Ignacio Cascos Fernández
Departamento de Estadística
Universidad Carlos III de Madrid

Vectores aleatorios

Estadística I — curso 2008–2009

En numerosas ocasiones estudiamos más de una variable asociada a un experimento aleatorio. Un *vector aleatorio* es una aplicación del espacio muestral E en $\mathbb{R}^n$. En el caso bidimensional ($n = 2$),

$$(X, Y) : E \longrightarrow \mathbb{R}^2 .$$

1. Distribución conjunta de un vector aleatorio

A la distribución de probabilidad que describe el comportamiento simultáneo de todas las variables que componen un vector aleatorio se le llama *distribución de probabilidad conjunta*.

1.1. Vectores aleatorios discretos

Si X e Y son dos variables aleatorias discretas, podemos definir

- *función de probabilidad conjunta*: $p(x, y) = P(X = x, Y = y)$, que debe cumplir, $p(x, y) \geq 0$ y $\sum_x \sum_y p(x, y) = 1$.
- *función de distribución conjunta*: $F(x_0, y_0) = P(X \leq x_0, Y \leq y_0) = \sum_{x \leq x_0} \sum_{y \leq y_0} p(x, y)$.

1.1.1. Distribución Multinomial

Dado un experimento aleatorio con k posibles resultados con probabilidades $p_1, p_2, \ldots, p_k$ constantes en distintas realizaciones y que se repite n veces en condiciones de independencia, un vector aleatorio $(X_1, X_2, \ldots, X_n)$ sigue distribución multinomial si cada X_i representa el número de veces (de entre las n realizaciones experimentales) en las que ocurre el resultado i-ésimo.

La función de probabilidad conjunta es

$$P(X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k) = \frac{n!}{x_1!x_2!\cdots x_k!} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k} ,$$

donde $x_1 + x_2 + \cdots + x_k = n$ y $p_1 + p_2 + \cdots + p_k = 1$.

1.2. Vectores aleatorios continuos

Si X e Y son dos variables aleatorias continuas, podemos definir

- *función de densidad conjunta*: $f(x, y)$ que debe cumplir $f(x, y) \geq 0$ y $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1$. Sirve para calcular cualquier probabilidad,

$$P(a \leq X \leq b, c \leq Y \leq d) = \int_a^b \int_c^d f(x, y) dy dx \,.$$

- *función de distribución conjunta*:

$$F(x_0, y_0) = P(X \leq x_0, Y \leq y_0) = \int_{-\infty}^{x_0} \int_{-\infty}^{y_0} f(x, y) dy dx \,.$$

Además tenemos que

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

2. Distribuciones marginales y condicionadas

Las distribuciones marginales y condicionadas son distribuciones unidimensionales asociadas a las de un vector aleatorio. Para ellas podemos calcular probabilidades, medias, varianzas, etc.

2.1. Distribuciones marginales

A la distribución, por separado, de cada una de las variables que componen el vector aleatorio, se le llama distribución marginal.

Variables discretas. Si X e Y son variables aleatorias discretas con función de probabilidad conjunta $p(x, y)$

- *función de probabilidad (marginal) de X:*

$$p_X(x) = P(X = x) = \sum_y P(X = x, Y = y)\,.$$

- *función de probabilidad (marginal) de Y:*

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

Variables continuas. Si X e Y son variables aleatorias continuas con función de densidad conjunta $f(x, y)$

- *función de densidad (marginal) de X:*

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

- *función de densidad (marginal) de Y:*

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

2.2. Distribuciones condicionadas

Variables discretas. Si X e Y son variables aleatorias discretas con función de probabilidad conjunta $p(x, y)$

- *función de probabilidad de Y condicionada a $X = x_0$ ($p_X(x_0) > 0$):*

$$p(y|x_0) = P(Y = y|X = x_0) = \frac{P(X = x_0, Y = y)}{P(X = x_0)} = \frac{p(x_0, y)}{p_X(x_0)}\,.$$

