

Lista 1 - MAE0327

Guilherme N^oUSP: 8943160 e Leonardo N^oUSP: 9793436

Exercício 1

Um pesquisador de um centro de reabilitação estudou a relação existente entre o status do condicionamento físico de pacientes antes de serem submetidos a uma cirurgia nos joelhos e o tempo (em dias) de terapia física necessário para a completa reabilitação. Seguem abaixo os dados de 24 pacientes do sexo masculino com idades entre 18 e 30 anos submetidos à cirurgia nos joelhos.

Abaixo da média	Na média	Acima da média
29	30	26
42	35	32
38	39	21
40	28	20
43	31	23
40	31	22
30	29	
42	35	
	29	
	33	

(a) Obtenha a tabela de Anova

Resolução

Como temos um conjunto de dados desbalanceado, foi ajustado um modelo de Anova levando este fato em consideração.

Table 1: Tabela de Anova

	Sum Sq	Df	F value	Pr(>F)
(Intercept)	11552	1	583.15385	0.00e+00
Medias	672	2	16.96154	4.13e-05
Residuals	416	21	NA	NA

(b) Teste a hipótese de igualdade de médias, utilizando $\alpha = 6\%$

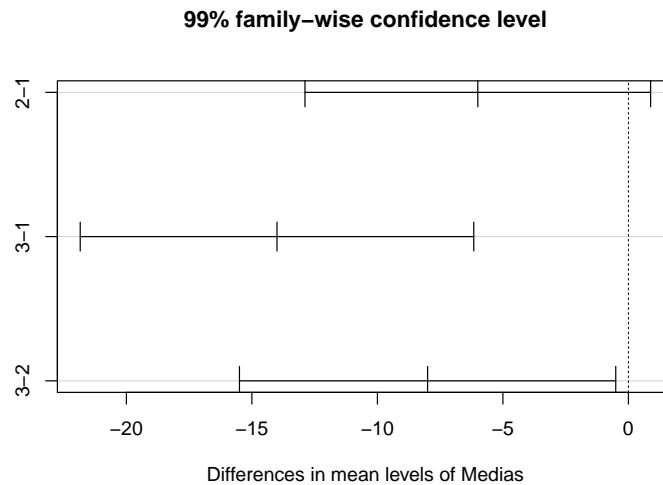
Resolução

Pela tabela 1 de Anova, temos que as médias não são iguais pois o p -value foi de < 0.0001 que ao nível de significância de 6%, nós temos evidências estatísticas para rejeitar a hipótese de igualdade de médias.

- (c) Caso a hipótese de igualdade de médias seja rejeitada, compare as médias 2 a 2 por meio de testes de hipóteses e da construção de intervalos de confiança.

Resolução

Sendo considerado um nível de confiança global de 94%, como temos 3 comparações o nível foi ajustado para 99%.



- 1: Abaixo da média
- 2: Na média
- 3: Acima da média

E como podemos observar apenas as médias de dias dos tratamentos 1 e 2 são iguais, enquanto a média do tratamento 1 é maior que a do tratamento 3 e a média do tratamento 2 é maior que a do tratamento 3.

Exercício 2

Pesquise sobre o teste de Welch e sobre o método de comparações múltiplas de Games-Howell.

- (a) Em qual situação ele deve ser aplicado?

Resolução

O teste de Welch é um teste equivalente a ANOVA, com o diferencial de ser para amostras com variâncias heterocedásticas. Logo as hipóteses para esse teste são:

$$\begin{cases} H_0 : \mu_1 = \mu_2 = \dots = \mu_k \\ H_1 : \text{Pelo menos uma media e diferente} \end{cases}$$

Se os dados são balanceados ($n_1 = n_2 = \dots = n_k$), o teste F da ANOVA é robusto em relação a desigualdade das variâncias e pode ser aplicado.

O método de Games-Howell foi criado para variações e tamanhos de amostras desiguais e é baseado na distribuição estatística q usando range studentizado de Tukey.

(b) O pacote R disponibiliza esses testes? Se sim, em qual pacote?

