

MAE 317

Planejamento e Pesquisa II

Profa. Júlia Maria Pavan Soler
pavan@ime.usp.br

IME/USP – 2º Semestre/2020

Modelos de ANOVA

Temos considerado modelos ANOVA para os seguintes delineamentos:

- **Estrutura de Tratamentos** (Fatores sob Estudo):
 - Um único Fator em J níveis
 - Fatorial Cruzado
 - Fatorial Cruzado 2^K (sem réplicas, com K elevado)
- **Estrutura das Unidades Experimentais:**
 - Delineamento Completamente Aleatorizado (DCA)
 - Delineamento Aleatorizado em Blocos Completos (DABC)

- 
- Para os esquemas de aleatorização DCA ou DABC, vamos agora considerar:
 - Estrutura de tratamentos: **Fatorial Hierárquico**
(fatores aninhados, embutidos, “*nested*”, supondo fatores FIXOS)

Delineamento Fatorial Hierárquico

Motivação: Há interesse em verificar se o desempenho dos alunos depende do tipo de Escola (A1, A2 e A3) e do Método de Ensino utilizado por elas (B1 a B6).

A1		A2		A3	
B1	B2	B3	B4	B5	B6
20	19	14	12	13	9
18	20	18	12	16	4
14	20	14	9	13	4

Os níveis do fator Método de Ensino (B1, B2, B3, B4, B5 e B6) está aninhado (embutido) dentro do fator Escola (A1, A2 e A3).

(Note que não é Fatorial cruzado!)

Estrutura do Experimento e Modelos ANOVA

Já vimos!

DCA com esquema de Um único fator em J níveis

$$y_{ij} = \mu + \tau_j + \varepsilon_{ij}; \quad \sum_{j=1}^J \tau_j = 0, \quad \varepsilon_{ij} \sim N(0, \sigma^2) \quad \text{iid} \quad \Rightarrow y_{ij} = \bar{y} + (\bar{y}_j - \bar{y}) + (y_{ij} - \bar{y}_j) \quad \text{útil!}$$

Identidade

Modelo estrutural e distribucional

$$\text{SQTotal} = \text{SQTrat} + \text{SQRes}$$

Esqueleto da ANOVA

DCA Fatorial Cruzado a x b

$$y_{ijk} = \mu + \tau_j + \beta_k + \gamma_{jk} + \varepsilon_{ijk}; \quad \sum_{j=1}^a \tau_j = \sum_{k=1}^b \beta_k = \sum_{j=1}^a \gamma_{jk} = \sum_{k=1}^b \gamma_{jk} = 0, \quad \varepsilon_{ijk} \sim N(0, \sigma^2) \quad \text{iid}$$

$$\Rightarrow y_{ijk} = \bar{y} + (\bar{y}_j - \bar{y}) + (\bar{y}_k - \bar{y}) + (\bar{y}_{jk} - \bar{y}_j - \bar{y}_k + \bar{y}) + (y_{ijk} - \bar{y}_{jk})$$

$$\text{SQTotal} = \text{SQ(A)} + \text{SQ(B)} + \text{SQ(A*B)} + \text{SQRes}$$

DABC com esquema de Um fator em J níveis

$$y_{ij} = \mu + \tau_j + \beta_i + \varepsilon_{ij}; \quad \sum_{j=1}^J \tau_j = \sum_{i=1}^b \beta_i = 0, \quad \varepsilon_{ij} \sim N(0, \sigma^2) \quad \text{iid}$$

$$\Rightarrow y_{ij} = \bar{y} + (\bar{y}_{.j} - \bar{y}) + (\bar{y}_i - \bar{y}) + (y_{ij} - \bar{y}_{.j} - \bar{y}_i + \bar{y})$$

$$\text{SQTotal} = \text{SQ(A)} + \text{SQ(B)} + \text{SQRes}$$

Modelo aditivo
Resíduo é a
interação.

Delineamento Fatorial Hierárquico

Identidades úteis para a construção das Somas de Quadrados:

Modelo de um único fator: $y_{ij} = \bar{y} + (\bar{y}_j - \bar{y}) + (y_{ij} - \bar{y}_j)$

Fatorial Cruzado: $y_{ijk} = \bar{y} + (\bar{y}_{j.} - \bar{y}) + (\bar{y}_{.k} - \bar{y}) + (\bar{y}_{jk} - \bar{y}_{j.} - \bar{y}_{.k} + \bar{y}) + (y_{ijk} - \bar{y}_{jk})$

O efeito de B dentro de A é a soma do efeito principal de B e do efeito de interação A*B

Fatorial Hierárquico: $y_{ijk} = \bar{y} + (\bar{y}_{j.} - \bar{y}) + (\bar{y}_{jk} - \bar{y}_{j.}) + (y_{ijk} - \bar{y}_{jk})$

$$\hat{\mu} \quad \hat{\tau}_j \quad \uparrow \hat{\beta}_{k(j)} \quad \hat{\epsilon}_{ij}$$

Efeito de B dentro de A: B(A)

Delineamento Fatorial Hierárquico

$$y_{ijk} = \mu + \tau_j + \beta_{k(j)} + e_{ijk}; \quad \sum_{j=1}^a \tau_j = \sum_{k=1}^b \beta_k = 0; \quad e_{ijk} \stackrel{iid}{\sim} N(0; \sigma^2)$$

Tabela de ANOVA:

F.V.	g. l.	SQ	QM	F
A	a-1	$\sum_{j=1}^a br(\bar{y}_j - \bar{y})^2$	SQ(A)/(a-1)	QM(A)/QMRes
B(A)	b(a-1)	$\sum_{k=1}^b \sum_{j=1}^a r(\bar{y}_{jk} - \bar{y}_j)^2$	SQ(B(A))/a(b-1)	QM(B(A))/QMRes
RESÍDUO	ab(r-1)	$\sum_{ijk} (y_{ijk} - \bar{y}_{jk})^2$	SQRes/ab(r-1)	
TOTAL	abr-1	$\sum_{ijk} (y_{ijk} - \bar{y})^2$		

Testes de interesse: $H_{01} : \tau_j = 0, \quad j = 1, \dots, a$

$H_{02} : \beta_{k(j)} = 0, \quad k = 1, \dots, b$

Delineamento Fatorial Hierárquico


Dados: Desempenho de alunos de acordo com o tipo de Escola (A1, A2 e A3) e do Método de Ensino utilizado por elas.

