Distribució normal bivariant i multivariant

Diplomatura d'Estadística

Estadística Matemàtica I

Jordi Ocaña Rebull

Departament d'Estadística

Divisió de Ciències Experimentals i Matemàtiques

UNIVERSITAT DE BARCELONA

Punts que tractarem:

- Construcció de la distribució normal multivariant.
- Funció de densitat normal multivariant.
- Propietats. Marginals. Condicionades. Independència i incorrelació.
- Distribució normal bivariant. Propietats.
- Distribució de combinacions lineals de les components.

UNIVERSITAT DE BARCELONA

Construcció de la normal multivariant. Idea bàsica.

- Idea guia: al cas univariant, a partir de $N(0,1)$ podem tenim qualsevol $N(\mu, \sigma)$:

$$Z \sim N(0,1) \Rightarrow X = \sigma Z + \mu \sim N\left(\mu, \sigma\right)$$

$$\left(\text{o } X \sim N\left(\mu, \sigma\right) \Rightarrow Z = \sigma^{-1}(X - \mu) \sim N(0,1)\ \right)$$

- Què seria la σ multivariant? Mètode de Cholesky permet la descomposició matricial amb **C** triangular inferior. **C** fa el paper de $\Sigma^{1/2}$.

$$\Sigma = \mathbf{C}\mathbf{C}'$$

UNIVERSITAT DE BARCELONA

La matriu **C** fa el paper d'una "desviació típica matricial"

Partim de $\mathbf{Z} = \left(Z_1,\ldots,Z_k\right)'$, $\quad Z_i\ iid$

amb $E(\mathbf{Z}) = \mathbf{0}\ \ \mathrm{cov}(\mathbf{Z}) = \mathbf{I}_k$.

Si $\boldsymbol{\mu} = \left(\mu_1,\ldots,\mu_k\right)'$ vector de mitjanes i $\boldsymbol{\Sigma} = \mathbf{CC}'$

fem el canvi $\mathbf{X} = \mathbf{CZ} + \boldsymbol{\mu}$ i tenim:

$$E(\mathbf{X}) = E\{\mathbf{CZ} + \boldsymbol{\mu}\} = \mathbf{C}E\{\mathbf{Z}\} + \boldsymbol{\mu} = \mathbf{0} + \boldsymbol{\mu} = \boldsymbol{\mu}$$

$$\mathrm{cov}(\mathbf{X}) = E\{(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})'\} = E\{(\mathbf{CZ})(\mathbf{CZ})'\}$$

$$= E\{\mathbf{CZZ'C'}\} = \mathbf{C}E\{\mathbf{ZZ}'\}\mathbf{C}' = \mathbf{CI}_k\mathbf{C}' = \mathbf{CC}' = \boldsymbol{\Sigma}$$

UNIVERSITAT DE BARCELONA

"Tipificat" multivariant

- Per tant, l'operació amb vectors aleatoris equivalent a "reduir" o "tipificar" una v.a. és:

$$\left\{\begin{array}{l}\mathbf{X}\\ E(\mathbf{X})=\mu\\ \operatorname{cov}(\mathbf{X})=\Sigma\end{array}\right\} \underset{\mathrm{CZ}+\mu}{\overset{\mathrm{C}^{-1}(\mathrm{X}-\mu)}{\rightleftarrows}} \left\{\begin{array}{l}\mathbf{Z}\\ E(\mathbf{Z})=\mathbf{0}\\ \operatorname{cov}(\mathbf{Z})=\mathbf{I}_k\end{array}\right\}$$

$$\Sigma = \mathrm{CC}' \qquad \mathrm{C} = \Sigma^{\frac{1}{2}}$$

UNIVERSITAT DE BARCELONA

Normal multivariant

- Si les Z_i fossin $N(0, 1)$, quina seria la distribució de $\mathbf{X} = \mathbf{CZ}+\boldsymbol{\mu}$?
- La distribució que surt d'aquest canvi de variable, de densitat conjunta

