

Capítulo 8

Distribución Normal Multivariante

Empezaremos definiendo la distribución normal bivalente o bivariada y luego pasaremos a un caso más general, el multivariante.

8.1. Normal bivalente

Definición 8.1 *Un par de variables aleatorias X_1 y X_2 tienen una **distribución normal bivalente** y se conocen como v.a. distribuidas normalmente en forma conjunta si y sólo si su densidad de probabilidad conjunta está dada por*

$$f_{X_1, X_2}(x_1, x_2) = \frac{1}{2\pi(1-\rho^2)^{1/2}\sigma_1\sigma_2} \cdot \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[\left(\frac{x_1 - \mu_1}{\sigma_1} \right)^2 - 2\rho \left(\frac{x_1 - \mu_1}{\sigma_1} \right) \left(\frac{x_2 - \mu_2}{\sigma_2} \right) + \left(\frac{x_2 - \mu_2}{\sigma_2} \right)^2 \right] \right\} I_{\mathbb{R}^2}(x_1, x_2)$$

y $\sigma_1, \sigma_2 \in \mathbb{R}^+$, $\rho \in \langle -1, 1 \rangle$.

Si dos v.a., X_1 y X_2 , tienen una distribución normal bivalente cuya f.d.p. está dada por la ecuación anterior, entonces

$$E(X_i) = \mu_i \quad \text{y} \quad \text{Var}(X_i) = \sigma_i^2 \quad \text{para } i = 1, 2.$$

Además

$$\rho_{X_1, X_2} = \rho.$$

Si X_1 y X_2 son v.a. independientes, entonces su coeficiente de correlación es cero. Por tanto, su f.d.p. conjunta está dada por:

$$f_{X_1, X_2}(x_1, x_2) = \left\{ \frac{1}{\sigma_1 \sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{x_1 - \mu_1}{\sigma_1} \right)^2 \right] \right\} \cdot \left\{ \frac{1}{\sigma_2 \sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{x_2 - \mu_2}{\sigma_2} \right)^2 \right] \right\} I_{\mathbb{R}^2}(x_1, x_2) .$$

Vemos que $f_{X_1, X_2}(x_1, x_2) = f_{X_1}(x_1) f_{X_2}(x_2)$.

Ahora vamos a ver que realmente:

$$f_{X_1, X_2}(x_1, x_2) = \frac{1}{2\pi(1-\rho^2)^{1/2}\sigma_1\sigma_2} \cdot \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[\left(\frac{x_1 - \mu_1}{\sigma_1} \right)^2 - 2\rho \left(\frac{x_1 - \mu_1}{\sigma_1} \right) \left(\frac{x_2 - \mu_2}{\sigma_2} \right) + \left(\frac{x_2 - \mu_2}{\sigma_2} \right)^2 \right] \right\} I_{\mathbb{R}^2}(x_1, x_2)$$

es f.d.p., demostrando que:

$$\int \int_{\mathbb{R}^2} f_{X_1, X_2}(x_1, x_2) dx_1 dx_2 = 1$$

Y, desde luego, se cumple que $f_{X_1, X_2}(x_1, x_2) \geq 0$, o mejor dicho $f_{X_1, X_2}(x_1, x_2) > 0$ para esta f.d.p.

Demostración. Hagamos el siguiente cambio de variable:

$$\begin{aligned} u &= \frac{x_1 - \mu_1}{\sigma_1} \\ v &= \frac{x_2 - \mu_2}{\sigma_2} \end{aligned}$$

así tenemos:

$$\int \int_{\mathbb{R}^2} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp \left[-\frac{1}{2(1-\rho^2)} (u^2 - 2\rho uv + v^2) \right] dv du$$

luego completando el cuadrado para u tenemos:

$$\int \int_{\mathbb{R}^2} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp \left[-\frac{1}{2(1-\rho^2)} [(u - \rho v)^2 + (1 - \rho^2)v^2] \right] dv du$$

y haciendo al cambio:

$$w = \frac{u - \rho v}{\sqrt{1-\rho^2}} \quad dw = \frac{du}{\sqrt{1-\rho^2}} ,$$

así tenemos:

$$\begin{aligned}
 \int \int_{\mathbb{R}^2} f_{X_1, X_2}(x_1, x_2) dx_1 dx_2 &= \int \int_{\mathbb{R}^2} \frac{1}{2\pi} \exp \left[-\frac{1}{2} (w^2 + v^2) \right] dv dw \\
 &= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} w^2 \right) dw \cdot \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} v^2 \right) dv \\
 &= 1 \cdot 1 \\
 &= 1
 \end{aligned}$$

