Vetores Aleatórios e Distribuição Normal Multivariada

Vicente G. Cancho
garibay@icmc.usp.br

Vetor e Matriz aleatório

- Um vetor aleatório é um vetor cujos elementos são variáveis aleatórias
- Da mesma forma, uma matriz aleatória é uma matriz cujos elementos são variáveis aleatórias
- O valor esperado de uma matriz aleatória (ou vetor) é a matriz (vetor) que consiste nos valores esperados de cada um de seus elementos,
- Especificamente, seja $\mathbf{X} = \{X_{ij}\}_{n\times p}$ uma matriz aleatória de ordem $n \times p$.
- Então, o valor esperado de $\boldsymbol{X}$, denotado por $E(\boldsymbol{X})$, é a matriz de números de ordem $n \times p$.

$$E[\mathbf{X}] = \begin{pmatrix} E[X_{12}] & E[X_{12}] & \dots & E[X_{1p}] \\ E[X_{21}] & E[X_{22}] & \dots & E[X_{2p}] \\ \vdots & \vdots & \ddots & \vdots \\ E[X_{n1}] & E[X_{n2}] & \dots & E[X_{np}] \end{pmatrix}$$

onde

$$\mu_{ij} = E[X_{ij}] = \begin{cases} \displaystyle\int_{-\infty}^{\infty} x_{ij} f_{ij}(x_{ij}) dx_{ij}, & \text{se } X_{ij} \text{ é v.a.c com f.d.p. } f_{ij}(x_{ij}) \\ \displaystyle\sum_{x_{ij}} x_{ij} p_{ij}(x_{ij}), & \text{se } X_{ij} \text{ é v.a.d. com f.p. } p_{ij}(x_{ij}) \end{cases}$$

Preposição

Sejam $\boldsymbol{X}$ *e* $\boldsymbol{Y}$ *matrizes aleatórias da mesma dimensão, e sejam* $\boldsymbol{A}$ *e* $\boldsymbol{B}$ *matrizes de constantes (adequadas). Então*

$$E[\boldsymbol{X} + \boldsymbol{Y}] = E[\boldsymbol{X}] + E[\boldsymbol{Y}]$$
$$E[\boldsymbol{AX} + \boldsymbol{BY}] = E[\boldsymbol{AX}] + E[\boldsymbol{BY}]$$

Vetor de Média e Matriz Covariância

Suponha $\boldsymbol{X}^\top = \begin{pmatrix} X_1, & X_2, & \dots, & X_p \end{pmatrix}$, é um **vetor aleatório** de ordem $p \times 1$. Então cada elemento de $\boldsymbol{X}$, é uma variável aleatória.
A **média** μ_i e **variância** $\sigma_i^2 = \sigma_{ii}$, para $i = 1, \dots, p$, são definidos por $\mu_i = E(X_i)$ e $\sigma_i^2 = Var(X_i) = E[(X_i - \mu_i)^2]$ respectivamente.
Especificamente,

$$\mu_i = \begin{cases} \displaystyle\int_{-\infty}^{\infty} x_i f_i(x_i) dx_i, & \text{se } X_i \text{ é v.a. contínua com f.d.p. } f_i(x_i) \\ \displaystyle\sum_{x_i} x_i p_i(x_i), & \text{se } X_i \text{ é v.a. discreta com f.p. } p_i(x_i) \end{cases}$$

Vetor de Média e Matriz Covariância

$$\sigma_i^2 = \begin{cases} \displaystyle\int_{-\infty}^{\infty} (x_i - \mu_i)^2 f_i(x_i) dx_i, & \text{se } X_i \text{ é v.a. contínua com f.d.p. } f_i(x_i) \\ \displaystyle\sum_{x_i} (x_i - \mu_i)^2 p_i(x_i), & \text{se } X_i \text{ é v.a. discreta com f.p. } p_i(x_i) \end{cases}$$

A **covariância** e a **correlação** (populacionais) entre duas variáveis X_i e X_k, com $i, k = 1, \ldots, p$ e $i \neq k$ são dadas, respectivamente, por

$$Cov(X_i, X_k) = \sigma_{ik} = E[(X_i - \mu_i)(X_k - \mu_k)] \text{ e}$$

$$Cor(X_i, X_k) = \rho_{ik} = \frac{\sigma_{ik}}{\sqrt{\sigma_i^2}\sqrt{\sigma_k^2}}$$

Mais especificamente,

$$\sigma_{ik} = \begin{cases} \displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x_i - \mu_i)(x_k - \mu_k)f_{ik}(x_i, x_k)dx_i dx_k, & \text{contínuo} \\ \displaystyle\sum_{x_i}\sum_{x_k}(x_i - \mu_i)(x_k - \mu_k)p_{ik}(x_i, x_k), & \text{discreto} \end{cases}.$$

Distribuição conjunta de um vetor aleatório

A distribuição conjunta das p variáveis aleatórias (v.a.'s) $X_1, \ldots, X_p$ ou do vetor $\boldsymbol{X} = (X_1, \ldots, X_p)^\top$ é descrito pela **função densidade de probabilidade conjunta** (ou função de probabilidade conjunta no caso discreto) de $X_1, \ldots, X_p$, dada por

$$f(x_1, \ldots, x_p) = f(\boldsymbol{x}),$$

com

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(u_1, \ldots, u_p) du_1 \ldots du_p = 1$$

(ou somatória no caso discreto).

