EYP 1025-1027 Métodos Probabilísticos

Clase 16

Profesor: Reinaldo B. Arellano-Valle

Departamento de Estadística
Pontificia Universidad Católica de Chile

Contenido I

Distribución de funciones de vectores aleatorios continuos

Suponga que $(X_1, \ldots, X_n)$ es un vector aleatorio continuo, y considere una transformación $(Y_1, \ldots, Y_n) = (g_1(X_1, \ldots, X_n), \ldots, g_n(X_1, \ldots, X_n))$ de modo que $(Y_1, \ldots, Y_n)$ también sea un vector aleatorio continuo.

Si $f_{\boldsymbol{X}}(\boldsymbol{x}) = f_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$ es la fdp de $(X_1, \ldots, X_n)$, suponga que se desea determinar la fdp $f_{\boldsymbol{Y}}(\boldsymbol{y}) = f_{Y_1,\ldots,Y_n}(y_1, \ldots, y_n)$ de $(Y_1, \ldots, Y_n)$.

Sea $\mathcal{X} = \{\boldsymbol{x} \in \mathbb{R}^n : f_{\boldsymbol{X}}(\boldsymbol{x}) > 0\}$ el soporte de $(X_1, \ldots, X_n)$. Escribiendo $\boldsymbol{y} = \boldsymbol{g}(\boldsymbol{y})$, donde $\boldsymbol{g} = (g_1, \ldots, g_n)$, entonces el recorrido de $(Y_1, \ldots, Y_n)$ queda determinado por,

$$\mathcal{Y} = \{\boldsymbol{y} \in \mathbb{R}^n : \boldsymbol{y} = \boldsymbol{g}(\boldsymbol{x}) \text{ para algún } \boldsymbol{x} \in \mathcal{X}\} \subseteq \mathbb{R}^n.$$

Suponga que $\mathcal{X} \subseteq \mathbb{R}^n$ e $\mathcal{Y} \subseteq \mathbb{R}^n$ son regiones abiertas, y que se distinguen dos casos según si la transformación $\boldsymbol{g} : \mathcal{X} \longrightarrow \mathcal{Y}$ establece o no una correspondencia biunívoca (biyección) entre $\mathcal{X}$ y $\mathcal{Y}$.

Distribución de funciones de vectores aleatorios continuos

Caso biunívoco

Suponga que $\boldsymbol{g} : \mathcal{X} \longrightarrow \mathcal{Y} \subseteq \mathbb{R}^n$ es una biyección entre $\mathcal{X}$ y $\mathcal{Y}$, donde

$$\boldsymbol{g}(x_1, \ldots, x_n) = (g_1(x_1, \ldots, x_n), \ldots, g_n(x_1, \ldots, x_n)) = (y_1, \ldots, y_n).$$

Entonces, existe una función inversa $\boldsymbol{h} = \boldsymbol{g}^{-1}$ en $\mathcal{Y}$, donde

$$x_1 = h_1(y_1, \ldots, y_n), \ldots, x_n = h_n(y_1, \ldots, y_n).$$

Suponga también que existen las derivadas

$$\frac{\partial x_i}{\partial y_j} = \frac{\partial h_i(y_1, \ldots, y_n)}{\partial y_j}, \quad i, j = 1, \ldots, n,$$

y que ellas sean continuas en $\mathcal{Y}$.

Distribución de funciones de vectores aleatorios continuos

Defina el *Jacobiano (de la transformación)* J como el determinante:

$$J = \det \begin{pmatrix} \frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial y_1} & \cdots & \frac{\partial x_n}{\partial y_n} \end{pmatrix},$$

el cual es una función de $(y_1, \ldots, y_n)$. Suponga, finalmente, que $J \neq 0$ para todo $(y_1, \ldots, y_n) \in \mathcal{Y}$.

