$$f_Z(z) = \int_{-\infty}^{\infty} f_{X,Y}(x, z-x)dx = \int_{-\infty}^{\infty} f_X(x)f_Y(z-x)dx \tag{21}$$

y se denomina el *producto convolución*, $f_X * f_Y$, de las densidades marginales f_X y f_Y.

Si las densidades marginales $f_X(x)$ y $f_Y(y)$ concentran la masa en $[0, \infty)$ la fórmula (21) del producto convolución es un poco más sencilla:

$$(f_X * f_Y)(z) = \int_0^{\infty} f_X(x)f_Y(z-x)dx = \int_0^z f_X(x)f_Y(z-x)dx. \tag{22}$$

□

Ejemplo 2.4 (Suma de exponenciales independientes de igual intensidad)**.** Sean X e Y variables aleatorias independientes con distribución exponencial de intensidad $\lambda > 0$. La densidad de la suma $X + Y$ es

$$f_{X+Y}(z) = \int_0^z \lambda e^{-\lambda x} \lambda e^{z-x} dx = \lambda^2 z e^{-\lambda z}. \tag{23}$$

En el lado derecho de la identidad (23) se puede reconocer la densidad de la distribución Gamma: $\Gamma(2, \lambda)$. □

2.1.2. Mínimo

Queremos caracterizar la función de distribución del mínimo entre dos variables aleatorias X e Y, $U := \min\{X, Y\}$. En primer lugar observamos que para cada $u \in \mathbb{R}$ vale que

$$\begin{aligned} F_U(u) &= \mathbb{P}(U \leq u) \\ &= \mathbb{P}(\min\{X,Y\} \leq u) = 1 - \mathbb{P}(\min\{X,Y\} > u) \\ &= 1 - \mathbb{P}(X > u, Y > u). \end{aligned} \tag{24}$$

Si (X, Y) es un vector aleatorio continuo con función de densidad conjunta $f_{X,Y}(x, y)$ tenemos que

$$F_U(u) = 1 - \int_u^\infty \int_u^\infty f_{X,Y}(x, y)dxdy. \tag{25}$$

Si (X, Y) es un vector aleatorio discreto con función de probabilidad conjunta $p_{X,Y}(x, y)$ tenemos que

$$F_U(u) = 1 - \sum_{x>u} \sum_{y>u} p_{X,Y}(x, y). \tag{26}$$

Si X e Y son independientes tenemos que

$$F_U(u) = 1 - \mathbb{P}(X > u)\mathbb{P}(Y > u). \tag{27}$$

Etcétera...

Ejemplo 2.5 (Mínimo de exponenciales independientes)**.** Sean X_1 e X_2 variables aleatorias exponenciales independientes de intensidades λ_1 y λ_2 respectivamente. De acuerdo con la identidad (27) tenemos que la función de distribución del mínimo $U = \text{mín}\{X_1, X_2\}$ es

$$F_U(u) = (1 - e^{-\lambda_1 u}e^{-\lambda_2 u})\mathbf{1}\{u \geq 0\} = (1 - e^{-(\lambda_1+\lambda_2)u})\mathbf{1}\{u \geq 0\}. \tag{28}$$

En palabras, *el mínimo de dos variables exponenciales independientes es una exponencial cuya intensidad es la suma de las intensidades de las variables originales.*

¿Cuál es la probabilidad de que $U = X_1$? Basta observar que $U = X_1$ si y solo si $X_1 \leq X_2$.

$$\mathbb{P}(U = X_1) = \mathbb{P}(X_1 \leq X_2) = \int_0^\infty \left(\int_0^{x_2} f_{X_1,X_2}(x_1, x_2)dx_1\right) dx_2,$$

donde $f_{X_1,X_2}(x_1, x_2)$ es la densidad conjunta de X_1 y X_2

$$f_{X_1,X_2}(x_1, x_2) = \lambda_1 e^{\lambda_1 x_1}\lambda_2 e^{\lambda_2 x_2}.$$

Por lo tanto,

$$\begin{aligned}
\mathbb{P}(U = X_1) &= \int_0^\infty \lambda_2 e^{\lambda_2 x_2}\left(\int_0^{x_2} \lambda_1 e^{\lambda_1 x_1}dx_1\right) dx_2 = \int_0^\infty \lambda_2 e^{\lambda_2 x_2}\left(1 - e^{\lambda_1 x_2}\right) dx_2 \\
&= \int_0^\infty \left(\lambda_2 e^{\lambda_2 x_2} - \lambda_2 e^{(\lambda_1+\lambda_2)x_2}\right) dx_2 = 1 - \frac{\lambda_2}{\lambda_1 + \lambda_2} \\
&= \frac{\lambda_1}{\lambda_1 + \lambda_2}.
\end{aligned}$$