Função distribuição acumulada

Se $\boldsymbol{X} = (X_1, \ldots, X_p)^\top$ é um vetor aleatório e tem f.d.p. $f(\boldsymbol{x}) = f(x_1, \ldots, x_p)$, então a **função distribuição acumulada** de $\boldsymbol{x}$ é dada por

$$F(\boldsymbol{x}) = P(\boldsymbol{X} \leq \boldsymbol{x}) = P(X_1 \leq x_1, \ldots, X_p \leq x_p)$$

que pode ser calculada no nosso contexto como

$$F(\boldsymbol{x}) = \int_{-\infty}^{x_1} \ldots \int_{-\infty}^{x_p} f(u_1, \ldots, u_p) du_1 \ldots du_p$$

(no caso discreto substituir por somatórias adequadas)

Função distribuição acumulada

Se $\boldsymbol{X} = (X_1, \ldots, X_p)^\top$ é um vetor aleatório e tem f.d.p. $f(\boldsymbol{x}) = f(x_1, \ldots, x_p)$, então a **função distribuição acumulada** de $\boldsymbol{x}$ é dada por

$$F(\boldsymbol{x}) = P(\boldsymbol{X} \leq \boldsymbol{x}) = P(X_1 \leq x_1, \ldots, X_p \leq x_p)$$

que pode ser calculada no nosso contexto como

$$F(\boldsymbol{x}) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_p} f(u_1, \ldots, u_p) du_1 \ldots du_p$$

(no caso discreto substituir por somatórias adequadas)

As variáveis aleatórias X_i e X_k, são estocasticamente independente se

$$P(X_i \leq x_i, X_k \leq x_k) = P(X_i \leq x_i)P(X_k \leq x_k)$$

para todos os pares possíveis (x_i, x_k),
No caso que X_i e X_k são variáveis aleatórias contínuas com f.d.p. $f_{ik}(x_i, x_k)$, a condição para a independência é

$$f_{ik}(x_i, x_k) = f_i(x_i)f_k(x_k)$$

para todo par (x_i, x_k).

Independência

Se tivermos p variáveis aleatórias contínuas $X_1, \ldots, X_p$, elas são estatisticamente independentes se a sua f.d.p. conjunta pode ser fatorada como

$$f(\boldsymbol{x}) = f_1(x_1) \ldots f_p(x_p), \forall (x_1, \ldots, x_p).$$

A independência tem uma implicação importante:

Preposição

Se X_i e X_k são independentes, então

$$Cov(X_i, X_k) = 0$$

Vetor de médias e matriz de variâncias e covariâncias

O **vetor de médias** (populacionais) de $\boldsymbol{X}$ é $\boldsymbol{\mu}_{p\times 1}$ e a **matriz de variâncias e covariâncias** (populacionais) de $\boldsymbol{X}$ é $\Sigma_{p\times p}$, em que

$$\boldsymbol{\mu} = E(\boldsymbol{X}) = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{pmatrix} = \begin{pmatrix} E(X_1) \\ E(X_2) \\ \vdots \\ E(X_p) \end{pmatrix} \text{ e } \Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1p} \\ & \sigma_2^2 & \dots & \sigma_{2p} \\ & & \ddots & \vdots \\ \text{sim.} & & & \sigma_p^2 \end{pmatrix}$$

Obs: Note que Σ é simétrica pois $Cov(X_i, X_k) = Cov(X_k, X_i)$, para $i, j = 1, \dots, p$.

A **matriz de correlações** (populacionais) de $\boldsymbol{X}$ é $\boldsymbol{\rho} = \{\rho\}_{p\times p}$, em que

$$\boldsymbol{\rho} = Cor(\boldsymbol{X}) = \begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1p} \\ & 1 & \cdots & \rho_{2p} \\ & & \ddots & \vdots \\ \text{sim.} & & & 1 \end{pmatrix}$$

- $\boldsymbol{\rho}$ é uma matriz simétrica
- $-1 \leq \rho_{ik} \leq 1$ para $i \neq k$ e $i, k = 1 \ldots, p$.

Matricialmente,

$$Cov(\boldsymbol{X}) = \Sigma = E[(\boldsymbol{X} - \boldsymbol{\mu})(\boldsymbol{X} - \boldsymbol{\mu})^\top].$$

Além disso,

$$Cor(\boldsymbol{X}) = \boldsymbol{\rho} = (V^{1/2})^{-1}\Sigma(V^{1/2})^{-1} = V^{-1/2}\Sigma V^{-1/2},$$

em que

$$V^{1/2} = \begin{pmatrix} \sqrt{\sigma_{11}} & 0 & \dots & 0 \\ 0 & \sqrt{\sigma_{22}} & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \sqrt{\sigma_{pp}} \end{pmatrix}$$

Considere a matriz de variância-covariâncias

$$\mathbf{\Sigma} = \begin{pmatrix} 4 & 1 & 2 \\ 1 & 9 & -3 \\ 2 & -3 & 25 \end{pmatrix} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{pmatrix}$$

A Matriz

$$V^{1/2} = \begin{pmatrix} \sqrt{4} & 0 & 0 \\ 0 & \sqrt{9} & 0 \\ 0 & 0 & \sqrt{25} \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

e

$$V^{-1/2} = \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & 1/5 \end{pmatrix}$$

Variância-Covariância: Exemplo

A matriz de correlações resulta

$$
\begin{aligned}
\boldsymbol{\rho} &= V^{-1/2}\boldsymbol{\Sigma}V^{-1/2}\\
&= \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & 1/5 \end{pmatrix} \begin{pmatrix} 4 & 1 & 2 \\ 1 & 9 & -3 \\ 2 & -3 & 25 \end{pmatrix} \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & 1/5 \end{pmatrix}
\end{aligned}
$$

$$
= \begin{pmatrix} 1 & 1/6 & 1/5 \\ 1/6 & 1 & -1/5 \\ 1/5 & -1/5 & 1 \end{pmatrix}
$$

Variância e covariância em forma matricial

Sejam $\boldsymbol{X}$ e $\boldsymbol{Y}$ vetores aleatórios p-dimensionais com $E(\boldsymbol{X}) = \boldsymbol{\mu}_X$ e $E(\boldsymbol{Y}) = \boldsymbol{\mu}_Y$.

Define-se

1.
$$\begin{aligned} Var(\boldsymbol{X}) = Cov(\boldsymbol{X}) &= E[(\boldsymbol{X} - \boldsymbol{\mu}_X)(\boldsymbol{X} - \boldsymbol{\mu}_X)^\top] \\ &= E[\boldsymbol{X}^\top \boldsymbol{X}] - \boldsymbol{\mu}_X^\top \boldsymbol{\mu}_X \end{aligned}$$

2.
$$\begin{aligned} Cov(\boldsymbol{X}, \boldsymbol{Y}) &= E[(\boldsymbol{X} - \boldsymbol{\mu}_X)(\boldsymbol{Y} - \boldsymbol{\mu}_Y)^\top] \\ &= E[\boldsymbol{X}^\top \boldsymbol{Y}] - \boldsymbol{\mu}_X^\top \boldsymbol{\mu}_Y \end{aligned}$$