Teorema 1.1

Sobre las condiciones dadas arriba, la fdp conjunta de $Y_1, \ldots, Y_n$ es,

$$f_{Y_1,\ldots,Y_n}(y_1, \ldots, y_n)$$
$$= \begin{cases} |J| f_{X_1,\ldots,X_n}(h_1(y_1, \ldots, y_n), \ldots, h_n(y_1, \ldots, y_n)), & \text{si } (y_1, \ldots, y_n) \in \mathcal{Y}, \\ & \text{si } (y_1, \ldots, y_n) \notin \mathcal{Y}. \end{cases}$$

Distribución de funciones de vectores aleatorios continuos

Idea de la demostración: Sea $\boldsymbol{h}(B) = \boldsymbol{g}^{-1}(B) = \{\boldsymbol{x} : \boldsymbol{g}(x) \in B\}$, $B \subset \mathbb{R}^n$. Entonces,

$$\begin{aligned}
P_{\boldsymbol{Y}}(B) &= P(\boldsymbol{Y} \in B) \\
&= P(\boldsymbol{g}(\boldsymbol{X}) \in B) \\
&= P(\boldsymbol{X} \in \boldsymbol{g}^{-1}(B)) \\
&= P(\boldsymbol{X} \in \boldsymbol{h}(B)) \\
&= \int_{\boldsymbol{h}(B)} f_{\boldsymbol{X}}(\boldsymbol{x}) d\boldsymbol{x} \quad (\boldsymbol{y} = \boldsymbol{g}(\boldsymbol{x}) \Longrightarrow \boldsymbol{x} = \boldsymbol{g}^{-1}(\boldsymbol{y}) = \boldsymbol{h}(\boldsymbol{y})) \\
&= \int_{B} f_{\boldsymbol{X}}(\boldsymbol{h}(\boldsymbol{y}))|J| d\boldsymbol{y} \quad (\boldsymbol{x} \in \boldsymbol{h}(B) \Longrightarrow \boldsymbol{y} = \boldsymbol{g}(\boldsymbol{x}) \in B),
\end{aligned}$$

de acuerdo con la formula para cambio de variables en integrales múltiples.

Por otro lado, como el v.a. $\boldsymbol{Y} = \boldsymbol{g}(\boldsymbol{X})$ también es (absolutamente) continuo, entonces

$$P_{\boldsymbol{Y}}(B) = P(\boldsymbol{Y} \in B) = \int_B f_{\boldsymbol{Y}}(\boldsymbol{y})d\boldsymbol{y}, \quad (**)$$

donde $f_{\boldsymbol{Y}}(\boldsymbol{y})$ es la fdp de $\boldsymbol{Y}$.

Considerando $(*)$ y $(**)$, se concluye que

$$\int_B f_{\boldsymbol{X}}(\boldsymbol{h}(\boldsymbol{y}))|J|d\boldsymbol{y} = \int_B f_{\boldsymbol{Y}}(\boldsymbol{y})d\boldsymbol{y}, \quad \forall\, B \subset \mathbb{R}^n,$$

y por lo tanto que $f_{\boldsymbol{X}}(\boldsymbol{h}(\boldsymbol{y}))|J| = f_{\boldsymbol{Y}}(\boldsymbol{y}), \quad \boldsymbol{y} \in \mathbb{R}^n$.

Distribución de funciones de vectores aleatorios continuos

Formalización para $n = 2$:

Teorema 1.2

Sea (X_1, X_1) un vector aleatorio continuo con fdp conjunta $f_{(X_1,X_2)}$. Sean, $\mathcal{X} = \{(x_1, x_2) : f_{(X_1,X_2)}(x_1, x_2) > 0\} \subseteq \mathbb{R}^2$. Suponga que,

i) $y_1 = g_1(x_1, x_2)$ e $y_2 = g_2(x_1, x_2)$ definen una transformación uno a uno de $\mathcal{X}$ en $\mathcal{Y} = \{(y_1, y_2) : y_1 = g_1(x_1, x_2),\ y_2 = g_2(x_1, x_2)$ para algún $(x_1, x_2) \in \mathcal{Y}\} \subseteq \mathbb{R}^2$.

ii) Las derivadas parciales de la trasnformación inversa $x_1 = h_1(y_1, y_2)$ y $x_2 = h_2(y_1, y_2)$ son continuas sobre $\mathcal{Y}$.

iii) El jacobiano de la transformación

$$J = \det \begin{pmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{pmatrix} \text{ es } \neq 0 \text{ para } (y_1, y_2) \ \in \ \mathcal{Y}.$$

