Modelos multivariados (marginais): Parte 2

Prof. Caio Azevedo

Exemplo 2: Distância do centro da glândula pituitária para a fissura pterigomaxilar (Potthoff and Roy (1964))

- Este conjunto de dados corresponde aos famosos dados de Potthoff-Roy, usado para demonstrar a utilização da MANOVA em dados de medidas repetidas (comparação entre grupos, embora comparação entre variáveis seja possível).
- O estudo considerou 16 meninos e 11 meninas, nos quais, nas idades 8, 10, 12 e 14 anos tiveram a distância (mm) do centro da glândula pituitária para a fissura pterigomaxilar medidas.

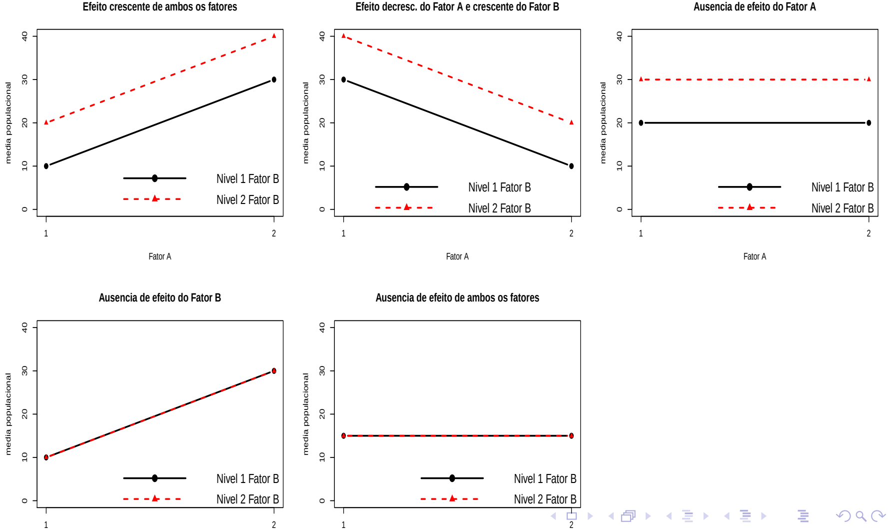
Um pouco sobre interação

- Em muitas situações, o pesquisador tem interesse em como dois ou mais fatores afetam o comportamento da variável resposta.
- Nem todos os fatores são, necessariamente, de interesse. Contudo, em princípio, todos devem ser controlados de alguma forma.
- Em nosso exemplo temos um fator intra-unidades (ano) e um entre-unidades (gênero).
- Pode haver fatores que funcionam como "bloco" (como na literatura de planejamento de experimentos).

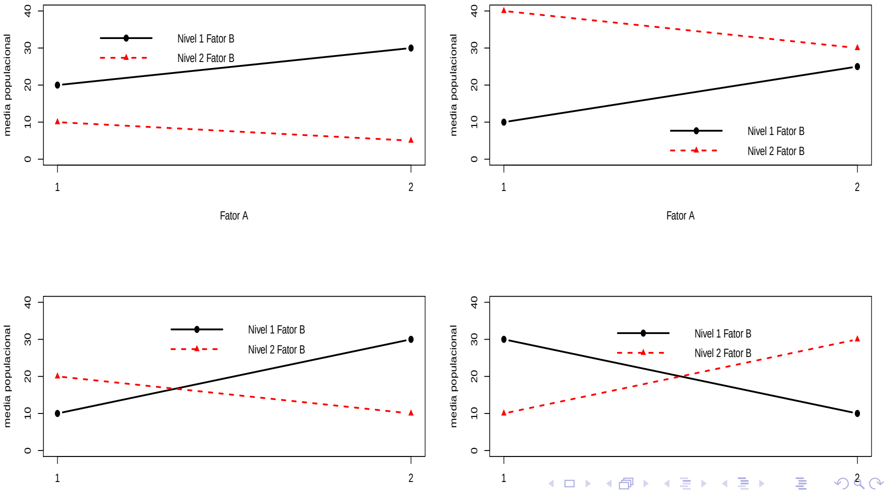
Exemplo

- Fator A: possui a níveis.
- Fator B: possui b níveis.
- Um deles pode ser quantitativo (não fator).
- Conceito importante: interação entre os fatores.
- Interação: a diferença entre as médias da resposta, entre dois níveis do Fator A, são iguais ao longo dos níveis do Fator B (vice-versa).
- Se uma das covariáveis não for um fator (as curvas em relação aos níveis do fator têm de ser paralelas).

