Definições

Patrícia de Siqueira Ramos

UNIFAL-MG, *campus* Varginha

7 de Abril de 2018

2 Definições

a) Vetor aleatório

- $\mathbf{X}$ é um vetor com p componentes, em que cada componente é uma variável aleatória (v.a.)
- X_i é uma v.a. $\forall i = 1, 2, \ldots, p$
- Cada variável pode ser analisada separadamente (pelo comportamento da sua distribuição de probabilidade), mas é melhor analisar o vetor como um todo (pode ser que haja relacionamentos entre as p variáveis)

$$\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix} = \begin{bmatrix} X_1 & X_2 & \ldots & X_p \end{bmatrix}^T.$$

a) Vetor aleatório

Ex.: Empresa de minério de ferro. Variáveis: X_1: teor de ferro, X_2: teor de umidade, X_3: granulometria.

a) Vetor aleatório

Ex.: Empresa de minério de ferro. Variáveis: X_1: teor de ferro, X_2: teor de umidade, X_3: granulometria.
$\mathbf{X} = \begin{bmatrix} X_1 & X_2 & X_3 \end{bmatrix}^T$ é um vetor aleatório com $p = 3$.

a) Vetor aleatório

Ex.: Empresa de minério de ferro. Variáveis: X_1: teor de ferro, X_2: teor de umidade, X_3: granulometria.
$\mathbf{X} = \begin{bmatrix} X_1 & X_2 & X_3 \end{bmatrix}^T$ é um vetor aleatório com $p = 3$.

Ex.: Pesquisa de mercado, amostra de um novo produto para n pessoas que darão notas de 1 a 7 para cada atributo.
X_1: nota para a cor da embalagem
X_2: nota para o sabor
X_3: nota para preço
X_4: nota para facilidade de compra

a) Vetor aleatório

Ex.: Empresa de minério de ferro. Variáveis: X_1: teor de ferro, X_2: teor de umidade, X_3: granulometria.
$\mathbf{X} = \begin{bmatrix} X_1 & X_2 & X_3 \end{bmatrix}^T$ é um vetor aleatório com $p = 3$.

Ex.: Pesquisa de mercado, amostra de um novo produto para n pessoas que darão notas de 1 a 7 para cada atributo.
X_1: nota para a cor da embalagem
X_2: nota para o sabor
X_3: nota para preço
X_4: nota para facilidade de compra
$\mathbf{X} = \begin{bmatrix} X_1 & X_2 & X_3 & X_4 \end{bmatrix}^T$ é um vetor aleatório com $p = 4$.

b) Vetor de médias

b) Vetor de médias

- O vetor $\mu = E(X)$ é chamado vetor de médias do vetor **X**.

$$\boldsymbol{\mu} = E(\mathbf{X}) = \begin{bmatrix} E(X_1) \\ E(X_2) \\ \vdots \\ E(X_p) \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix}$$

- $\boldsymbol{\mu}_i = E(\mathbf{X}_i)$ é a esperança da variável X_i
- A média $\boldsymbol{\mu}_i$ é uma das medidas mais utilizadas para sintetizar a informação de tendência central da distribuição de valores da variável X_i

c) Matriz de covariâncias

c) Matriz de covariâncias

- A matriz de covariâncias do vetor aleatório **X** é uma matriz simétrica dada por

$$\mathbf{\Sigma}_{p\times p} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{bmatrix},$$

- em que a diagonal é composta pelas variâncias e os outros elementos são as covariâncias entre as variáveis.

Variância

- A variância do i-ésimo componente de $\mathbf{X}$ é

$$V(X_i) = E[(X_i - \mu_i)^2] = \sigma_i^2 = \sigma_{ii}$$

- O desvio padrão é σ_i ou $\sqrt{\sigma_{ii}}$
- O desvio padrão informa a dispersão dos valores da v.a. X_i em relação a μ_i (estão próximos ou distantes de μ_i)
- Grandes valores de σ_i indicam maior dispersão de valores em relação à média da distribuição

Covariância

A covariância entre os valores da i-ésima e j-ésima variáveis do vetor aleatório $\mathbf{X}$ é

$$Cov(X_i, X_j) = E[(X_i - \mu_i)(X_j - \mu_j)] = \sigma_{ij}.$$

Covariância

- Quando $i = j$, $Cov(X_i, X_i) = V(X_i)$
- A covariância mede o grau de relacionamento linear entre duas variáveis

Covariância

- Quando $i = j$, $Cov(X_i, X_i) = V(X_i)$
- A covariância mede o grau de relacionamento linear entre duas variáveis
- Se valores de X_i acima da média μ_i estiverem associados a valores X_j acima da média μ_j, σ_{ij} tende a ser positiva (e o mesmo para valores abaixo da média para X_i e X_j)
- Se valores de X_i acima da média μ_i estiverem associados a valores X_j abaixo da média μ_j, ou vice-versa, σ_{ij} tende a ser negativa