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{k}{2}}|\Sigma|^{\frac{1}{2}}}\exp\left\{-\frac{1}{2}(\mathbf{x}-\mu)'\Sigma^{-1}(\mathbf{x}-\mu)\right\}$$

s'anomena normal multivariant, i l'indicarem

$$\mathbf{X} \sim \mathcal{N}\left(\mu, \Sigma\right)$$

Distribució normal multivariant

Propietats i comentaris

- Definició anterior solament vàlida per $|\Sigma| \neq 0$ (existeix inversa de $\boldsymbol{\Sigma}$).
- Existeix definició més general,

$$\mathbf{X} \sim \mathcal{N}\left(\mu, \mathrm{B}^{-1}\right)$$

per forma quadràtica de matriu $\mathbf{B}$ (cal el concepte de g-inversa):

$$f(\mathbf{x}) = \frac{|\mathrm{B}|^{\frac{1}{2}}}{(2\pi)^{\frac{k}{2}}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\mu)'\mathrm{B}(\mathbf{x}-\mu)\right\}$$

UNIVERSITAT DE BARCELONA

Distribució normal multivariant

Propietats i comentaris

- Funció característica:

$$\varphi_{\mathbf{X}}(\mathbf{t}) = E\left\{e^{i(\mathbf{t}'\mathbf{X})}\right\} = \exp\left\{i\boldsymbol{\mu}'\mathbf{t} - \tfrac{1}{2}\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}\right\}$$

$$= \exp\left\{i\sum_{j=1}^{k}\mu_j t_j - \tfrac{1}{2}\sum_{j=1}^{k}\sum_{l=1}^{k} t_j t_l \sigma_{jl}\right\}$$

- Marginals: són normals $N(\mu_j, \sigma_j)$

$$\varphi\left(0,\ldots,0,t_j,0,\ldots,0\right) = \varphi_{X_j}\left(t_j\right) = e^{i\mu_j t_j - \frac{1}{2}t_j^2\sigma_j^2}$$

- Estabilitat de la distribució sota transformacions lineals.

Distribució normal multivariant

Propietats i comentaris

- La distribució marginal per un subvector qualsevol (un subconjunt de m variables, $m < k$) també és normal multivariant.
- La condicionada d'un subvector a un altre és normal multivariant.
- Per aquesta distribució, independència equival a incorrelació:

$$\Sigma \text{ diagonal} \quad \Leftrightarrow \quad X_1, \ldots, X_k \ \text{ est. indep.}$$

UNIVERSITAT DE BARCELONA

Distribució normal bivariant

- Pel cas bivariant, k=2, s'indica

$$(X,Y) \sim N(\mu_1,\mu_2,\sigma_1,\sigma_2,\rho)$$

amb densitat

$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2}Q(x,y)\right\}$$

on

$$Q(x,y) =$$

$$\frac{1}{1-\rho^2}\left(\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right) + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right)$$

UNIVERSITAT DE BARCELONA

Normal bivariant
propietats. I.

- Funció característica:

$$\varphi(t_1,t_2)=$$

$$\exp\left\{i(\mu_1 t_1+\mu_2 t_2)-\tfrac{1}{2}\left(\sigma_1^2 t_1^2+2\rho\sigma_1\sigma_2 t_1 t_2+\sigma_2^2 t_2^2\right)\right\}$$

- Marginals:

$$X\sim N(\mu_1,\sigma_1)\qquad Y\sim N(\mu_2,\sigma_2)$$

- Condicionades:

$$Y|X=x\sim N(\mu_x,\sigma)$$

$$\mu_x=\mu_2+\frac{\rho\sigma_2}{\sigma_1}(x-\mu_1)\qquad \sigma=\sigma_2\sqrt{1-\rho^2}$$

UNIVERSITAT DE BARCELONA

Normal bivariant
propietats. i II.

- Tota combinació lineal de X i Y és normal (encara que siguin dependents):

$$(X,Y) \sim N(\mu_1,\mu_2,\sigma_1,\sigma_2,\rho) \quad \Rightarrow$$
$$\beta_1 X + \beta_2 Y + \alpha \sim N(\mu,\sigma) \qquad (\beta_1 \text{ o } \beta_2 \neq 0)$$
$$\mu = \beta_1\mu_1 + \beta_2\mu_2 + \alpha$$
$$\sigma = \sqrt{\beta_1^2\sigma_1^2 + \beta_2^2\sigma_2^2 + 2\beta_1\beta_2\sigma_1\sigma_2\rho}$$

- Incorrelació equival a independència:

$$\rho = 0 \Leftrightarrow f(x,y) = f_1(x) f_2(y)$$