

Planejamento e Análise Estatística de Experimentos fatoriais: análise de dados de experimentos completamente aleatorizados - Parte 3

Prof. Caio Azevedo

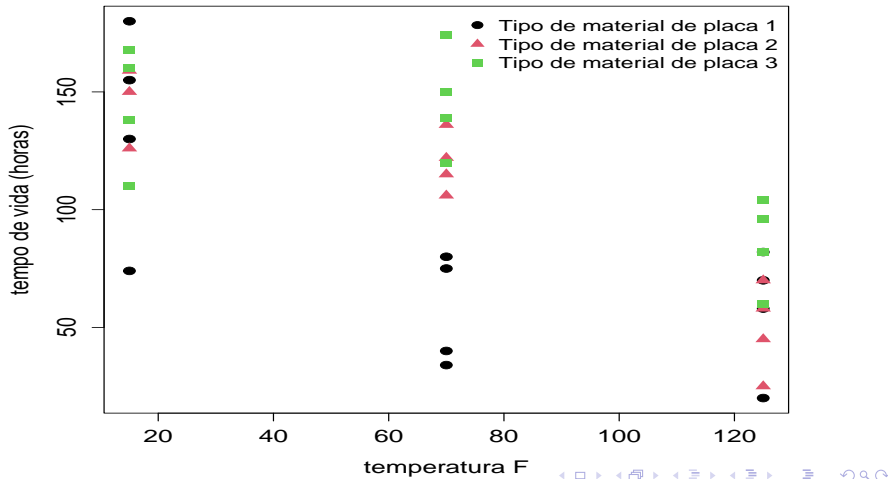
Exemplo 4: Resistência de materiais (Continuação)

- Um engenheiro está desenvolvendo um tipo de bateria para ser usado em um dispositivo eletrônico sujeito à variações extremas de temperatura.
- Fatores de interesse:
 - Tipo de material da placa: 1, 2 e 3.
 - Temperatura: 15°F, 70°F e 125°F. Equivalente à -9,44°C, 21,11°C e 51,67 °C, respectivamente
- Para cada combinação (tipo de material da placa × temperatura) 4 baterias foram feitas.
- Variável resposta: tempo de vida em horas de cada bateria .

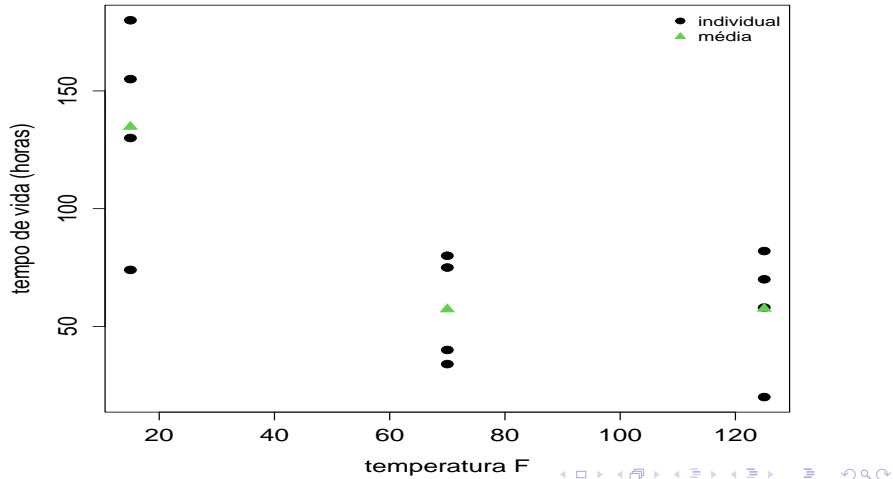
Resultados oriundos do Modelo Fatorial com interação

- Identificamos um padrão de diferença entre as médias através da **análise anterior**.
- Podemos ajustar um modelo de regressão (linear ou quadrático) para quantificar o impacto no tempo de vida com o aumento dos níveis dos fatores.
- Naturalmente, podemos levar em consideração os resultados anteriores para propor um modelo de regressão que leve em consideração o padrão de diferenças entre as médias.
- Entretanto, vamos começar com um modelo mais simples.

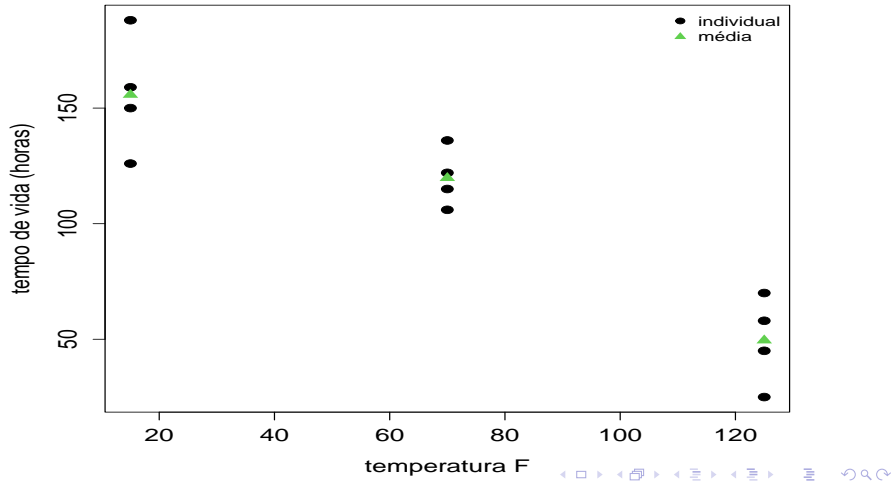
Diagrama de Dispersão



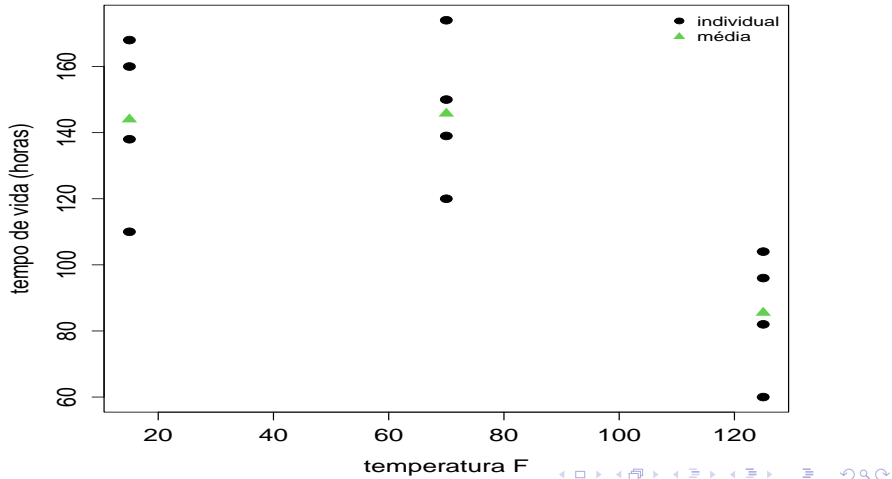
Diagramas de Dispersão (Tipo de Material 1)



Diagramas de Dispersão (Tipo de Material 2)



Diagramas de Dispersão (Tipo de Material 3)



Modelo linear 1: retas (M1)

$$Y_{ij} = (\alpha + \delta_i) + (\beta + \lambda_i) x_{ij} + \xi_{ij},$$
$$i = 1, 2, 3, j = 1, \dots, 12, \delta_1 = \lambda_1 = 0,$$
$$\xi_{ij} \stackrel{iid}{\sim} N(0, \sigma^2).$$

- Y_{ij} : tempo de vida (horas) do j -ésimo componente construído com o tipo de material i foi submetido.
- x_{ij} : temperatura ao qual o j -ésimo componente construído com o tipo de material i foi submetido.

Cont.

- $\alpha + \delta_i$: valor esperado (média) do tempo de vida de componentes construídos com o tipo de material i submetidos à uma temperatura de $0^\circ F$.
- $\beta + \lambda_i$: incremento no valor esperado do tempo de vida componentes construídos com o tipo de material i quando a temperatura de submissão aumenta em $1^\circ F$.
- Obs: note que os parâmetros δ_i e λ_i diferenciam as retas entre si.