Covariância

- Quando $i = j$, $Cov(X_i, X_i) = V(X_i)$
- A covariância mede o grau de relacionamento linear entre duas variáveis
- Se valores de X_i acima da média μ_i estiverem associados a valores X_j acima da média μ_j, σ_{ij} tende a ser positiva (e o mesmo para valores abaixo da média para X_i e X_j)
- Se valores de X_i acima da média μ_i estiverem associados a valores X_j abaixo da média μ_j, ou vice-versa, σ_{ij} tende a ser negativa
- Covariância tem informação sobre o relacionamento linear entre duas variáveis, mas é difícil julgar se a relação é forte ou não apenas pelo seu valor (correlação é mais útil para isso)

Covariância

- Em formato matricial

$$Cov(\mathbf{X}) = E(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})^T =$$

$$= \begin{bmatrix} E(X_1-\mu_1)^2 & \dots & E(X_1-\mu_1)(X_p-\mu_p) \\ E(X_2-\mu_2)(X_1-\mu_1) & \dots & E(X_2-\mu_2)(X_p-\mu_p) \\ \vdots & \vdots & \vdots \\ E(X_p-\mu_p)(X_1-\mu_1) & \dots & E(X_p-\mu_p)^2 \end{bmatrix}$$

Covariância

Forma equivalente da covariância:

$$Cov(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j) = E(X_i X_j) - \mu_i \mu_j.$$

Covariância

Forma equivalente da covariância:

$$Cov(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j) = E(X_i X_j) - \mu_i \mu_j.$$

Matricialmente

$$Cov(\mathbf{X}) = E(\mathbf{X}\mathbf{X}^T) - \boldsymbol{\mu}\boldsymbol{\mu}^T$$

Covariância

Se X_i e X_j são independentes, $Cov(X_i, X_j) = 0$ (mas o inverso não é verdadeiro):

- X_i e X_j independentes, $E(X_i \cdot X_j) = E(X_i)E(X_j)$
- Como $Cov(X_i, X_j) = E[(X_i - \mu_i)(X_j - \mu_j)]$,

$$\begin{aligned} Cov(X_i, X_j) &= E(X_i X_j - X_i \mu_j - X_j \mu_i + \mu_i \mu_j) \\ &= \ldots \\ &= E(X_i)E(X_j) - \mu_i \mu_j = 0 \end{aligned}$$

d) Matriz de correlações

Correlação

O coeficiente de correlação entre as i-ésima e j-ésima variáveis do vetor $\mathbf{X}$ é definido por

$$\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}}.$$

- Medida de covariação entre as variáveis em uma escala padronizada

Correlação

O coeficiente de correlação entre as i-ésima e j-ésima variáveis do vetor $\mathbf{X}$ é definido por

$$\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}}.$$

- Medida de covariação entre as variáveis em uma escala padronizada
- $-1 \leq \rho_{ij} \leq 1, \qquad i,j = 1,2,\ldots,p.$ (a covariância possui domínio $]-\infty, +\infty[$)
- Na escala padronizada, um valor muito próximo de 1 ou -1 indica que as variáveis estão fortemente associadas
- Quando $i = j$, $\rho_{ij} = \rho_{ii} = 1$
- Se $\rho_{ij} = 0$, as variáveis não possuem associação linear

Matriz de correlações

A matriz de correlações do vetor aleatório $\mathbf{X}$ é uma matriz simétrica dada por

$$\boldsymbol{\rho}_{p\times p} = \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{21} & 1 & \cdots & \rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p1} & \rho_{p2} & \cdots & 1 \end{bmatrix}$$

* Exemplo

Matriz de correlações

Matricialmente, se tomarmos $\mathbf{V} = diag(\mathbf{\Sigma}) = diag(\sigma_{ii})$ e definirmos $\mathbf{V}^{-1/2} = diag(1/\sqrt{\sigma_{ii}})$, teremos

$$\boldsymbol{\rho} = \mathbf{V}^{-1/2}\mathbf{\Sigma}\mathbf{V}^{-1/2} = \begin{bmatrix} 1 & \rho_{12} & \dots & \rho_{1p} \\ \rho_{21} & 1 & \dots & \rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p1} & \rho_{p2} & \dots & 1 \end{bmatrix}$$