Como corolario se obtiene que

$$\mathbb{P}(U = X_2) = \frac{\lambda_2}{\lambda_1 + \lambda_2}.$$

□

2.2. El método del Jacobiano

Teorema 2.6 (Cambio de variables en la integral múltiple). *Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función integrable. Sean $G_0 \subset \mathbb{R}^n$ y $G \subset \mathbb{R}^n$ regiones abiertas y sea $h : G_0 \to G$, $h = (h_1, \dots, h_n)$ una biyección entre G_0 y G, cuyas componentes tienen derivadas parciales de primer orden continuas. Esto es, para todo $1 \leq i, j \leq n$, las funciones $\frac{\partial h_i(\mathbf{y})}{\partial y_j}$ son continuas. Si el Jacobiano de h es diferente de cero en casi todo punto, entonces,*

$$\int_A f(\mathbf{x}) d\mathbf{x} = \int_{h^{-1}(A)} f(h(\mathbf{y})) |J_h(\mathbf{y})| d\mathbf{y},$$

para todo conjunto abierto $A \subset G$, donde

$$J_h(\mathbf{y}) = \det \left(\left(\frac{\partial h_i(\mathbf{y})}{\partial y_j} \right)_{i,j} \right).$$

El siguiente resultado, que caracteriza la distribución de un cambio de variables aleatorias, es una consecuencia inmediata del Teorema 2.6.

Corolario 2.7. Sea $\mathbf{X}$ un vector aleatorio n-dimensional con función densidad de probabilidad $f_{\mathbf{X}}(\mathbf{x})$. Sean $G_0 \subset \mathbb{R}^n$ y $G \subset \mathbb{R}^n$ regiones abiertas y sea $g : G \to G_0$ una biyección cuya función inversa $h = g^{-1}$ satisface las hipótesis del Teorema 2.6. Si $\mathbb{P}(\mathbf{X} \in G) = 1$, entonces, el vector aleatorio $\mathbf{Y} = g(\mathbf{X})$ tiene función densidad de probabilidad $f_{\mathbf{Y}}(\mathbf{y})$ de la forma:

$$f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{X}}(g^{-1}(\mathbf{y})) |J_{g^{-1}}(\mathbf{y})|. \tag{29}$$

Demostración. Cualquiera sea el conjunto abierto $B \subset G_0$ tenemos

$$\mathbb{P}(\mathbf{Y} \in B) = \mathbb{P}(g(\mathbf{X}) \in B) = \mathbb{P}(\mathbf{X} \in g^{-1}(B)) = \int_{g^{-1}(B)} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}.$$

Poniendo $f = f_{\mathbf{X}}$ y $h = g^{-1}$ en el Teorema 2.6 se obtiene

$$\int_{g^{-1}(B)} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} = \int_B f_{\mathbf{X}}(g^{-1}(\mathbf{y})) |J_{g^{-1}}(\mathbf{y})| d\mathbf{y}.$$

En consecuencia,

$$\mathbb{P}(\mathbf{Y} \in B) = \int_B f_{\mathbf{X}}(g^{-1}(\mathbf{y})) |J_{g^{-1}}(\mathbf{y})| d\mathbf{y}.$$

Por lo tanto, el vector aleatorio $\mathbf{Y}$ tiene función densidad de probabilidad de la forma $f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{X}}(g^{-1}(\mathbf{y})) |J_{g^{-1}}(\mathbf{y})|$. □

Nota Bene. Operativamente, la fórmula (29) para hallar la densidad conjunta de $\mathbf{Y} = g(\mathbf{X})$ involucra los siguientes pasos: 1. Invertir las variables (i.e., despejar las x's en función de las y's). 2. Calcular el Jacobiano de la inversa de g (i.e., calcular el determinante de la matriz formada por las derivadas parciales de las x_i respecto de las y_j). 3. Substituir los resultados obtenidos en los pasos 1. y 2. en la fórmula (29). **Aunque mecánico, el método del jacobiano es un método de naturaleza analítica muy poderoso**. $\square$

Nota Bene. Con frecuencia es más fácil obtener el jacobiano de $\mathbf{y}$ en relación a $\mathbf{x}$, pues $\mathbf{Y}$ es dado como función de $\mathbf{X}$, es bueno recordar que los dos jacobianos son recíprocos y se puede obtener $J_{g^{-1}}(\mathbf{y})$ a partir de $J_g(\mathbf{x})$, invirtiendo este último y substituyendo $\mathbf{x}$ por $g^{-1}(\mathbf{y})$. Esta regla es análoga a la regla para la derivada de una función inversa en el caso unidimensional:

$$\frac{dg^{-1}(y)}{dy} = \left.\frac{1}{g'(x)}\right|_{x=g^{-1}(y)} = \frac{1}{g'(g^{-1}(y))}.$$

$\square$

Ejemplo 2.8 (Transformaciones lineales)**.** Si $(X_1, X_2) = (aY_1 + bY_2, cY_1 + dY_2)$. Entonces,

$$g_{Y_1,Y_2}(y_1, y_2) = |ad - bc| f_{X_1,X_2}(ay_1 + by_2, cy_1 + dy_2).$$

En general, si $\mathbf{X} = A\mathbf{Y}$, donde $A \in \mathbb{R}^{n\times n}$ es una matriz inversible, se obtiene

$$g_{\mathbf{Y}}(\mathbf{y}) = |\det(A)| f_{\mathbf{X}}(A\mathbf{y}).$$

Ejemplo 2.9 (Suma y resta de normales independientes)**.** Sean X_1 y X_2 dos variables aleatorias independientes con distribuciones normales $\mathcal{N}(\mu_1, \sigma^2)$ y $\mathcal{N}(\mu_2, \sigma^2)$, respectivamente. Su densidad conjunta es

$$f_{X_1,X_2}(x_1, x_2) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{1}{2\sigma^2}\left((x_1-\mu_1)^2 + (x_2-\mu_2)^2\right)\right) \tag{30}$$