Así queda demostrado que f_{X_1, X_2} es realmente f.d.p. \triangleleft

La f.g.m. de la normal bivalente es:

$$\psi_{X_1, X_2}(t_1, t_2) = \exp \left[t_1 \mu_1 + t_2 \mu_2 + \frac{1}{2} (t_1^2 \sigma_1^2 + 2\rho t_1 t_2 \sigma_1 \sigma_2 + t_2^2 \sigma_2^2) \right] .$$

De la f.g.m. tenemos:

$$\begin{aligned}
 E(X_1) &= \left. \frac{\partial \psi_{X_1, X_2}(t_1, t_2)}{\partial t_1} \right|_{t_1, t_2=0} = \mu_1 \\
 E(X_1^2) &= \left. \frac{\partial^2 \psi_{X_1, X_2}(t_1, t_2)}{\partial t_1^2} \right|_{t_1, t_2=0} = \mu_1^2 + \sigma_1^2 .
 \end{aligned}$$

Por tanto:

$$\text{Var}(X_1) = E(X_1^2) - [E(X_1)]^2 = \sigma_1^2 .$$

Del mismo modo obtenemos los momentos para X_2 .

Otro valor es:

$$\begin{aligned}
 \text{Cov}(X_1, X_2) &= E[(X_1 - \mu_1)(X_2 - \mu_2)] \\
 &= E(X_1 X_2 - X_1 \mu_2 - X_2 \mu_1 + \mu_1 \mu_2) \\
 &= E(X_1 X_2) - \mu_1 \mu_2 \\
 &= \rho \sigma_1 \sigma_2 .
 \end{aligned}$$

Teorema 8.1 Sea una v.a. bidimensional (X_1, X_2) con distribución normal bivariada. La f.d.p. marginal de X_1 es la normal con media μ_1 y varianza σ_1^2 . La f.d.p. marginal de X_2 es la normal con media μ_2 y varianza σ_2^2 .

Obtenemos la f.d.p. condicional de una de las variables mediante la relación:

$$f_{X_1|X_2}(x_1|x_2) = \frac{f_{X_1, X_2}(x_1, x_2)}{f_{X_2}(x_2)} .$$

Otro teorema de importancia es:

Teorema 8.2 Sea la v.a. bidimensional (X_1, X_2) con distribución normal bivariada. La f.d.p. condicional de X_1 , dado $X_2 = x_2$, es normal con media $\mu_1 + (\rho\sigma_1/\sigma_2)(x_2 - \mu_2)$ y varianza $\sigma_1^2(1 - \rho^2)$. La f.d.p. condicional de X_2 , dado $X_1 = x_1$, es normal con media $\mu_2 + (\rho\sigma_2/\sigma_1)(x_1 - \mu_1)$ y varianza $\sigma_2^2(1 - \rho^2)$.

Así:

$$f_{X_1|X_2}(x_1|x_2) = \frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2\sigma_1^2(1-\rho^2)} \left[x_1 - \mu_1 - \frac{\rho\sigma_1}{\sigma_2}(x_2 - \mu_2) \right]^2 \right\} I_{\mathbb{R}}(x_1)$$

$$f_{X_2|X_1}(x_2|x_1) = \frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2\sigma_2^2(1-\rho^2)} \left[x_2 - \mu_2 - \frac{\rho\sigma_2}{\sigma_1}(x_1 - \mu_1) \right]^2 \right\} I_{\mathbb{R}}(x_2).$$

En la distribución normal bivariada condicional, a la media le llamamos *regresión*. Así

$$E(X_1|x_2) = \mu_1 + \rho\sigma_1 \left(\frac{x_2 - \mu_2}{\sigma_2} \right)$$

es la regresión para X_1 , que es una función lineal de X_2 .

