



Tema 4 - Variables aleatorias continuas multidimensionales

Ricardo Alberich, Juan Gabriel Gomila y Arnau Mir

Variables aleatorias bidimensionales continuas

Variables aleatorias bidimensionales continuas Introducción

DEFINICIÓN DE VARIABLE ALEATORIA BIDIMENSIONAL CONTINUA.

Recordemos que una v.a. bidimensional **continua** cuando su conjunto de valores en \mathbb{R}^2 , $(X, Y)(\Omega)$ es un producto de intervalos.

DEFINICIÓN FUNCIÓN DE DISTRIBUCIÓN CONJUNTA

La función de distribución acumulada conjunta o simplemente distribución conjunta se define como

$$F_{XY}(x, y) = P(X \leq x, Y \leq y).$$

Función de distribución acumulada, función de densidad

DEFINICIÓN FUNCIÓN DE DENSIDAD CONJUNTA

Sea $f_{XY} : \mathbb{R} \times \mathbb{R} \mapsto [0, +\infty)$ diremos que es una densidad bidimensional del vector aleatorio bidimensional (X, Y) si

$$F_{XY}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{XY}(t_x, t_y) dt_x dt_y.$$

Llamaremos dominio de la variable conjunta a

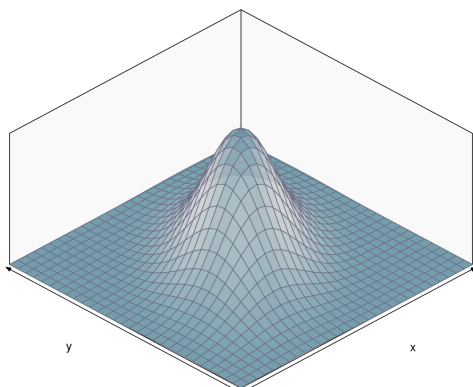
$$D_{XY} = \{(x, y) \in \mathbb{R}^2 \mid f_{XY}(x, y) > 0\}.$$

Es decir es el conjunto de valores posibles que toma la v.a. (X, Y) .

3/39

4/39

Gráfica de una función de densidad



Gráfica

library(tidyverse)

5/39

Propiedades de la función de densidad conjunta

Sea (X, Y) una **variable aleatoria bidimensional continua** con dominio $D_{XY} \subset \mathbb{R}^2$.

Su **función de densidad conjunta** verifica las siguientes propiedades:

·

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_{XY}(x, y) \, dx dy = 1.$$

·

Sea B un subconjunto cualquiera del dominio D_{XY} . El valor de la probabilidad $P((X, Y) \in B)$ se puede calcular de la forma siguiente:

$$P((X, Y) \in B) = \int \int_B f_{XY}(x, y) \, dx dy.$$

Es decir, la probabilidad de que la variable bidimensional tome valores en B es igual al volumen que genera la densidad conjunta sobre el recinto B .

6/39

Variables aleatorias marginales y su distribución

Consideremos una variable aleatoria **bidimensional continua** (X, Y) con **función de densidad conjunta** $f_{XY}(x, y)$ y con dominio D_{XY} .

La de la **función de densidad conjunta** contiene suficiente información para obtener las **funciones de densidad** de las variables X e Y .

Dichas variables X e Y se denominan **variables marginales** y sus correspondientes **funciones de densidad**, **funciones de densidad marginales** f_X de la variable X con dominio D_X y f_Y de la variable Y con dominio D_Y .

Veamos cómo obtener f_X y f_Y a partir de la densidad conjunta f_{XY} .

Distribuciones marginales

8/39

Funciones de probabilidad marginales

PROPOSICIÓN. CÁLCULO DE LAS FUNCIONES DE DENSIDAD MARGINALES.

Sea (X, Y) una variable aleatoria **bidimensional continua** con **función de densidad conjunta** $f_{XY}(x, y)$, con $(x, y) \in D_{XY}$.

Las **funciones de densidad marginales** $f_X(x)$ y $f_Y(y)$ se calculan usando las expresiones siguientes:

- $f_X(x) = \int_{-\infty}^{+\infty} f_{XY}(x, y) \, dy.$
- $f_Y(y) = \int_{-\infty}^{+\infty} f_{XY}(x, y) \, dx.$

9/39

Independencia de variables aleatorias continuas

Recordemos que dos sucesos A y B son independientes si

$$P(A \cap B) = P(A) \cdot P(B).$$

¿Cómo trasladar dicho concepto al caso de variables aleatorias continuas?

Dada una variable aleatoria bidimensional continua (X, Y) con dominio D_{XY}

Así que al menos todos los sucesos de la forma $P(X \leq x, Y \leq y)$ deberán ser independientes.

