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

Prof. Caio Azevedo

Contexto

- 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.
- Analisaremos uma situação com dois fatores dentro de uma estrutura balanceada.

Descrição

- Fator A: possui a níveis.
- Fator B: possui b níveis.
- 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 observações (balanceado). Cada uma das n observações são alocadas aleatoriamente à cada uma das combinações (fatores). Temos uma PCA (planejamento completamente casualizado).

Descrição (Cont.)

- Note que tem-se um total de $n \times a \times b$ observações.
- 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).

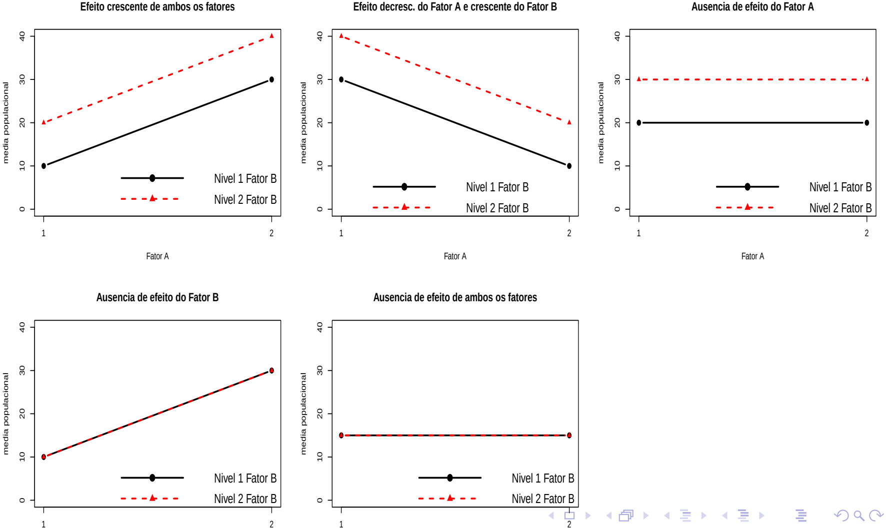
Exemplo 10: Resistência de materiais

- 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^{o}F, 70^{o}F e 125^{o}F. Equivalente à -9,44^{o}C, 21,11^{o}C e 51,67 ^{o}C, respectivamente
- Para cada combinação (tipo de material da placa $\times$ temperatura) 4 baterias foram feitas.
- Variável resposta: tempo de vida em horas de cada bateria .

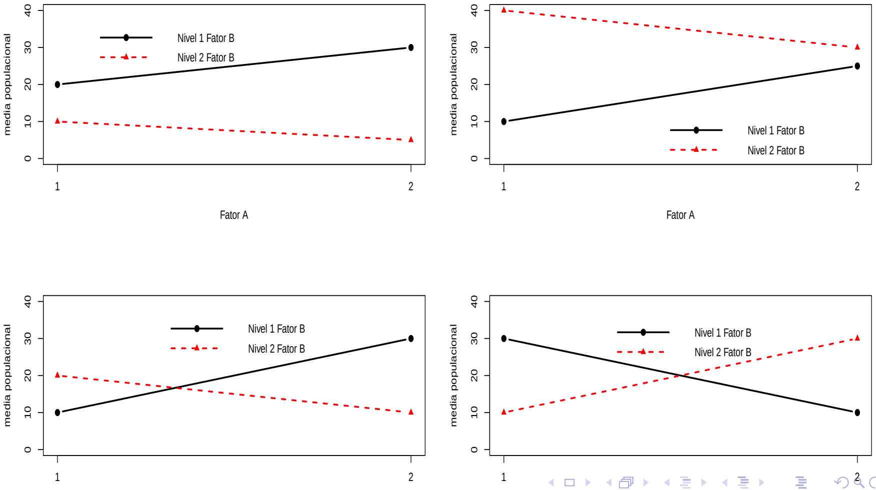
Exemplo 10: continuação

- Experimento balanceado: 4 observações por tratamento.
- Um fator quantitativo (temperatura) e um fator qualitativo (tipo de material da placa).
- Como analisar o experimento?
- Qual seria um modelo apropriado?
- Como estimar os parâmetros e comparar as médias de interesse?
- Como verificar as suposições do modelo?