Entonces la fdp conjunta de $Y_1 = g_1(X_1, X_2)$ y $Y_2 = g_2(X_1, X_2)$, está dada por,

$$f_{Y_1,Y_2}(y_1, y_2) = \begin{cases} |J| f_{X_1,X_2}(h_1(y_1, y_2), h_2(y_1, y_2), & \text{si } (y_1, y_2) \in \mathcal{Y}, \\ 0, & \text{si } (y_1, y_2) \notin \mathcal{Y}. \end{cases}$$

Por ejemplo, para $n = 2$, la transformación,

$$y_1 = x_1 + x_2 \longrightarrow g_1(x_1, x_2) = x_1 + x_2$$
$$y_2 = x_1 - x_2 \longrightarrow g_2(x_1, x_2) = x_1 - x_2$$

es uno-a-uno sobre todo $\mathbb{R}^2$. La transformación inversa es,

$$x_1 = (y_1 + y_2)/2 \longrightarrow h_1(x_1, x_2) = (y_1 + y_2)/2$$
$$x_2 = (y_1 - y_2)/2 \longrightarrow h_2(x_1, x_2) = (y_1 - y_2)/2,$$

y el *Jacobiano de la transformación* es,

$$J = \det \begin{pmatrix} \partial x_1/\partial y_1 & \partial x_1/\partial y_2 \\ \partial x_2/\partial y_1 & \partial x_2/\partial y_2 \end{pmatrix} = \det \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix} = -1/2$$

$\Longrightarrow |J| = +1/2$. Asi,

$$f_{Y_1,Y_2}(y_1, y_2) = \begin{cases} \frac{1}{2} f_{X_1,X_2}((y_1 + y_2)/2, (y_1 - y_2)/2), & \text{si } (y_1, y_2) \in \mathcal{Y}, \\ 0, & \text{eoc.} \end{cases}$$

Aplicaciones:

1) Sean $X_1, X_2 \overset{\text{iid}}{\sim} \exp(1)$, es decir,

$$f_{X_1,X_2}(x_1,x_2) = f_{X_1}(x_1) f_{X_2}(x_2) = \begin{cases} e^{-x_1}e^{-x_2}, & \text{si } x_1 > 0,\ x_2 > 0, \\ 0, & \text{eoc.} \end{cases}$$

Queremos la fdp de $(Y_1, Y_2) = (X_1 + X_2, X_1 - X_2)$. Note que $x_1, x_2 > 0 \Longrightarrow y_1 = x_1 + x_2 > 0$. Luego,

$$f_{Y_1,Y_2}(y_1,y_2) = \begin{cases} \frac{1}{2}e^{-\frac{y_1+y_2}{2}}e^{-\frac{y_1-y_2}{2}}, & \text{si } \frac{y_1+y_2}{2} > 0,\ \frac{y_1-y_2}{2} > 0, \\ 0, & \text{eoc,} \end{cases}$$

$$= \begin{cases} \frac{1}{2}e^{-y_1}, & \text{si } y_2 > y_1,\ y_2 < -y_1,\ y_1 > 0, \\ 0, & \text{eoc,} \end{cases}$$

$$= \begin{cases} \frac{1}{2}e^{-y_1}, & \text{si } |y_2| < y_1,\ y_1 > 0, \\ 0, & \text{eoc.} \end{cases}$$

2) Sean $X_1, X_2 \overset{\text{iid}}{\sim} U(0,1)$, es decir,

$$f_{X_1,X_2}(x_1,x_2) = f_{X_1}(x_1) f_{X_2}(x_2) = \begin{cases} 1, & \text{si } 0 < x_1 < 1,\, 0 < x_2 < 1, \\ 0, & \text{eoc.} \end{cases}$$

Entonces, la fdp de $(Y_1, Y_2) = (X_1 + X_2, X_1 - X_2)$ es,

$$\begin{aligned} f_{Y_1,Y_2}(y_1,y_2) &= \begin{cases} \frac{1}{2}, & \text{si } 0 < \frac{y_1+y_2}{2} < 1,\, 0 < \frac{y_1-y_2}{2} < 1, \\ 0, & \text{eoc,} \end{cases} \\ &= \begin{cases} \frac{1}{2}, & \text{si } -y_1 < y_2 < 2 - y_1,\, 0 < y_1 < 1, \\ & -(2-y_1) < y_2 < y_1,\, 1 < y_1 < 2, \\ 0, & \text{eoc.} \end{cases} \end{aligned}$$

Tarea: En 1) y 2), bosqueje los recorridos $\mathcal{X}$ e $\mathcal{Y}$, determine las distribuciones marginales de Y_1 e Y_2, calcule Cov(Y_1, Y_2), y estudie si las nuevas variables son independientes.