Resolução

O teste de Welch é encontrado no pacote *onewaytests* e função **welch.test**. O método de Games Howell é encontrado no pacote *userfriendlyscience* e na função **oneway**.

c) Escolha o conjunto de dados adequado (rim direito, rim esquerdo ou coração) e aplique os testes. Não há necessidade de realizar a análise de resíduos novamente. Adote $\alpha = 10\%$

Resolução

Escolhendo o conjunto de dados “Coração”, pois nos slides o teste de levene apresentou um *p-value* muito pequeno, ou seja, rejeitou-se a hipótese de homocedasticidade, assim:

```
##
##   Welch's Heteroscedastic F Test (alpha = 0.1)
## -----
##   data : Coracao and Grupo
##
##   statistic   : 2.22125
##   num df      : 5
##   denom df    : 19.51097
##   p.value     : 0.09333001
##
##   Result      : Difference is statistically significant.
## -----
```

Com um $\alpha = 10\%$ fixado, temos que o teste acima apresentou resultado significativo, iremos realizar comparações múltiplas pelo método de Games Howell:

Table 2: Comparações múltiplas de Games Howell

groups	Mean Difference	Standard Error	t	df	p	upper limit	lower limit
1 : 2	0.020	0.020	0.736	8.923	0.972	0.119	-0.078
1 : 3	-0.015	0.012	0.921	14.472	0.935	0.039	-0.069
1 : 4	0.165	0.045	2.594	6.474	0.223	0.412	-0.082
1 : 5	-0.030	0.013	1.579	16.000	0.623	0.031	-0.090
1 : 6	0.006	0.018	0.244	13.611	1.000	0.092	-0.079
2 : 3	-0.036	0.019	1.317	8.350	0.770	0.062	-0.133
2 : 4	0.145	0.047	2.157	7.831	0.350	0.392	-0.102
2 : 5	-0.050	0.020	1.757	9.798	0.530	0.049	-0.149
2 : 6	-0.014	0.024	0.418	13.843	0.998	0.096	-0.124
3 : 4	0.180	0.045	2.846	6.370	0.168	0.428	-0.067
3 : 5	-0.014	0.013	0.804	16.456	0.963	0.043	-0.071
3 : 6	0.022	0.018	0.855	12.886	0.951	0.105	-0.062
4 : 5	-0.195	0.045	3.042	6.606	0.130	0.052	-0.442
4 : 6	-0.159	0.047	2.391	7.604	0.267	0.087	-0.405
5 : 6	0.036	0.019	1.344	14.969	0.758	0.123	-0.051

A partir das comparações múltiplas observamos que não é possível tomar conclusões se uma média é maior ou menor que outro tratamento, pois todas as comparações possuem o 0 em seu intervalo de confiança.

Exercício 3

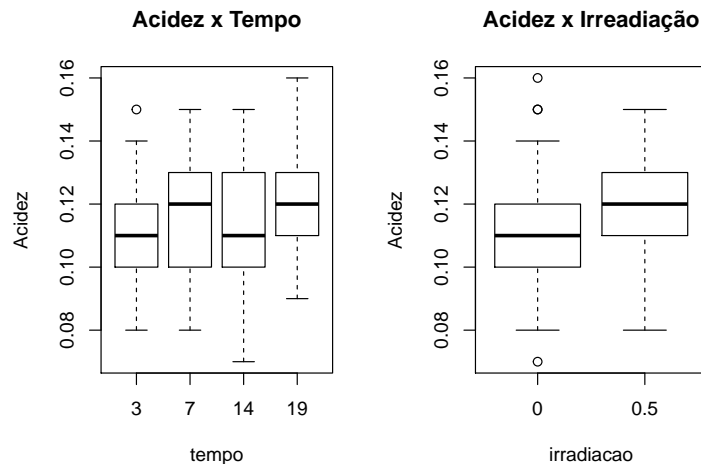
Um estudo (RAE-CEA-04P07) foi planejado com o objetivo de verificar a eficácia da aplicação de distintos níveis de irradiação ionizante em mamões ao longo do tempo. Entre várias variáveis coletadas em cada mamão, destacamos as seguintes:

- Variável resposta de interesse: Acidez titulável (percentual de ácido cítrico em g/mL).
- Fatores ou variáveis explicativas de interesse: irradiação (dosagem de irradiação utilizada no mamão, em kGy: 0 e 0.5) e tempo (número de dias após a irradiação: 3, 7, 14 e 19).