Particionando a matriz de covariância

Seja vetor $\boldsymbol{X}$ particionada com $\boldsymbol{X} = (\boldsymbol{X}_1^\top, \boldsymbol{X}_2^\top)^\top$ em $\boldsymbol{X}_1$ e $\boldsymbol{X}_2$ são vetores de dimensões q e $p - q$ respectivamente, ou seja,

$$\boldsymbol{X} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_q \\ \dots \\ X_{q+1} \\ \vdots \\ X_p \end{pmatrix} = \begin{pmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{pmatrix} \text{ e } \boldsymbol{\mu} = \begin{pmatrix} E[X_1] \\ E[X_2] \\ \vdots \\ E[X_q] \\ \dots \\ E[X_{q+1}] \\ \vdots \\ E[Xp] \end{pmatrix} = \begin{pmatrix} E[\boldsymbol{X}_1] \\ E[\boldsymbol{X}_2] \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}$$

Particionando a matriz de covariância

A matriz de variância-covariância $\boldsymbol{\Sigma}$ particionada resulto

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}$$

onde

- $\boldsymbol{\Sigma}_{11}$ é a covariância de $\boldsymbol{X}_1$ de ordem $q \times q$
- $\boldsymbol{\Sigma}_{22}$ é é a covariância de $\boldsymbol{X}_2$ de ordem $(p-q) \times p-q$.
- $\boldsymbol{\Sigma}_{12}$ é de ordem $q \times (p-1)$,
- $\boldsymbol{\Sigma}_{21}$ é de ordem $(p-q) \times p$
- $\boldsymbol{\Sigma}_{12} = \boldsymbol{\Sigma}_{21}^{\top}$ e são covariância de $\boldsymbol{X}_1$ e $\boldsymbol{X}_2$

Vetor de Médias e Matriz de Covariância de uma Combinação Linear de Variáveis Aleatórias

Combinação Linear de p variáveis aleatórias

Se $\boldsymbol{X}$ é um vetor de aleatória p-dimensional e $\boldsymbol{c}$ é um vetor de constantes reais de ordem $p \times 1$, então a combinação linear

$$\boldsymbol{c}^\top \boldsymbol{X} = c_1 X_1 + \cdots + c_p X_p$$

tem média e variância

$$\mu = E[\boldsymbol{c}^\top \boldsymbol{X}] = \boldsymbol{c}^\top E[\boldsymbol{X}] \text{ e}$$
$$\sigma^2 = Var[\boldsymbol{c}^\top \boldsymbol{X}] = \boldsymbol{c}^\top Var[\boldsymbol{X}] \boldsymbol{c}$$

respectivamente.

Exercício (mostre)

Vetor de Médias e Matriz de Covariância de uma Combinação Linear de Variáveis Aleatórias

Considere as q combinações lineares de p variáveis aleatórias,

$$
\begin{aligned}
Z_1 &= c_{11}X_1 + \cdots + c_{1p}X_p \\
Z_2 &= c_{21}X_1 + \cdots + c_{2p}X_p \\
\vdots &= \qquad\qquad \vdots \\
Z_q &= c_{q1}X_1 + \cdots + c_{qp}X_p
\end{aligned}
$$

ou

$$
\boldsymbol{Z} = \begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_q \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1p} \\ c_{21} & c_{22} & \dots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{q1} & c_{q2} & \dots & c_{qp} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_q \end{pmatrix} = \boldsymbol{CX}
$$

A combinação linear $\boldsymbol{Z} = \boldsymbol{CX}$ tem

$$
\begin{aligned}
\boldsymbol{\mu_Z} &= E[\boldsymbol{Z}] = \boldsymbol{C}\,E[\boldsymbol{X}] = \boldsymbol{C}\boldsymbol{\mu_X} \\
\boldsymbol{\Sigma_Z} &= Cov[\boldsymbol{Z}] = \boldsymbol{C}\,Cov[\boldsymbol{X}]\boldsymbol{C}^\top = \boldsymbol{C}\boldsymbol{\Sigma_X}\boldsymbol{C}^\top
\end{aligned}
$$

Seja $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ a.a. de uma população com $E(\boldsymbol{X}) = \boldsymbol{\mu}$ e $Cov(\boldsymbol{X}) = \boldsymbol{\Sigma}$, sendo $\boldsymbol{X}_i \in R^p$. Então

1. $\bar{\boldsymbol{X}}$ (vetor de medias amostrais) é um estimador não-viesado para $\boldsymbol{\mu}$, com $Cov(\bar{\boldsymbol{X}}) = \dfrac{\boldsymbol{\Sigma}}{n}$ e
2. $\dfrac{n}{n-1}\boldsymbol{S}$ é um estimador não-viesado para $\boldsymbol{\Sigma}$, ou seja $E\left(\dfrac{n}{n-1}\boldsymbol{S}\right) = \boldsymbol{\Sigma}$.

onde $\boldsymbol{S}$ a matriz de variância amostral.

A Distribuição Normal Multivariada

Distribuição Normal Multivariada

Uma variável aleatória contínua X tem distribuição normal com média μ e variância σ^2 se sua função densidade é dada por

$$f(x) = (2\pi\sigma^2)^{-1/2} \exp\left\{-\frac{1}{2}(x-\mu)\sigma^{-2}(x-\mu)\right\}, \quad x \in R.$$

- O expoente da densidade normal univariada: $(x-\mu)(\sigma^2)^{-1}(x-\mu)$ mede a distância quadrada de x em relação à μ em unidade de desvido padrão.
- Esta distância pode ser generalizada para o caso multivariado, com vetor $\boldsymbol{X}$ de ordem $p \times 1$ dada por,

  $$(\boldsymbol{x}-\boldsymbol{\mu})^\top \Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})$$

  onde $\boldsymbol{\mu}$ é o vetor valores esperados do vetor $\boldsymbol{X}$ e Σ a matriz de covariâncias de $\boldsymbol{X}$.

Definição

Dizemos que um vetor aleatório $\boldsymbol{X}$ de ordem $(p \times 1)$ tem distribuição normal multivariada (ou p-variada) com vetor de médias $\boldsymbol{\mu}$ e matriz de variâncias e covariâncias Σ de ordem $p \times p$, se sua função de densidade conjunta é dada por

$$f(\boldsymbol{x}) = (2\pi|\Sigma|)^{-1/2} \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\top}\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\}, \quad \boldsymbol{x} \in R^{p}. \qquad (1)$$

onde $\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix}$ e $\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \cdots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{pmatrix}$.