Consideramos el cambio de variables $(y_1, y_2) = g(x_1, x_2) = (x_1 + x_2, x_1 - x_2)$ cuya inversa es $(x_1, x_2) = g^{-1}(y_1, y_2) = \left(\frac{y_1+y_2}{2}, \frac{y_1-y_2}{2}\right)$. Es fácil ver que el jacobiano de g^{-1} no depende de (y_1, y_2) y es igual a $-\frac{1}{2}$. De acuerdo con la identidad (29) tenemos que

$$\begin{aligned}
f_{Y_1,Y_2}(y_1, y_2) &= \frac{1}{4\pi\sigma^2} \exp\left(-\frac{1}{2\sigma^2}\left(\left(\frac{y_1+y_2}{2} - \mu_1\right)^2 + \left(\frac{y_1-y_2}{2} - \mu_2\right)^2\right)\right) \\
&\propto \exp\left(-\frac{1}{4\sigma^2}\left(y_1^2 - 2(\mu_1+\mu_2)y_1\right)\right) \exp\left(-\frac{1}{4\sigma^2}\left(y_2^2 - 2(\mu_1-\mu_2)y_2\right)\right) \\
&\propto \exp\left(-\frac{(y_1-(\mu_1+\mu_2))^2}{2(2\sigma^2)}\right) \exp\left(-\frac{(y_2-(\mu_1-\mu_2))^2}{2(2\sigma^2)}\right) \qquad (31)
\end{aligned}$$

La identidad (31) permite concluir que las variables Y_1 e Y_2 son independientes y que se distribuyen de la siguiente manera: $Y_1 \sim \mathcal{N}(\mu_1 + \mu_2, 2\sigma^2)$, $Y_2 \sim \mathcal{N}(\mu_1 - \mu_2, 2\sigma^2)$. En pocas palabras, *si X_1 y X_2 son dos variables aleatorias independientes con distribuciones normales $\mathcal{N}(\mu_1, \sigma^2)$ y $\mathcal{N}(\mu_2, \sigma^2)$, entonces $X_1 + X_2$ y $X_1 - X_2$ son independientes y $X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, 2\sigma^2)$ y $X_1 - X_2 \sim \mathcal{N}(\mu_1 - \mu_2, 2\sigma^2)$* $\square$

Nota Bene. Sean X_1 y X_2 dos variables aleatorias independientes con distribuciones normales $\mathcal{N}(\mu_1, \sigma_1^2)$ y $\mathcal{N}(\mu_2, \sigma_2^2)$, respectivamente. Cálculos similares permiten deducir que $X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$ *y* $X_1 - X_2 \sim \mathcal{N}(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$*. Más aún,* $X_1 + X_2$ *y* $X_1 - X_2$ *son independientes si y solo si* $\sigma_1^2 = \sigma_2^2$. □

Ejemplo 2.10 (Persistencia de la mala suerte)**.** Sean X_1 y X_2 variables aleatorias independientes con distribución común exponencial de intensidad λ. Vamos a hallar la densidad conjunta de (Y_1, Y_2) donde

$$(Y_1, Y_2) = (X_1 + X_2, X_1/X_2).$$

Para ello consideramos la transformación

$$g(x_1, x_2) = (x_1 + x_2, x_1/x_2) = (y_1, y_2).$$

La transformación inversa de g es

$$x_1 = \frac{y_1 y_2}{1 + y_2}, \qquad x_2 = \frac{y_1}{1 + y_2} \tag{32}$$

y se obtiene resolviendo un sistema de dos ecuaciones en las variables x_1 y x_2:

$$\left\{ \begin{array}{l} x_1 + x_2 = y_1 \\ x_1/x_2 = y_2 \end{array} \right. \iff \left\{ \begin{array}{l} x_1 + x_2 = y_1 \\ x_1 = y_2 x_2 \end{array} \right. \iff \left\{ \begin{array}{l} (1 + y_2)x_2 = y_1 \\ x_1 = y_2 x_2 \end{array} \right. \iff \left\{ \begin{array}{l} x_2 = \frac{y_1}{1+y_2} \\ x_1 = \frac{y_1 y_2}{1+y_2} \end{array} \right.$$

El Jacobiano de la transformación inversa $J_{g^{-1}}(y_1, y_2) = \det\left(\left(\frac{\partial x_i}{\partial y_j}\right)_{i,j}\right)$ es

$$\begin{aligned} J_{g^{-1}}(y_1, y_2) &= \frac{\partial x_1}{\partial y_1}\frac{\partial x_2}{\partial y_2} - \frac{\partial x_1}{\partial y_2}\frac{\partial x_2}{\partial y_1} \\ &= \left(\frac{y_2}{1 + y_2}\right)\left(\frac{-y_1}{(1 + y_2)^2}\right) - \left(\frac{y_1}{(1 + y_2)^2}\right)\left(\frac{1}{1 + y_2}\right) \\ &= \frac{-y_1 y_2}{(1 + y_2)^3} - \frac{y_1}{(1 + y_2)^3} = -\frac{y_1(1 + y_2)}{(1 + y_2)^3} = -\frac{y_1}{(1 + y_2)^2} \end{aligned} \tag{33}$$