- *función de probabilidad de X condicionada a $Y = y_0$ ($p_Y(y_0) > 0$):*

$$p(x|y_0) = P(X = x|Y = y_0) = \frac{P(X = x, Y = y_0)}{P(Y = y_0)} = \frac{p(x, y_0)}{p_Y(y_0)}\,.$$

Variables continuas. Si X e Y son variables aleatorias continuas con función de densidad conjunta $f(x,y)$

- *función de densidad de* Y *condicionada a* $X = x_0$ *(*$f_X(x_0) > 0$*):*

$$f(y|x_0) = \frac{f(x_0, y)}{f_X(x_0)} \,.$$

- *función de densidad de* X *condicionada a* $Y = y_0$ *(*$f_Y(y_0) > 0$*):*

$$f(x|y_0) = \frac{f(x, y_0)}{f_Y(y_0)} \,.$$

3. Independencia entre variables aleatorias

Dos variables aleatorias son independientes si el valor que toma una no aporta información sobre el valor que tomará la otra.

- *Variables discretas:* X e Y son independientes si para todo x, y se cumple alguna de estas condiciones

$$\begin{aligned} p(y|x) &= p_Y(y) \\ p(x|y) &= p_X(x) \\ p(x,y) &= p_X(x)p_Y(y) \,. \end{aligned}$$

- *Variables continas:* X e Y son independientes si para todo x, y se cumple alguna de estas condiciones

$$\begin{aligned} f(y|x) &= f_Y(y) \\ f(x|y) &= f_X(x) \\ f(x,y) &= f_X(x)f_Y(y) \,. \end{aligned}$$

En general X y Y son independientes si su función de distribución conjunta puede escribirse como producto de las marginales, $F(x,y) = F_X(x)F_Y(y)$ para todos x, y.

4. Características de un vector aleatorio

Trabajamos con un vector aleatorio con n componentes $X_1, X_2, \ldots, X_n$ representándolo como un vector columna,

$$X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}.$$

4.1. Esperanza

El vector de medias de un vector aleatorio X es aquel cuyas componentes son las esperanzas de cada componente de X,

$$\mu = \mathbb{E}[X] = \begin{pmatrix} \mathbb{E}[X_1] \\ \mathbb{E}[X_2] \\ \vdots \\ \mathbb{E}[X_n] \end{pmatrix}.$$

Dado un vector aleatorio bidimensional (X, Y), podemos hallar la esperanza de una transformación suya como:

$$\mathbb{E}[h(X,Y)] = \begin{cases} \sum_x \sum_y h(x,y)p(x,y) & \text{si } X, Y \text{ discretas} \\ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(x,y)f(x,y)dxdy & \text{si } X, Y \text{ continuas} \end{cases}.$$

4.2. Covarianza

La covarianza es una medida de la relación lineal entre dos variables,

$$\mathrm{Cov}(X,Y) = \mathbb{E}\big[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])\big] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]\,.$$

Propiedades de la covarianza:

- si X e Y son independientes, $\mathrm{Cov}(X,Y) = 0$ ($\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$).
- $\mathrm{Cov}(X,Y) = 0$ no implica que X e Y sean independientes.
- $\mathrm{Cov}(aX + b, cY + d) = ac\mathrm{Cov}(X,Y)$.

4.3. Correlación

La correlación es una medida adimensional de la relación lineal entre dos variables,

$$\rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}[X]\mathrm{Var}[Y]}}\,.$$

Propiedades de la correlación:

- si X e Y son independientes, $\rho(X,Y) = 0$.
- $-1 \leq \rho(X,Y) \leq 1$.
- $|\rho(X, aX+b)| = 1$.