Notação: $\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

Exemplo (Normal bivariada)

Para $p = 2$, a f.d.p. dada em (1) é dada em termos dos parâmetros individuais

$$f(\boldsymbol{x}) = (2\pi)^{-1}|\Sigma|^{-1/2}\exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\top\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\}, \ \boldsymbol{x} \in R^2. \quad (2)$$

onde $\boldsymbol{x} = \binom{x_1}{x_2}$, $\boldsymbol{\mu} = \binom{\mu_1}{\mu_2}$ *e* $\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}$.

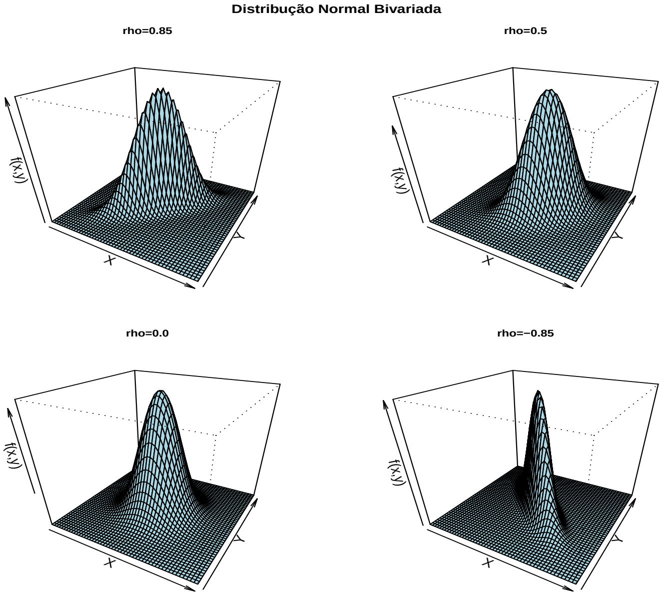
Introduzindo o coeficiente de correlação de X_1 e X_2, $\rho = \frac{\sigma_{12}}{\sqrt{\sigma_{11}\sigma_{22}}} \to \sigma_{12} = \rho\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}$. Daí temos que,

$$\Sigma^{-1} = \frac{1}{\sigma_{11}\sigma_{22}(1-\rho^2)}\begin{pmatrix} \sigma_{22} & -\rho\sqrt{\sigma_{11}}\sqrt{\sigma_{22}} \\ -\rho\sqrt{\sigma_{11}}\sqrt{\sigma_{22}} & \sigma_{11} \end{pmatrix}$$

Logo, a f.d.p. em (2) se reduz a

$$f(x_1, x_2) = \frac{1}{2\pi\sqrt{\sigma_{11}\sigma_{22}(1-\rho^2)}} \exp\left\{ \frac{-1}{2(1-\rho^2)} \left[\frac{(x_1-\mu_1)^2}{\sigma_{11}} + \frac{(x_2-\mu_2)^2}{\sigma_{22}} - 2\rho \left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}} \right) \left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}} \right) \right] \right\} \tag{3}$$

Distribuição normal multivariada



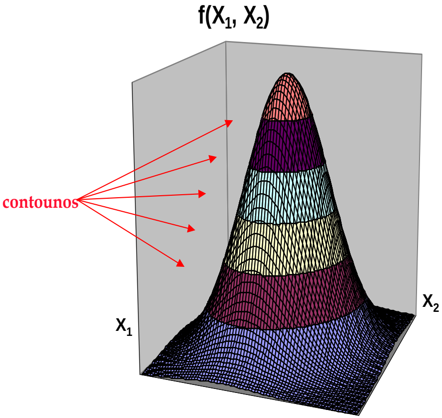
Distribuição normal multivariada

A densidade normal é constante em superfícies cujas distâncias quadráticas $(\boldsymbol{x}-\boldsymbol{\mu})^\top\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})$ são constantes.
Esses padrões são chamados de contornos ou curvas de nível

$$E_c = \{\boldsymbol{x}; (\boldsymbol{x}-\boldsymbol{\mu})^\top\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) = c^2\}. \tag{4}$$

Além disso, considere a decomposição espectral de $\Sigma = \Gamma\Lambda\Gamma^\top$ onde $\Gamma = (\gamma_1, \ldots, \gamma_p)$ e $\Lambda = diag(\lambda_1, \ldots, \lambda_p)$ com $\lambda_1 \geq \ldots \geq \lambda_p$.
O E_c, para qualquer c é um elipsoide centrada em $\boldsymbol{\mu}$ e tem eixos, $\pm c\sqrt{\lambda_i}\gamma_i$, onde $\Sigma\gamma_i = \lambda_i\gamma_i$, para $\imath = 1, \ldots, p$.

Contornos da distribuição normal multivariada



Exemplo

Considere a distribuição normal bivariada com $\sigma_{11} = \sigma_{22}$. Os eixos da elipsóide dados por (4) sao fornecidos pelos autovalores e autovetores de Σ. Portanto, para obte-los, a equação $|\Sigma - \boldsymbol{I}\lambda| = 0$ deve ser resolvida

$$\begin{vmatrix} \sigma_{11} - \lambda & \sigma_{12} \\ \sigma_{12} & \sigma_{11} - \lambda \end{vmatrix} = (\sigma_{11} - \lambda)^2 - \sigma_{12}^2 = 0$$

Consequentemente os autovalores resultam: $\lambda_1 = \sigma_{11} + \sigma_{12}$ e $\lambda_2 = \sigma_{11} - \sigma_{12}$. Os autovetores são determinados por :

$$\Sigma \boldsymbol{e}_i = \lambda_i \boldsymbol{e}_i, \; i = 1, 2.$$

Continuação do exemplo

Para $i = 1$ tem-se

$$\begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{11} \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} = (\sigma_{11} + \sigma_{12}) \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

ou

$$\begin{aligned} \sigma_{11} e_1 + \sigma_{12} e_2 &= (\sigma_{11} + \sigma_{12}) e_1 \\ \sigma_{12} e_1 + \sigma_{11} e_2 &= (\sigma_{11} + \sigma_{12}) e_2 \end{aligned}$$

Essas equações levam ao resultado de que $e_1 = e_2$, e após normalização, o primeiro autovetor é: $\boldsymbol{e}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$