Para cada tratamento, constituído de cada combinação dos níveis dos dois fatores estudados, foram alocados 32 mamões papaia colhidos no mesmo dia. No entanto, alguns mamões foram inutilizados, sendo que o experimento foi realizado com um total de 237 mamões.

- (a) Faça uma análise descritiva das variáveis Acidez titulável levando em consideração os fatores descritos acima.

Resolução



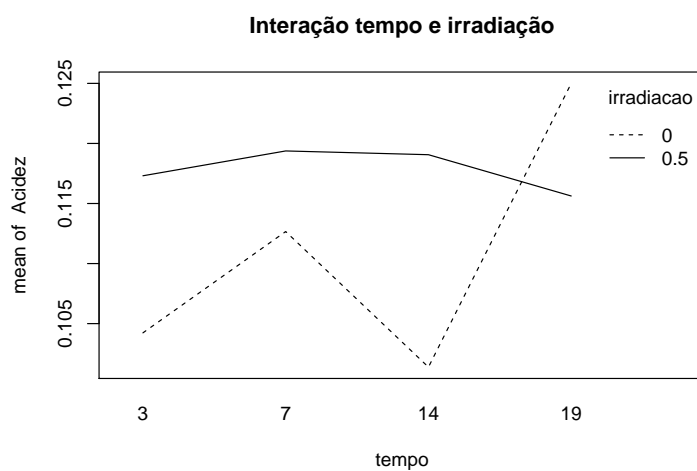
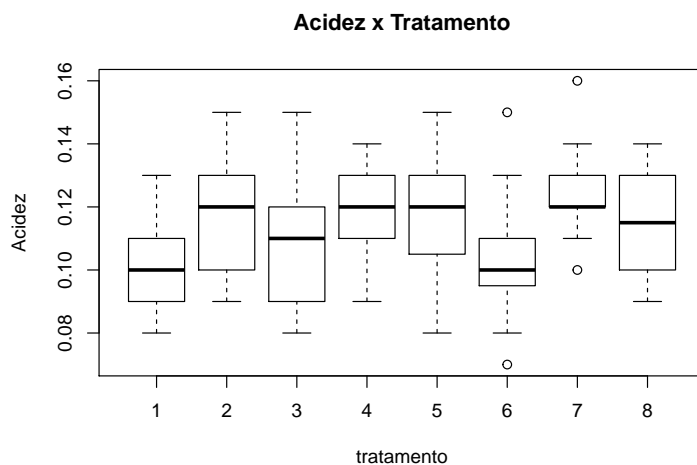


Table 3: Média da Acidez por irradiação

0	0.5
0.111	0.118

Table 4: Média da Acidez por tempo

3	7	14	19
0.111	0.116	0.111	0.12

Table 5: Média da Acidez por tratamento

1	2	3	4	5	6	7	8
0.104	0.117	0.108	0.119	0.119	0.104	0.125	0.116

Table 6: Desvio Padrão da Acidez por irradiação

0	0.5
0.018	0.015

Table 7: Desvio Padrão da Acidez por tempo

3	7	14	19
0.017	0.016	0.019	0.015

Table 8: Desvio Padrão da Acidez por tratamento

1	2	3	4	5	6	7	8
0.015	0.016	0.019	0.012	0.018	0.019	0.013	0.015

Table 9: Mediana da Acidez por irradiação

0	0.5
0.11	0.12

Table 10: Mediana da Acidez por tempo

3	7	14	19
0.11	0.12	0.11	0.12

Table 11: Mediana da Acidez por tratamento

1	2	3	4	5	6	7	8
0.1	0.12	0.11	0.12	0.12	0.1	0.12	0.115

Com a análise descritiva dos dados, a acidez aparentemente, possui menor valor mediano no tempo de 3 e 14 que nos tempos de 7 e 19, a irradiação tem um valor mediano maior para uma dosagem de 0.5 kGy. No gráfico de perfis, as retas não aparecem paralelas, representando um indício que há interação entre os fatores. Pelas demais tabelas observa-se que as médias e as medianas por irradiação, tempo e interação apesar de apresentarem uma diferença, são valores bem próximos um dos outros e os desvios padrões são pequenos, com isso é necessário uma análise mais completa para a tomada de decisões.

