UNIVERSITAT DE BARCELONA

Vectors aleatoris

Diplomatura d'Estadística

Estadística Matemàtica I

Jordi Ocaña Rebull

Departament d'Estadística

Divisió de Ciències Experimentals i Matemàtiques

Punts que tractarem:

- Vectors aleatoris. Cas general.
  - Funció de distribució conjunta. Propietats.
- Cas discret.
  - Funció de densitat. Propietats.
  - Densitat condicional.
  - Independència estocàstica.
- Cas absolutament continu.
- Vector de mitjanes i matriu de covariàncies.

UNIVERSITAT DE BARCELONA

Concepte de vector aleatori

- Sigui $(\Omega, \mathcal{A}, P)$ un espai de probabilitat.
- Un vector aleatori és una funció

  $\mathbf{X} = (X_1,\ldots,X_k)$ definida com:

  $$\Omega \to \mathbb{R}^k$$
  $$\omega \mapsto \mathbf{X}(\omega) = (X_1(\omega),\ldots,X_k(\omega))$$
- Tal que, per tot $\mathbf{x}=(x_1,\ldots,x_k)$, verifica:

  $$\{\omega \in \Omega : X_1(\omega) \leq x_1,\ldots,X_k(\omega) \leq x_k\}$$
  $$= [X_1 \leq x_1,\ldots,X_k \leq x_k] \in \mathcal{A}$$

Funció de distribució conjunta

$$F_{\mathbf{X}} : \mathbb{R}^k \to [0,1]$$

$$F_{\mathbf{X}}(\mathbf{x}) = F_{X_1 \ldots X_k}(x_1, \ldots, x_k) = F(x_1, \ldots, x_k)$$

$$= P\left(X_1 \le x_1, X_2 \le x_2, \ldots, X_k \le x_k \right)$$

- Probabilitat acumulada fins al punt **x**.
- Definida sempre gràcies a la condició exigida en la definició de vector aleatori.

UNIVERSITAT DE BARCELONA

Propietats de la funció de distribució conjunta. I

Límits i marginals:

1. $F(x_1, \ldots, x_{i-1}, -\infty, x_{i+1}, \ldots, x_k) = 0$
per qualsevol component i
o combinació de components.

2. $F(+\infty, +\infty, \ldots, +\infty) = 1$

3. $F_{X_2 X_3 \ldots X_k}(x_2, x_3, \ldots, x_k) = F(+\infty, x_2, x_3, \ldots, x_k)$
$F_{X_3 \ldots X_m}(x_3, \ldots, x_k) = F(+\infty, +\infty, x_3, \ldots, x_k)$
etc, i així per qualsevol combinació de variables.

Propietats de la funció de distribució conjunta. i II

- No decreixent:

$$x_i < x_i' \;\Rightarrow\; F(x_1,\ldots,x_i,\ldots,x_k) \le F(x_1,\ldots,x_i',\ldots,x_k)$$

- Probabilitat sobre un paral·lelepípede:

$$\mathbf{a} = (a_1,\ldots,a_k),\;\; \mathbf{b} = (b_1,\ldots,b_k), \qquad a_i \le b_i, i = 1,\ldots,k$$

$$P(\mathbf{X} \in (\mathbf{a},\mathbf{b}]) = P(a_1 < X_1 \le b_1,\cdots,a_k < X_k \le b_k) =$$
$$F(b_1,b_2,b_3,\ldots,b_k)$$
$$-F(a_1,b_2,\ldots,b_k) - F(b_1,a_2,\ldots,b_k) - F(b_1,b_2,a_3,\ldots,b_k) - \ldots$$
$$+F(a_1,a_2,b_3,\ldots,b_k) + F(a_1,b_2,a_3,\ldots,b_k) + \ldots$$
$$\cdots + (-1)^k F(a_1,a_2,a_3,\ldots,a_k)$$

Vectors aleatoris discrets. Funció de densitat discreta

- El seu recorregut és finit o numerable.
- Funció de densitat conjunta discreta

$$f_{\mathbf{X}}(\mathbf{x}) = f_{X_1 \dots X_k}(x_1, \dots, x_k) = f(x_1, \dots, x_k) =$$

$$P\left(X_1 = x_1, X_2 = x_2, \dots, X_k = x_k \right)$$

- És una probabilitat:

$$0 \leq f_{\mathbf{X}}(\mathbf{x}) \leq 1$$

$$\sum_{x_1} \cdots \sum_{x_m} f(x_1, \dots x_m) = 1$$

UNIVERSITAT DE BARCELONA

Altres propietats de la densitat discreta

- Relació amb la funció de distribució:

$$F(x_1,\dots x_k) = \sum_{t_1 \le x_1} \cdots \sum_{t_k \le x_k} f(t_1,\dots t_k)$$

- Densitats marginals:

$$f_{2\dots k}(x_2,\dots,x_k) = \sum_{x_1} f(x_1,x_2,\dots x_k)$$

$$f_{3\dots k}(x_3,\dots,x_k) = \sum_{x_1}\sum_{x_2} f(x_1,x_2,\dots x_k) \quad \text{etc.}$$

- Probabilitat sobre qualsevol recinte:

$$P\{\mathbf{X} \in S\} = \sum_{\mathbf{x}\in S} f(\mathbf{x})$$

UNIVERSITAT DE BARCELONA

Densitat condicionada i independència estocàstica

- Funció de densitat condicionada:

$$f(x_{l+1},\ldots,x_k \mid x_1,\ldots,x_l) = \frac{f(x_1,\ldots x_l, x_{l+1},\ldots x_k)}{f_{1\ldots l}(x_1,\ldots,x_l)}$$