De forma similar foi obtido o segundo autovetor, o qual é:

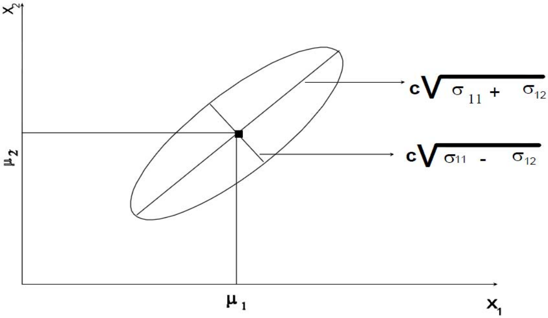
$\boldsymbol{e}_2 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$

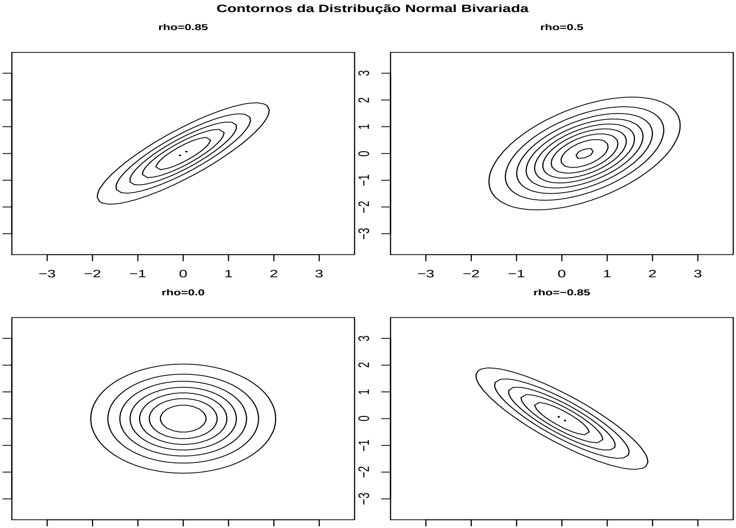
Continuação do exemplo

Se a covariância é positiva, $\lambda_1 = \sigma_{11} + \sigma_{12}$ e o maior autovalor e seu autovetor associado se posiciona ao longo de uma linha de 45^0 através do ponto $\boldsymbol{\mu} = [\mu_1, \mu_2]$. , para qualquer $\sigma_{12} > 0$. Os eixos são fornecidos por $\pm c\sqrt{\lambda_i}\boldsymbol{e}_i \quad (i = 1, 2)$ e estão representados na Figura seguinte.

Continuação do exemplo

Figura: Contornos (curva de nível) de densidade constante para a distribuicão normal bivariada com $\sigma_{11} = \sigma_{22}$ e $\sigma_{12} > 0$




Contornos da Distribução Normal Bivariada
rho=0.85
rho=0.5
rho=0.0
rho=−0.85
−3 −2 −1 0 1 2 3
−3 −2 −1 0 1 2 3

Definição (Distribuição normal p-variada padrão)

Um vetor aleatório $\boldsymbol{Z}$ de ordem $p \times 1$ tem distribuição normal p-variada padrão se somente se a f.d.p de é dada por

$$f(z) = (2\pi)^{-p/2} \exp\left\{-\frac{1}{2}\boldsymbol{z}'\boldsymbol{z}\right\}, \quad \boldsymbol{z} \in R^p \tag{5}$$

Notação: $\boldsymbol{Z} \sim N_p(\boldsymbol{0}, \boldsymbol{I})$

Preposição (1)

Se $\boldsymbol{Z} \sim N_p(\mathbf{0}, \boldsymbol{I})$, então a função geradora de momentos de $\boldsymbol{Z}$ é dada por

$$m_{\boldsymbol{Z}}(\boldsymbol{t}) = e^{\frac{\boldsymbol{t}^\top \boldsymbol{t}}{2}}, \ t \in R^p. \tag{6}$$

Prova: Da definição da f.g.m.

$$\begin{aligned} m_{\boldsymbol{Z}}(\boldsymbol{t}) = E(e^{\boldsymbol{t}^\top \boldsymbol{Z}}) &= \int \cdots \int \frac{1}{(2\pi)^{p/2}} e^{\boldsymbol{t}^\top \boldsymbol{z} - \frac{\boldsymbol{z}^\top \boldsymbol{z}}{2}} d\underline{z} \\ &= e^{\sum_{i=1}^p t_i^2/2} \int \cdots \int \frac{1}{(2\pi)^{p/2}} e^{-\frac{1}{2}\sum_{i=1}^p (z_i - t_i)^2} dz_1, \ldots, dz \\ &= e^{\frac{\boldsymbol{t}^\top \boldsymbol{t}}{2}}. \end{aligned}$$

Preposição (2)

Considere o vetor aleatório $\boldsymbol{Z}$, *tal que,* $\boldsymbol{Z} \sim N_p(\boldsymbol{0}, \boldsymbol{I})$. *Seja* γ *um vetor de constantes reais de ordem* $p \times 1$ *e* γ_0 *qualquer constante real. Então a variável aleatória,*

$$X = \gamma' \boldsymbol{Z} + \gamma_0$$

tem distribuição normal com média γ_0 *e variância* $\gamma'\gamma$.

Prova: Como $X = \gamma' \boldsymbol{Z} + \gamma_0$ é uma v.a, vamos determinar a f.g.m. Por definição,

$$\begin{aligned} m_X(t) = E(e^{tX}) = E(e^{t(\gamma^\top \boldsymbol{Z} + \gamma_0)}) &= e^{t\gamma_0} E(e^{t(\gamma')}) \\ &= \int \cdots \int \frac{1}{(2\pi)^{p/2}} e^{t(\gamma' \boldsymbol{z}) - \frac{\boldsymbol{z}^\top \boldsymbol{z}}{2}} d\underline{z} \\ &= e^{t\gamma_0 + \frac{(\gamma'\gamma)t^2}{2}}, \ t \in R. \end{aligned}$$

Exemplo

Seja $Z \sim N_3(\mathbf{0}, \boldsymbol{I})$. *Determinar a distribuição de* $Y = Z_1 - 2Z_2 + 4Z_3 + 6$. *Neste caso temos que*

$$Y = [1, -2, 4]\begin{bmatrix} Z_1 \\ Z_2 \\ Z_3 \end{bmatrix} + 6 = \boldsymbol{\gamma}'\boldsymbol{Z} + \gamma_0.$$

Do resultado anterior temos que $Y \sim N(6, 21)$.