Modelo linear 2: parábolas (M2)

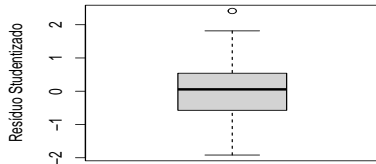
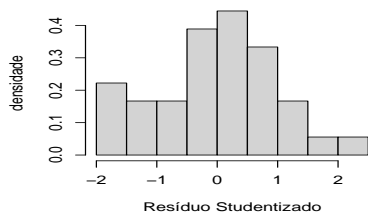
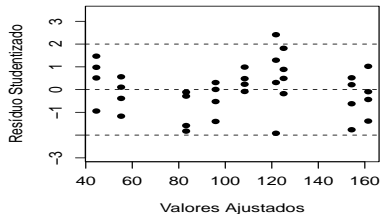
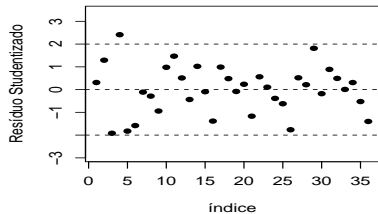
$$Y_{ij} = (\alpha + \delta_i) + (\beta + \lambda_i) x_{ij} + (\gamma + \tau_j) x_{ij}^2 + \xi_{ij},$$
$$i = 1, 2, 3, j = 1, \dots, 12, \delta_1 = \lambda_1 = \tau_1 = 0,$$
$$\xi_{ij} \stackrel{iid}{\sim} N(0, \sigma^2).$$

- Y_{ij} : tempo de vida (horas) do j -ésimo componente construído com o tipo de material i foi submetido.
- x_{ij} : temperatura ao qual o j -ésimo componente construído com o tipo de material i foi submetido.

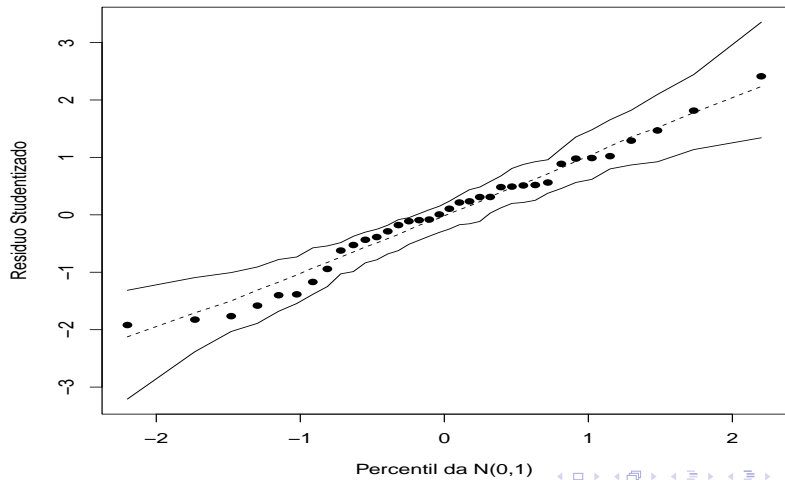
Cont.

- $\alpha + \delta_i$: valor esperado (média) do tempo de vida de componentes construídos com o tipo de material i submetidos à uma temperatura de $0^\circ F$.
- $-\frac{\beta + \lambda_i}{2(\alpha + \delta_i)}$: Valor da temperatura que fornece o tempo de vida mínimo (ou máximo).
- Obs: note que os parâmetros δ_i , λ_i e τ_i diferenciam as parábolas entre si.

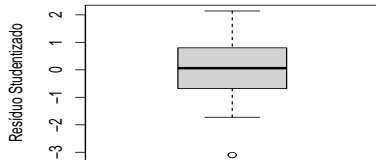
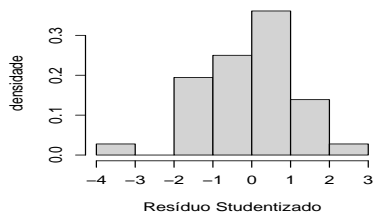
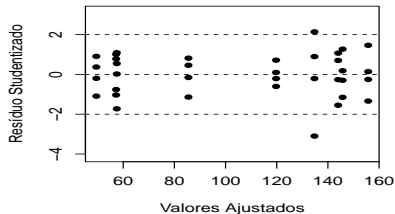
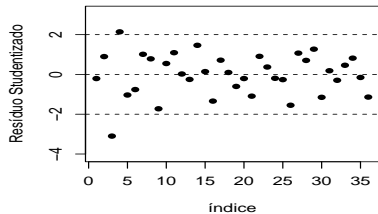
Análise de resíduos (M1)



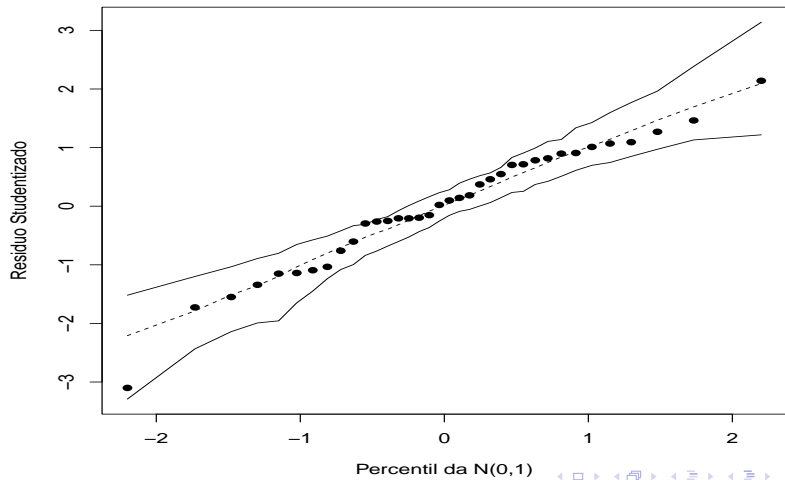
QQ plot com envelope (M1)



Análise de resíduos (M2)



QQ plot com envelope (M2)



Anova M1

- Que hipóteses cada uma das linhas abaixo estão testando ? (veja o respectivo código para o significado de cada “Fator”)

Fator	gl	SQ	QM	Estat. F	p-valor
tmaterfac	2	10683,72	5341,86	6,26	0,0053
temp	1	39042,67	39042,67	45,74	< 0,0001
tmaterfac:temp	2	2315,08	1157,54	1,36	0,2730
Residuals	30	25605,50	853,52		

Anova M2

- Que hipóteses cada uma das linhas abaixo estão testando ? (veja o respectivo código para o significado de cada “Fator”)

Fator	gl	SQ	QM	Estat. F	p-valor
tmaterfac	2	10683,72	5341,86	7,91	0,0020
temp	1	39042,67	39042,67	57,82	< 0,0001
temp2	1	76,06	76,06	0,11	0,7398
tmaterfac:temp	2	2315,08	1157,54	1,71	0,1991
tmaterfac:temp2	2	7298,69	3649,35	5,40	0,0106
Residuals	27	18230,75	675,21		

Comentários

- Parece que as suposições de normalidade e homocedasticidade para cada um dos modelos não são válidas para o conjunto de dados em questão.
- Uma alternativa: modelos de regressão com distribuição positiva e assimétrica para a variável resposta, que permita variâncias diferentes entre os grupos e com diferentes coeficientes de variação.
- Distribuições positivas: família gama (mãe não a tradicional), família normal inversa, família Weibull, família lognormal, família Birbaun-Saunders, normal assimétrica (apesar de ter suporte na reta).

Estimativas dos parâmetros M1

Parâmetros	Est.	EP	IC(95%)	Estat. t	p-valor
α	132,33	15,62	[100,43;164,22]	8,47	<0,0001
δ_2	43,62	22,09	[-1,49;88,73]	1,97	0,0576
δ_3	29,98	22,09	[-15,13;75,10]	1,36	0,1847
β	-0,70	0,19	[-1,09;-0,32]	-3,74	0,0008
λ_2	-0,26	0,27	[-0,81;0,28]	-0,99	0,3288
λ_3	0,17	0,27	[-0,37;0,71]	0,64	0,5259

Os resultados indicam que, talvez, um modelo que considere uma única reta seja apropriado para modelar os dados.

Estimativas dos parâmetros M2

Parâmetros	Est.	EP	IC(95%)	Estat. t	p-valor
α	169,38	20,57	[127,18:211,58]	8,24	<0,0001
δ_2	-9,76	29,09	[-69,44:49,93]	-0,34	0,7399
δ_3	-36,62	29,09	[-96,30:23,06]	-1,26	0,2188
β	-2,50	0,76	[-4,05:-0,95]	-3,31	0,0026
λ_2	2,33	1,07	[0,14:4,52]	2,18	0,0382
λ_3	3,40	1,07	[1,21:5,60]	3,19	0,0036
γ	0,01	0,01	[>0,00:0,02]	2,44	0,0214
τ_2	-0,02	0,01	[-0,03:<0,00]	-2,49	0,0193
τ_3	-0,02	0,01	[-0,04;-0,01]	-3,11	0,0044

Estimativas dos parâmetros M2

- Os resultados indicam que, talvez, um modelo que considere parábolas diferentes, mas com os mesmos “interceptos” seja apropriado.
- Critérios de Informação:

Modelo	AIC	BIC	AICc	SABIC	HQCIC	CAICe
M1	352,58	363,66	355,47	341,33	353,89	366,08
M2	346,35	362,18	353,27	330,48	349,32	367,60

- Pelo resultados acima, o modelo M2 é mais apropriado do que o modelo M1. Vamos continuar a análise com o modelo M2.