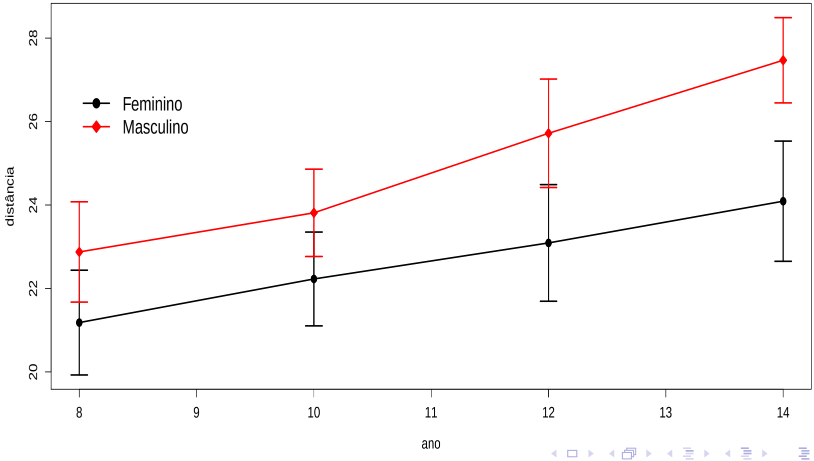
Perfis médios: ausência de interação



Perfis médios: presença de interação



Perfil médio

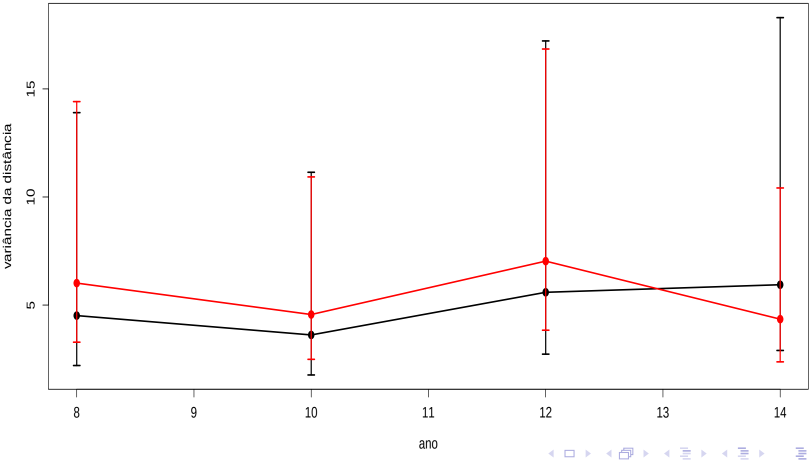


Var. (diagonal), correlações (acima) e covar. (abaixo)

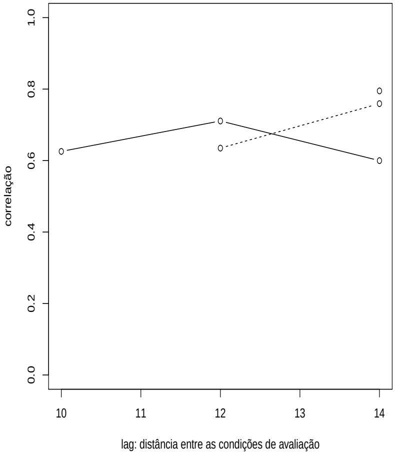
Ano (feminino)			
8	10	12	14
4,51	0,83	0,86	0,84
3,35	3,62	0,90	0,88
4,33	4,03	5,59	0,95
4,36	4,08	5,47	5,94

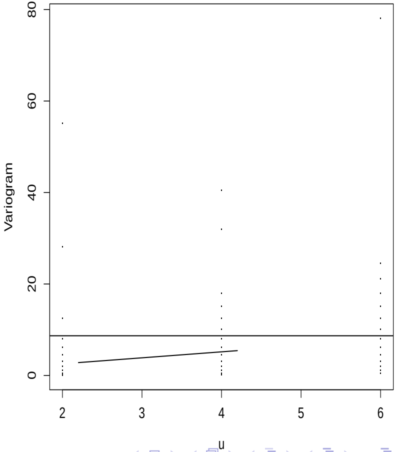
Ano (masculino)			
8	10	12	14
6,02	0,44	0,56	0,32
2,29	4,56	0,39	0,63
3,63	2,19	7,03	0,59
1,61	2,81	3,24	4,35