Esto implicará que cualesquiera dos sucesos de cada variables con independientes.

10/39

Independencia de variables aleatorias continuas

CONDICIONES PARA INDEPENDENCIA DE VARIABLES ALEATORIAS BIDimensionALES CONTINUAS

Dada (X, Y) una **variable aleatoria bidimensional continua** con **función de densidad** f_{XY} y **funciones de probabilidad marginales** f_X y f_Y .

Diremos que X e Y son independientes si se cumple al menos una de las siguientes condiciones:

- $f_{XY}(x, y) = f_X(x) \cdot f_Y(y)$ para todo $(x, y) \in D_{XY}$
- $F_{XY}(x, y) = F_X(x) \cdot F_Y(y)$ para todo $(x, y) \in D_{XY}$

11/39

Esperanza y varianza de las distribuciones marginales

- $E(X) = \int_{-\infty}^{+\infty} x \cdot f_X(x) \, dx.$
- $E(Y) = \int_{-\infty}^{+\infty} y \cdot f_Y(y) \, dy.$
- $\sigma_X^2 = Var(X) = E((X - E(X))^2) = E(X^2) - E(X)^2.$
- $\sigma_Y^2 = Var(Y) = E((Y - E(Y))^2) = E(Y^2) - E(Y)^2.$

12/39

Distibuciones condicionales

- Dado un valor fijo $y \in D_Y$ definimos la distribución condicional de la v.a. X condicionada a que $Y = y$ como

$$f_{X|Y=y}(x) = \frac{f_{XY}(x, y)}{f_Y(y)}, \text{ para todo } x \in D_X.$$

- Dado un valor fijo $y \in D_Y$ definimos la distribución condicional de la v.a. Y condicionada a que $X = x$ como

$$f_{Y|X=x}(y) = \frac{f_{XY}(x, y)}{f_X(x)}, \text{ para todo } Y \in D_Y.$$

13/39

Esperanzas condicionales

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

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

PROPIEDAD

Si las variables X e Y son independientes se cumple que

1. $E(X|Y = y) = E(X)$
2. $E(Y|X = x) = E(Y)$

15/39

Esperanzas de funciones de v.a. continuas bidimensionales

DEFINICIÓN:

Sea (X, Y) una variable aleatoria bidimensional continua y $g(X, Y) : \mathbb{R}^2 \mapsto \mathbb{R}$ una función de esa variable bidimensional entonces

$$E(g(X, Y)) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) \cdot f_{XY}(x, y) \, dxdy.$$

17/39

Distibuciones condicionales e independencia

PROPIEDAD

Si las variables X e Y son independientes se cumple que

- $f_{X|Y=y}(x) = f_X(x)$
- $f_{Y|X=x}(y) = f_Y(y)$

14/39

Esperanzas de funciones de v.a. continuas bidimensionales. Covarianza y correlación

Esperanzas de funciones de v.a. continuas bidimensionales

PROPIEDAD: En particular:

$$E(X + Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x + y) \cdot f_{XY}(x, y) \, dxdy = \mu_X + \mu_Y.$$

$$\begin{aligned} Var(X + Y) &= E\left((X + Y - E(X + Y))^2\right) \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x + y - (\mu_X + \mu_Y))^2 \cdot f_{XY}(x, y) \, dxdy. \end{aligned}$$

18/39

Esperanzas de funciones de v.a. continuas bidimensionales

PROPIEDAD: Sea (X, Y) una variable aleatoria bidimensional entonces se cumple que:

- $E(X + Y) = E(X) + E(Y) = \mu_X + \mu_Y$.
- Si X e Y son independientes entonces $E(X \cdot Y) = E(X) \cdot E(Y) = \mu_X \cdot \mu_Y$.
- Si X e Y son independientes entonces $Var(X + Y) = Var(X) + Var(Y) = \sigma_X^2 + \sigma_Y^2$.

19/39

Medida de la variación conjunta: covarianza

Se denomina **covarianza** entre las variables X e Y :

$$\sigma_{XY} = Cov(X, Y) = E((X - \mu_X)(Y - \mu_Y)).$$

La covarianza puede calcularse también con:

$$Cov(X, Y) = E(X \cdot Y) - E(X) \cdot E(Y) = E(X \cdot Y) - \mu_X \cdot \mu_Y,$$

PROPIEDAD. Si las variables X e Y son **independientes**, entonces $Cov(X, Y) = 0$.

Es una consecuencia de que si X e Y son independientes entonces que vimos que $E(X \cdot Y) = E(X) \cdot E(Y) = \mu_X \cdot \mu_Y$.