Además:

$$F_{X_1, X_2}(x_1, x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f_{X_1, X_2}(s, t) \, ds \, dt \quad .$$

Gráfica de la distribución normal bivalente

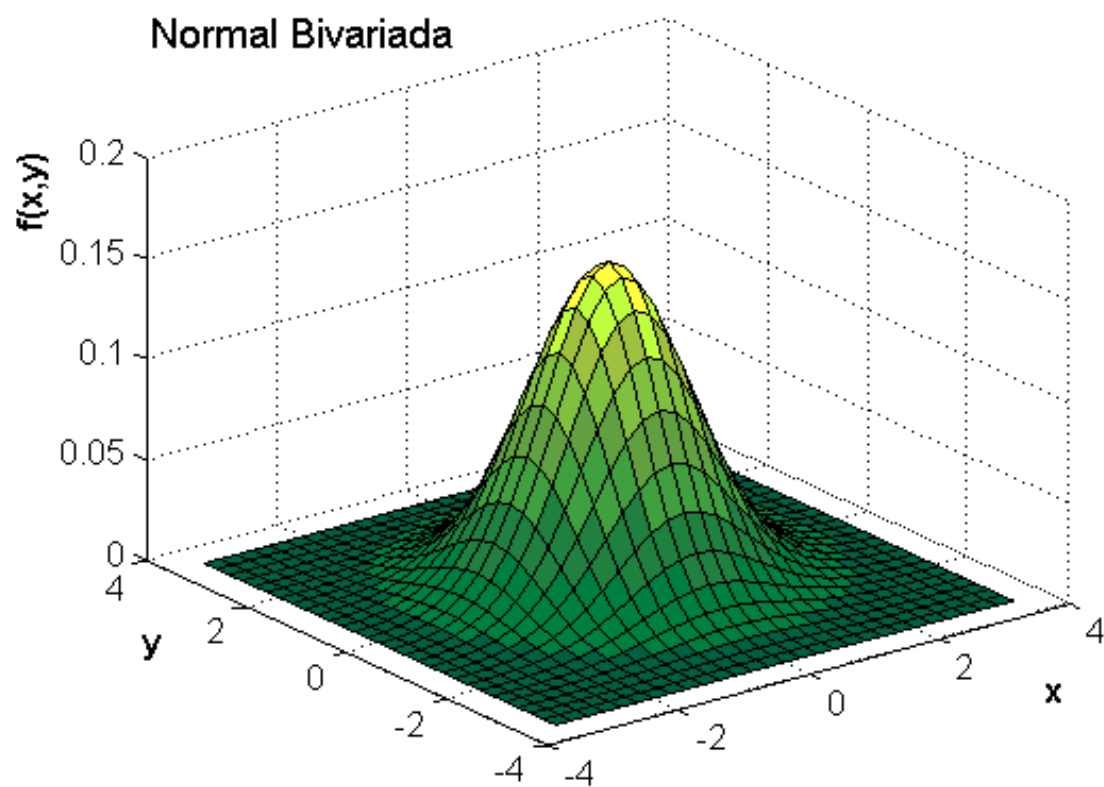


Figura 8.1: La f.d.p. para la normal bivariada.

8.2. Normal multivariante

Ya hemos definido la distribución normal bivalente; ahora definiremos y daremos las propiedades para la distribución normal multivariante.

Antes definiremos algunas expresiones que utilizamos en esta sección.

$$\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \quad \text{Un vector aleatorio.}$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{Valor del anterior vector aleatorio.}$$

$$\mathbf{X}^t \mathbf{A} \mathbf{X} = \sum_{i,j} a_{ij} x_i x_j \quad \text{Forma cuadrática definida positiva.}$$

$$\mathbf{X}^t = [X_1, X_2, \dots, X_n] \quad \text{Vector transpuesto de } \mathbf{X}.$$

$$\mathbf{A} \quad \text{Una matriz cuadrada simétrica y definida positiva.}$$

De lo anterior, $\mathbf{X}^t \mathbf{A} \mathbf{X} > 0$ para todo \mathbf{X} .

Definición 8.2 Se dice que el vector aleatorio \mathbf{X} sigue una **distribución normal n -variante** si su f.d.p. conjunta está dada por

$$f_{\mathbf{X}}(x_1, x_2, \dots, x_n) = \frac{\sqrt{|\mathbf{A}|}}{(2\pi)^{n/2}} \exp \left[-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^t \mathbf{A}(\mathbf{x} - \boldsymbol{\mu}) \right] \mathbf{I}_{\mathbb{R}^n}(\mathbf{x})$$

donde $\boldsymbol{\mu}^t = [\mu_1, \mu_2, \dots, \mu_n] \in \mathbb{R}^n$ es fijo y \mathbf{A} es una matriz simétrica y definida positiva.