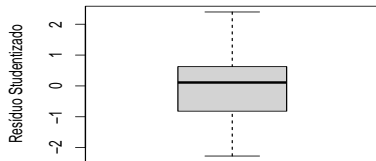
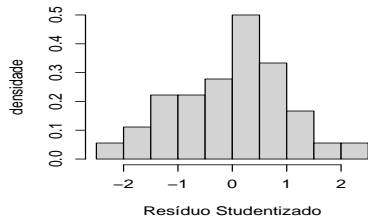
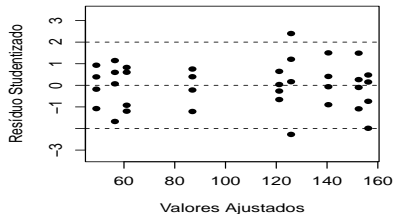
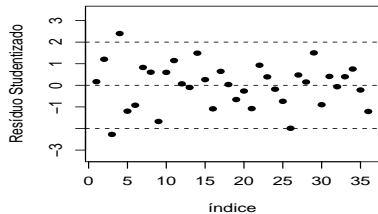
Comentários

- Com efeito, o teste $\mathbf{C}\boldsymbol{\beta} = \mathbf{0}$ para testar $H_0 : \delta_2 = \delta_3 = 0$ resultou em p-valor = 0,4385, o que indica que podemos considerar os mesmos interceptos.
- O modelo 3 é dado por:

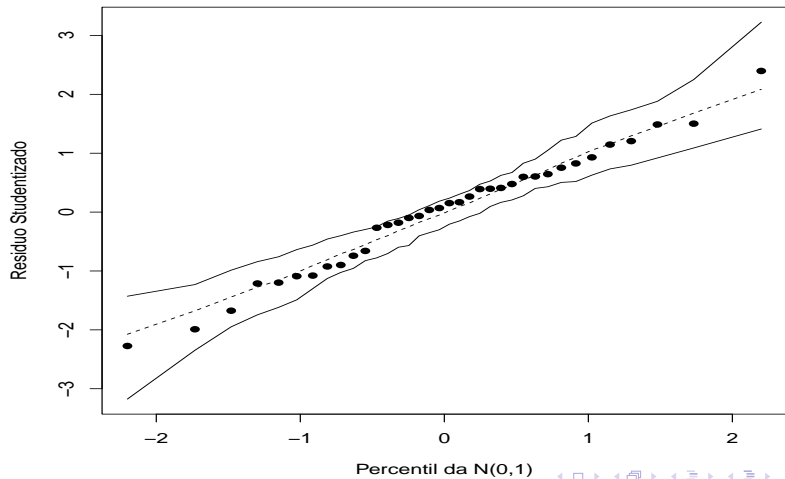
$$Y_{ij} = \alpha + (\beta + \lambda_i) x_{ij} + (\gamma + \tau_j) x_{ij}^2 + \xi_{ij},$$
$$i = 1, 2, 3, j = 1, \dots, 12, \lambda_1 = \tau_1 = 0,$$
$$\xi_{ij} \stackrel{iid}{\sim} N(0, \sigma^2).$$

em que cada um de seus elementos são como descritos para o modelo M2.

Análise de resíduos (M3)



QQ plot com envelope (M3)



Comentários

- Parece que as suposições de normalidade e homocedasticidade não são válidas para o conjunto de dados em questão.
- Uma alternativa: modelos de regressão com distribuição positiva e assimétrica para a variável resposta, que permita variâncias diferentes entre os grupos e com diferentes coeficientes de variação.
- Distribuições positivas: família gama (mas não a tradicional), família normal inversa, família Weibull, família lognormal, família Birbaun-Saunders, normal assimétrica (apesar de ter suporte na reta).
- Contudo, seguiremos com ele por questões pedagógicas.

Anova M1

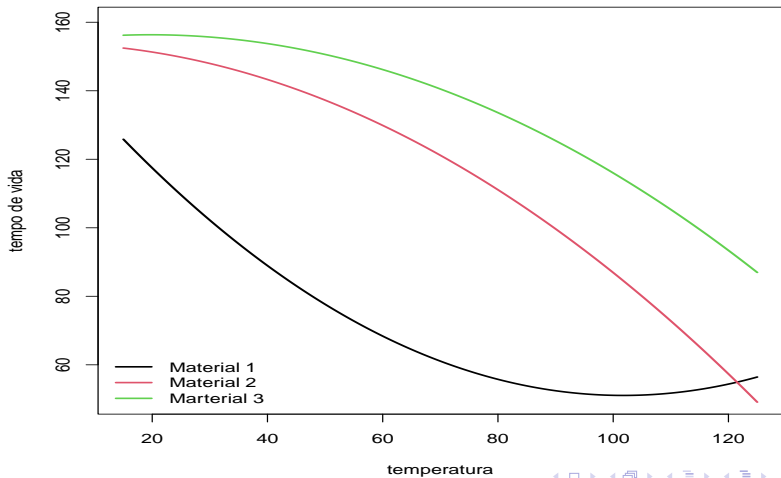
- Que hipóteses cada uma das linhas abaixo estão testando ? (veja o respectivo código para o significado de cada “Fator”)

Fator	gl	SQ	QM	Estat. F	p-valor
mX3	7	459168,31	65595,47	98,16	< 0,0001
Residuals	29	19378,69	668,23		

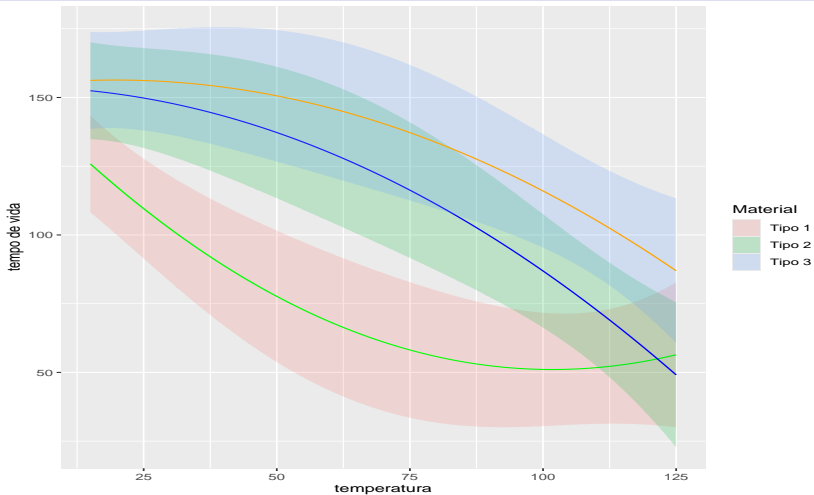
Estimativas dos parâmetros do modelo

Parâmetro	Estimativa	EP	IC(95%)	Estat. t	p-valor
α	153,92	11,81	[129,76;178,08]	13,03	<0,0001
β	-2,02	0,54	[-3,13;-0,91]	-3,72	0,0009
δ_2	2,03	0,57	[0,86;3,19]	3,56	0,0013
δ_3	2,27	0,57	[1,11;3,43]	3,99	0,0004
γ	0,01	<0,01	[0,00;0,02]	2,38	0,0242
τ_2	-0,02	0,00	[-0,03;-0,01]	-3,34	0,0023
τ_3	-0,02	0,00	[-0,03;-0,01]	-3,24	0,0030