21/39

Propiedades de la covarianza

- Sea (X, Y) una variable aleatoria bidimensional. Entonces la **varianza de la suma/resta** se calcula usando la expresión siguiente:

$$Var(X \pm Y) = Var(X) + Var(Y) \pm 2 \cdot Cov(X, Y).$$

- Sea (X, Y) una variable aleatoria bidimensional donde las variables X e Y son **independientes**. Entonces:

$$Var(X + Y) = Var(X) + Var(Y).$$

23/39

Covarianza y correlación

Covarianza entre las variables

La **covarianza** es una medida de lo relacionadas están las variables X e Y :

- Si cuando $X \geq \mu_X$, también ocurre que $Y \geq \mu_Y$ o viceversa, cuando $X \leq \mu_X$, también ocurre que $Y \leq \mu_Y$, el valor $(X - \mu_X)(Y - \mu_Y)$ será positivo y la **covarianza** será positiva.
- Si por el contrario, cuando $X \geq \mu_X$, también ocurre que $Y \leq \mu_Y$ o viceversa, cuando $X \leq \mu_X$, también ocurre que $Y \geq \mu_Y$, el valor $(X - \mu_X)(Y - \mu_Y)$ será negativo y la **covarianza** será negativa.
- En cambio, si a veces ocurre una cosa y a veces ocurre otra, la **covarianza** va cambiando de signo y puede tener un valor cercano a 0.

22/39

Coefficiente de correlación entre las variables

DEFINICIÓN DEL COEFICIENTE DE CORRELACIÓN. Sea (X, Y) una variable aleatoria bidimensional. Se define el **coeficiente de correlación** entre las variables X e Y como:

$$\rho_{XY} = \frac{Cov(X, Y)}{\sqrt{Var(X)} \cdot \sqrt{Var(Y)}} = \frac{E(X \cdot Y) - \mu_X \cdot \mu_Y}{\sqrt{E(X^2) - \mu_X^2} \cdot \sqrt{E(Y^2) - \mu_Y^2}}.$$

24/39

Coeficiente de correlación entre las variables

OBSERVACIÓN: Si las variables X e Y son independientes, su **coeficiente de correlación** $\rho_{XY} = 0$ es nulo ya que su **covarianza** lo es.

Notemos también que la **correlación** no tiene unidades y es invariante a cambios de escala.

Además, la **covarianza** de las **variables tipificadas** $\frac{X-\mu_X}{\sigma_X}$ y $\frac{Y-\mu_Y}{\sigma_Y}$ coincide con la **correlación** de X e Y .

El **coeficiente de correlación** es un valor normalizado ya que siempre está entre -1 y 1: $-1 \leq \rho_{XY} \leq 1$.

25/39

Matriz de varianzas-covarianzas y matriz de correlaciones

Sea (X, Y) una variable bidimensional. Notemos que

- $Cov(X, X) = \sigma_{XX} = \sigma_X^2$.
- $Cov(Y, Y) = \sigma_{YY} = \sigma_Y^2$.
- $\sigma_{XY} = Cov(X, Y) = Cov(Y, X) = \sigma_{YX}$.

Se denomina matriz de varianzas-covarianzas y se suele denotar como Σ a

$$\Sigma = \begin{pmatrix} Cov(X, X) & Cov(X, Y) \\ Cov(Y, X) & Cov(Y, Y) \end{pmatrix} = \begin{pmatrix} \sigma_{XX} & \sigma_{XY} \\ \sigma_{YX} & \sigma_{YY} \end{pmatrix} = \begin{pmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{YX} & \sigma_Y^2 \end{pmatrix}$$

27/39

Matriz de varianzas-covarianzas y matriz de correlaciones

Se denomina matriz de correlaciones a

$$R = \begin{pmatrix} Cor(X, X) & Cor(X, Y) \\ Cor(Y, X) & Cor(Y, Y) \end{pmatrix} = \begin{pmatrix} 1 & \rho_{XY} \\ \rho_{YX} & 1 \end{pmatrix} = \begin{pmatrix} 1 & \rho_{XY} \\ \rho_{XY} & 1 \end{pmatrix}.$$

29/39

Coeficiente de correlación entre las variables

OBSERVACIÓN: Si las variables X e Y tienen dependencia lineal, por ejemplo si $Y = a \cdot X + b$ para algunas constantes $a, b \in \mathbb{R}$, entonces su **coeficiente de correlación** $\rho_{XY} = \pm 1$, es decir toma el valor 1 si la pendiente $a > 0$ y -1 si $a < 0$.

De forma similar:

- si $Cor(X, Y) = +1$ X e Y tienen relación lineal con pendiente positiva.
- si $Cor(X, Y) = -1$ X e Y tienen relación lineal con pendiente negativa.