- (b) Para estudar as diferenças entre as irradiações e os tempos quanto à média das variáveis Acidez Titulável, foi utilizado um modelo de ANOVA com dois fatores fixos (irradiação e tempo). Escreva o modelo de análise de variância para este problema, obtenha o modelo ajustado e realize as comparações múltiplas de interesse.

Resolução

O modelo de anova para este problema é: $Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk}$

E sua tabela de Anova:

Table 12: Tabela de Anova

	Sum Sq	Df	F value	Pr(>F)
(Intercept)	0.2824654	1	1149.446846	0.0000000
irradiacao	0.0022231	1	9.046449	0.0029260
tempo	0.0098369	3	13.343266	0.0000000
irradiacao:tempo	0.0063649	3	8.633651	0.0000188
Residuals	0.0562745	229	NA	NA

Table 13: Comparações múltiplas de Games Howell

groups	Mean Difference	Standard Error	t	df	p	upper limit	lower limit
1 : 2	0.013	0.003	3.025	49.741	0.070	0.027	-0.001
1 : 3	0.004	0.003	1.015	61.263	0.970	0.017	-0.009
1 : 4	0.015	0.003	4.218	46.402	0.003	0.027	0.004
1 : 5	0.015	0.003	3.419	55.955	0.024	0.028	0.001
1 : 6	0.000	0.004	0.087	26.879	1.000	0.018	-0.019
1 : 7	0.021	0.003	5.571	48.916	0.000	0.033	0.009
1 : 8	0.011	0.003	2.838	54.166	0.107	0.024	-0.001
2 : 3	-0.009	0.003	2.107	58.535	0.423	0.004	-0.022
2 : 4	0.002	0.003	0.548	44.181	0.999	0.014	-0.010
2 : 5	0.002	0.003	0.391	55.434	1.000	0.016	-0.012
2 : 6	-0.014	0.004	2.413	28.486	0.274	0.005	-0.032
2 : 7	0.008	0.003	1.970	46.898	0.512	0.020	-0.005
2 : 8	-0.002	0.003	0.403	52.561	1.000	0.012	-0.015
3 : 4	0.011	0.002	3.152	71.214	0.046	0.022	0.000
3 : 5	0.011	0.003	2.511	67.883	0.208	0.024	-0.003
3 : 6	-0.005	0.004	0.851	26.822	0.988	0.013	-0.022
3 : 7	0.017	0.003	4.577	70.957	0.000	0.028	0.005
3 : 8	0.007	0.003	1.847	71.991	0.591	0.020	-0.005
4 : 5	0.000	0.003	0.082	52.978	1.000	0.012	-0.012
4 : 6	-0.016	0.004	3.077	21.057	0.087	0.001	-0.033
4 : 7	0.006	0.002	1.830	58.845	0.603	0.015	-0.004
4 : 8	-0.004	0.002	1.098	57.598	0.955	0.007	-0.014
5 : 6	-0.015	0.004	2.719	29.308	0.157	0.003	-0.034
5 : 7	0.006	0.003	1.514	55.438	0.797	0.018	-0.006
5 : 8	-0.003	0.003	0.819	60.568	0.991	0.010	-0.017
6 : 7	0.021	0.004	4.105	22.475	0.009	0.038	0.004
6 : 8	0.012	0.004	2.205	25.669	0.381	0.030	-0.006
7 : 8	-0.009	0.003	2.633	58.820	0.165	0.002	-0.021

Na comparações múltiplas, nota-se que as combinações 1:4, 1:5, 1:7, 3:4, 3:7 e 6:7 possuem resultados interessantes, as demais comparações não são possíveis de tomar conclusões se uma média é maior ou menor que outro tratamento. A média do tratamento 1 é menor que os tratamentos 4,5 e 7; o tratamento 3 por sua vez é menor que o 4 e o 7 e por fim o tratamento 6 também é menor que o tratamento 7.

