

MAE 5776

# ANÁLISE MULTIVARIADA

Júlia M Pavan Soler  
[pavan@ime.usp.br](mailto:pavan@ime.usp.br)

1º Semestre/2022

# Dados Multivariados

$$Y_{n \times p} = \begin{pmatrix} Y_{11} & Y_{12} & \dots & Y_{1p} \\ Y_{21} & Y_{22} & \dots & Y_{2p} \\ \dots & \dots & \dots & \dots \\ Y_{n1} & Y_{n2} & \dots & Y_{np} \end{pmatrix}$$

Y: Matriz de Dados Multivariados

$$S_{p \times p} = \begin{pmatrix} s_{11} & & & \\ s_{21} & s_{22} & \sim & \\ \dots & \dots & \dots & \\ s_{p1} & s_{p2} & \dots & s_{pp} \end{pmatrix}$$

S: Matriz de Covariâncias entre as p variáveis

$$D_{n \times n} = \begin{pmatrix} 0 & & & \\ d_{21} & 0 & \sim & \\ \dots & \dots & \dots & \\ d_{n1} & d_{n2} & \dots & 0 \end{pmatrix}$$

D: Matriz de Distâncias entre os n indivíduos

Já vimos ☺

# MAE5776

$$Y_{n \times p} = (Y_{ij}) \in \mathbb{R}^{n \times p}$$

- Estatísticas Descritivas Multivariadas:  $Y (\mathbb{R}^{n \times p}) \Rightarrow S (\mathbb{R}^{p \times p})$   
 $D (\mathbb{R}^{n \times n})$
- Inferência sobre  $\mu \in \mathbb{R}^p$ :

Caso de Uma Única População

Caso de Duas Populações (Pareadas Independentes)

Testes  $T^2$

Caso de Duas ou Mais Populações (MANOVA):

- DCA com Um Único Fator:  $T = H + E$
- DCA Fatorial Cruzado:  $T = H_1 + H_2 + H_1:2 + E$
- DCA Fatorial Hierárquico:  $T = H_1 + H_2(1) + E$
- Delineamento Aleatorizado em Blocos:  $T = H_1 + B + E$

Tabelas MANOVA:

**Decomposição da SQPC ( $\mathbb{R}^{p \times p}$ )**

Note que:  $S_{p \times p} = \frac{T}{n}$ ;  $S_{c \times p \times p} = \frac{E}{n - G}$

$$T = S_T; \quad H = S_B; \quad E = S_W$$

notação

# Análises em Espaços Duais

## MANOVA ( $\mathfrak{R}^{p \times p}$ ), ASCA ( $\mathfrak{R}^{n \times p}$ )

- DCA com Um Único Fator:  $Y_{ig \ p \times 1} = \bar{Y} + (\bar{Y}_g - \bar{Y}) + (Y_{ig} - \bar{Y}_g)$

$$S_{T \ p \times p} = S_{B \ p \times p} + S_{W \ p \times p} \Leftrightarrow Y_{n \times p} = Y_{M \ n \times p} + Y_{H \ n \times p} + Y_{E \ n \times p}$$

- DCA Fatorial Cruzado:  $Y_{ijk} = \bar{Y} + (\bar{Y}_{j.} - \bar{Y}) + (\bar{Y}_{.k} - \bar{Y}) + (\bar{Y}_{jk} - \bar{Y}_{j.} - \bar{Y}_{.k} + \bar{Y}) + (Y_{ijk} - \bar{Y}_{jk})$

$$S_{T \ p \times p} = S_{H1} + S_{H2} + S_{H12} + S_W \Leftrightarrow Y_{n \times p} = Y_M + Y_{H1} + Y_{H2} + Y_{H12} + Y_E$$

- DCA Fatorial Hierárquico:  $Y_{ijk \ p \times 1} = \bar{Y} + (\bar{Y}_{j.} - \bar{Y}) + (\bar{Y}_{jk} - \bar{Y}_{j.}) + (Y_{ijk} - \bar{Y}_{jk})$

$$S_{T \ p \times p} = S_{H1} + S_{H2(l)} + S_W \Leftrightarrow Y_{n \times p} = Y_M + Y_{H1} + Y_{H2(l)} + Y_E$$