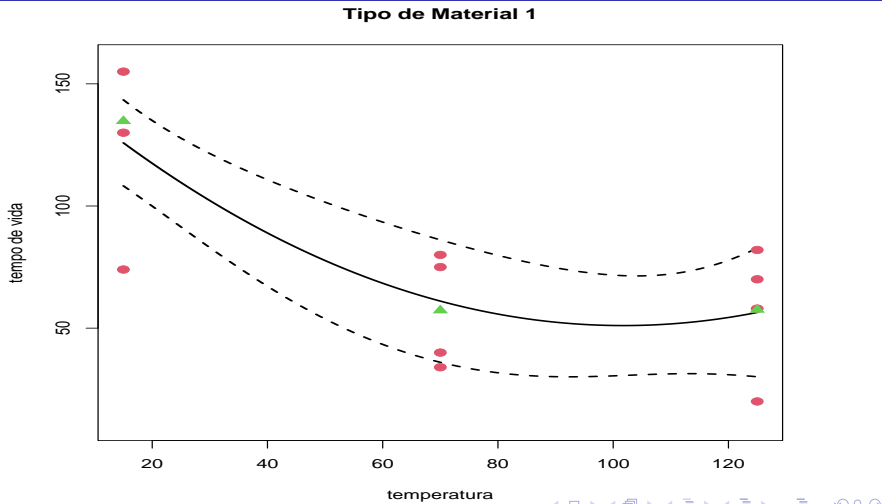
Curvas ajustadas



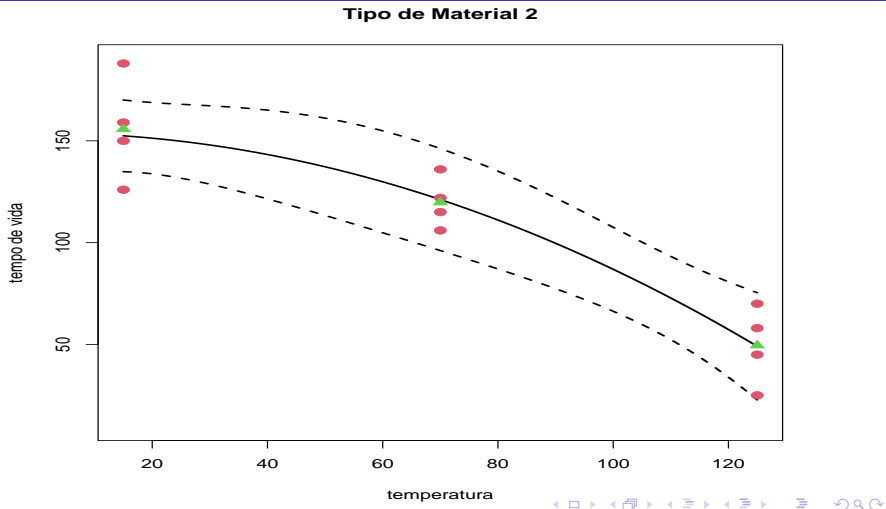
Curvas ajustadas e intervalos de confiança



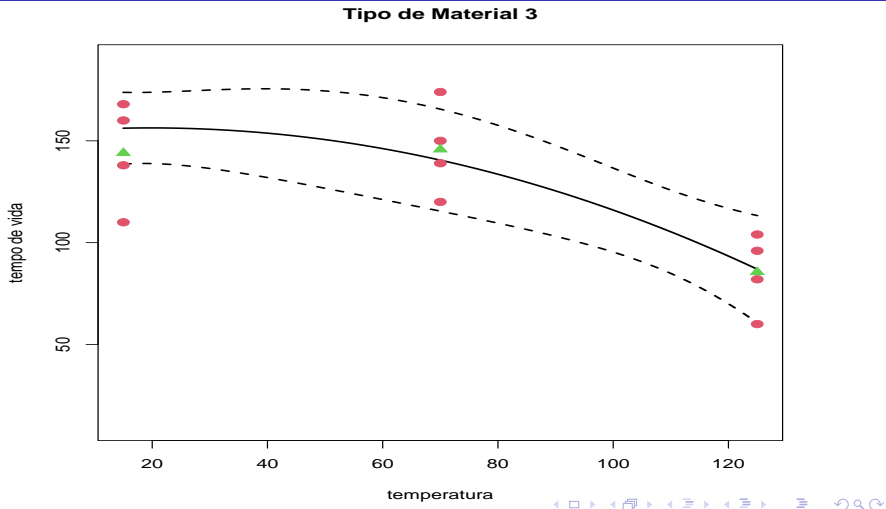
Estimativas pontuais e intervalares e valores observados



Estimativas pontuais e intervalares e valores observados



Estimativas pontuais e intervalares e valores observados



Comentários

- Apesar de mal ajustado, segundo a análise residual, do ponto de vista preditivo, o modelo apresentou um comportamento muito bom.
- Exercício: estudar a obtenção dos valores das temperaturas que levam ao tempo de vida máximo (mínimo) assim como os respectivos tempos de vida.

PCA desbalanceado com dois fatores

- Considere um experimento desbalanceado com dois fatores e um número geral de níveis.
- Os conceitos permanecem inalterados bem como as formas de análise.
- Mudanças nas fórmulas das somas de quadrados.
- Desbalanceamento desproporcional (tamanhos amostrais muito diferentes) pode comprometer a ortogonalidade das somas de quadrados.

Descrição

- Fator A: possui a níveis, $i=1,\dots,a$.
- Fator B: possui b níveis, $j=1,\dots,b$.
- Grupos: há um total de $a \times b$ grupos (tratamentos), que são definidos pelas interseções dos níveis de cada grupo.
- Para cada grupos vamos considerar um total de n_{ij} observações. Cada uma das n_{ij} observações são alocadas aleatoriamente à cada uma das combinações (fatores). Temos uma PCA (planejamento completamente casualizado).

- Neste caso número total de observações $n = \sum_{i=1}^a \sum_{j=1}^b n_{ij}$.

Modelo com interação (casela de referência)

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \xi_{ijk},$$

(Fator A), $i = 1, \dots, a$; (Fator B), $j = 1, \dots, b$;

(unidades experimentais), $k = 1, \dots, n_{ij}$.

- Erros $\xi_{ijk} \stackrel{i.i.d}{\sim} N(0, \sigma^2)$, $\mu, \alpha_i, \beta_j, (\alpha\beta)_{ij}$ não aleatórios.
- Restrições : $\alpha_1 = \beta_1 = (\alpha\beta)_{1j} = (\alpha\beta)_{i1} = 0, \forall i, j$.

Modelo sem interação (casela de referência)

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \xi_{ijk},$$

(Fator A), $i = 1, \dots, a$; (Fator B), $j = 1, \dots, b$;

(unidades experimentais), $k = 1, \dots, n_{ij}$.

- Erros $\xi_{ijk} \stackrel{i.i.d}{\sim} N(0, \sigma^2)$, $\mu, \alpha_i, \beta_j, (\alpha\beta)_{ij}$ não aleatórios.
- Restrições : $\alpha_1 = \beta_1 = (\alpha\beta)_{1j} = (\alpha\beta)_{i1} = 0, \forall i, j$.

Somas de quadrados

- Decomposição da soma de quadrados total:

$$\begin{aligned} SQT &= \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^{n_{ij}} (Y_{ijk} - \bar{Y}_{...})^2 = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^{n_{ij}} \left[(\bar{Y}_{i..} - \bar{Y}_{...}) \right. \\ &\quad \left. + (\bar{Y}_{.j.} - \bar{Y}_{...}) + (\bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...}) + (Y_{ijk} - \bar{Y}_{ij.}) \right]^2 \\ &= \sum_{i=1}^a \sum_{j=1}^b n_{ij} (\bar{Y}_{i..} - \bar{Y}_{...})^2 + \sum_{i=1}^a \sum_{j=1}^b n_{ij} (\bar{Y}_{.j.} - \bar{Y}_{...})^2 \\ &\quad + \sum_{i=1}^a \sum_{j=1}^b n_{ij} (\bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...})^2 + \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^{n_{ij}} (Y_{ijk} - \bar{Y}_{ij.})^2. \end{aligned}$$

Somas de quadrados

- Em que

$$\bar{Y}_{i..} = \frac{1}{n_i} \sum_{j=1}^b \sum_{k=1}^{n_{ij}} Y_{ijk}; \quad \bar{Y}_{.j.} = \frac{1}{n_{.j}} \sum_{i=1}^a \sum_{k=1}^{n_{ij}} Y_{ijk}$$

$$\bar{Y}_{ij.} = \frac{1}{n} \sum_{k=1}^{n_{ijk}} Y_{ijk}; \quad \bar{Y}_{...} = \frac{1}{n} \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^{n_{ijk}} Y_{ijk}$$

$$n_i = \sum_{j=1}^b n_{ij}; \quad n_{.j} = \sum_{i=1}^a n_{ij}; \quad n = \sum_{i=1}^a \sum_{j=1}^b n_{ij}.$$

Tabela de análise de variância

- Para testar a significância de cada efeito

FV	SQ	GL	QM	Estadística F	pvalor
Fator A	SQF_A	a-1	$QMF_A = \frac{SQF_A}{(a-1)}$	$F_A = \frac{QMF_A}{QMR}$	$\min(F(f_A H_0), S(f_A H_0))$
Fator B	SQF_B	b-1	$QMF_B = \frac{SQF_B}{(b-1)}$	$F_B = \frac{QMF_B}{QMR}$	$\min(F(f_B H_0), S(f_B H_0))$
Interação	SQInt	(a-1)(b-1)	$QMInt = \frac{SQInt}{[(a-1)(b-1)]}$	$F_{Int} = \frac{QMInt}{QMR}$	$\min(F(f_{Int} H_0), S(f_{Int} H_0))$
Resíduo	SQR	n-ab	$QMR = \frac{SQR}{(n-ab)}$		
Total	SQT	n-1			

FV: fonte de variação, SQ: soma de quadrados, GL: graus de liberdade, QM: quadrado médio. $F(x|H_0), S(x|H_0)$ fda e fds no ponto x sob H_0 , respectivamente.