Exercício 4

Foi administrado hormônio de crescimento a crianças com baixa estatura que não atingiram a puberdade. O pesquisador tinha interesse nos efeitos dos fatores A: sexo (masculino, M e feminino, F) e B: desenvolvimento ósseo, sobre a taxa de crescimento induzida pela administração do hormônio. Os níveis considerados do fator B foram: severamente deprimido (SD), moderadamente deprimido (MD) e pouco deprimido (PD). A variável resposta foi a taxa de crescimento, definida como a diferença entre a altura ao final e antes do tratamento, expressa em cm por mês. Três crianças foram selecionadas de cada grupo sexo versus desenvolvimento ósseo. Quatro das 18 crianças não puderam terminar o experimento realizado durante um ano.

Table 14. Taxa de crescimento (cm)

Sexo	Desenvolvimento ósseo					
	SD		MD		PD	
M	1,4	2,4	2,2	2,1	1,7	0,7 1,1
F	2,4		2,5	1,8	2,0	0,5 0,9 1,3

(a) Escreva o modelo para o experimento com suposições e restrições, explicando os termos do mesmo.

Resolução

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk};$$

$$Y_{ijk} \sim N(\mu_{ij}; \sigma^2), \text{ independentes}$$

$$e_{ijk} \sim N(0; \sigma^2), \text{ independentes}$$

Sendo $i = 1, 2, 3$ nível do tratamento de Desenvolvimento ósseo, $j = 1, 2$ nível do tratamento de sexo, α_i efeito do nível i do fator de Desenvolvimento ósseo, β_j efeito do nível j do fator de sexo, γ_{ij} efeito da interação ij entre os fatores, e_{ijk} é o erro aleatório.

b) O modelo parece bem ajustado? Justifique.

Resolução

Para o exercício foi utilizado 1 para o sexo masculino e 2 para o sexo feminino; 1 para severamente deprimido, 2 para moderadamente deprimido e 3 para pouco deprimido. Primeiramente foi feito o modelo completo:

Table 14: Tabela de Anova

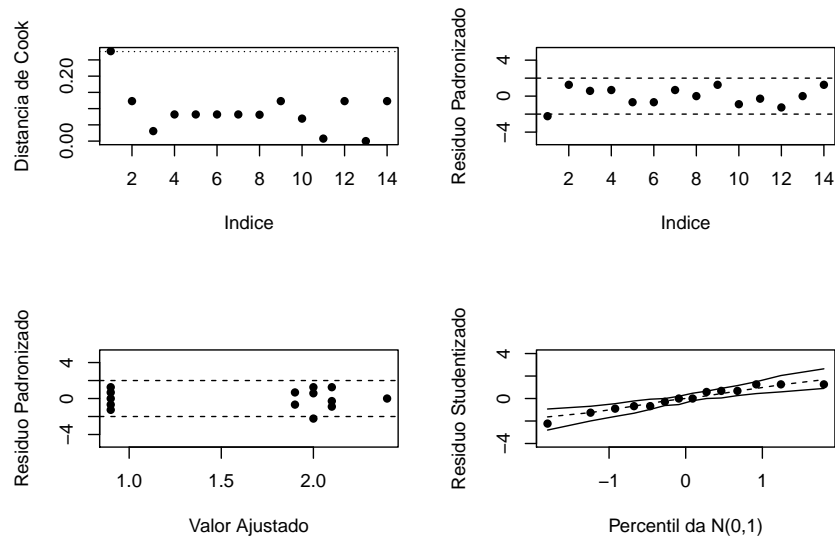
	Sum Sq	Df	F value	Pr(>F)
(Intercept)	12.0000000	1	73.8461538	0.0000260
sexo	0.1200000	1	0.7384615	0.4151604
desenvolvimentoosseo	1.6171429	2	4.9758242	0.0394405
sexo:desenvolvimentoosseo	0.0754286	2	0.2320879	0.7980337
Residuals	1.3000000	8	NA	NA

Neste modelo, o efeito de interação e de sexo aparentam não serem significantes. Portanto foi feito um modelo reduzido com apenas o desenvolvimento ósseo:

Table 15: Tabela de Anova modelo reduzido

	Sum Sq	Df	F value	Pr(>F)
(Intercept)	17.640000	1	132.1798	0.0000002
desenvolvimentoosseo	4.306286	2	16.1339	0.0005355
Residuals	1.468000	11	NA	NA