Exercício

Seja $Z \sim N_4(\mathbf{0}, \boldsymbol{I})$. *Determinar a distribuição de* $X = Z_1 + Z_2 + 4Z_3 + Z_4 + 1$.

Preposição (3)

Se $\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \Sigma)$, *com f.d.p. dada em (1) e seja*

$$\boldsymbol{Y} = \Sigma^{-1/2}(\boldsymbol{X} - \boldsymbol{\mu}), \tag{7}$$

onde $\Sigma^{-1/2}$ *é a matriz raiz quadrada de* Σ^{-1}. *Então* $Y_1, \ldots, Y_p$ *são variáveis aleatórias independentes com distribuição normal padrão.*

Prova: De (7), tem-se que

$$(\boldsymbol{x} - \boldsymbol{\mu})^\top \Sigma^{-1}(\boldsymbol{x} - \boldsymbol{\mu}) = \boldsymbol{Y}^\top \boldsymbol{Y}.$$

Observe em (7), $\boldsymbol{x} = \Sigma^{1/2}\boldsymbol{y} + \boldsymbol{\mu}$, daí o Jacobiano da transformação é $|\Sigma|^{1/2}$. Considerando (1), a f.d.p. é dada por

$$g(y) = \frac{1}{(2\pi)^{p/2}} e^{-\sum_{i=1}^{p} y_i^2/2}, \; y_i \in R.$$

Preposição (4)

Se $\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \Sigma)$, *então a função geradora de momentos de* $\boldsymbol{X}$ *é dada por*

$$\phi_{\boldsymbol{X}}(\boldsymbol{t}) = e^{\boldsymbol{t}'\boldsymbol{\mu} + \frac{\boldsymbol{t}'\Sigma\boldsymbol{t}}{2}}, \; t \in R^p. \tag{8}$$

Prova: Considerando o resultado (3), $\boldsymbol{X} = \Sigma^{1/2}\boldsymbol{Y} + \boldsymbol{\mu}$ determinamos que a função geradora de momentos de $\boldsymbol{X}$ é dada por

$$\phi_{\boldsymbol{X}}(\boldsymbol{t}) = E(e^{\boldsymbol{t}'X}) = e^{\boldsymbol{t}'\mu}E(e^{\boldsymbol{u}'y}),$$

onde

$$\boldsymbol{u}' = \boldsymbol{t}'\Sigma^{1/2}.$$

Continuação da demonstração do resultado 4

Como os Y_i´s são independentes $N_1(0,1)$ tem-se

$$E(e^{\boldsymbol{u}'y}) = \prod_{i=1}^{p} \phi_{y_i}(u_i) = \prod_{i=1}^{p} e^{u_i^2/2} = e^{\boldsymbol{u}'\boldsymbol{u}/2}$$
$$= e^{\boldsymbol{t}'\mu + \boldsymbol{t}'\Sigma^{1/2}\Sigma^{1/2}\boldsymbol{t}/2}$$
$$= e^{\boldsymbol{t}'\boldsymbol{\mu} + \frac{\boldsymbol{t}'\Sigma\boldsymbol{t}}{2}}.$$

$$\phi_{\boldsymbol{X}}(\boldsymbol{t}) = e^{\boldsymbol{t}'\mu + \boldsymbol{u}'\boldsymbol{u}/2} = e^{\boldsymbol{t}'\mu + \boldsymbol{t}'\Sigma^{1/2}\Sigma^{1/2}\boldsymbol{t}/2}$$
$$= e^{\boldsymbol{t}'\boldsymbol{\mu} + \frac{\boldsymbol{t}'\Sigma\boldsymbol{t}}{2}}.$$

Preposição (5)

Se $\boldsymbol{X}$ é um vetor aleatório normal multivariado de ordem $p \times 1$ com vetor médias $\boldsymbol{\mu}$ e matriz de covariância Σ e seja A uma matriz de constantes de ordem $m \times p$ e $\boldsymbol{b}$ vetor de constante de ordem $m \times 1$, então o vetor aleatório

$$Y = AX + \boldsymbol{b},$$

tem distribuição normal m-variada com vetor de médias $A\boldsymbol{\mu} + \boldsymbol{b}$ e covariância $A\Sigma A^\top$.

Prova: (Exercício)

Se $\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \Sigma)$, então:

1. Combinações lineares das componentes de $\boldsymbol{X}$ são normalmente distribuidos. Ou seja, a v.a.

$$Y = a_1X_1 + a_2X_2 + ...a_pX_p = \begin{bmatrix} a_1 & a_2 & \cdots & a_p \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \boldsymbol{X} = \boldsymbol{aX},$$

   terá distribuição $N_1(\boldsymbol{a}^\top \boldsymbol{\mu}, \boldsymbol{a}^\top \Sigma \boldsymbol{a})$, onde $\boldsymbol{a} = (a_1, \ldots, a_p)$.

2. Todo os componentes de $\boldsymbol{X}$ tem distribuição normal. Fazendo

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \boldsymbol{X} = X_1,$$

   Assim, $x_1 \sim N(\mu_1, \sigma_{11})$.

Propriedades da distribuição normal multivariada

1. Considere a partição $\boldsymbol{X} = \begin{bmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{bmatrix}$ onde $\boldsymbol{X}_1$ e $\boldsymbol{X}_2$ são vetores aleatórios de ordem $q \times 1$ e $(p-q) \times 1$ respectivamente. Sejam as correpondentes partições no vetor de média e covariância, dadas por

$$\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \quad \text{e} \quad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

   Seja a matriz ,$A = [\boldsymbol{I}_{q \times q} \mid \boldsymbol{O}_{(p-q) \times p}]$, do Resultado (5) tem-se, $A\boldsymbol{X} = \boldsymbol{X}_1 \sim N_q(\boldsymbol{\mu}_1, \Sigma_{11})$

2. Se os componentes de covariância forem zero entre dois subconjuntos de $\boldsymbol{X}$, implica em dizer que eles são independentemente distribuídos. Esta propriedade só é válida se $\boldsymbol{X}$ tiver distribuição normal multivariada.