Distribución de funciones de vectores aleatorios continuos

3) Sean $X_1, X_2 \overset{\text{iid}}{\sim} N(0,1)$, es decir, con fdp's marginales

$$f_{X_1}(x) = f_{X_2}(x) = \phi(x) = (2\pi)^{-1/2} e^{-x^2/2}, \quad x \in \mathbb{R},$$

y fdp conjunta,

$$f_{X_1,X_2}(x_1,x_2) = \phi(x_1)\phi(x_2) = (2\pi)^{-1} e^{-(x_1^2+x_2^2)/2}, \quad (x_1,x_2) \in \mathbb{R}^2.$$

Entonces,

$$Y_1 = X_1 + X_2 \sim N(0,2), \quad f_{Y_1}(y_1) = \frac{1}{\sqrt{2}}\phi\left(\frac{y_1}{\sqrt{2}}\right),\ y_1 \in \mathbb{R},$$
$$Y_2 = X_1 - X_2 \sim N(0,2), \quad f_{Y_2}(y_2) = \frac{1}{\sqrt{2}}\phi\left(\frac{y_2}{\sqrt{2}}\right),\ y_2 \in \mathbb{R}.$$

Bosqueje los recorridos en este caso. Obtenga la fdp conjunta de Y_1 e Y_2. ¿Son Y_1 e Y_2 variables aleatorias independientes? Justifique.

Ejemplo 1.1

Sean X_1 e X_2 variables aleatorias continuas independientes. Se desea la fdp de $X_1 + X_2$.

Paso 1: Sean $Y_1 = X_1 + X_2$ (variable de interés) e $Y_2 = X_2$ (variable auxiliar arbitraria). Entonces, $y_1 = g_1(x_1, x_2) = x_1 + x_2$ e $y_2 = g_1(x_1, x_2) = x_2$, con inversos $x_1 = h_1(y_1, y_2) = y_1 - y_2$ y $x_2 = h_2(y_1, y_2) = y_2$. Luego,

$$J = \begin{pmatrix} \partial x_1/\partial y_1 & \partial x_1/\partial y_2 \\ \partial x_2/\partial y_1 & \partial x_2/\partial y_2 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = 1,$$

y

$$f_{Y_1,Y_2}(y_1, y_2) = f_{X_1,X_2}(y_1 - y_2, y_2) = f_{X_1}(y_1 - y_2) f_{X_2}(y_2).$$

Paso 2: $f_{Y_1}(y_1) = \int_{\{y_2 : (y_1 - y_2, y_2) \in \mathcal{X}\}} f_{X_1}(y_1 - y_2) f_{X_2}(y_2) dy_2$.

Distribución de funciones de vectores aleatorios continuos

Por ejemplo, si $Y = X_1 + X_2$, donde $X_1, X_2 \overset{\text{iid}}{\sim} \exp(\lambda)$, entonces

$$f_Y(y) = \begin{cases} \int_0^y \lambda^2 e^{-\lambda(y-z+z)} dz & \text{si } y > 0, \\ 0, & \text{eoc}, \end{cases}$$
$$= \begin{cases} \lambda^2 y e^{-\lambda y} & \text{si } y > 0, \\ 0, & \text{eoc}, \end{cases}$$

es decir, $Y = X_1 + X_2 \sim Gama(2, \lambda)$

Extensión: Si $X_i \sim Gama(\alpha_i, \lambda)$, $i = 1, \ldots, n$ son va's independientes, entonces $Y = \sum_{i=1}^n X_i \sim Gama\left(\sum_{i=1}^n \alpha_i, \lambda\right)$.

Distribución de funciones de vectores aleatorios continuos

Caso no-biunívoco

En muchos casos donde $\boldsymbol{g} : \mathcal{X} \longrightarrow \mathcal{Y}$ no es uno-a-uno, también podemos determinar la distribución de $\boldsymbol{Y} = \boldsymbol{g}(\boldsymbol{X})$ mediante el método del jacobiano descrito anteriormente.