Perfis médios: ausência de interação



Perfis médios: presença de interação



Voltando ao Exemplo 10

- Vamos considerar, inicialmente, somente os dois primeiros níveis de cada fator.
- Dados:

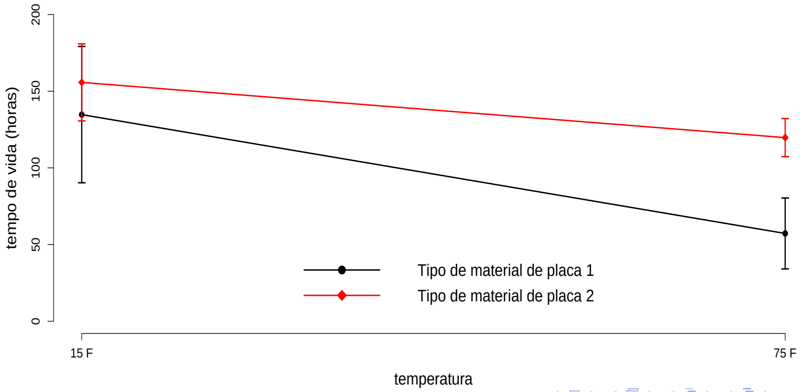
Material	**Temperatura (^oF)**			
	15		**70**	
1	130	155	34	40
	74	180	80	75
2	150	188	136	122
	159	126	106	112

Análise descritiva

Não há sentido em construir box-plots ou histogramas (poucas observações por grupo).

Material	Temp.	Medida descritiva					
		Média	DP	Var.	CV%	Máximo	Mínimo
1	15 F	134,75	45,35	2056,92	74,00	180,00	33,66
	70 F	57,25	23,60	556,92	34,00	80,00	41,22
2	15 F	155,75	25,63	656,25	126,00	188,00	16,45
	70 F	119,75	12,66	160,25	106,00	136,00	10,57

Gráfico de perfis médios



Modelo (casela de referência)

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

(Fator A), $i = 1, 2$; (Fator B), $j = 1, 2$; (unidades experimentais), $k = 1, 2, 3, 4$

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

Interpretações dos parâmetros

- Neste caso

$$
\begin{aligned}
\mu_{11} &= \mu \\
\mu_{21} &= \mu + \alpha_2 \\
\mu_{12} &= \mu + \beta_2 \\
\mu_{22} &= \mu + \alpha_2 + \beta_2 + (\alpha\beta)_{22}
\end{aligned}
$$

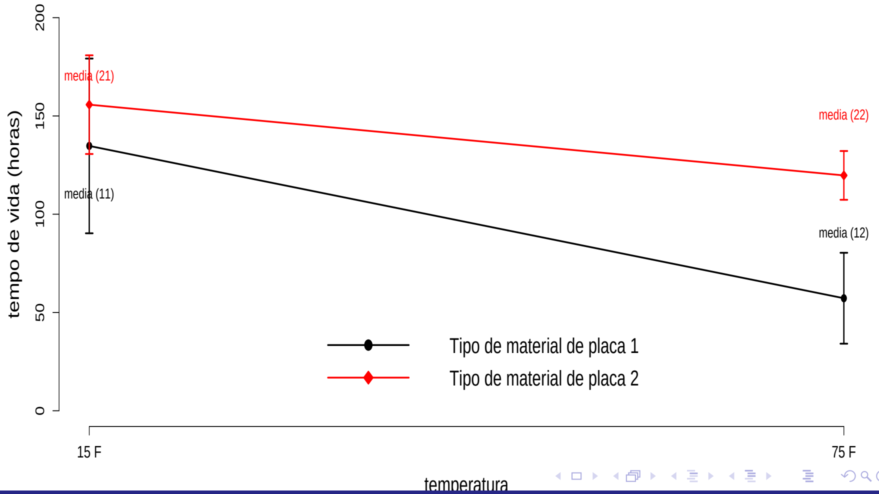
Interpretações dos parâmetros

- Fator A (1: material 1, 2: material 2).
- Fator B (1: 15^oF, 2: 75^oF).
- Parâmetros $\boldsymbol{\beta} = (\mu, \alpha_2, \beta_2, (\alpha\beta)_{22})$ (modelo identificado).
- Se $(\alpha\beta)_{22} = 0$.
  - α_2: incremento na vida média de baterias feitas com material 2 em relação àquelas feitas com material 1 submetidas à qualquer uma das duas temperaturas.
  - β_2: incremento na vida média de baterias submetidas à temperatura de 75^oF em relação submetidas à temperatura de 15^oF, feitas com qualquer um dos dois tipos de material.

