

MAT-042: Distribuciones bivariadas

Felipe Osorio

fosorios.mat.utfsm.cl

Departamento de Matemática, UTFSM



Distribuciones bivariadas

Para un par de variables aleatorias X e Y se tiene la función de densidad conjunta como

$$f(x, y) = P(X = x, Y = y)$$

Notación:

Escribiremos

$$P(X = x, Y = y) = P(\{X = x\} \cap \{Y = y\}).$$

Ejemplo:

Distribución bivariada con X e Y tomando valores 0 o 1

$X \setminus Y$	0	1	total
0	1/9	2/9	1/3
1	3/9	4/9	2/3
total	1/3	2/3	1

En este caso,

$$f(1, 1) = P(X = 1, Y = 1) = \frac{4}{9}.$$



Definición 1 (Densidad conjunta):

Decimos que $f(x, y)$ es la función de densidad del vector aleatorio (X, Y) si:

(a) $f(x, y) \geq 0$ para todo (x, y) .

(b) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1$,

(c) Para cualquier conjunto $A \subset \mathbb{R}^2$,

$$P((X, Y) \in A) = \int \int_A f(x, y) dx dy.$$

Observación:

Para el caso de distribuciones discretas, la definición es análoga.



Ejemplo:

Sea (X, Y) con densidad,

$$f(x, y) = \begin{cases} x + y, & 0 \leq x \leq 1, 0 \leq y \leq 1, \\ 0 & \text{en otro caso.} \end{cases}$$

Entonces,

$$\begin{aligned} \int_0^1 \int_0^1 (x + y) \, dx \, dy &= \int_0^1 \left[\int_0^1 x \, dx \right] dy + \int_0^1 \left[\int_0^1 y \, dx \right] dy \\ &= \int_0^1 \frac{1}{2} \, dy + \int_0^1 y \, dy = \frac{1}{2} + \frac{1}{2} \\ &= 1 \end{aligned}$$

es decir, $f(x, y)$ es una densidad.



Definición 2 (Distribución acumulada conjunta):

La función de distribución acumulada conjunta (CDF) $F(x, y)$ es definida como:

$$F(x, y) = P(X \leq x, Y \leq y).$$

Propiedades:

La CDF tiene las siguientes propiedades:

- (a) $F(x, y)$ es función monótona creciente y continua a la derecha en (x, y) .
- (b) $0 \leq F(x, y) \leq 1$.
- (c) $F(-\infty, y) = F(x, -\infty) = 0$.
- (d) $F(+\infty, +\infty) = 1$.



Para el caso continuo, tenemos:

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(r, s) dr ds$$

Además, por el Teorema fundamental del Cálculo,

$$\frac{\partial^2 F(x, y)}{\partial x \partial y} = f(x, y).$$



Resultado 1 (Densidad marginal):

Sea (X, Y) un vector aleatorio bivariado con densidad conjunta $f_{X,Y}(x, y)$. Entonces, las densidades marginales de X e Y , $f_X(x) = P(X = x)$ y $f_Y(y) = P(Y = y)$ son, respectivamente dadas por:

(a) Para el caso discreto:

$$f_X(x) = \sum_y f(x, y) = \sum_y P(X = x, Y = y),$$

$$f_Y(y) = \sum_x f(x, y) = \sum_x P(X = x, Y = y).$$

(b) Para variables continuas:

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



Ejemplo:

Suponga que

$$f_{X,Y}(x,y) = e^{-(x+y)}, \quad x, y \geq 0.$$

De este modo,

$$f_X(x) = e^{-x} \int_0^{\infty} e^{-y} dy = e^{-x}, \quad x \geq 0.$$



Observación:

Para variables aleatorias continuas podemos usar la siguiente relación:

$$\begin{aligned}F_X(x) &= \mathbf{P}(X \leq x) = \int_{-\infty}^x f_X(t) dt \\&= \int_{-\infty}^x \int_{-\infty}^{+\infty} f_{X,Y}(t, y) dy dt \\&= F_{X,Y}(x, +\infty).\end{aligned}$$

Además,

$$f_X(x) = \frac{dF_X(x)}{dx},$$

y análogamente

$$F_Y(y) = F_{X,Y}(+\infty, y).$$



Definición 3 (Densidad condicional):

La función de densidad condicional para el caso discreto es definida como:

$$\begin{aligned} f_{X|Y}(x|y) &= P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} \\ &= \frac{f_{X,Y}(x, y)}{f_Y(y)}, \end{aligned}$$

si $f_Y(y) > 0$. Para el caso continuo, tenemos

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}, \quad f_Y(y) > 0.$$

Observación:

A pesar de la notación $f_{X|Y}(x|y)$ se debe destacar que $f_{X|Y}(x|y)$ es función de X .