Variâncias em cada condição



Perfis da matriz de correlação e variograma





Modelagem para os dados do Exemplo 2

$$Y_{ijk} = \beta_0 + \alpha_k + (\beta_1 + \gamma_k)(x_{ijk} - 8) + \xi_{ijk},$$

$j = 1, 2, ..., n_{ik}$, (indivíduo), $i = 1, 2, 3, 4$ (ano (condição de avaliação)),

$k = 1, 2$(gênero - 1: feminino, 2:masculino), $n_{i1} = 11; n_i 2 = 16, \forall i$

- $\alpha_1 = \gamma_1 = 0$.
- x_{ijk} : é o ano (8,10,12,14), em que a distância, correspondente ao instante i, foi medida no indivíduo j do grupo k.
- Y_{ijk} : é a distância no instante i do indivíduo j do grupo k.
- $E(Y_{ijk}|x_{ijk} = 8) = \beta_0$ é a distância esperada no oitavo ano de vida para indivíduos do gênero feminino.
- α_2 é o incremento na distância esperada no oitavo ano de vida para indivíduos do gênero masculino em relação aos do gênero feminino.

Modelagem para os dados do Exemplo 2 (cont.)

- β_1 : é o incremento na distância esperada no intervalo de um ano para indivíduos do gênero feminino.
- γ_2 : é o incremento na distância esperada no intervalo de um ano para indivíduos do gênero masculino em relação ao incremento para indivíduos do gênero feminino.
- (1) : $\mathcal{V}(Y_{ijk}) = \sigma^2$ (homocedástico); (2)$\mathcal{V}(Y_{ijk}) = \sigma_i^2 = \sigma^2\delta_k^2$, $\delta_1 \equiv 1$; (3)$\mathcal{V}(Y_{ijk}) = \sigma_i^2 = \sigma^2\delta_k^2 \exp(x_{ijk}\gamma)$, $\delta_1 \equiv 1$ (heterocedástico).
- $Corre(Y_{ijk}, Y_{i'jk})$ (1) Uniforme, (2) AR(1), (3) (ARMA(1,1)).

Modelos

Modelo	Variância	Correlação
HU	Homocedástico	U
HAR1	Homocedástico	AR(1)
HARMA11	Homocedástico	ARMA(1,1)
HE2U	Heterocedástico (2)	U
HE2AR1	Heterocedástico (2)	AR(1)
HE2ARMA11	Heterocedástico (2)	ARMA(1,1)
HE3U	Heterocedástico (3)	U
HE3AR1	Heterocedástico (3)	AR(1)
HE3ARMA11	Heterocedástico (3)	ARMA(1,1)

Modelos

Modelo	AIC	BIC
HU	445,76	461,62
HAR1	456,59	472,45
HARMA11	447,39	465,90
HE2U	**436,19**	**454,70**
HE2AR1	447,00	465,52
HE2ARMA11	437,41	458,57
HE3U	438,02	459,17
HE3AR1	448,95	470,11
HE3ARMA11	439,33	463,13

Estimativas dos parâmetros (modelo completo)

Parâmetro	Estimativa	EP	IC(95%)	Estatística	p-valor
β_0	21,21	0,51	[20,21 ; 22,21]	41,90	$<$ 0,0001
α_2	1,41	0,88	[-0,34 ; 3,15]	1,60	0,1104
β_1	0,48	0,06	[0,36 ; 0,60]	7,89	$<$ 0,0001
γ_2	0,30	0,11	[0,09 ; 0,51]	2,88	0,0040

Ajustar um modelo reduzido (U com heterocedasticidade (2)) sem o parâmetro α_2, ou seja:

$Y_{ijk} = \beta_0 + (\beta_1 + \gamma_k)(x_{ijk} - 8) + \xi_{ijk},$

Estimativas dos parâmetros

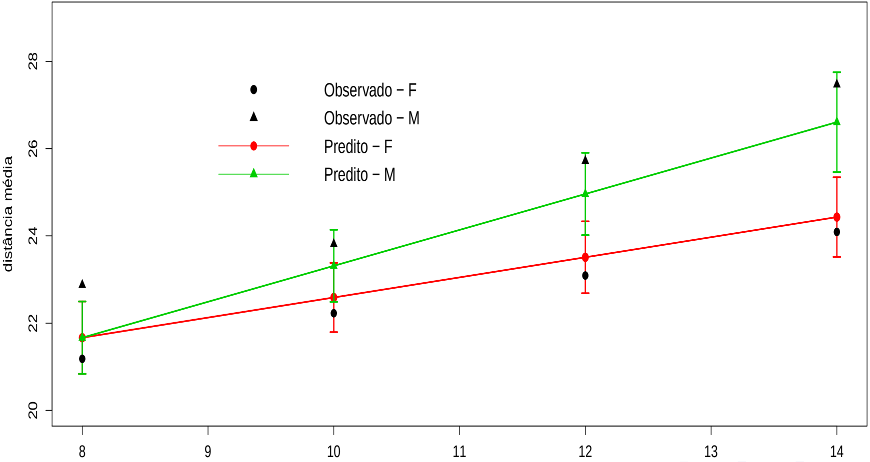
Parâmetro	Estimativa	EP	IC(95%)	Estatística	p-valor
β_0	21,67	0,42	[20,83 ; 22,51]	51,14	$< 0{,}0001$
β_1	0,46	0,06	[0,34 ; 0,58]	7,79	$< 0{,}0001$
γ_2	0,36	0,10	[0,17 ; 0,56]	3,66	$< 0{,}0001$

Estimativas dos parâmetros

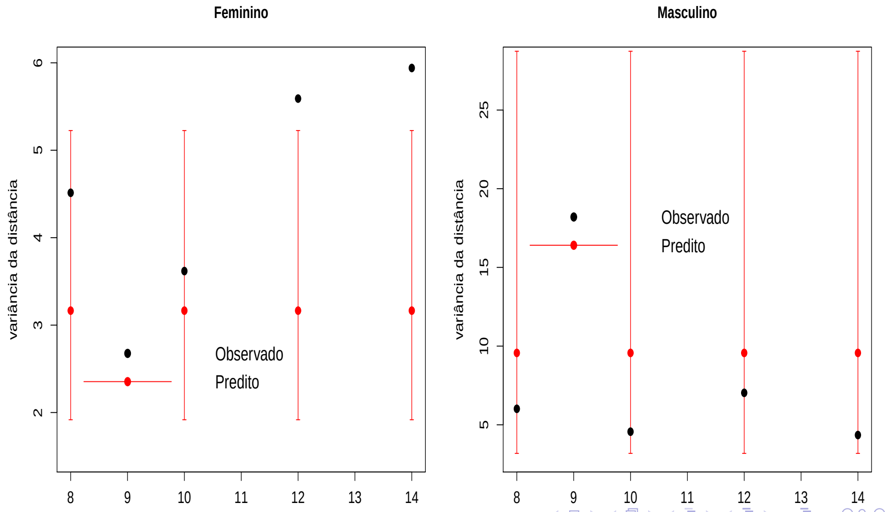
Parâmetro	Estimativa	IC(95%)
σ^2	3,17	[1,92 ; 5,23]
δ	1,74	[1,29 ; 2,34]

Parâmetro	Estimativa	IC(95%)
ρ	0,75	[0,58 ; 0,86]