26/39

Matriz de varianzas-covarianzas y matriz de correlaciones

Sea (X, Y) una variable bidimensional. Notemos que

- $Cor(X, X) = \rho_{XX} = 1$.
- $Cor(Y, Y) = \rho_{YY} = 1$.
- $\rho_{XY} = Cor(X, Y) = Cor(Y, X) = \rho_{YX}$.

28/39

La distribución normal bivalente

Definición de distribución normal bivalente

Sea (X, Y) una variable continua bidimensional con $E(X) = \mu_X$, $E(Y) = \mu_Y$

$$\sigma_X^2 = \text{Var}(X), \sigma_Y^2 = \text{Var}(Y), \sigma_{XY} = \text{Cov}(X, Y).$$

Y si denotamos por

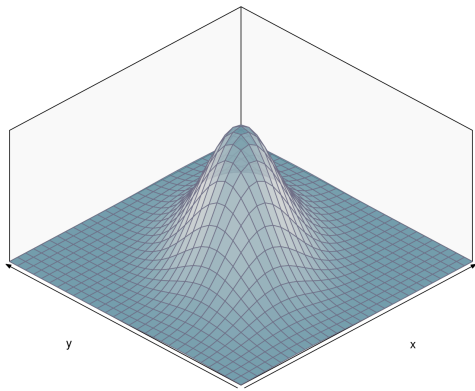
$$\mu = \begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}$$

y por

$$\Sigma = \begin{pmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{pmatrix}.$$

31/39

Gráfica de la distribución gaussiana (X, Y) .



33/39

Conceptos básicos. Función de probabilidad y de distribución.

Consideremos un vector compuesto de n variables aleatorias continuas (X_1, X_2, \dots, X_n)

Su **función de densidad de probabilidad** es una función $f_{X_1, X_2, \dots, X_n} : \mathbb{R}^n \mapsto [0, +\infty)$ tal que

$$\begin{aligned} F_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) &= P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n) \\ &= \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \dots \int_{-\infty}^{x_n} f(t_1, t_2, \dots, t_n) dt_1 dt_2 \dots dt_n. \end{aligned}$$

35/39

Definición de distribución normal bivalente

Diremos que el vector $\begin{pmatrix} X \\ Y \end{pmatrix}$ sigue una ley **normal o gaussiana bidimensional**

$$N\left(\mu = \begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}, \Sigma = \begin{pmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{pmatrix}\right)$$

si su densidad es

$$f_{XY}(x, y) = \frac{1}{\sqrt{(2\pi)^2 \cdot \det(\Sigma)}} \cdot e^{-\frac{1}{2}((x, y) - \mu)^t \cdot \Sigma^{-1} \cdot ((x, y) - \mu)}.$$

32/39

Distribuciones multidimensionales

Independencia

DEFINICIÓN INDEPENDENCIA

Diremos que las variables continuas X_1, X_2, \dots, X_n son **INDEPENDIENTES** cuando

$$f_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) = f_{X_1}(x_1) \cdot f_{X_2}(x_2) \cdot \dots \cdot f_{X_n}(x_n).$$

PROPIEDAD

Las variables X_1, X_2, \dots, X_n son **INDEPENDIENTES** si y solo si

$$F_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) = F_{X_1}(x_1) \cdot F_{X_2}(x_2) \cdot \dots \cdot F_{X_n}(x_n).$$

36/39

Conceptos básicos

VECTOR DE MEDIAS

Si denotamos $E(X_i) = \mu_i$ para $i = 1, 2, \dots, n$ el **vector de medias** es

$$E(X_1, X_2, \dots, X_n) = (E(X_1), E(X_2), \dots, E(X_n)) = (\mu_1, \mu_2, \dots, \mu_n).$$

COVARIANZA Y VARIANZAS

Si denotamos $\sigma_{ij} = Cov(X_i, X_j)$ para todo i, j en $1, 2, \dots, n$ entonces tenemos que

- $\sigma_{ii} = Cov(X_i, X_i) = \sigma_{ii} = \sigma_i^2.$
- $\sigma_{ij} = Cov(X_i, X_j) = Cov(X_j, X_i) = \sigma_{ji}.$

Conceptos básicos

Si denotamos $\rho_{ij} = Cor(X_i, X_j)$ para todo i, j en $1, 2, \dots, n$ entonces tenemos que

- $\rho_{ii} = Cor(X_i, X_i) = 1.$
- $\rho_{ij} = Cor(X_i, X_j) = Cor(X_j, X_i) = \rho_{ji}.$

Matrices de varianzas-covarianzas y de correlaciones

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \dots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \dots & \sigma_n^2 \end{pmatrix}, \quad R = \begin{pmatrix} 1 & \rho_{12} & \dots & \rho_{1n} \\ \rho_{21} & 1 & \dots & \rho_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n1} & \rho_{n2} & \dots & 1 \end{pmatrix}.$$