Interpretações dos parâmetros (cont.)

- A não nulidade de $(\alpha\beta)_{22}$ faz com que os incrementos anteriores não dependam somente de α_2 e β_2. Neste caso:
  - Dependendo da temperatura, a diferença entre a vida média de baterias feitas com os materiais 1 e 2 não é a mesma.
  - Dependendo do tipo de material, a diferença entre a vida média de baterias submetidas as temperaturas $15^o F$ e $75^o F$ não é a mesma.
- O parâmetro $(\alpha\beta)_{22}$ determina a existência ou não de interação.

Visualização dos significados dos parâmetros



Interpretações dos parâmetros (cont.)

- Não existe interação, neste naso, $\leftrightarrow H_0 : \mu_{21} - \mu_{11} = \mu_{22} - \mu_{12}$ for verdadeira.
- Por outro lado, a hipótese acima equivale à:

$$H_0 : \alpha_2 = \alpha_2 + \beta_2 + (\alpha\beta)_{22} - \beta_2 \leftrightarrow (\alpha\beta)_{22} = 0$$

- Portanto, inexiste interação $\leftrightarrow (\alpha\beta)_{22} = 0$.

Interpretações dos parâmetros (cont.)

- Se existe interação, portanto se $(\alpha\beta)_{22} \neq 0$, temos que:
- α_2: incremento na vida média de baterias feitas com material 2 em relação àquelas feitas com material 1 submetidas à temperatura de $15^o F$.
- β_2: incremento na vida média de baterias submetidas à temperatura de $75^o F$ em relação submetidas à temperatura de $15^o F$, feitas com material do tipo 1.
- $(\alpha\beta)_{22}$: interação entre os fatores.

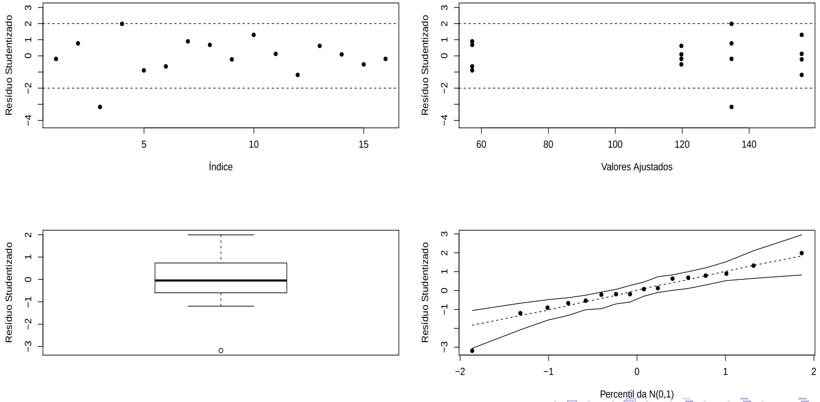
Hipótese de interesse

- Comparar simultaneamente todas as médias deixa de ter sentido prático.
- Primeira hipótese (ausência de interação): $H_0 : (\alpha\beta_{22}) = 0$ vs $H_1 : (\alpha\beta_{22}) \neq 0$
- Se a hipótese acima (H_0) não for rejeitada, então:
  - Ausência de efeito principal de material: $H_0 : \alpha_2 = 0$ vs $H_1 : \alpha_2 \neq 0$.
  - Ausência de efeito principal de temperatura: $H_0 : \beta_2 = 0$ vs $H_1 : \beta_2 \neq 0$.
- Eventualmente, algum tipo de comparação entre as médias remanescentes.