Preposição (6)

Se $\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ *e* $\boldsymbol{X}_1$ *e* $\boldsymbol{X}_2$ *é a partição de* $\boldsymbol{X}$ *de ordem* $q \times 1$ *e* $(p-q) \times 1$ *respectivamente e suponha que* $\boldsymbol{\mu}$ *e* $\boldsymbol{\Sigma}$ *são particionados de acordo. Isto é*

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{pmatrix} \sim N_p\left(\begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\right).$$

Então

(i) *Os vetores aleatórias* $\boldsymbol{X}_1$ *e* $\boldsymbol{X}_2$ *são independentes se e somente se* $Cov(\boldsymbol{X}_1, \boldsymbol{X}_2) = \Sigma_{12} = \boldsymbol{0}$

(ii) *A distribuição condicional* $\boldsymbol{X}_1$ *dado* $\boldsymbol{X}_2 = \boldsymbol{x}_2$ *é*

$$N_p\left(\boldsymbol{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\boldsymbol{x}_2 - \boldsymbol{\mu}_2), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right)$$

Propriedades da distribuição normal multivariada

Prova: Considere o vetor aleatório $\boldsymbol{X}_1^* = \boldsymbol{X}_1 + \Sigma_{12}\Sigma_{22}^{-1}\boldsymbol{X}_2$ e $\boldsymbol{X}_2^* = \boldsymbol{X}_2$.

$$\boldsymbol{X}^* = \begin{pmatrix} \boldsymbol{X}_1^* \\ \boldsymbol{X}_2^* \end{pmatrix} = A\boldsymbol{X}, \ \boldsymbol{A} = \begin{pmatrix} \boldsymbol{I}_p & \Sigma_{12}\Sigma_{22}^{-1} \\ \boldsymbol{0}_{p\times(p-q)} & \boldsymbol{I}_{p-q} \end{pmatrix}$$

Pela Preposição 5, o vetor aleatório $\boldsymbol{X}^*$ tem distribuição normal. Uma inspeção da matriz de covariância de $\boldsymbol{X}^*$ leva que $\boldsymbol{X}_1^*$ e $\boldsymbol{X}_2^*$ são independentes. Daí

$$\boldsymbol{X}_1 = \boldsymbol{X}_1^* - \Sigma_{12}\Sigma_{22}^{-1}\boldsymbol{X}_2$$

e que a distribuição (em lei) de $\boldsymbol{X}_1$ dado $\boldsymbol{X}_2 = x_2$ é

$$\mathcal{L}(\boldsymbol{X}_1|\boldsymbol{X}_2 = x_2) = \mathcal{L}(\boldsymbol{X}_1^* - \Sigma_{12}\Sigma_{22}^{-1}\boldsymbol{X}_2|\boldsymbol{X}_2 = x_2) = \mathcal{L}(\boldsymbol{X}_1^* - \Sigma_{12}\Sigma_{22}^{-1}\boldsymbol{x}_2|\boldsymbol{X}_2 = x_2)$$

a qual é normal multivariada de dimensão q.

Teorema do Limite Central Multivariado

Teorema

Se $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots$ é uma sequência vetores aleatórios i.i.d, de dimensão p com $E[\boldsymbol{X}] = \boldsymbol{\mu}$ e $Cov(\boldsymbol{X}) = \boldsymbol{\Sigma}$ (positiva definida) então

$$n^{-1/2} \sum_{j=1}^{n} (\boldsymbol{X}_j - \boldsymbol{\mu}) \xrightarrow{d} N_p(\mathbf{0}, \boldsymbol{\Sigma}), \; n \longrightarrow \infty.$$

ou

$$n^{1/2} \left(\bar{\boldsymbol{X}}_{-}\boldsymbol{\mu}\right) \xrightarrow{d} N_p(\mathbf{0}, \boldsymbol{\Sigma}), \; n \longrightarrow \infty.$$

Corolário

(Método Delta) Sob as condições do TCL, se $h(\cdot)$ é uma função $h : R^p \longrightarrow R$, então

$$n^{1/2} \left(h(\bar{\boldsymbol{X}}) - h(\boldsymbol{\mu})\right) \xrightarrow{d} N_p(\mathbf{0}, (\Delta h(\boldsymbol{\mu}))^\top \boldsymbol{\Sigma} \Delta h(\boldsymbol{\mu})), \; n \longrightarrow \infty.$$

onde $\Delta h(\boldsymbol{x})$ denota a gradiente de h em $\boldsymbol{x}$.

Seja $\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \Sigma)$. Uma forma quadrática em $\boldsymbol{X}$ é uma variável aleatória da forma

$$Y = \boldsymbol{X}^\top \boldsymbol{A} \boldsymbol{X} = \sum_{i=1}^{p} \sum_{i=1}^{p} X_i a_{ij} X_j$$

onde $\boldsymbol{A} = \{a_{ij}\}$ é uma matriz simétrica de ordem $p \times p$ e X_j é o j-ésimo elemento de $\boldsymbol{X}$.

Estamos interessados na distribuição das formas quadráticas e nas condições em que duas formas quadráticas são independentes.

Exemplo

Se $\boldsymbol{X} \sim N_p(\mathbf{0}, \boldsymbol{I}_p)$ *e* $\boldsymbol{A} = \boldsymbol{I}_p$, *Então*

$$Y = \boldsymbol{X}^\top \boldsymbol{A}\boldsymbol{X} = \boldsymbol{X}^\top \boldsymbol{X} = \sum_{i=1}^{p} X_i^2 \sim \chi^2_{(p)}$$

Exemplo

Se $\boldsymbol{X} \sim N_p(\mathbf{0}, \boldsymbol{I}_p)$ *e* $\boldsymbol{A} = \boldsymbol{I}_p$, *Então*

$$Y = \boldsymbol{X}^\top \boldsymbol{A}\boldsymbol{X} = \boldsymbol{X}^\top \boldsymbol{X} = \sum_{i=1}^{p} X_i^2 \sim \chi^2_{(p)}$$

Teorema

Se $\boldsymbol{X} \sim N_p(0, \sigma^2 \boldsymbol{I}_p)$ *e* $\boldsymbol{A}$ *é uma matriz simétrica de ordem* $p \times p$. *Então*

$$\frac{\boldsymbol{X}^\top \boldsymbol{A}\boldsymbol{X}}{\sigma^2} \sim \chi^2_{(m)}$$

se e somente $\boldsymbol{A}$ *idempotente de posto* $m < p$.