Observación:

Evidentemente,

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{f_{Y|X}(y|x)f_X(x)}{f_Y(y)},$$

y

$$\begin{aligned} f_{X|Y}(x|y) &= \frac{f_{Y|X}(y|x)f_X(x)}{\int_{-\infty}^{\infty} f_{X,Y}(x,y) dx} \\ &= \frac{f_{Y|X}(y|x)f_X(x)}{\int_{-\infty}^{\infty} f_{Y|X}(y|x)f_X(x) dx}, \end{aligned}$$

que corresponde al Teorema de Bayes para densidades.



En efecto, la densidad condicional debe satisfacer las condiciones de una densidad.

$$\begin{aligned}\int_{-\infty}^{\infty} f_{X|Y}(x|y) \, dx &= \int_{-\infty}^{\infty} \frac{f_{X,Y}(x,y)}{f_Y(y)} \, dx \\ &= \frac{1}{f_Y(y)} \int_{-\infty}^{\infty} f_{X,Y}(x,y) \, dx \\ &= \frac{f_Y(y)}{f_Y(y)} = 1.\end{aligned}$$

Además, es obvio que

$$f_{X|Y}(x|y) \geq 0.$$



Los momentos de una variable aleatoria condicional se definen de forma análoga:

$$E(X|Y) = \int_{-\infty}^{\infty} x f_{X|Y}(x|y) dx,$$

$$E(Y|X) = \int_{-\infty}^{\infty} y f_{Y|X}(y|x) dy,$$

y

$$\text{var}(Y|X) = E(Y^2|X) - \{E(Y|X)\}^2.$$



Definición 4 (Independencia):

Sea (X, Y) un vector aleatorio con densidad conjunta $f(x, y)$ y densidades marginales $f_X(x)$, $f_Y(y)$, respectivamente. Entonces X e Y se dicen independientes si, para todo $x, y \in \mathbb{R}$,

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

Observación:

Análogamente, si $F(x, y)$ representa la CDF del vector (X, Y) y $G(x)$, $H(y)$ las CDFs de X e Y , respectivamente. Entonces X e Y son independientes, si

$$F(x, y) = G(x) \cdot H(y).$$

Además, si X e Y son independientes, tenemos que:

$$f(y|x) = \frac{f(x, y)}{f_X(x)} = \frac{f_X(x) \cdot f_Y(y)}{f_X(x)} = f_Y(y).$$



Observación:

Diremos que $\mathbf{X} = (X_1, \dots, X_k)^\top$ es un vector aleatorio k -dimensional.¹ Usaremos la notación $f(\mathbf{x})$ y $F(\mathbf{x})$ para la densidad y CDF del vector aleatorio \mathbf{X} . En este caso

$$F(\mathbf{x}) = P(\mathbf{X} \leq \mathbf{x}) = P(X_1 \leq x_1, \dots, X_k \leq x_k),$$

y asumiremos que existe f no negativa, tal que

$$F(\mathbf{x}) = \int_{-\infty}^{\mathbf{x}} f(\mathbf{u}) d\mathbf{u}, \quad \mathbf{x} \in \mathbb{R}^k.$$

Además,

$$f(\mathbf{x}) = \frac{\partial^k F(\mathbf{x})}{\partial x_1 \cdots \partial x_k},$$

con $\int_{\mathbb{R}^k} f(\mathbf{x}) d\mathbf{x} = 1$.

¹donde X_1, \dots, X_k son variables aleatorias.

Definición 5 (Variables IID):

Si X_1, \dots, X_n son variables independientes y cada una tiene la misma distribución marginal con CDF F , decimos que X_1, \dots, X_n son independientes e idénticamente distribuidas (IID) y escribimos

$$X_1, \dots, X_n \sim F.$$

Si F tiene densidad f podemos escribir $X_1, \dots, X_n \sim f$.