4.4. Matriz de varianzas y covarianzas

Se trata de una matriz cuadrada $n \times n$ que viene dada por

$$M_X = \mathbb{E}\big[(X-\mu)(X-\mu)^{\mathrm{t}}\big] = \begin{pmatrix} \mathrm{Var}[X_1] & \mathrm{Cov}(X_1,X_2) & \cdots & \mathrm{Cov}(X_1,X_n) \\ \mathrm{Cov}(X_1,X_2) & \mathrm{Var}[X_2] & \cdots & \mathrm{Cov}(X_2,X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}(X_1,X_n) & \mathrm{Cov}(X_2,X_n) & \cdots & \mathrm{Var}[X_n] \end{pmatrix}.$$

5. Transformaciones de vectores aleatorios

Dado un vector aleatorio $X = (X_1, \ldots, X_n)^{\mathrm{t}}$ con función de densidad conjunta $f_X(x_1, \ldots, x_n)$ lo transformamos en otro vector aleatorio $Y = (Y_1, \ldots, Y_n)^{\mathrm{t}}$ con la misma dimensión

$$\begin{aligned} Y_1 &= g_1(X_1, \ldots, X_n) \\ \vdots &\qquad \vdots \\ Y_n &= g_n(X_1, \ldots, X_n) \end{aligned}$$

de tal modo que existan transformadas inversas.

La función de densidad del nuevo vector aleatorio Y es

$$f_Y(y_1, \ldots, y_n) = f_X(g^{-1}(y_1, \ldots, y_n)) \left| \det \begin{pmatrix} \frac{\mathrm{d}x_1}{\mathrm{d}y_1} & \cdots & \frac{\mathrm{d}x_1}{\mathrm{d}y_n} \\ \vdots & \ddots & \vdots \\ \frac{\mathrm{d}x_n}{\mathrm{d}y_1} & \cdots & \frac{\mathrm{d}x_n}{\mathrm{d}y_n} \end{pmatrix} \right|.$$

5.1. Convolución

Si X_1 y X_2 son variables aleatorias independientes con funciones de densidad $f_{X_1}(x_1)$ y $f_{X_2}(x_2)$, la función de densidad de $Y = X_1 + X_2$ es

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{X_1}(y-x) f_{X_2}(x) dx \, .$$

5.2. Transformaciones lineales

Si $X = (X_1, X_2, \ldots, X_n)^{\mathrm{t}}$ es un vector aleatorio n-dimensional y A es una matriz de dimensión $m \times n$, el vector aleatorio $Y = AX$ (m-dimensional) satisface:

$$\mathbb{E}[Y] = \mathbb{E}[AX] = A\mathbb{E}[X] \, ,$$
$$M_Y = \mathbb{E}\big[(AX - A\mathbb{E}[X])(AX - A\mathbb{E}[X])^{\mathrm{t}}\big] = AM_X A^{\mathrm{t}} \, .$$

6. Distribución Normal multivariante

Decimos que $X = (X_1, X_2)^{\mathrm{t}}$ sigue distribución *Normal bivariante* con vector de medias $\mu = (\mu_1, \mu_2)^{\mathrm{t}}$ y matriz de varianzas y covarianzas

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

si tiene función de densidad

$$f(x,y) = \frac{1}{2\pi|\Sigma|^{1/2}} \exp\left\{ -\frac{1}{2}(x_1-\mu_1, x_2-\mu_2)\Sigma^{-1} \begin{pmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{pmatrix} \right\}, \quad x_1, x_2 \in \mathbb{R} \, .$$

6.1. Propiedades de la distribución Normal bivariante

- Si $\rho = 0$, entonces X_1 y X_2 son independientes.
- dados $a_1, a_2 \in \mathbb{R}$, la variable aleatoria $a_x X_1 + a_2 X_2$ sigue distribución Normal,

  $$a_1X_1 + a_2X_2 \sim \mathrm{N}\Big(a_1\mu_1 + a_2\mu_2, \sqrt{a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + 2a_1a_2\rho\sigma_1\sigma_2}\Big) \, .$$

  En particular, X_1 y X_2 siguen distribución Normal.
- Las variables aleatorias $X_1|_{X_2=x_2}$ y $X_2|_{X_1=x_1}$ siguen distribución Normal.