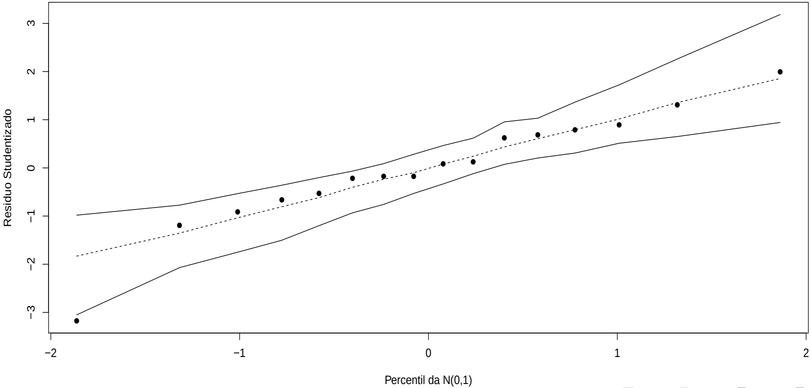
Hipótese de interesse (cont.)

- Se a hipótese acima de ausência de interação não for rejeitada, então não faz sentido estudar os efeitos principais isoladamente
- Portanto, deve-se efetuar algum tipo de comparação entre as médias.

Análise de resíduos



Gráficos de evelopes



Hipótese de interesse (cont.)

- Aparenten ausência de correlação.
- Aparente normalidade dos resíduos (não se tem subsíduos para se concluir o contrário).
- Presença de heterocedasticidade.

Estimativas dos parâmetros do modelo

Parâmetro	Estimativa	EP	IC(95%)	Estat. t	pvalor
μ	134,75	14,64	[102,85; 166,65]	9,20	$<$ 0,0001
α_2	21,00	20,71	[-24,12 ; 66,12]	1,01	0,3305
β_2	-77,50	20,71	[-122,62 ;-32,38]	-3,74	0,0028
$(\alpha\beta)_{22}$	41,50	29,28	[-22,31 ;105,31]	1,42	0,1819

Os resultados indicam inxistência de interação e de efeito do tipo de material. Contudo, vamos proceder de um modo “backward”, ajustando um modo somente sem interação.

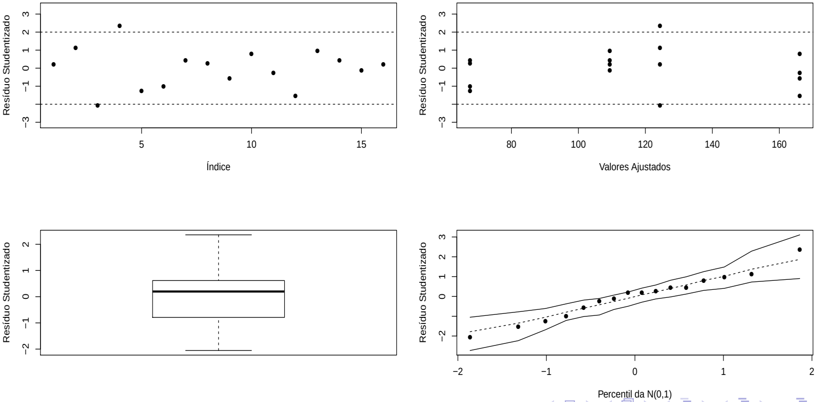
Modelo reduzido (casela de referência)

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

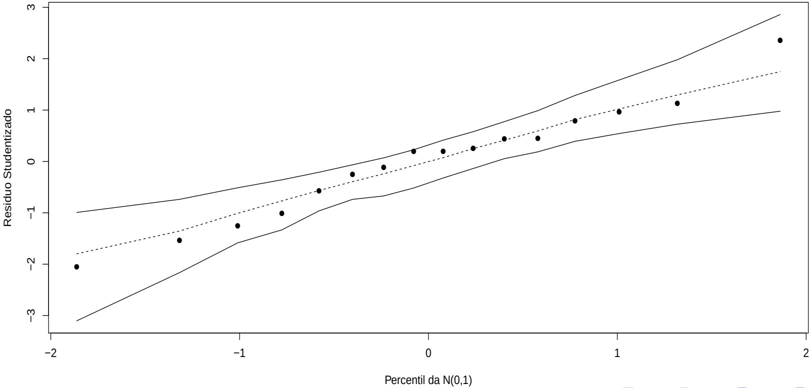
(Fator A), $i = 1, 2$; (Fator B), $j = 1, 2$; (unidades experimentais), $k = 1, 2, 3, 4$