	A1		A2		A3	
	B1	B2	B3	B4	B5	B6
	20	19	14	12	13	9
	18	20	18	12	16	4
	14	20	14	9	13	4

Médias

B(A)	17.33	19.67	15.33	11.0	14.0	5.67
A	18.5		13.17		9.83	
Geral	13.83					

Estimativas das médias de B estão confundidas com as de B(A)



Delineamento Fatorial Hierárquico

Dados: Desempenho de alunos de acordo com o tipo de Escola (A1, A2 e A3) e do Método de Ensino utilizado por elas.

Tabela de ANOVA:

	Df	Sum Sq	Mean Sq	F value	Pr (>F)
A	2	229.333	114.667	23.4545	7.145e-05 H_{01}
B(A)	3	140.500	46.833	9.5795	0.001656 H_{02}
Residuals	12	58.667	4.889		

$$H_{01} : \tau_j = 0, \quad j = 1, \dots, a$$

$$H_{02j} : \beta_{k(j)} = 0; \quad k = 1, \dots, b_j, \quad j = 1, \dots, a$$

Conclusão?

Delineamento Fatorial Hierárquico

Modelo de ANOVA: Fatorial Hierárquico

$$y_{ijk} = \mu + \tau_j + \beta_{k(j)} + \varepsilon_{ijk}; \quad \sum_{j=1}^a \tau_j = \sum_{k=1}^b \beta_{k(j)} = 0, \quad \varepsilon_{ijk} \stackrel{iid}{\sim} N(0, \sigma^2)$$

$$y_{ijk} = \underbrace{\bar{y}}_{\hat{\mu}} + \underbrace{(\bar{y}_{.j} - \bar{y})}_{\hat{\tau}_j} + \underbrace{(\bar{y}_{jk} - \bar{y}_{.j})}_{\hat{\beta}_{k(j)}} + \underbrace{(y_{ijk} - \bar{y}_{jk})}_{\hat{e}_{ij}}$$

Parametrização de desvios em relação à média geral. Escreva o modelo na forma matricial.

Formule e ajuste o modelo via a parametrização casela de referência (adotada no R)

	A1		A2		A3	
	M1	M2	M3	M4	M5	M6
$\hat{\beta}_{k(j)}$	-1.17	1.17	2.17	-2.17	4.17	-4.17
$\hat{\tau}_j$	4.67		-0.67		-4.00	
$\hat{\mu}$			18.83			

Delineamento Fatorial Hierárquico

Dados: Desempenho de alunos de acordo com tipo de Escola e Método de Ensino.

Comparações (Tukey_HSD a 95%) entre Médias de desempenho escolar de alunos

Efeito Principal de Escola (A): (Tukey_HSD a 95%)
 Correção para as 3 comparações

Tukey multiple comparisons of means 95% family-wise confidence level

	diff	lwr	upr	p adj
E2-E1	-5.333333	-8.739047	-1.92762003	0.0033897
E3-E1	-8.666667	-12.072380	-5.26095336	0.0000533
E3-E2	-3.333333	-6.739047	0.07237997	0.0552280

Conclusão?

$$\mu_1 > (\mu_2 = \mu_3)$$

Correção para as 3 comparações

Efeito Principal de Métodos de Ensino dentro de Escola - B(A): (Tukey_HSD a 95%)

	dif	li	ls	sig
(M2-M1) E1	2.333333	-1.600169	6.266835	ns
(M4-M3) E2	-4.333333	-8.266835	-0.3998313	*
(M6-M5) E3	-8.333333	-12.26684	-4.399831	*

Conclusão?

$$\mu_{1(1)} = \mu_{2(1)}$$

$$\mu_{3(2)} = \mu_{4(2)}$$

$$\mu_{5(3)} = \mu_{6(3)}$$

Delineamento Fatorial Hierárquico

Um estudo investiga a eficiência de tratamentos a uma doença. De acordo com a gravidade da doença, tratamentos específicos são prescritos. Para o estudo, três níveis de gravidade da doença foram considerados e 20 pacientes, com a doença em cada um desses níveis, foram selecionados de um Centro de Atendimento à Saúde e aleatorizados a tipos específicos de tratamento da doença. Os dados de uma resposta de cura são apresentados a seguir.

Gravidade da Doença	1				2				3			
	1	2	3	4	5	6	7	8	9	10	11	12
Pacientes	65	68	56	45	74	69	52	73	69	63	81	67
	58	62	65	56	81	76	56	78	83	70	72	79
	63	75	58	54	76	80	62	83	74	72	73	73
	57	64	70	48	80	78	58	75	78	68	76	77
	66	70	64	60	68	73	51	76	80	75	70	71

Descreva a estrutura do experimento em termos dos fatores sob estudo e de sua atribuição (aleatória ou não) às unidades experimentais. Proponha um modelo de ANOVA para análise destes dados? Há efeito da Gravidade da Doença na resposta de cura? Há diferença entre os tratamentos na resposta de cura de acordo com a gravidade da doença?

Como construir Testes de Aleatorização neste caso?