Para hacer esto, es suficiente que $\boldsymbol{g} = (g_1, \ldots, g_n)$ sea uno-a-uno cuando se restringe a cada una de k regiones (abiertas) disjuntas cuya unión contenga el valor de $\boldsymbol{X}$ con probabilidad 1.

A continuación, se ilustra el procedimiento en el caso bivariado.

Distribución de funciones de vectores aleatorios continuos

Considere el caso bivariado con $n = 2$; la extensión para $n > 2$ es similar.

Si la transformación no es uno-a-uno, sea $\{B_0, B_1, \ldots, B_k\}$ una partición de $\mathcal{X} = \{(x_1, x_2) : f_{X_1,X_2}(x_1, x_2) > 0\}$.

Suponga que:
a) El conjunto B_0 (posiblemente vacío) satisface $P\left((X_1, X_2) \in B_0\right) = 0$.

b) Para cada $i = 1, \ldots, k$, la transformación

$$(y_1, y_2) = (g_1(x_1, x_2), g_2(x_1, x_2)),$$

es uno-a-uno desde B_i sobre $\mathcal{Y}$.

Entonces, para cada $i = 1, \ldots, k$, existe la transformación inversa desde $\mathcal{Y}$ a B_i; denote la $i-$ésima transformación inversa por,

$$(x_{1i}, x_{2i}) = (h_{1i}(y_1, y_2), h_{2i}(y_1, y_2)) \quad i = 1, \ldots, k.$$

Distribución de funciones de vectores aleatorios continuos

Sea J_i el jacobiano calculado desde la $i-$ésima transformación inversa, es decir,

$$J_i = \det \begin{pmatrix} \frac{\partial x_{1i}}{\partial y_1} & \frac{\partial x_{1i}}{\partial y_2} \\ \frac{\partial x_{2i}}{\partial y_1} & \frac{\partial x_{2i}}{\partial y_2} \end{pmatrix}, \quad i = 1, \ldots, k.$$

Suponga, fialmente, que estos jacobianos no son cero en $\mathcal{Y}$.

Entonces, bajo las condiciones anteriores, la fdp de (Y_1, Y_2) esta dada por,

$$\begin{aligned} & f_{Y_1,Y_2}(y_1, y_2) \\ & \qquad = \begin{cases} \sum_{i=1}^{k} |J_i| \, f_{X_1,X_2}\left(h_{1i}(y_1, y_2), h_{2i}(y_1, y_2)\right), & \text{si } (y_1, y_2) \in \mathcal{Y}, \\ 0, & \text{si } (y_1, y_2) \notin \mathcal{Y}. \end{cases} \end{aligned}$$

Nota: La extensión del resultado al caso n-dimensional es inmediata.

Ejemplo 1.2

Sean $X_1, X_2 \overset{\text{iid}}{\sim} N(0,1)$. Considere la transformación

$$Y_1 = \frac{X_1}{X_2} \quad \text{y} \quad Y_2 = |X_2|.$$

Esta es una transformación desde $\mathcal{X} = \mathbb{R}^2$ a $\mathcal{Y} = \{(y_1, y_2) : y_2 > 0\} \subset \mathbb{R}^2$ que no es uno-a-uno, ya que los puntos (x_1, x_2) y $(-x_1, -x_2)$ de $\mathcal{X}$ son ambos mapeado en el mismo punto (y_1, y_2) de $\mathcal{Y}$.

Pero si nos restringimos a los valores positivos o negativos de x_2, entonces la transformación restringida es uno-a-uno en cada caso.

Distribución de funciones de vectores aleatorios continuos

Sean $B_0 = \{(x_1, x_2) : x_2 = 0\}$, $B_1 = \{(x_1, x_2) : x_2 > 0\}$ y $B_2 = \{(x_1, x_2) : x_2 < 0\}$.

Entonces, B_0, B_1 y B_2 forman una partición de $\mathcal{X} = \mathbb{R}^2$ tal que

a) $P\left((X_1, X_2) \in B_0\right) = P(X_2 = 0) = 0$

b) $\mathcal{Y} = \{(y_1, y_2) : y_2 > 0\}$ es la imagen tanto de B_1 como de B_2 bajo la transformación $(g_1(x_1, x_2), g_2(x_1, x_2)) = (x_1/x_2, |x_2|)$.