Substituyendo los resultados (32) y (33) en la fórmula (29) se obtiene:

$$f_{Y_1,Y_2}(y_1, y_2) = f_{X_1,X_2}\left(\frac{y_1 y_2}{1 + y_2}, \frac{y_1}{1 + y_2}\right)\frac{|y_1|}{(1 + y_2)^2}. \tag{34}$$

Por hipótesis,

$$\begin{aligned} f_{X_1,X_2}(x_1, x_2) &= f_{X_1}(x_1) f_{X_2}(x_2) = \lambda e^{-\lambda x_1}\mathbf{1}\{x_1 > 0\}\lambda e^{-\lambda x_2}\mathbf{1}\{x_2 > 0\} \\ &= \lambda^2 e^{-\lambda(x_1 + x_2)}\mathbf{1}\{x_1 > 0, x_2 > 0\}. \end{aligned} \tag{35}$$

De (34) y (35) se obtiene

$$\begin{aligned} f_{Y_1,Y_2}(y_1, y_2) &= \lambda^2 e^{-\lambda y_1}\frac{y_1}{(1 + y_2)^2}\mathbf{1}\{y_1 > 0, y_2 > 0\} \\ &= \left(\lambda^2 y_1 e^{-\lambda y_1}\mathbf{1}\{y_1 > 0\}\right)\left(\frac{1}{(1 + y_2)^2}\mathbf{1}\{y_2 > 0\}\right). \end{aligned} \tag{36}$$

De (36) se deduce que las variables Y_1 e Y_2 son independientes.

Nota Bene sobre la persistencia de la mala suerte. De (36) se deduce que la densidad del cociente $Y_2 = X_1/X_2$ de dos variables exponenciales independientes de igual intensidad es de la forma

$$f_{Y_2}(y_2) = \frac{1}{(1+y_2)^2}\mathbf{1}\{y_2 > 0\}. \tag{37}$$

En consecuencia, *la variable Y_2 tiene esperanza infinita*. Se trata de un hecho notable que ofrece una explicación probabilística de un fenómeno conocido por cualquiera que haya entrado en una fila de espera denominado la *persistencia de la mala suerte*[1]

¿Por qué? Supongamos que la variable X_1 representa el tiempo de espera para ser atendidos en la fila elegida (a la que llamaremos la fila 1) y que X_2 representa el tiempo de espera en otra fila que estamos observando mientras esperamos ser atendidos (a la que llamaremos la fila 2). El cociente X_1/X_2 representa la proporción del tiempo esperado en la fila 1 en en relación al tiempo de espera en fila 2. Por ejemplo, $X_1/X_2 \geq 3$ significa esperamos por lo menos el triple del tiempo que hubiésemos esperado en la otra fila.

Integrando (37) se deduce que

$$\mathbb{P}(Y_2 \leq y_2) \;=\; \int_0^{y_2} \frac{1}{(1+y)^2}dy = 1 - \frac{1}{1+y_2} = \frac{y_2}{1+y_2}, \qquad y_2 \geq 0$$

Equivalentemente,

$$\mathbb{P}(Y_2 > y_2) = \frac{1}{1+y_2}, \qquad y_2 \geq 0$$

En particular, la probabilidad de que tengamos que esperar por lo menos el triple del tiempo que hubiésemos esperado en la otra fila es 1/4. Aunque de acuerdo con este modelo, en promedio, la mitad de las veces esperamos menos tiempo que en la otra fila, en la práctica, el fenómeno de la *mala suerte* se ve sobredimensionado porque no le prestamos atención a los tiempos cortos de espera.

Para percibir qué significa prácticamente el resultado $\mathbb{E}[X_1/X_2] = +\infty$ basta simular algunos valores de la variable X_1/X_2. Por ejemplo, en 10 simulaciones obtuvimos la siguiente muestra:

1.2562, 0.8942, 0.9534, 0.3596, **29.3658**, 1.2641, 3.3443, 0.3452, **13.5228**, 7.1701.

El lector puede extraer sus propias conclusiones. □

Ejemplo 2.11 (Gammas y betas)**.** Sean X_1 y X_2 variables aleatorias independientes con distribuciones $\Gamma(\nu_1, \lambda)$ y $\Gamma(\nu_2, \lambda)$. Vamos a hallar la densidad conjunta de (Y_1, Y_2) donde

$$Y_1 = X_1 + X_2, \qquad \text{e} \qquad Y_2 = \frac{X_1}{X_1 + X_2}.$$

[1]Basta elegir una fila en las múltiples cajas de un supermercado para sufrir este fenómeno y observar que en la fila elegida el tiempo de espera es el doble o el triple que el tiempo de espera en las otras filas.