Obtenção das somas de quadrados de forma matricial

- De forma semelhante ao caso de um único fator com vários níveis, as somas de quadrados associadas à cada fator (fatores principais e interação), são obtidas através das diferenças das somas de quadrados de resíduos de um modelo sem o fator em questão e com o fator em questão (excluindo-se os fatores de “ordem” superior).
- Lembrando que, a forma geral da SQR é dada por

$$SQR = \mathbf{Y}' (\mathbf{I} - \mathbf{H}) \mathbf{Y},$$

em que

$$\mathbf{H} = \mathbf{X} (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}'.$$

Obtenção das SQ's sob a forma matricial (cont.)

- No caso de um experimento com dois fatores temos os seguintes modelos possíveis:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \xi_{ijk}, \text{ (Modelo 1),}$$

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \xi_{ijk}, \text{ (Modelo 2),}$$

$$Y_{ijk} = \mu + \beta_j + \xi_{ijk}, \text{ (Modelo 3),}$$

$$Y_{ijk} = \mu + \alpha_i + \xi_{ijk}, \text{ (Modelo 4)}$$

Obtenção das somas de quadrados de forma matricial (cont.)

- Exercício: escrever a matriz \mathbf{X} para cada modelo acima. Elas serão nomeadas \mathbf{X}_1 , \mathbf{X}_2 e \mathbf{X}_3 , respectivamente
- Seja SQR_i associada ao modelo i e $\mathbf{H}_i = \mathbf{X}_i(\mathbf{X}_i'\mathbf{X}_i)^{-1}\mathbf{X}_i'$.
- Assim, temos

$$SQF_A = SQR_3 - SQR_1 = \mathbf{Y}'(\mathbf{H}_2 - \mathbf{H}_3)\mathbf{Y},$$

$$SQF_B = SQR_4 - SQR_1 = \mathbf{Y}'(\mathbf{H}_2 - \mathbf{H}_4)\mathbf{Y},$$

$$SQInt = SQR_2 - SQR_1 = \mathbf{Y}'(\mathbf{H}_1 - \mathbf{H}_2)\mathbf{Y}.$$

Esperanças dos Quadrados Médios

■ Expressões:

$$\mathcal{E}(QMF_A) = \sigma^2 + \frac{\sum_{j=1}^b n_{ij}(\bar{\mu}_{i.} - \bar{\mu}_{..})^2}{a-1}$$

$$\mathcal{E}(QMF_B) = \sigma^2 + \frac{\sum_{i=1}^a n_{ij}(\bar{\mu}_{.j} - \bar{\mu}_{..})^2}{a-1}$$

$$\mathcal{E}(QMF_{Int}) = \sigma^2 + \frac{\sum_{i=1}^a \sum_{j=1}^b n_{ij}[(\mu_{ij} + \bar{\mu}_{..}) - (\bar{\mu}_{i.} + \bar{\mu}_{.j})]^2}{a-1}$$

$$\mathcal{E}(QMR) = \sigma^2.$$

Esperanças dos Quadrados Médios

- Expressões:
- Prova: basta colocar cada uma delas em forma matricial e usar o resultado ([aqui](#)) slide 27.
- Exercício: obter as esperanças par o caso balanceado.
- Exercício: provar que sob a validade de H_0 ,

$$\mathcal{E}(QMF_A) = \mathcal{E}(QMF_B) = \mathcal{E}(QMF_{Int}) = \sigma^2.$$

Exemplo 5: O problema de engarrafamento de refrigerantes

- Uma empresa está interessada em que a quantidade de refrigerante colocada em cada garrafa seja a mais uniforme entre os vasilhames.
- Fatores de interesse:
 - Percentual de carbonatação (CARB): 10%, 12% e 14%.
 - Pressão de operação no enchimento (PRE): 25 e 30 psi.
 - Velocidade na linha de produção (VELOC): 200 e 250 bpm.
- Para cada combinação dos fatores (temos um total de $2 \times 2 \times 3 = 12$ tratamentos) foram medidas as diferenças entre o a quantidade de refrigerante inserida no vasilhame menos o valor padrão de dois refrigerantes escolhidos ao acaso.

Exemplo 5: continuação

- Variável resposta: diferença de preenchimento (DIP) .
- Planejamento completamente aleatorizado (PCA) balanceado, com 3 fatores, sendo dois com 2 níveis e 1 com três níveis.
- Fator A: CARB
- Fator B: PRE.
- Fator C: VELOC.
- Temos um total de 12 tratamentos.

Modelo (casela de referência)

$$Y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + \xi_{ijkl},$$

(Fator A), $i = 1, 2, 3$; (Fator B), $j = 1, 2$; (Fator C), $k = 1, 2$;

(unidades experimentais), $l = 1, 2$

- Erros $\xi_{ijkl} \stackrel{iid}{\sim} N(0, \sigma^2)$, $\mu, \alpha_i, \beta_j, \gamma_k, (\alpha\beta)_{ij}, (\alpha\gamma)_{ik}, (\beta\gamma)_{jk}, (\alpha\beta\gamma)_{ijk}$ não aleatórios.
- Restrições : $\alpha_1 = \beta_1 = \gamma_1 = (\alpha\beta)_{1j} = (\alpha\beta)_{i1} = (\alpha\gamma)_{i1} = (\alpha\gamma)_{1k} = (\beta\gamma)_{j1} = (\beta\gamma)_{1k} = 0, \forall i, j, k, (\alpha\beta\gamma)_{ijk} = 0$, se $i = 1$, e/ou, $j = 1$, e/ou, $k = 1$.

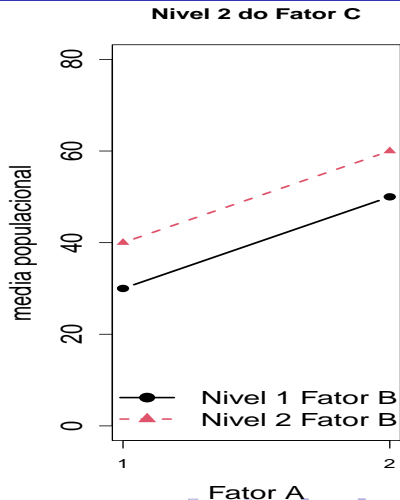
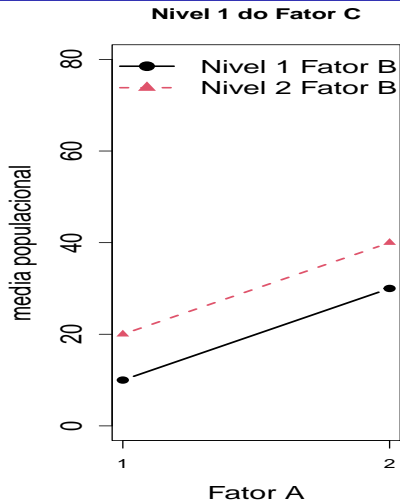
Interpretações dos parâmetros

- Agora tem-se: interações de segunda ordem (3 fatores simultaneamente), interação de primeira ordem (entre pares de fatores) e efeitos dos fatores principais.
- Se $(\alpha\beta\gamma)_{ijk} = 0, \forall i, j, k$ não há interação de segunda ordem.
- Em não havendo interação de segunda ordem
 - Se $(\alpha\beta)_{ij} = 0, \forall i, j$ não há interação de primeira ordem entre os Fatores A e B.
 - Se $(\alpha\gamma)_{ik} = 0, \forall i, k$ não há interação de primeira ordem entre os Fatores A e C.
 - Se $(\beta\gamma)_{jk} = 0, \forall j, k$ não há interação de primeira ordem entre os Fatores B e C.