- Erros $\xi_{ijk} \overset{i.i.d}{\sim} N(0, \sigma^2)$, $\mu, \alpha_i, \beta_j, (\alpha\beta)_{ij}$ não aleatório.
- $\mathcal{E}_{\xi_{ijk}}(Y_{ijk}) = \mu_{ij} = \mu + \alpha_i + \beta_j, \mathcal{V}_{\xi_{ijk}}(Y_{ijk}) = \sigma^2$.
- Restrições : $\alpha_1 = \beta_1 = 0, \forall i, j$.
- $Y_{ijk} \overset{ind.}{\sim} N(\beta_j, \sigma^2)$.

Análise de resíduos: modelo reduzido



Gráficos de evelopes



Hipótese de interesse (cont.)

- Aparenten ausência de correlação.
- Aparente normalidade dos resíduos (não se tem subsíduos para se concluir o contrário).
- Presença de heterocedasticidade.

Estimativas dos parâmetros do modelo

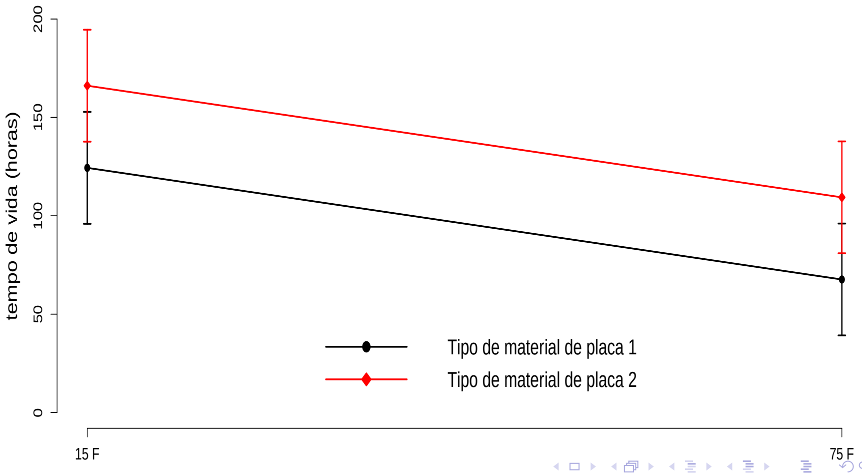
Parâmetro	Estimativa	EP	IC(95%)	Estat. t	pvalor
μ	124,38	13,16	[95,94 ; 152,81]	9,45	<0,0001
α_2	41,75	15,20	[8,91 ; 74,59]	2,75	0,0166
β_2	-56,75	15,20	[-89,59 ; -23,91]	-3,73	0,0025

As estimativas dos parâmetros (α_2, β_2)' confirmam a existência de efeito dos fatores tipo de material e temperatura. Estimaremos as médias via modelo reduzido.

Estimativas finais das médias

Tratamento	Estimativa	EP	IC(95%)
Mater. do tipo 1 - temp. de 15^oF	124,38	13,16	[95,94 ; 152,81]
Mater. do tipo 1 - temp. de 75^oF	67,62	13,16	[39,19 ; 96,06]
Mater. do tipo 2 - temp. de 15^oF	166,12	13,16	[137,69 ; 194,56]
Mater. do tipo 2 - temp. de 75^oF	109,38	13,16	[80,94 ; 137,81]

Perfis médios ajustados pelo modelo reduzido



Comentários

- Apesar dos intervalos de confianças para as médias se interceptarem, eles não incluem as estimativas pontuais do outro grupo.
- Os resultados indicam que, apesar do fato anterior, as médias não são iguais.
- Os reduzidos tamanhos amostrais, por grupo, não permitiram a obtenção de IC's com comprimentos menores. Ou seja, do ponto de visgta estatístico, as diferenças entre as médias não são tão significativas.