Para ello consideramos la transformación

$$g(x_1, x_2) = \left(x_1 + x_2, \frac{x_1}{x_1 + x_2} \right) = (y_1, y_2).$$

La transformación inversa de g es

$$x_1 = y_1 y_2, \qquad x_2 = y_1(1 - y_2). \tag{38}$$

El Jacobiano de la transformación inversa es

$$J_{g^{-1}}(y_1, y_2) = \frac{\partial x_1}{\partial y_1}\frac{\partial x_2}{\partial y_2} - \frac{\partial x_1}{\partial y_2}\frac{\partial x_2}{\partial y_1} = y_2(-y_1) - y_1(1 - y_2) = -y_1 \tag{39}$$

Substituyendo los resultados (38) y (39) en la fórmula (29) se obtiene:

$$f_{Y_1,Y_2}(y_1, y_2) = f_{X_1,X_2}\left(y_1 y_2, y_1(1 - y_2)\right)|y_1|. \tag{40}$$

Por hipótesis,

$$\begin{aligned} f_{X_1,X_2}(x_1, x_2) &= = \frac{\lambda^{\nu_1} x_1^{\nu_1 - 1} e^{-\lambda x_1}}{\Gamma(\nu_1)} \mathbf{1}\{x_1 > 0\} \frac{\lambda^{\nu_2} x_2^{\nu_2 - 1} e^{-\lambda x_2}}{\Gamma(\nu_2)} \mathbf{1}\{x_2 > 0\} \\ &= \frac{\lambda^{\nu_1+\nu_2} x_1^{\nu_1 - 1} x_2^{\nu_2 - 1} e^{-\lambda(x_1 + x_2)}}{\Gamma(\nu_1)\Gamma(\nu_2)} \mathbf{1}\{x_1 > 0, x_2 > 0\}. \end{aligned} \tag{41}$$

De (40) y (41) se obtiene

$$\begin{aligned} f_{Y_1,Y_2}(y_1, y_2) &= \frac{\lambda^{\nu_1+\nu_2}(y_1 y_2)^{\nu_1 - 1}(y_1(1 - y_2))^{\nu_2 - 1} e^{-\lambda y_1}}{\Gamma(\nu_1)\Gamma(\nu_2)} \mathbf{1}\{y_1 y_2 > 0, y_1(1 - y_2) > 0\}|y_1| \\ &= \left(\frac{\lambda^{\nu_1+\nu_2} y_1^{\nu_1+\nu_2-1} e^{-\lambda y_1}}{\Gamma(\nu_1 + \nu_2)} \mathbf{1}\{y_1 > 0\} \right) \\ &\quad \times \left(\frac{\Gamma(\nu_1 + \nu_2) y_2^{\nu_1 - 1}(1 - y_2)^{\nu_2 - 1}}{\Gamma(\nu_1)\Gamma(\nu_2)} \mathbf{1}\{0 < y_2 < 1\} \right). \end{aligned} \tag{42}$$

Por lo tanto, Y_1 e Y_2 son independientes y sus distribuciones son $Y_1 \sim \Gamma(\nu_1 + \nu_2, \lambda)$, $Y_2 \sim \beta(\nu_1, \nu_2)$:

$$\begin{aligned} f_{Y_1}(y_1) &= \frac{\lambda^{\nu_1+\nu_2}}{\Gamma(\nu_1 + \nu_2)} y_1^{\nu_1+\nu_2-1} e^{-\lambda y_1} \mathbf{1}\{y_1 > 0\}, \\ f_{Y_2}(y_2) &= \frac{\Gamma(\nu_1 + \nu_2)}{\Gamma(\nu_1)\Gamma(\nu_2)} y_2^{\nu_1 - 1}(1 - y_2)^{\nu_2 - 1} \mathbf{1}\{0 < y_2 < 1\}. \end{aligned}$$

□

Nota Bene. Algunos autores utilizan (y promueven!) el método del Jacobiano como una herramienta para obtener la densidad de variables aleatorias de la forma $Y_1 = g_1(X_1, X_2)$. Hacen lo siguiente: 1. Introducen una variable auxiliar de la forma $Y_2 = g_2(X_1, X_2)$ para obtener un cambio de variables $(g_1, g_2) : \mathbb{R}^2 \to \mathbb{R}^2$. 2. Utilizan la fórmula del Jacobiano (29) para obtener la densidad conjunta de (Y_1, Y_2) a partir de la densidad conjunta de (X_1, X_2). 3. Obtienen la densidad de Y_1 marginando (i.e., integrando la densidad conjunta de (Y_1, Y_2) con respecto de y_2). Por ejemplo,

Suma: $(X_1, X_2) \to (X_1 + X_2, X_2) =: (Y_1, Y_2)$. En tal caso, $(x_1, x_2) = (y_1 - y_2, y_2)$ y el Jacobiano tiene la forma $J(y_1, y_2) = \frac{\partial x_1}{\partial y_1}\frac{\partial x_2}{\partial y_2} - \frac{\partial x_1}{\partial y_2}\frac{\partial x_2}{\partial y_1} = 1$. De donde se obtiene

$$f_{Y_1}(y_1) = \int_{\mathbb{R}} f_{X_1,X_2}(y_1 - y_2, y_2)dy_2.$$

Producto: $(X_1, X_2) \to (X_1X_2, X_1) =: (Y_1, Y_2)$. En tal caso, $(x_1, x_2) = (y_2, y_1/y_2)$ y el Jacobiano tiene la forma $J(y_1, y_2) = \frac{\partial x_1}{\partial y_1}\frac{\partial x_2}{\partial y_2} - \frac{\partial x_1}{\partial y_2}\frac{\partial x_2}{\partial y_1} = -\frac{1}{y_2}$. De donde se obtiene