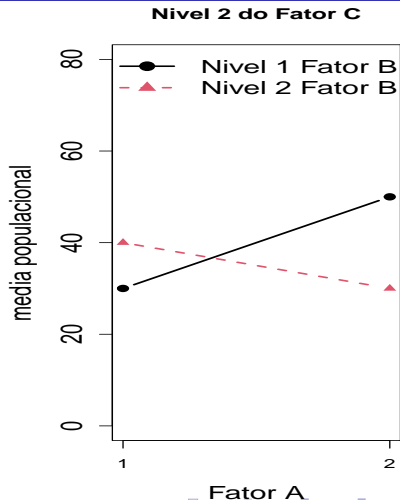
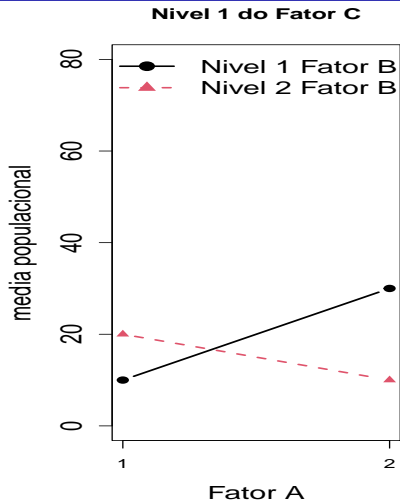
Interpretações dos parâmetros (cont.)

- Em não havendo interação de nenhuma ordem
 - Se $(\alpha)_{ij} = 0, \forall i$ não há efeito do fator principal A.
 - Se $(\beta)_j = 0, \forall j$ não há efeito do fator principal B.
 - Se $(\gamma)_k = 0, \forall k$ não há efeito do fator principal C.

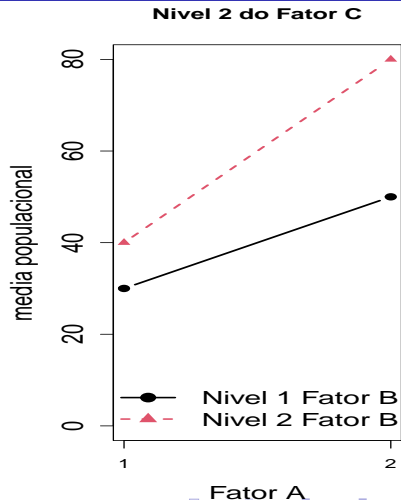
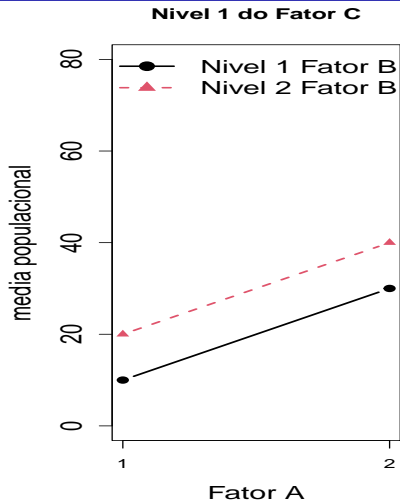
Perfis médios: ausência de interação de segunda ordem



Perfis médios: ausência de interação de segunda ordem



Perfis médios: presença de interação de segunda ordem



Observações

- Se existe interação de segunda ordem, existem interações de ordens inferiores (e, conseqüentemente existem efeitos dos fatores individuais).
- Se não existe interação de segunda ordem, a existência das interações de primeira ordem podem ser verificadas da forma usual, bem como a existência dos efeitos dos fatores individuais (ver exemplos apresentados).

Médias

■ Neste caso

$$\mu_{111} = \mu, \mu_{211} = \mu + \alpha_2, \mu_{311} = \mu + \alpha_3$$

$$\mu_{121} = \mu + \beta_2, \mu_{221} = \mu + \alpha_2 + \beta_2 + (\alpha\beta)_{22},$$

$$\mu_{321} = \mu + \alpha_3 + \beta_2 + (\alpha\beta)_{32}$$

$$\mu_{112} = \mu + \gamma_2, \mu_{212} = \mu + \alpha_2 + \gamma_2 + (\alpha\gamma)_{22},$$

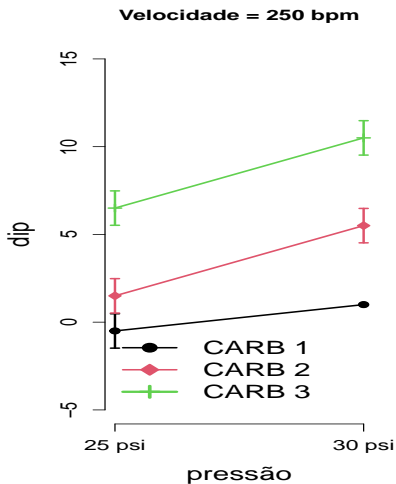
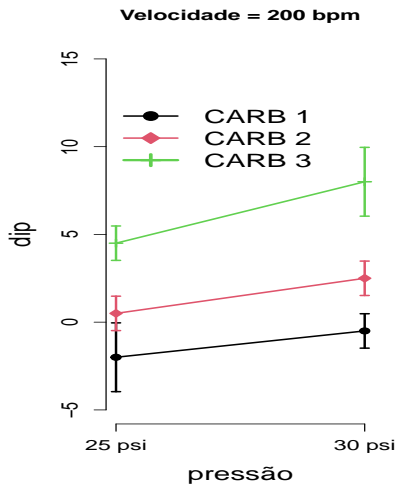
$$\mu_{312} = \mu + \alpha_3 + \gamma_2 + (\alpha\gamma)_{32}$$

$$\mu_{122} = \mu + \beta_2 + \gamma_2 + (\beta\gamma)_{22},$$

$$\mu_{222} = \mu + \alpha_2 + \beta_2 + \gamma_2 + (\alpha\beta)_{22} + (\beta\gamma)_{22} + (\alpha\gamma)_{22} + (\alpha\beta\gamma)_{222},$$

$$\mu_{322} = \mu + \alpha_3 + \beta_2 + \gamma_2 + (\alpha\beta)_{32} + (\beta\gamma)_{22} + (\alpha\gamma)_{32} + (\alpha\beta\gamma)_{322}.$$

Voltando ao exemplo: Gráfico de perfis médios



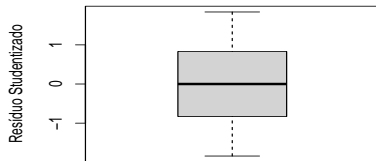
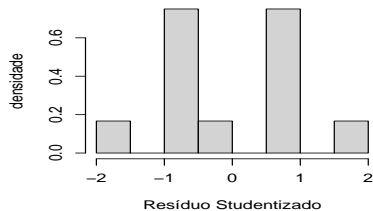
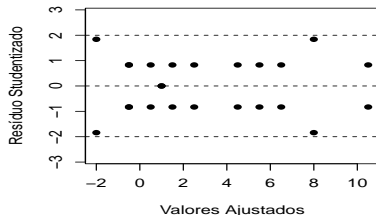
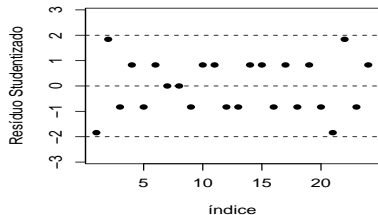
Análise descritiva

CARB	PRE	VELOC	Medida descritiva					
			Média	DP	Var.	CV%	Mínimo	Máximo
10	25 psi	200 bpm	-2,00	1,41	2,00	-70,71	-3,00	-1,00
10	25 psi	250 bpm	-0,50	0,71	0,50	-141,42	-1,00	0,00
10	30 psi	200 bpm	-0,50	0,71	0,50	-141,42	-1,00	0,00
10	30 psi	250 bpm	1,00	0,00*	0,00*	0,00*	1,00	1,00
12	25 psi	200 bpm	0,50	0,71	0,50	141,42	0,00	1,00
12	25 psi	250 bpm	1,50	0,71	0,50	47,14	1,00	2,00
12	30 psi	200 bpm	2,50	0,71	0,50	28,28	2,00	3,00
12	30 psi	250 bpm	5,50	0,71	0,50	12,86	5,00	6,00
14	25 psi	200 bpm	4,50	0,71	0,50	15,71	4,00	5,00
14	25 psi	250 bpm	6,50	0,71	0,50	10,88	6,00	7,00
14	30 psi	200 bpm	8,00	1,41	2,00	17,68	7,00	9,00
14	30 psi	250 bpm	10,50	0,71	0,50	6,73	10,00	11,00

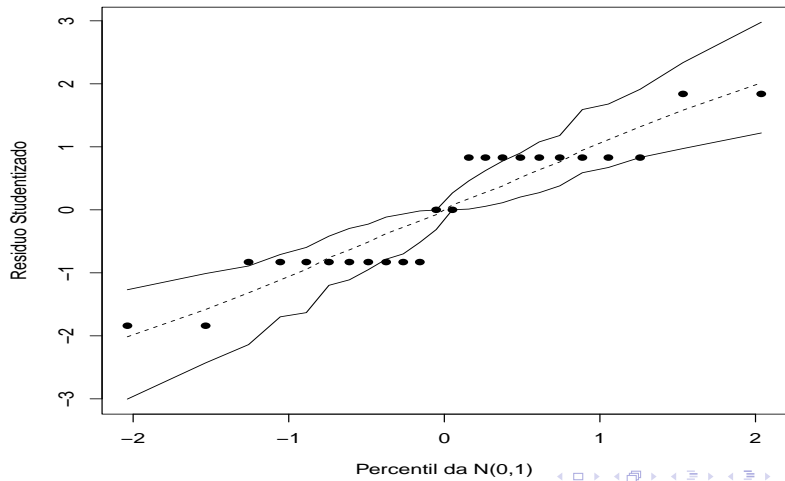
Análise descritiva

- Temos duas observações por grupo.
- Não há sentido em construir box-plots ou histogramas, nem mesmo realizar nenhum teste acerca de normalidade ou homocedasticidade, nem mesmo calcular assimetria ou curtose.