Ahora demostraremos que la integral de $f_{\mathbf{X}}$ sobre \mathbb{R}^n es 1.

Demostración. Vemos que $f_{\mathbf{X}}(x_1, x_2, \dots, x_n) \geq 0 \forall \mathbf{x} \in \mathbb{R}^n$. Tomando $\mathbf{Y} = \mathbf{X} - \boldsymbol{\mu}$ tenemos

$$\frac{\partial x_i}{\partial y_j} = 0 \text{ si } i \neq j \text{ y } \frac{\partial x_i}{\partial y_i} = 1 \quad \forall i = 1, 2, \dots, n$$

así el Jacobiano de la transformación es la unidad, luego

$$\int_{\mathbb{R}^n} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} = \int_{\mathbb{R}^n} \frac{\sqrt{|\mathbf{A}|}}{(2\pi)^{n/2}} \exp \left[-\frac{1}{2}\mathbf{y}^t \mathbf{A} \mathbf{y} \right] d\mathbf{y}$$

Dado que A es definida positiva, existe una matriz ortogonal P tal que

$$P'AP = D \quad \text{y} \quad |D| = |P'| |A| |P| = |A| = \lambda_1 \cdot \lambda_2 \cdots \lambda_n$$

donde $\lambda_1, \lambda_2, \dots, \lambda_n$ son las raíces características de A que además son positivas. Así tenemos:

$$\begin{aligned} \int_{\mathbb{R}^n} f_{\mathbf{X}}(\mathbf{x}) \, d\mathbf{x} &= \frac{\sqrt{|P^t A P|}}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \exp \left[-\frac{1}{2} (P\mathbf{y})^t A (P\mathbf{y}) \right] d\mathbf{y} \\ &= \frac{\sqrt{|A|}}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \exp \left[-\frac{1}{2} \mathbf{y}^t D \mathbf{y} \right] d\mathbf{y} \\ &= \frac{\sqrt{|A|}}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \exp \left[-\frac{1}{2} [y_1 \ y_2 \ \cdots \ y_n] \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \right] d\mathbf{y} \\ &= \frac{\prod_{i=1}^n \sqrt{\lambda_i}}{(2\pi)^{n/2}} \prod_{i=1}^n \int_{\mathbb{R}} \exp \left[-\frac{1}{2} \lambda_i y_i^2 \right] dy_i \\ &= \frac{\prod_{i=1}^n \sqrt{\lambda_i}}{(\sqrt{2\pi})^n} \prod_{i=1}^n \frac{\sqrt{2\pi}}{\sqrt{\lambda_i}} \\ &= 1 \end{aligned}$$

donde al final hicimos el cambio de variable $t_i = y_i \sqrt{\lambda_i}$, $dy_i = \frac{dt_i}{\sqrt{\lambda_i}}$ y así para $i = \overline{1, n}$

$$\int_{\mathbb{R}} \exp \left[-\frac{1}{2} \lambda_i y_i^2 \right] dy_i = \int_{\mathbb{R}} \frac{1}{\sqrt{\lambda_i}} \exp \left[-\frac{1}{2} t_i^2 \right] dt_i = \frac{\sqrt{2\pi}}{\sqrt{\lambda_i}}$$

A $(\mathbf{x} - \boldsymbol{\mu})^t A (\mathbf{x} - \boldsymbol{\mu})$ es llamada *forma cuadrática de la normal multivariante*.

◁

Teorema 8.3 *Sea el vector aleatorio \mathbf{X} que sigue una distribución normal n -variante, entonces:*

$$E(\mathbf{X}) = \boldsymbol{\mu} \quad .$$

Teorema 8.4 *En la distribución normal n -variante, A es la inversa de la matriz de covarianzas $\text{Cov}(\mathbf{X})$; es decir, $\text{Cov}(\mathbf{X})^{-1} = A$, y $A^{-1} = \text{Cov}(\mathbf{X})$.*