Además, las transformaciones inversas de $\mathcal{Y}$ a B_1 e $\mathcal{Y}$ a B_2 están dadas por

$$\boldsymbol{h}_1 : \mathcal{Y} \to B_1 \Longrightarrow \; x_{11} = h_{11}(y_1, y_2) = y_1 y_2, \quad x_{21} = h_{21}(y_1, y_2) = y_2,$$
$$\boldsymbol{h}_2 : \mathcal{Y} \to B_1 \Longrightarrow \; x_{12} = h_{12}(y_1, y_2) = -y_1 y_2, \quad x_{22} = h_{22}(y_1, y_2) = -y_2.$$

Note que el primer inverso produce valores positivos de x_2 mientras que el segundo inverso da valores negativos de x_2.

Note también que los jacobianos de las dos transformaciones inversas son $|J_1| = |J_2| = y_2$ para todo $(y_1, y_2) \in \mathcal{Y}$.

Distribución de funciones de vectores aleatorios continuos

Luego, usando que $f_{X_1,X_2}(x_1,x_2) = \phi(x_1)\phi(x_2) = (2\pi)^{-1}e^{-x_1^2/2}e^{-x_2^2/2}$, para $-\infty < x_1 < \infty$ y $-\infty < x_2 < \infty$, se obtiene que,

$$
\begin{aligned}
&f_{Y_1,Y_2}(y_1,y_2)\\
&\quad = \begin{cases} \frac{|y_2|}{2\pi}e^{-(y_1y_2)^2/2}e^{-y_2^2/2} + \frac{|y_2|}{2\pi}e^{-(-y_1y_2)^2/2}e^{-(-y2)^2/2}, \\ \qquad\qquad\qquad\qquad \text{si } -\infty < y_1 < \infty,\ 0 < y_2 < \infty, \\ 0, & \text{eoc}, \end{cases}\\
&\quad = \begin{cases} \frac{y_2}{\pi}e^{-\left(1+y_1^2\right)y_2^2/2}, & \text{si } -\infty < y_1 < \infty,\ 0 < y_2 < \infty \\ 0, & \text{eoc}. \end{cases}
\end{aligned}
$$

La fdp marginal de Y_1 puede calcularse como,

$$\begin{aligned} f_{Y_1}(y_1) &= \int_0^\infty \frac{y_2}{\pi} e^{-\left(1+y_1^2\right)y_2^2/2} dy_2 \\ &= \frac{1}{2\pi}\int_0^\infty e^{-\left(1+y_1^2\right)z/2} dz \\ &= \frac{1}{\pi\left(1+y_1^2\right)}, \quad -\infty < y_1 < \infty. \end{aligned}$$

De aquí, se concluye que la razón de dos variables aleatorias normales estándar independientes, es una variable aleatoria con distribución de Cauchy.

Ademas, es fácil ver que $f_{Y_2}(y_2) = 2\phi(y_2)$ para $y_2 > 0$; es decir, si $X_2 \sim N(0,1)$ entonces $Y_2 = |X_2| \sim HN(0,1)$ (half-normal).

Ejemplo 1.3

Asumiendo nuevamente que $X_1, X_2 \overset{\text{iid}}{\sim} N(0,1)$, pruebe que

$$\begin{aligned} Y_1 &= \frac{X_1 + X_2}{2} \sim N(0, 1/2) \\ Y_2 &= \frac{(X_1 - X_2)^2}{2} \sim \chi^2_1 := Gama(1/2, 1/2), \end{aligned}$$

y que Y_1 e Y_2 son variables aleatorias independientes.

Extensión: Sean $X_1, \ldots, X_n \overset{\text{iid}}{\sim} N(\mu, \sigma^2)$. Entonces, la media muestral,

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

y la varianza muestral,

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2,$$

son variables aleatorias independientes, con

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \quad \text{y} \quad \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1},$$

para cada $n \geq 2$.

Distribución de funciones de vectores aleatorios continuos

Ejemplo 1.4

Suponga que $X_1 \sim N(0,1)$ y $X_2 \sim \chi^2_\nu$ $(\nu > 0)$ son independientes. Entonces, la distribución de la variable aleatoria

$$Y = \frac{X_1}{\sqrt{X_2/\nu}},$$

se llama t (de Student) con ν grados de libertad, y se denota como $Y \sim t_\nu := t(0,1,\nu)$.