Análise de resíduos



QQ plot para os resíduos



Comentários

- Parece que as suposições do modelo não são válidas para o conjunto de dados em questão (embora o ajuste tenha melhorado em relação à situação anterior).
- Ausência de homocedasticidade e normalidade.
- Uma alternativa: modelos de regressão com distribuição assimétrica/caudas pesadas para a variável resposta, que permita variâncias diferentes: normal e t assimétricas.

Comentários

- Entretanto, parece (não se tem mais informações) que resposta varia, de forma discreta (no sentido probabilístico), num intervalo de amplitude “pequena” (talvez o use de uma transformação da binomial seja mais apropriado).
- O modelo de regressão normal linear, aparentemente, não é adequado para analisar os dados em questão, apesar do ajuste ter melhorado em relação à situação anterior (considerando apenas dois fatores).
- Contudo, seguiremos com ele por questões pedagógicas.

Tabela ANOVA

FV	SQ	GL	QM	Estatística F	pvalor
CARB	2	252,750	126,375	178,4118	<0,0001
PRE	1	45,375	45,375	64,0588	<0,0001
VELOC	1	22,042	22,042	31,1176	0,0001
CARB x PRE	2	5,250	2,625	3,7059	0,0558
CARB x VELOC	2	0,583	0,292	0,4118	0,6714
PRE x VELOC	1	1,042	1,042	1,4706	0,2486
CARB x PRE x VELOC	2	1,083	0,542	0,7647	0,4869
Resíduo	12	8,500	0,708		

Tabela ANOVA

- Ausência de interação de segunda ordem e de interação de primeira entre CARB e VELOC e PRE e VELOC.
- **OBS: Pesquisar as formas escalares e matriciais das somas de quadrados**

Estimativas dos parâmetros do modelo

Parâmetro	Estimativa	EP	IC(95%)	Estat. t	pvalor
μ	-2,00	0,60	[-3,17;-0,83]	-3,36	0,0057
α_2	2,50	0,84	[0,85; 4,15]	2,97	0,01117
α_3	6,50	0,84	[4,85;8,15]	7,72	<0,0001
β_2	1,50	0,84	[-0,15; 3,15]	1,78	0,1000
γ_2	1,50	0,84	[-0,15;3,15]	1,78	0,1000
$(\alpha\beta)_{22}$	0,50	1,19	[-1,83; 2,83]	0,42	0,6818
$(\alpha\beta)_{32}$	2,00	1,19	[-0,33;4,33]	1,68	0,1187
$(\alpha\gamma)_{22}$	-0,50	1,19	[-2,83; 1,83]	-0,42	0,6818
$(\alpha\gamma)_{32}$	0,50	1,19	[-1,83;2,83]	0,42	0,6818
$(\beta\gamma)_{22}$	-0,00	1,19	[-2,33;2,33]	-0,00	> 0,9999
$(\alpha\beta\gamma)_{222}$	2,00	1,68	[-1,30;5,30]	1,19	0,2577
$(\alpha\beta\gamma)_{322}$	0,50	1,68	[-2,80;3,80]	0,30	0,7715

Modelo reduzido (MR) (casela de referência)

$$Y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + \xi_{ijkl},$$

(Fator A), $i = 1, 2, 3$; (Fator B), $j = 1, 2$; (Fator C), $k = 1, 2$;
(unidades experimentais), $l = 1, 2$

- Erros $\xi_{ijkl} \stackrel{i.i.d}{\sim} N(0, \sigma^2)$, $\mu, \alpha_i, \beta_j, \gamma_k, (\alpha\beta)_{ij}$ não aleatórios.
- Restrições : $\alpha_1 = \beta_1 = \gamma_1 = (\alpha\beta)_{1j} = (\alpha\beta)_{i1} = 0, \forall i, j$.

Médias: modelo reduzido

- Neste caso

$$\mu_{111} = \mu, \mu_{211} = \mu + \alpha_2, \mu_{311} = \mu + \alpha_3$$

$$\mu_{121} = \mu + \beta_2, \mu_{221} = \mu + \alpha_2 + \beta_2 + (\alpha\beta)_{22}$$

$$\mu_{321} = \mu + \alpha_3 + \beta_2 + (\alpha\beta)_{32}$$

$$\mu_{112} = \mu + \gamma_2, \mu_{212} = \mu + \alpha_2 + \gamma_2$$

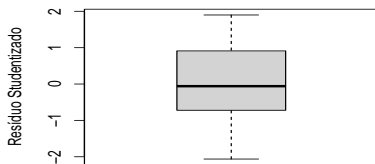
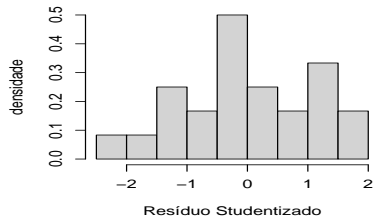
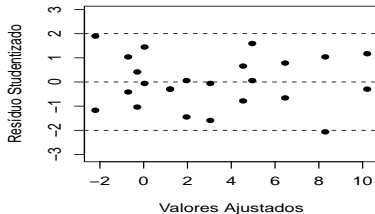
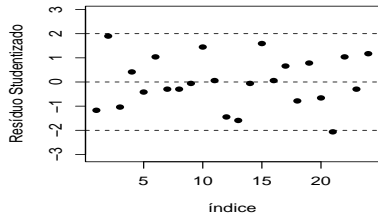
$$\mu_{312} = \mu + \alpha_3 + \gamma_2$$

$$\mu_{122} = \mu + \beta_2 + \gamma_2$$

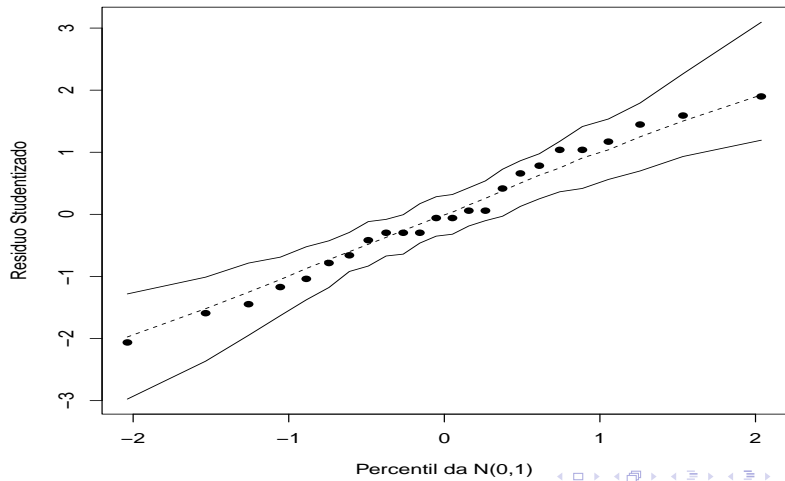
$$\mu_{222} = \mu + \alpha_2 + \beta_2 + \gamma_2 + (\alpha\beta)_{22}$$

$$\mu_{322} = \mu + \alpha_3 + \beta_2 + \gamma_2 + (\alpha\beta)_{32}$$

Análise de resíduos (MR)



QQ plot para os resíduos (MR)



Comentários

- O ajuste do modelo parece ter melhorado bastante, embora persista certo nível de assimetria negativa e uma eventual falta de homocedasticidade possa ser aventada.
- Entretanto, a qualidade do ajuste pode ser considerada aceitável, embora outros modelos possam e devam ser testados.