- Delineamento Aleatorizado em Blocos:  $Y_{jk} = \bar{Y} + (\bar{Y}_{j.} - \bar{Y}) + (\bar{Y}_{.k} - \bar{Y}) + (Y_{jk} - \bar{Y}_{j.} - \bar{Y}_{.k} + \bar{Y})$

$$S_{T \ p \times p} = S_{H1} + S_{Bloco} + S_W \Leftrightarrow Y_{n \times p} = Y_M + Y_{H1} + Y_B + Y_E$$

# Análises em Espaços Duais

## Estatística de Mantel ( $\mathbb{R}^{n \times n}$ )

$$Y_{n \times p} = (Y_{ij}) \in \mathbb{R}^{n \times p};$$

$$Y_{n \times p} = X_{n \times q} \beta_{q \times p} + E_{n \times p}$$

Matriz de  
Planejamento  
ou Regressores

Matriz de  
Parâmetros

Matriz de Erros  
(aleatório)

Modelo Linear da  
MANOVA (ou Regressão  
Multivariada)

$Y_1$	$Y_2$	...	$Y_p$		$X_1$	$X_2$	...	$X_q$
$Y_{n \times p}$					$X_{n \times q}$			

Matriz X: definida de acordo com o estudo (parametrização de médias, casela de referência, regressores)

$$\hat{Y}_{n \times p} = X \hat{\beta} = X (X'X)^{-1} X' Y = PY$$

$$\hat{e}_{n \times p} = Y - \hat{Y} = (I_n - P)Y$$

# MANOVA: DCA com Um Único Fator

$$H_0 : \mu_g = \mu_{p \times p}, \quad g = 1, \dots, q$$

Fonte de Variação	Número de graus de liberdade	Soma de Quadrados e Produtos Cruzados (SQPC)
Tratamento	$q-1$	$H = S_B = Y' \left[ P - \frac{1}{n} 1_{n \times n} \right] Y$
Resíduo	$n-q$	$E = S_W = Y' [I_n - P] Y$
TOTAL	$n-1$	$T = S_T = Y' \left[ I_n - \frac{1}{n} 1_{n \times n} \right] Y$

Caso assintótico: 
$$-\left( N - 1 - \frac{p + g}{2} \right) \ln \left( \frac{|E|}{|H + E|} \right) \stackrel[n \rightarrow \infty]{(n-p) \rightarrow \infty} \chi^2_{p(g-1)}(\alpha)$$

# Análises em Espaços Duais

## Estatística de Mantel ( $\mathfrak{R}^{n \times n}$ )

$$Y_{n \times p} = (Y_{ij}) \in \mathfrak{R}^{n \times p};$$

$$Y_{n \times p} = X_{n \times q} \beta_{q \times p} + E_{n \times p}$$

Matriz de  
Planejamento  
ou Regressores

Matriz de  
Parâmetros

Matriz de Erros  
(aleatório)

Modelo Linear da  
MANOVA (ou Regressão  
Multivariada)

$$\hat{Y}_{n \times p} = X \hat{\beta} = X (X'X)^{-1} X' Y = PY$$

Matriz de Gower

$$D_{n \times n} = (d_{ik}) \rightarrow A_{n \times n} = \left( -\frac{1}{2} d_{ik}^2 \right) \rightarrow G_{n \times n} = \left( I_n - \frac{1}{n} 1_{n \times n} \right) A \left( I_n - \frac{1}{n} 1_{n \times n} \right)$$

$H_0$ :  $\notin$  efeito das variáveis preditoras (na distância entre indivíduos)

$$F = \frac{tr(PGP) / (q-1)}{tr[(I_n - P)G(I_n - P)] / (n-q)}$$

Distribuição obtida por  
Permutação das linhas e  
colunas de G