- Independència estocàstica entre k v.a. discretes:

$X_1,\ldots,X_k$ est. independents sii,
per tot $\mathbf{x} = (x_1,\ldots,x_k)$,
$f(x_1,\ldots,x_k) = f_1(x_1)\, f_2(x_2)\cdots f_k(x_k)$

- I pel cas general (discret, continu, etc.):

$$F(x_1,\ldots,x_k) = F_1(x_1)\, F_2(x_2)\cdots F_k(x_k)$$

UNIVERSITAT DE BARCELONA

Cas absolutament continu

- En general són aplicables totes les coses dites pel cas discret, amb els canvis habituals:
  - Integrals en lloc de sumatoris
  - $P\{\mathbf{X} = \mathbf{x}\} = 0$
  - Funció de densitat **no** és probabilitat sobre cada punt, és densitat de probabilitat, per tant $f(\mathbf{x}) \geq 0$ (i no valor sempre entre 0 i 1).
  - Finalment, $$f(x_1, \ldots, x_k) = \frac{\partial^k F(x_1, \ldots x_k)}{\partial x_1 \ldots \partial x_k}$$

UNIVERSITAT DE BARCELONA

Moments de vectors aleatoris

- Útil expressió vectorial o matricial.

  Convenció: vectors aleatoris com a **vectors columna**

$$\mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_k \end{pmatrix} \qquad \mathbf{X}' = \left(X_1, X_2, \ldots, X_k \right)$$

- Els dos més importants són:
  – Vector de mitjanes
  – Matriu de variàncies-covariàncies

Vector de mitjanes

- Esperança aplicada a cada component del vector aleatori:

$$\mu = E(\mathbf{X}) = E\begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_k \end{pmatrix} = \begin{pmatrix} E(X_1) \\ E(X_2) \\ \vdots \\ E(X_k) \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_k \end{pmatrix}$$

Matriu de variàncies i covariàncies:

- Matriu formada per les covariàncies entre cada parell de components:

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1k} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{k1} & \sigma_{k2} & \cdots & \sigma_{kk} \end{pmatrix}$$

$$\sigma_{ij} = \operatorname{cov}\left(X_i, X_j\right) \qquad \sigma_{ii} = \sigma_i^2 = \operatorname{var}\left(X_i\right)$$

Matriu de covariàncies com a producte matricial

- Producte matricial de dos vectors,

$$\begin{pmatrix} a_1 \\ \vdots \\ a_k \end{pmatrix} \begin{pmatrix} b_1, \dots, b_k \end{pmatrix} = \begin{pmatrix} a_1 b_1 & \dots & a_1 b_k \\ \vdots & \ddots & \vdots \\ a_k b_1 & \dots & a_k b_k \end{pmatrix}$$

- sobre el vector aleatori centrat

$$\Sigma = \operatorname{cov}(\mathbf{X}) = E\left\{(\mathbf{X} - \mu)(\mathbf{X} - \mu)'\right\} =$$

$$\begin{pmatrix} E\{(X_1 - \mu_1)(X_1 - \mu_1)\} & \dots & E\{(X_1 - \mu_1)(X_k - \mu_k)\} \\ \vdots & \ddots & \vdots \\ E\{(X_k - \mu_k)(X_1 - \mu_1)\} & \dots & E\{(X_k - \mu_k)(X_k - \mu_k)\} \end{pmatrix}$$

Propietats de la matriu de covariàncies

- Simètrica: $\sigma_{ij} = \sigma_{ji}$
- Semidefinida positiva: $\det(\Sigma) \geq 0$
- Definida positiva, $\det(\Sigma) > 0$, si les components del vector aleatori són linealment independents.
- Diagonal si v.a. independents:

$$X_1,\ldots,X_k \text{ est. independents } \begin{Bmatrix} \Rightarrow \\ \not\Leftarrow \end{Bmatrix} \Sigma = \begin{pmatrix} \sigma_1^2 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \sigma_k^2 \end{pmatrix}$$

UNIVERSITAT DE BARCELONA

Matriu de correlacions

$$\mathrm{P} = \begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1k} \\ \rho_{21} & 1 & \cdots & \rho_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{k1} & \rho_{k2} & \cdots & 1 \end{pmatrix}$$

$$\rho_{ij} = \rho\left(X_i, X_j\right) = \frac{\sigma_{ij}}{\sqrt{\sigma_i^2\,\sigma_j^2}} = \frac{\sigma_{ij}}{\sigma_i\,\sigma_j}$$

$$\sigma_{ij} = \rho_{ij}\,\sigma_i\,\sigma_j$$

Relació entre la matriu de correlacions i la de covariàncies

si

$$\mathbf{D} = \begin{pmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_k \end{pmatrix} \qquad \mathbf{D}^{-1} = \begin{pmatrix} \frac{1}{\sigma_1} & 0 & \dots & 0 \\ 0 & \frac{1}{\sigma_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sigma_k} \end{pmatrix}$$

aleshores:

$$\mathbf{\Sigma} = \mathbf{DPD} \qquad \mathrm{P} = \mathbf{D}^{-1}\mathbf{\Sigma}\mathbf{D}^{-1}$$