$$f_{Y_1}(y_1) = \int_{\mathbb{R}} f_{X_1,X_2}(y_2, y_1/y_2)|y_2|^{-1}dy_2.$$

Cociente: $(X_1, X_2) \to (X_1/X_2, X_2) =: (Y_1, Y_2)$. En tal caso, $(x_1, x_2) = (y_1y_2, y_2)$ y el Jacobiano tiene la forma $J(y_1, y_2) = \frac{\partial x_1}{\partial y_1}\frac{\partial x_2}{\partial y_2} - \frac{\partial x_1}{\partial y_2}\frac{\partial x_2}{\partial y_1} = y_2$. De donde se obtiene

$$f_{Y_1}(y_1) = \int_{\mathbb{R}} f_{X_1,X_2}(y_1y_2, y_2)|y_2|dy_2.$$

$\square$

Ejercicios adicionales

3. [James p.97] Si X, Y, Z tienen densidad conjunta

$$f_{X,Y,Z}(x, y, z) = \frac{6}{(1 + x + y + z)^4}\mathbf{1}\{x > 0, y > 0, z > 0\}.$$

Hallar la densidad de la variable aleatoria $W = X + Y + Z$ de dos maneras diferentes (método básico y método del Jacobiano)

2.3. Funciones k a 1

Para determinar la distribución $\mathbf{Y} = g(\mathbf{X})$ podemos utilizar el método del jacobiano en muchos casos en que la función g no es 1 a 1, bastando para ello que g sea 1 a 1 cuando se la restringe a una de k regiones abiertas disjuntas cuya unión contiene al valor de $\mathbf{X}$ con probabilidad 1.

Supongamos que G, $G_1, \ldots, G_k$ sean regiones abiertas de $\mathbb{R}^n$ tales que $G_1, \ldots G_k$ sean disjuntas y valga

$$\mathbb{P}\left(\mathbf{X} \in \bigcup_{\ell=1}^{k} G_\ell\right) = 1$$

y tales que la función $g|G_\ell$, la restricción de g a G_ℓ, sea una correspondencia 1 a 1 entre G_ℓ y G, para todo $\ell = 1, \ldots, k$. Además de eso, supongamos que la función inversa de $g|G_\ell$, denotada por $h^{(\ell)}$, satisfaga todas las condiciones de la función h del caso anterior, e indiquemos con $J_\ell(\mathbf{x}, \mathbf{y})$ el jacobiano de la función $h^{(\ell)}$. (Este jacobiano es función de $\mathbf{y} \in G$. Notemos que $h^{(\ell)} : G \to G_\ell$ es una biyección.)

Teorema 2.12. *Bajo las condiciones enunciadas más arriba, si* $\mathbf{X}$ *tiene densidad* $f_{\mathbf{X}}(\mathbf{x})$*, entonces* $\mathbf{Y}$ *tiene densidad*

$$f_{\mathbf{Y}}(\mathbf{y}) = \sum_{\ell=1}^{k} f_{\mathbf{X}}(h^{(l)}(\mathbf{y}))|J_\ell(\mathbf{x}, \mathbf{y})|\mathbf{1}\{\mathbf{y} \in G\}$$

Demostración. Sea $B \subset G$,

$$\begin{aligned}
\mathbb{P}(\mathbf{Y} \in B) &= \mathbb{P}(g(\mathbf{X}) \in B) = \sum_{\ell=1}^{k} \mathbb{P}(g(\mathbf{X}) \in B, X \in G_\ell) = \sum_{\ell=1}^{k} \mathbb{P}(X \in h^{(\ell)}(B)) \\
&= \sum_{\ell=1}^{k} \int_{h^{(\ell)}(B)} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} = \text{(cambio de variables en la integral)} \\
&= \sum_{\ell=1}^{k} \int_B f_{\mathbf{X}}(h^{(\ell)}(\mathbf{y}))|J_\ell(\mathbf{x}, \mathbf{y})| d\mathbf{y} = \int_B \left(\sum_{\ell=1}^{k} f_{\mathbf{X}}(h^{(\ell)}(\mathbf{y}))|J_\ell(\mathbf{x}, \mathbf{y})|\right) d\mathbf{y}.
\end{aligned}$$

□

Ejemplo 2.13. Sean X e Y independientes con distribución común $\mathcal{N}(0, 1)$. Mostrar que $Z = X^2 + Y^2$ y $W = X/Y$ son independientes y hallar sus distribuciones.

Solución. La función $g : \mathbb{R}^2 \to \mathbb{R}^2$, definida por $g(x, y) = (x^2 + y^2, x/y) = (z, w)$, es 2 a 1.

Sean $G = \{(x, y) : x > 0\}$, $G_1 = \{(x, y) : y > 0\}$, $G_2 = \{(x, y) : y < 0\}$. Entonces, $g|G_1$ y $g|G_2$ son correspondencias 1 a 1 entre las regiones abiertas G_i y G, $i = 1, 2$, y $\mathbb{P}((X, Y) \in G_1 \cup G_2) = 1$ (la probabilidad que (X, Y) tome valores en la recta $\{(x, y) : y = 0\}$ es $\mathbb{P}(Y = 0) = 0$).