Observación:

Cuando $X_1, \dots, X_n \stackrel{\text{IID}}{\sim} F$ también decimos que X_1, \dots, X_n es una muestra aleatoria de tamaño n desde F (o f).²

Si X_1, \dots, X_n son independientes, es fácil notar que

$$f(\mathbf{x}) = \prod_{i=1}^n f_{X_i}(x_i).$$

²Y anotamos que X_1, \dots, X_n es una m.a.(n) desde F .

A partir de ahora, consideraremos que $f(\cdot)$ es un miembro de una familia paramétrica:

$$\mathcal{P} = \{P_{\theta} : \theta \in \Theta\},$$

donde $\Theta \subset \mathbb{R}^p$ se denomina espacio paramétrico.

En cuyo caso escribimos $X \sim f(x; \theta)$. De este modo,

$$f(x_1, \dots, x_n; \theta) = \prod_{i=1}^n f(x_i; \theta).$$

Para abreviar anotaremos

$$f(\mathbf{x}; \theta) = \prod_{i=1}^n f(x_i; \theta).$$

Evidentemente,

$$\log f(\mathbf{x}; \theta) = \sum_{i=1}^n \log f(x_i; \theta).$$



Ejemplo:

Sea X_1, \dots, X_n m.a.(n) desde $\text{Exp}(\theta)$. De este modo, su distribución conjunta es dada por:

$$\begin{aligned} f(x_1, \dots, x_n; \theta) &= \prod_{i=1}^n f(x_i; \theta) = \prod_{i=1}^n \theta e^{-\theta x_i} \\ &= \theta^n \exp\left(-\theta \sum_{i=1}^n x_i\right) \\ &= \theta^n e^{-\theta n \bar{x}} \end{aligned}$$

Es decir, en este caso la densidad conjunta de X_1, \dots, X_n es función de θ y \bar{X} .

Además,

$$\log f(\mathbf{x}; \theta) = n \log \theta - n\theta \bar{x}.$$



Recordatorio 1:

Sabemos que la función de densidad de $X \sim N(\mu, \sigma^2)$, es dada por:

$$f(x; \mu, \sigma^2) = (2\pi\sigma^2)^{-1/2} \exp \left\{ -\frac{1}{2} \left(\frac{y - \mu}{\sigma} \right)^2 \right\}$$

donde $\mu \in \mathbb{R}$ y $\sigma^2 > 0$. Además,

$$E(X) = \mu, \quad \text{var}(X) = \sigma^2.$$

Evidentemente, si $Z \sim N(0, 1)$ entonces

$$\phi(z) = (2\pi)^{-1/2} \exp \left(-\frac{1}{2} z^2 \right).$$

Propiedades:

- (a) Si $X \sim N(\mu, \sigma^2) \Leftrightarrow Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$
- (b) Si $X \sim N(\mu, \sigma^2)$, entonces $Y = aX + b \sim N(a\mu + b, a^2\sigma^2)$.



Distribución normal multivariada

Sea $\mathbf{X} = (X_1, \dots, X_p)^\top$, la distribución normal multivariada es caracterizada por un vector de medias $\boldsymbol{\mu} = (\mu_1, \dots, \mu_p)^\top$ y una matriz de covarianzas $\boldsymbol{\Sigma} = (\sigma_{ij}) \geq \mathbf{0}$. En cuyo caso escribimos $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

Considere Z_1, \dots, Z_p variables IID desde $N(0, 1)$ y suponga $\mathbf{Z} = (Z_1, \dots, Z_p)^\top$. Entonces,

$$\begin{aligned} f_{\mathbf{Z}}(\mathbf{z}) &= \prod_{i=1}^p f_{Z_i}(z_i) = \prod_{i=1}^p (2\pi)^{-1/2} \exp\left(-\frac{1}{2}z_i^2\right) \\ &= (2\pi)^{-p/2} \exp\left(-\frac{1}{2} \sum_{i=1}^p z_i^2\right) \\ &= (2\pi)^{-p/2} \exp\left(-\frac{1}{2} \|\mathbf{z}\|^2\right), \end{aligned}$$

y anotamos $\mathbf{Z} \sim N_p(\mathbf{0}, \mathbf{I})$.