* Exemplo

e) Outras variâncias

e) Outras variâncias

- A matriz de covariâncias $\mathbf{\Sigma}(p \times p)$ de um vetor aleatório $\mathbf{X}$ possui $p(p+1)/2$ parâmetros referentes a
  - p variâncias e
  - $p(p-1)/2$ covariâncias

e) Outras variâncias

- A matriz de covariâncias $\boldsymbol{\Sigma}(p \times p)$ de um vetor aleatório $\mathbf{X}$ possui $p(p+1)/2$ parâmetros referentes a
  - p variâncias e
  - $p(p-1)/2$ covariâncias
- Se p for grande, o número de parâmetros pode ser muito grande
- É interessante usar medidas resumo tais como determinante, traço ou o conjunto de autovalores de $\boldsymbol{\Sigma}$

Variância total

A variância total do vetor aleatório **X** é definida como

$$\text{Variância total} = \text{traço}(\boldsymbol{\Sigma}) = tr(\boldsymbol{\Sigma}) = \sigma_{11} + \sigma_{22} + \cdots + \sigma_{pp}.$$

Variância total

A variância total do vetor aleatório **X** é definida como

$$\text{Variância total} = \text{traço}(\boldsymbol{\Sigma}) = tr(\boldsymbol{\Sigma}) = \sigma_{11} + \sigma_{22} + \cdots + \sigma_{pp}.$$

- Forma de sintetizar a variância global da distribuição multivariada por ser a soma das variâncias de todas as variáveis do vetor **X**
- Altos valores de variâncias totais indicam maior dispersão global das variáveis X_i, $i = 1, 2, \ldots, p$.

Obs.: $\mathrm{tr}(\boldsymbol{\Sigma}) = \mathrm{tr}(\boldsymbol{\Lambda}) = \lambda_1 + \cdots + \lambda_p$ (soma dos autovalores).

Variância generalizada

A variância generalizada do vetor aleatório $\mathbf{X}$ é definida como o determinante da matriz $\mathbf{\Sigma}$, ou seja,

$$\text{Variância generalizada} = |\mathbf{\Sigma}|$$

$$\text{Desvio padrão generalizado} = |\mathbf{\Sigma}|^{1/2}$$

- Também fornece uma noção da dispersão global da distribuição multivariada
- Altos valores de variâncias generalizadas também indicam maior dispersão global das variáveis
- Mas, ao contrário da variância total, a variância generalizada é influenciada pelas covariâncias (ou correlações) entre as variâncias (ver Ferreira(2008) para explicação geométrica detalhada)

* Exemplo

Variância generalizada

A variância generalizada do vetor aleatório **X** é definida como o determinante da matriz **Σ**, ou seja,

$$\text{Variância generalizada} = |\boldsymbol{\Sigma}|$$

$$\text{Desvio padrão generalizado} = |\boldsymbol{\Sigma}|^{1/2}$$

- Também fornece uma noção da dispersão global da distribuição multivariada
- Altos valores de variâncias generalizadas também indicam maior dispersão global das variáveis
- Mas, ao contrário da variância total, a variância generalizada é influenciada pelas covariâncias (ou correlações) entre as variâncias (ver Ferreira(2008) para explicação geométrica detalhada)

* Exemplo

f) Combinações lineares

Combinações lineares

- Seja $\mathbf{X}$ um vetor aleatório, composto por p variáveis, com vetor de médias $\boldsymbol{\mu}$ e matriz de covariâncias $\boldsymbol{\Sigma}$.
- Seja $\mathbf{a}_{p\times 1}$ um vetor de escalares (conhecidos), isto é,

$$\mathbf{a} = \begin{bmatrix} a_1 & a_2 & \dots & a_p \end{bmatrix}^T$$

Combinações lineares

- Seja $\mathbf{X}$ um vetor aleatório, composto por p variáveis, com vetor de médias $\boldsymbol{\mu}$ e matriz de covariâncias $\boldsymbol{\Sigma}$.
- Seja $\mathbf{a}_{p\times 1}$ um vetor de escalares (conhecidos), isto é,

$$\mathbf{a} = \begin{bmatrix} a_1 & a_2 & \dots & a_p \end{bmatrix}^T$$

- Seja Z a variável definida por

$$Z = \mathbf{a}^T\mathbf{X} = a_1X_1 + a_2X_2 + \dots a_pX_p.$$

- Então, Z é uma v.a. e uma combinação linear das variáveis pertencentes a $\mathbf{X}$.

Combinações lineares

- A média (esperança matemática) de Z é dada por

$$\mu_Z = E(Z) = a_1\mu_1 + a_2\mu_2 + \cdots + a_p\mu_p,$$

em que $\mu_i = E(X_i)$.

- A variância de Z é dada por

$$\sigma_Z^2 = a^T \boldsymbol{\Sigma} a = \sum_{i=1}^{p} a_i^2\sigma_i^2 + \sum_{i \neq j} a_i a_j \sigma_{ij}.$$