Necesitamos, entonces, obtener los jacobianos de las funciones inversas $h^{(1)}$ y $h^{(2)}$ en G. Pero para ello, basta obtener los jacobianos de las funciones $g|G_1$ y $g|G_2$, que son recíprocos de los jacobianos de las inversas, y substituimos el valor (x, y) por el valor $h^{(1)}(z, w)$ o $h^{(2)}(z, w)$. Tenemos

$$J_1(z, w) = \left(\begin{vmatrix} 2x & 2y \\ \frac{1}{y} & -\frac{x}{y^2} \end{vmatrix}\right)^{-1} = \left(-2\left(\frac{x^2}{y^2} + 1\right)\right)^{-1} = -\frac{1}{2(w^2 + 1)}$$

y

$$J_2(z,w) = -\frac{1}{2(w^2+1)}.$$

Por lo tanto, la densidad de (Z, W) es

$$f_{Z,W}(z,w) = \left(f(h^{(1)}(z,w)) + f(h^{(2)}(z,w))\right)\frac{1}{2(w^2+1)}\mathbf{1}\{(z,w)\in G\}.$$

Como

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

tenemos

$$\begin{aligned} f_{Z,W}(z,w) &= 2\left(\frac{1}{2\pi}e^{-z/2}\right)\frac{1}{2(w^2+1)}\mathbf{1}\{z>0, w\in\mathbb{R}\} \\ &= \left(\frac{1}{2}e^{-z/2}\mathbf{1}\{z>0\}\right)\frac{1}{\pi(w^2+1)}. \end{aligned}$$

Como la densidad conjunta es el producto de dos densidades, concluimos que Z y W son independientes, $Z \sim Exp(1/2)$ y $W \sim$ Cauchy. □

2.4. Mínimo y máximo

Sean X e Y dos variables aleatorias con función de distribución conjunta continua $F_{X,Y}(x,y)$. Sean $U := \text{mín}\{X,Y\}$ y $V := \text{máx}\{X,Y\}$. Queremos hallar las funciones de distribución conjunta y las marginales de U de V.

Sean u y v dos números reales tales que $u \le v$, es fácil ver que la función de distribución conjunta de U y V es de la forma

$$F_{U,V}(u,v) = \mathbb{P}(U \le u,\, V \le v) = F_{X,Y}(u,v) + F_{X,Y}(v,u) - F_{X,Y}(u,u). \tag{43}$$

Mínimo. La función de distribución marginal del mínimo se puede obtener de la distribución conjunta fijando u y haciendo $v \to \infty$:

$$F_U(u) = \lim_{v\to\infty} F_{U,V}(u,v) = F_X(u) + F_Y(u) - F_{X,Y}(u,u). \tag{44}$$

Cuando las variables X e Y son independientes la identidad (44) adopta la forma

$$F_U(u) = F_X(u) + F_Y(u) - F_X(u)F_Y(u). \tag{45}$$

Si además de independientes, las variables X e Y tienen función de distribución común $F_X = F_Y = F$, resulta

$$F_U(u) = 2F(u) - F(u)^2. \tag{46}$$

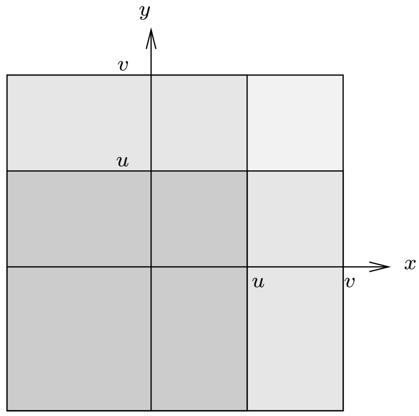

Figura 3: Observar que el evento $\{U \leq u, V \leq v\}$ se puede representar en la forma $\{X \leq u, Y \leq v\} \cup \{Y \leq u, X \leq v\}$.

En cualquiera de los tres casos, derivando respecto de u obtenemos la densidad del mínimo. Por ejemplo, si las variables X e Y son independientes, derivando la identidad (45) se obtiene

$$\begin{aligned} f_U(u) &= f_X(u) + f_Y(u) - f_X(u)F_Y(u) - F_X(u)f_Y(u) \\ &= f_X(u)(1 - F_Y(u)) + f_Y(u)(1 - F_X(u)). \end{aligned} \tag{47}$$

Máximo. La función de distribución marginal del máximo se puede obtener observando que el evento $\{V \leq v\} = \{X \leq v, Y \leq v\}$. Por lo tanto,

$$F_V(v) = F_{X,Y}(v, v) \tag{48}$$

Cuando las variables X e Y son independientes la identidad (48) adopta la forma

$$F_V(v) = F_X(v)F_Y(v) \tag{49}$$

Si además de independientes, las variables X e Y tienen función de distribución común $F_X = F_Y = F$, resulta

$$F_V(v) = F(v)^2. \tag{50}$$

En cualquiera de los tres casos, derivando respecto de v obtenemos la densidad del máximo. Por ejemplo, si las variables X e Y son independientes, derivando la identidad (49) se obtiene

$$f_V(v) = f_X(v)F_Y(v) - F_X(v)f_Y(v). \tag{51}$$

□

Ejercicios adicionales

4. [James p.99] Sean $X_1, \dots, X_n$ variables aleatorias independientes e idénticamente distribuidas, con densidad común f. Mostrar que la densidad conjunta de

$$U = \min_{1 \leq i \leq n} X_i \qquad \text{y} \qquad V = \max_{1 \leq i \leq n} X_i$$

es

$$f_{U,V}(u, v) = n(n-1)[F(v) - F(u)]^{n-2} f(u) f(v) \mathbf{1}\{u < v\}.$$

(*Sugerencia*. Primero hallar $\mathbb{P}(u < U, V \leq v)$. Después, calcular las derivadas parciales cruzadas de la distribución conjunta.)