Más generalmente, $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ tiene densidad

$$f_{\mathbf{X}}(\mathbf{x}) = |2\pi\boldsymbol{\Sigma}|^{-1/2} \exp \left\{ -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}) \right\}.$$

Dado que $\boldsymbol{\Sigma}$ es simétrica y definida positiva tenemos que $\boldsymbol{\Sigma} = \mathbf{B}\mathbf{B}^\top$ esto lleva al resultado siguiente.

Resultado 2:

Si $\mathbf{Z} \sim N_p(\mathbf{0}, \mathbf{I})$, entonces

$$\mathbf{X} = \boldsymbol{\mu} + \mathbf{B}\mathbf{Z} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}).$$

Asimismo si $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, entonces

$$\mathbf{B}^{-1}(\mathbf{X} - \boldsymbol{\mu}) \sim N_p(\mathbf{0}, \mathbf{I}).$$



Ejemplo:

Sea $\mathbf{Z} \sim N_2(\mathbf{0}, \Sigma)$, donde

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \quad -1 < \rho < 1.$$

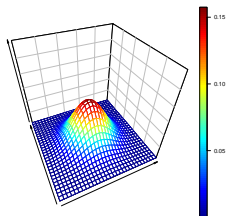
En cuyo caso, la función de densidad adopta la forma:

$$f(\mathbf{x}) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2(1-\rho^2)} (x_1^2 + x_2^2 - 2\rho x_1 x_2) \right\}.$$

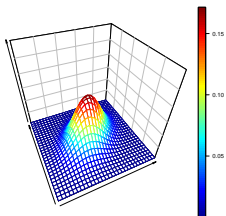
En la siguiente slide se presenta la densidad para $\rho = 0.0, 0.4$ y 0.8 .



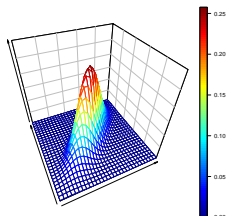
Distribución normal multivariada



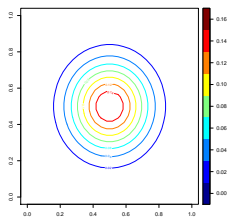
(a) $\rho = 0.0$



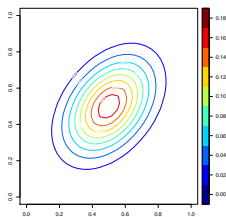
(b) $\rho = 0.4$



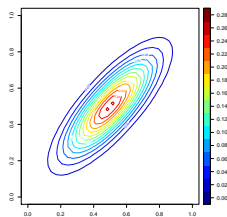
(c) $\rho = 0.8$



(d) contornos



(e) contornos



(f) contornos

Resultado 3:

Suponga $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, entonces

$$\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b} \sim N_k(\mathbf{A}\boldsymbol{\mu} + \mathbf{b}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^\top),$$

donde $\mathbf{A} \in \mathbb{R}^{k \times p}$ con $\text{rg}(\mathbf{A}) = k$ y $\mathbf{b} \in \mathbb{R}^k$.

Resultado 4:

Suponga $\mathbf{Z} = (Z_1, \dots, Z_p)^\top \sim N_p(\mathbf{0}, \mathbf{I})$,³ entonces

$$U = \mathbf{Z}^\top \mathbf{Z} = \sum_{i=1}^p Z_i^2 \sim \chi^2(p),$$

donde $\chi^2(p)$ denota una distribución chi-cuadrado con p grados de libertad.⁴

³Es decir, Z_1, \dots, Z_p son variables aleatorias IID

⁴ $U \sim \chi^2(p) \stackrel{d}{=} \text{Gama}(p/2, 1/2)$.

Resultado 5:

Sea $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ y considere la partición:

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix}, \quad \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}.$$

Entonces,

- (a) La distribución marginal de $\mathbf{X}_i \sim N_{p_i}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_{ii})$, $i = 1, 2$.
- (b) La distribución condicional de $\mathbf{X}_2 | \mathbf{X}_1 = \mathbf{x}_1$ es

$$\mathbf{X}_2 | \mathbf{X}_1 = \mathbf{x}_1 \sim N_{p_2}(\boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{22} \boldsymbol{\Sigma}_{11}^{-1}(\mathbf{x}_1 - \boldsymbol{\mu}_1), \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21} \boldsymbol{\Sigma}_{11}^{-1} \boldsymbol{\Sigma}_{21}).$$

- (c) $U = (\mathbf{X} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{X} - \boldsymbol{\mu}) \sim \chi^2(p)$.