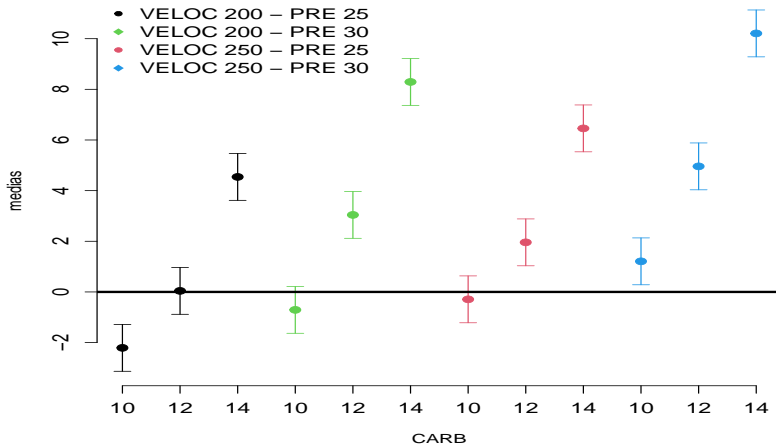
Tabela ANOVA: modelo reduzido

FV	SQ	GL	QM	Estatística F	pvalor
CARB	2	252,75	126,38	191,68	<0,0001
PRE	1	45,38	45,38	68,82	<0,0001
VELOC	1	22,04	22,04	33,43	<0,0001
CARB x PRE	2	5,25	2,63	3,98	0,03818
Resíduos	17	11,21	0,66		
Total	22	336,63			

Estimativas dos parâmetros do modelo

Parâmetro	Estimativa	EP	IC(95%)	Estat. t	pvalor
μ	-2,21	0,44	[-3,07; -1,35]	-5,04	0,0001
α_2	2,25	0,57	[1,12;3,38]	3,92	0,0011
α_3	6,75	0,57	[5,62;7,88]	11,76	<0,0001
β_2	1,50	0,57	[0,37;2,63]	2,61	0,0182
γ_2	1,92	0,33	[1,27;2,57]	5,78	<0,0001
$(\alpha\beta)_{22}$	1,50	0,81	[-0,09;3,09]	1,85	0,0822
$(\alpha\beta)_{32}$	2,25	0,81	[0,66;3,84]	2,77	0,0131

Médias estimadas: modelo reduzido



Comparações entre as médias

- Perguntas de interesse : quais tratamento fornecem DIP's médios:
 - Em torno do 0?
 - Negativos?
 - Positivos?
- Quais são os padrões de igualdade entre as médias.
- Após realizadas algumas comparações (em termos de $C\beta = \mathbf{0}$) discutidas em classe (veja também o arquivo com os programas relacionados), chegou-se ao padrão a seguir.
- Considere que os tratamentos (de acordo com o gráfico anterior) estão numeradas de 1 a 12 (conforme o padrão apresentado).

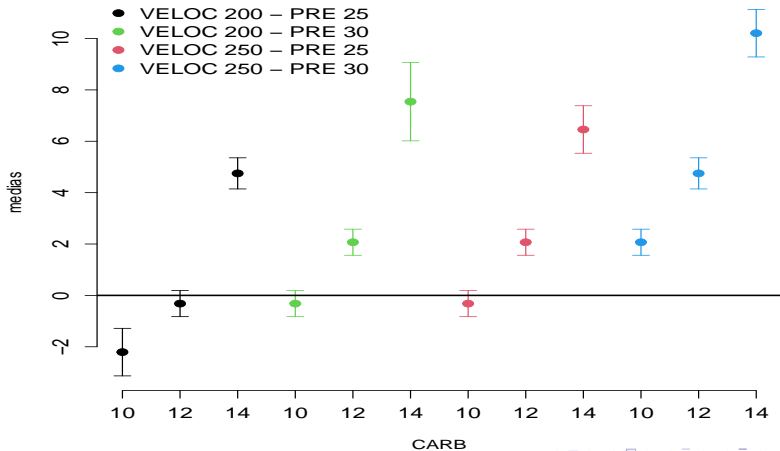
Padrões

- Identificou-se 8 grupos, a saber:
 - Grupo 1: tratamento 1.
 - Grupo 2: tratamentos 2, 4 e 7.
 - Grupo 3: tratamentos 3 e 11.
 - Grupo 4: tratamento 5, 8 e 10.
 - Grupo 5: tratamento 6.
 - Grupo 6: tratamento 9.
 - Grupo 7: tratamento 12.

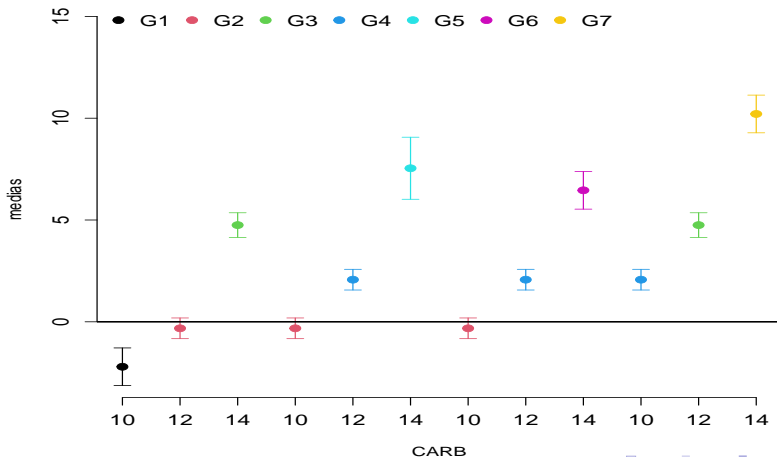
Padrões

- Resultados dos testes $C\beta = 0$:
 - Igualdade entre as médias dos grupos 2, 4 e 7 ($p = 0,4425$).
 - Igualdade entre as médias dos grupos 5, 8 e 10 ($p = 0,4425$).
 - Igualdade entre as médias dos grupos 3, 9 e 11 ($p = 0,0494$).
 - Igualdade entre as médias dos grupos 3 e 9 ($p < 0,0001$).
 - Igualdade entre as médias dos grupos 3 e 11 ($p = 0,5380$).
 - Igualdade entre as médias dos grupos 9 e 11 ($p = 0,0182$).
- Os três gráficos referem-se as médias previstas, sob os grupos encontrados, para o modelo reduzido (anterior).

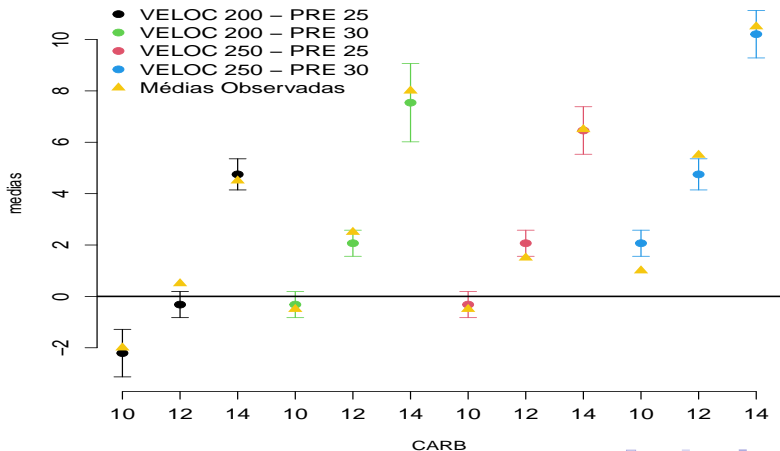
Médias estimadas via MR



Médias estimadas via MR (por grupo)



Médias estimadas via MR (com valores observados)



Estimativas finais das médias

Grupo	Estimativa	EP	IC(95%)
1	-2,21	0,44	[-3,13;-1,28]
2	-0,32	0,24	[-0,83;0,19]
3	4,75	0,29	[4,14;5,36]
4	2,07	0,24	[1,56;2,58]
5	7,54	0,72	[6,02;9,07]
6	6,46	0,44	[5,53;7,38]
7	10,21	0,44	[9,28;11,13]

Comentários

- Em geral, é melhor estimar as médias com o modelo reduzido (todas as médias são estimadas com maior precisão).
- Entretanto, também é válido estimar as médias, com o modelo “atual” .
- Com efeito, nesse caso, seria muito complicado (ou mesmo inviável) inserir as restrições no ajuste do modelo, a fim de obter um modelo mais reduzido, com os grupos já considerados em seu ajuste.

- Conclusões finais:
 - Grupo 1: DIP médio negativo.
 - Grupo 2: DIP médio em torno do zero.
 - Grupos 3 a 7: DIP médio positivo.