Tarea: Pruebe que

$$f_Y(y) = \frac{\Gamma[(\nu+1)/2]}{\Gamma(\nu/2)\sqrt{\pi\nu}} \left(1 + \frac{y^2}{\nu}\right)^{-(\nu+1)/2}, \quad -\infty < y < \infty.$$

Nota: $P(Y \leq y) = \int_{-\infty}^{y} f_Y(z)dz$ se encuentra tabulada para varios valores de y y ν.

Distribución de funciones de vectores aleatorios continuos

Aplicación: Sean $X_1, \dots, X_n \overset{\text{iid}}{\sim} N(\mu, \sigma^2)$. Vimos que $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$ y $S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2$ son variables aleatorias independientes, con

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \quad \text{y} \quad \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}.$$

Es decir,

$$\frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \sim N(0,1) \quad \text{y} \quad \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1},$$

y son independientes. Luego,

$$\frac{\sqrt{n}(\bar{X} - \mu)/\sigma}{\sqrt{(n-1)S^2/(n-1)\sigma^2}} = \frac{\sqrt{n}(\bar{X} - \mu)}{S} \sim t_{n-1}.$$

Ejemplo 1.5

Sean $X_1 \sim \chi^2_r$ y $X_2 \sim \chi^2_s$ $(r,\ s > 0)$variables aleatorias independientes. La distribución de la variable aleatoria

$$Y = \frac{X_1/r}{X_2/s},$$

se llama F (de Fisher) con r y s grados de libertad, y se denota como $Y \sim F_{r,s}$. Note que $Y^{-1} \sim F_{s,r}$.

Tarea: Pruebe que

$$f_Y(y) = \begin{cases} \frac{\Gamma[(r+s)/2]}{\Gamma(r/2)\Gamma(s/2)} \left(\frac{r}{s}\right)^{r/2} y^{(r/2)-1} \left(1 + \frac{r}{s}\ y\right)^{-(r+s)/2}, & \text{si } y > 0, \\ 0, & \text{eoc.} \end{cases}$$

Nota: $P(Y \leq y) = \int_{-\infty}^{y} f_Y(z)dz$ se encuentra tabulada para varios valores de y, r y s.

Distribución de funciones de vectores aleatorios continuos

Aplicación: Suponga que

$$X_1, \ldots, X_n \overset{\text{iid}}{\sim} N(\mu_X, \sigma^2)$$

e

$$Y_1, \ldots, Y_m \overset{\text{iid}}{\sim} N(\mu_Y, \sigma^2),$$

donde los X_i's son independientes de los Y_i's. Entonces,

$$(n-1)S_X^2/\sigma^2 = \sum_{i=1}^{n}(X_i - \bar{X})^2/\sigma^2 \sim \chi^2_{n-1}$$

y

$$(m-1)S_Y^2/\sigma^2 = \sum_{i=1}^{m}(Y_i - \bar{Y})^2/\sigma^2 \sim \chi^2_{m-1},$$

y son variables aleatorias independientes. Luego,

$$\frac{S_X^2}{S_Y^2} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2/(n-1)}{\sum_{i=1}^{m}(Y_i - \bar{Y})^2/(m-1)} \sim F_{n-1,m-1}.$$

Ejemplo 1.6

Sean $X_1 \sim Gama(\alpha_1, \lambda)$ y $X_2 \sim Gama(\alpha_2, \lambda)$ $(\alpha_1,\ \alpha_2,\ \lambda > 0)$ variables aleatorias independientes. La distribución de la variable aleatoria

$$Y = \frac{X_1}{X_1 + X_2},$$

se llama beta con parámetros α_1 y α_2, denotada por $Y \sim Beta(\alpha_1, \alpha_2)$.

Tarea: Pruebe que

$$f_Y(y) = \begin{cases} \frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} y^{\alpha_1-1}(1-y)^{\alpha_2-1}, & \text{si } \ 0 < y < 1, \\ 0, & \text{eoc.} \end{cases}$$

Tarea: Pruebe que $Y_1 = \frac{X_1}{X_1+X_2}$ e $Y_2 = X_1 + X_2$ son independientes; usando este resultado pruebe que $\mathrm{E}(Y_1) = \frac{\mathrm{E}(X_1)}{\mathrm{E}(Y_2)} = \frac{\alpha_1}{\alpha_1+\alpha_2}$.