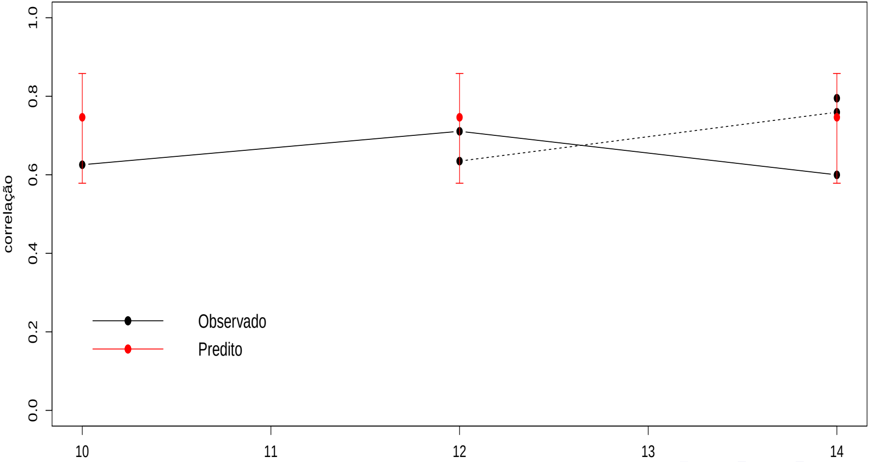
Perfis médios: observados e preditos

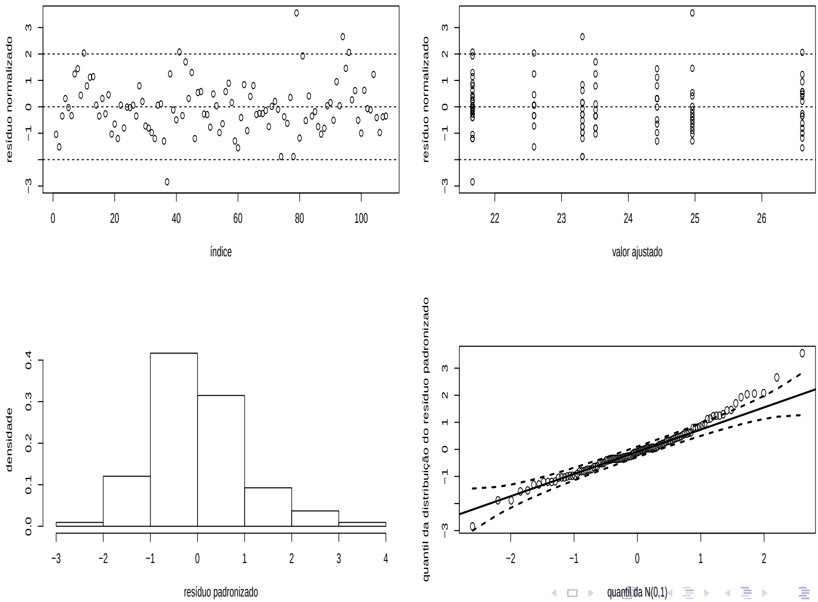


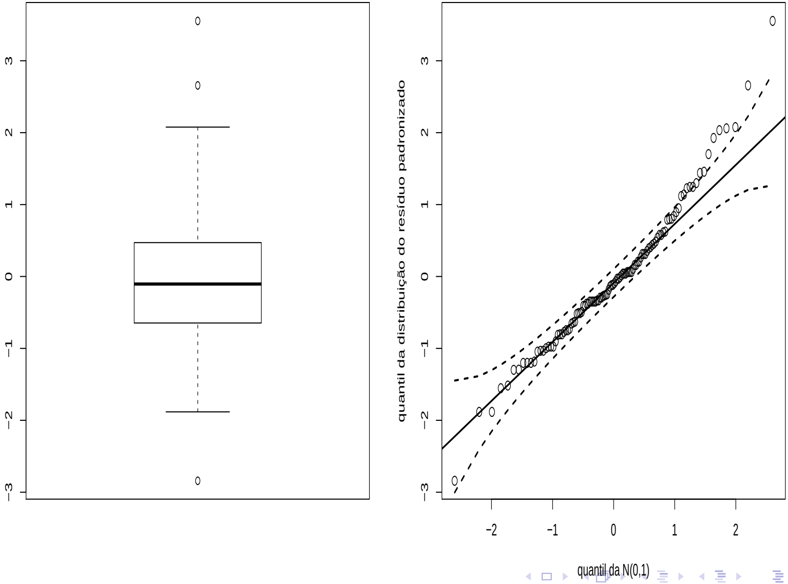
Variâncias: observadas e preditas



Correlações: observadas e preditas




resíduo normalizado
índice
resíduo normalizado
valor ajustado
densidade
resíduo padronizado
quantil da distribuição do resíduo padronizado
quantil da N(0,1)



3
2
1
0
−1
−2
−3
quantil da distribuição do resíduo padronizado
-2
-1
0
1
2
quantil da N(0,1)