5. [James p.99] Sean $X_1, \dots, X_n$ variables aleatorias independientes e idénticamente distribuidas, con distribución uniforme en $[0, \theta]$, donde $\theta > 0$. Sean

$$U = \min_{1 \leq i \leq n} X_i \qquad \text{y} \qquad V = \max_{1 \leq i \leq n} X_i$$

(**a**) Mostrar que la densidad conjunta de (U, V) es

$$f_{U,V}(u, v) = n(n-1)(v-u)^{n-2}/\theta^n \mathbf{1}\{0 \leq u < v \leq \theta\}.$$

(**b**) Mostrar que la densidad de $W = V - U$ es

$$f_W(w) = \frac{n(n-1)w^{n-2}}{\theta^{n-1}} \left(1 - \frac{w}{\theta}\right) \mathbf{1}\{0 \leq w \leq \theta\}.$$

---

3. Funciones regulares e independencia

Definición 3.1. *Una función g se dice* regular *si existen números $\cdots < a_1 < a_0 < a_1 < \cdots$, con $a_i \to \infty$ y $a_{-i} \to -\infty$, tales que g es continua y monótona sobre cada intervalo (a_i, a_{i+1}).*

Ejemplo 3.2. La función $\operatorname{sen} x$ es regular; todos los polinomios son funciones regulares. Un ejemplo de una función que no es regular es $\mathbf{1}\{x \in \mathbb{Q}\}$. □

Teorema 3.3. *Sean $X_1, \dots, X_n$ variables aleatorias independientes, y sean $g_1, \dots, g_n$ funciones regulares. Entonces $g(X_1), \dots, g(X_n)$ son variables aleatorias independientes.*

Demostración. Para simplificar la prueba supondremos que $n = 2$. De la regularidad se deduce que para todo $x \in \mathbb{R}$ podemos escribir

$$A_1 := \{y : g_1(y) \leq x\} = \cup_i A_{1,i}(x) \qquad \text{y} \qquad A_2 := \{y : g_2(y) \leq x\} = \cup_i A_{2,i}(x),$$

como uniones de intervalos disjuntos de a pares. Por lo tanto, podemos escribir

$$\begin{aligned}
\mathbb{P}(g_1(X_1) \leq x, g_1(X_2) \leq y) &= \sum_i \sum_j \mathbb{P}(X_1 \in A_{1,i}, X_2 \in A_{2,i}) \\
&= \sum_i \sum_j \mathbb{P}(X_1 \in A_{1,i})\mathbb{P}(X_2 \in A_{2,i}) \\
&= \sum_i \mathbb{P}(X_1 \in A_{1,i}) \sum_j \mathbb{P}(X_2 \in A_{2,i}) \\
&= \mathbb{P}(g_1(X_1) \leq x)\mathbb{P}(g_2(X_2) \in A_{2,i}).
\end{aligned}$$

□

En rigor de verdad, vale un resultado mucho más general.

Teorema 3.4. *Si para* $1 \leq i \leq n$, $1 \leq j \leq m_i$, $X_{i,j}$ *son independientes y* $f_i : \mathbb{R}^{m_i} \to \mathbb{R}$ *son medibles entonces* $f_i(X_{i,1}, \dots, X_{i,m_i})$ *son independientes.*

Demostración. Durrett(1996), p.25-27. □

Un caso concreto que usaremos permanentemente al estudiar sumas es el siguiente: si $X_1, \dots, X_n$ son independientes, entonces $X = X_1 + \cdots + X_{n-1}$ y X_n son independientes.

Ejercicios adicionales

6. (**Fragmentaciones aleatorias.**) Si $U_1, \dots, U_n$ son independientes con distribución común $\mathcal{U}(0,1)$, entonces

$$-\log \prod_{i=1}^{n} U_i \sim \Gamma(n, 1).$$

7. Una varilla de 1 metro de longitud es sometida a un proceso de fragmentación aleatoria. En la primera fase se elige un punto al azar de la misma y se la divide por el punto elegido en dos varillas de longitudes L_1 y L_2. En la segunda fase se elige un punto al azar de la varilla de longitud L_1 y se la divide por el punto elegido en dos varillas de longitudes $L_{1,1}$ y $L_{1,2}$. Calcular la probabilidad de que $L_{1,1}$ sea mayor que 25 centímetros.

---

4. Bibliografía consultada

Para redactar estas notas se consultaron los siguientes libros:

1. Durrett R.:Probability. Theory and Examples. Duxbury Press, Belmont. (1996).

2. Feller, W.: An introduction to Probability Theory and Its Applications. Vol. 2. John Wiley & Sons, New York. (1971).

3. James, B. R.: probabilidade: um curso em nível intermediario. IMPA, Rio de Janeiro. (2002).

4. Meester, R.: A Natural Introduction to Probability Theory. Birkhauser, Berlin. (2008).

5. Meyer, P. L.: Introductory Probability and Statistical Applications. Addison-Wesley, Massachusetts. (1972).

6. Ross, S.: Introduction to Probability Models. Academic Press, San Diego. (2007)

7. Soong, T. T.: Fundamentals of Probability and Statistics for Engineers. John Wiley & Sons Ltd. (2004).