Extensión: Si $X_i \overset{\text{ind}}{\sim} Gama(\alpha_i, \lambda)$, $i = 1, \ldots, n$, entonces,

$$\frac{X_1 + \cdots + X_m}{X_1 + \cdots + X_m + X_{m+1} + \cdots + X_n} \sim Beta(\alpha_1 + \cdots + \alpha_m, \alpha_{m+1} + \cdots + \alpha_n),$$

donde hay que recordar que $\sum_{i=1}^{k} X_i \sim Gama\left(\sum_{i=1}^{k} \alpha_i, \lambda\right)$ para cada $k = 1, \ldots, n$.

Nota: La función beta se define como,

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}, \quad \alpha, \beta > 0.$$

Transformación de variables aleatorias independientes

Nota: Si $X_1, \ldots, X_n$ variables aleatorias independientes (arbitrarias), entonces funciones de subvectores disjuntos de los X_i's también son independientes. Por ejemplo, $Y_1 = X_1 + X_2 + X_3$, $Y_2 = (X_5 - X_4)^2$, $Y_3 = \max\{X_6, X_7\}$ e $Y_4 = e^{-Y_8}$, son variables aleatorias independientes.

Para ilustrar el procedimiento, se probará el siguiente caso especial mediante la fda conjunta.

Teorema 2.1

Sean $X_1, \ldots, X_n$ variables aleatorias independientes. Para cada $i = 1, \ldots, n$, suponga que $g_i(x_i)$ es una función solo de x_i. Entonces, las transformaciones $Y_1 = g_1(X_1), \ldots, Y_n = g_n(X_n)$ también son variables aleatorias independientes.

Distribución de funciones de vectores aleatorios continuos

Demostración 2.1

La fda conjunta de $Y_1 = g_1(X_1), \ldots, Y_n = g_n(X_n)$ es

$$\begin{aligned}
F_{Y_1,\ldots,Y_n}(y_1,\ldots,y_n) &= P(Y_1 \leq y_1, \ldots, Y_n \leq y_n) \\
&= P\{g_1(X_1) \leq y_1, \ldots, g_n(X_n) \leq y_n\} \\
&= P\{g_1(X_1) \in (-\infty, y_1], \ldots, g_n(X_n) \in (-\infty, y_n]\} \\
&= P\{X_1 \in g_1^{-1}((-\infty, y_1]), \ldots, X_n \in g_n^{-1}((-\infty, y_n])\} \\
&= P\{X_1 \in g_1^{-1}((-\infty, y_1])\} \times \cdots \times P\{X_n \in g_n^{-1}((-\infty, y_n])\} \\
&= P\{g_1(X_1) \in (-\infty, y_1]\} \times \cdots \times P\{g_n(X_n) \in (-\infty, y_n]\} \\
&= P\{g_1(X_1) \leq y_1\} \times \cdots \times P\{g_n(X_n) \leq y_n\} \\
&= P\{Y_1 \leq y_1\} \times \cdots \times P\{Y_n \leq y_n\} \\
&= F_{Y_1}(y_1) \times \cdots \times F_{Y_n}(y_n).
\end{aligned}$$

Ejemplo 2.1

Sean $X_1, \ldots, X_n \overset{\text{iid}}{\sim} N(0,1)$. Entonces:

1) $X_1^2, \ldots, X_n^2 \overset{\text{iid}}{\sim} \chi_1^2$; lo que permite probar también (tarea) que $\sum_{i=1}^n X_i^2 \sim \chi_n^2$.

2) $I_{(0,\infty)}(X_1), \ldots, I_{(0,\infty)}(X_n) \overset{\text{iid}}{\sim} Ber(1/2)$, donde

$$I_A(x) = \begin{cases} 1 & \text{si } x \in A \text{ ocurre,} \\ 0 & \text{si } x \in A^c \text{ ocurre.} \end{cases}$$

3) $\Phi(X_1), \ldots, \Phi(X_n) \overset{\text{iid}}{\sim} U(0,1)$, donde $\Phi(z)$, $z \in \mathbb{R}$, es la fda de la distribución $N(0,1)$.