Corolário

Se $\boldsymbol{X} \sim N_p(\mathbf{0}, \boldsymbol{\Sigma})$ *e* $\boldsymbol{A}$ *é uma matriz simétrica de ordem* $p \times p$. *Então*

$$\boldsymbol{X}^\top \boldsymbol{A} \boldsymbol{X} \sim \chi^2_{(m)}$$

se e somente $\boldsymbol{A\Sigma}$ *idempotente (ou* $\boldsymbol{\Sigma A}$ *idempotente) de posto* $m < p$.

Exemplo

Se $\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, *então* $(\boldsymbol{X} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{X} - \boldsymbol{\mu}) \sim \chi^2_{(p)}$

Teorema

Seja $\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{A}$ *uma matriz simétrica de posto m. Então*

$$\boldsymbol{X}^\top \boldsymbol{A}\boldsymbol{X} \sim \chi^2_{(m)}(\lambda)$$

se somente se $\boldsymbol{A}\boldsymbol{\Sigma}$ *é idempotente, sendo* $\lambda = \frac{1}{2}\boldsymbol{\mu}^\top \boldsymbol{A}\boldsymbol{\mu}$, *o parâmetro de não centralidade)*

Exemplo

Se $\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, *então* $(\boldsymbol{X} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{X} - \boldsymbol{\mu}) \sim \chi^2_{(p)}(0)$

pois $\lambda = \frac{1}{2}(E(\boldsymbol{X} - \boldsymbol{\mu}))^\top \Sigma^{-1} E(\boldsymbol{X} - \boldsymbol{\mu}) = 0$

Teorema

Seja $\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ $\boldsymbol{A}$ *uma matriz simétrica de ordem* $p \times p$ *e* $\boldsymbol{B}$ *uma matriz de ordem* $k \times p$*. Se* $\boldsymbol{B\Sigma A} = \boldsymbol{0}$ *então* $\boldsymbol{BX}$ *e* $\boldsymbol{X}^\top \boldsymbol{AX}$ *são independentes.*

Exemplo

Seja $X_1, \ldots, X_n$ *é uma amostra aleatória de* $N(\mu, \sigma^2)$ *A média amostral* $\bar{X}_n$ *e a variância amostral* $S_n^2 = (n-1)^{-1}\sum_{i=1}^{2}(X_i - \bar{X}_n)^2$ *são independentes.*

Note que : $\boldsymbol{X} \sim N_n(\mu \mathbf{1}_n, \Sigma = \sigma^2 I_n)$, $\bar{X}_n = (\frac{1}{n}, \ldots, \frac{1}{n}) \begin{pmatrix} X_1 \\ \ldots \\ X_n \end{pmatrix} = \frac{1}{n}\mathbf{1}^\top \boldsymbol{X} =$ $\boldsymbol{BX}$ e $(n-1)^{-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2 = (n-1)^{-1}\boldsymbol{X}^\top(\boldsymbol{I}_n - \frac{1}{n}\mathbf{1}\frac{1}{n}\mathbf{1}^\top)\boldsymbol{X} = \boldsymbol{X}^\top A\boldsymbol{X}$. Daí tem-se $\boldsymbol{B\Sigma A} = \boldsymbol{0}$

Teorema

Seja $\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. *Suponha que* $\boldsymbol{A}$ *e* $\boldsymbol{B}$ *são matrizes simétricas de ordem* $p \times p$. *Se* $\boldsymbol{B\Sigma A} = \boldsymbol{0}$ *então* $\boldsymbol{X}^\top \boldsymbol{BX}$ *e* $\boldsymbol{X}^\top \boldsymbol{AX}$ *são independentes.*

Exercício

Considere o modelo de regressão múltipla

$$X_i = \beta_0 + \sum_{j=1}^{p} z_{ij}\beta_j + \varepsilon_i, \ i = 1, \ldots, n$$

com $\varepsilon_i \overset{i.i.d}{\sim} N(0, \sigma^2)$. *Mostre que a soma de quadrados do residual e soma de quadrados da regressão são independentes.*

Teorema

Se $\boldsymbol{A}$ *é uma matriz de ordem* $p \times p$ *e* $\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. *Então*

- $E[\boldsymbol{X}^\top \boldsymbol{A}\boldsymbol{X}] = tr(\boldsymbol{A}\Sigma) + \boldsymbol{\mu}^\top \boldsymbol{A}\boldsymbol{\mu}$
- $Var[\boldsymbol{X}^\top \boldsymbol{A}\boldsymbol{X}] = tr(\boldsymbol{A}\Sigma\boldsymbol{A}\Sigma) + 4\boldsymbol{\mu}^\top \boldsymbol{A}\Sigma\mu$

Prova Como $\Sigma = E[\boldsymbol{X}\boldsymbol{X}^\top] - \boldsymbol{\mu}\boldsymbol{\mu}^\top$, tem-se $E[\boldsymbol{X}\boldsymbol{X}^\top] = \Sigma + \boldsymbol{\mu}\boldsymbol{\mu}^\top$.

$$\begin{aligned}
E[\boldsymbol{X}^\top \boldsymbol{A}\boldsymbol{X}] &= E[tr(\boldsymbol{X}^\top \boldsymbol{A}\boldsymbol{X})] = E[tr(\boldsymbol{A}\boldsymbol{X}\boldsymbol{X}^\top)] = tr(E[\boldsymbol{A}\boldsymbol{X}\boldsymbol{X}^\top]) \\
&= tr(\boldsymbol{A}E[\boldsymbol{X}\boldsymbol{X}^\top]) = tr(\boldsymbol{A}(\Sigma + \boldsymbol{\mu}\boldsymbol{\mu}^\top)) \\
&= tr(\boldsymbol{A}\Sigma + A\boldsymbol{\mu}\boldsymbol{\mu}^\top) = tr(\boldsymbol{A}\Sigma) + tr(A\boldsymbol{\mu}\boldsymbol{\mu}^\top) \\
&= tr(\boldsymbol{A}\Sigma) + \boldsymbol{\mu}^\top \boldsymbol{A}\boldsymbol{\mu}.
\end{aligned}$$