No modelo reduzido, o efeito de desenvolvimento ósseo realmente possui efeito no crescimento. Realizando o diagnóstico:



```
## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value Pr(>F)
## group 5  0.2373 0.9351
##      8
```

```
##
## Shapiro-Wilk normality test
##
## data:  ts
## W = 0.93523, p-value = 0.3606
```

```
##
## Anderson-Darling normality test
##
## data:  ts
## A = 0.33483, p-value = 0.4567
```

Os resíduos aparentam dentro do esperado, com as suposições de normalidade, independência e homocedasticidade satisfeitas.

c) Quais conclusões você pode extrair da tabela de ANOVA?

Resolução

Como já visto no item anterior é possível afirmar que a taxa de crescimento depende do desenvolvimento ósseo.

- d) Utilize o método de comparações múltiplas de Bonferroni para comparações múltiplas. Quais as suas conclusões?

Resolução

Como apenas o fator desenvolvimento osseo é estatisticamente significativo, iremos realizar as comparações múltiplas de bonferroni para os três níveis com coeficiente de confiança global iguala 95%, assim:

Para contruir o Intervalo de confiança utilizou-se:

$$t_{[1-\alpha/(2g);gl_C]} = t_{[1-0.05/(2*3);11]} = 2.82$$

Para $\mu_1 - \mu_2$ pode ser estimado por $\hat{\alpha}_2 = -0.08$ com desvio padrão de $\sqrt{var(\hat{\alpha}_2)} = 0.245 \Rightarrow IC(\mu_1 - \mu_2) = -0.08 \pm 2.82 * 0.245 \Rightarrow -0.08 \pm 0.691$

Para $\mu_1 - \mu_3$ pode ser estimado por $\hat{\alpha}_3 = -1.2$ com desvio padrão de $\sqrt{var(\hat{\alpha}_3)} = 0.245 \Rightarrow IC(\mu_1 - \mu_3) = -1.2 \pm 2.82 * 0.245 \Rightarrow -1.2 \pm 0.691$

Para $\mu_2 - \mu_3$ pode ser estimado por $\hat{\alpha}_2 - \hat{\alpha}_3 = 1.12$ com desvio padrão de $\sqrt{var(\hat{\alpha}_2 - \hat{\alpha}_3)} = 0.43 \Rightarrow IC(\mu_2 - \mu_3) = 1.12 \pm 2.82 * 0.43 \Rightarrow 1.12 \pm 1.2126$

Exercício 5

Os dados abaixo representam as produções de milho antes (X) e depois (Y) da utilização de 3 tipos de fertilizantes: A_1 , A_2 e A_3 . As unidades experimentais são 21 localidades. Para cada tipo de fertilizante foram escolhidas aleatoriamente 7 localidades.

A_1		A_2		A_3	
X	Y	X	Y	X	Y
3	6	4	8	3	6
1	4	5	9	2	7
3	5	5	7	2	7
1	3	4	9	3	7
2	4	3	8	4	8
1	3	1	5	1	5
4	6	2	7	4	7

- (a) Escreva o modelo de covariância com suposições e restrições.

Resolução

O modelo de Ancova usual é dado por:

$$Y_{ij} = \mu + \alpha_i + \gamma(X_{ij} - \bar{X}) + e_{ij}$$

onde $i = 1, \dots, a$ e $j = 1, \dots, n$. Adotando a parametrização de casela de referência temos:

y_{ij} : valor da variável resposta associada à unidade experimental j submetida ao tratamento i;

μ : Média do grupo 1 (casela de referência);

α_i : efeito fixo de tratamento i, sujeito à restrição $\alpha_1 = 0$;

γ : é o coeficiente de regressão para a relação linear entre Y e X;

X_{ij} : valor da covariável X para a unidade experimental j submetida ao tratamento i;

e_{ij} : erro aleatório associado à unidade experimental j submetida ao tratamento i. Supomos que e_{ij} são independentes com distribuição $N(0, \sigma^2)$.

- (b) Escreva o modelo de regressão equivalente ao modelo de covariância para este problema. Especifique os vetores β e Y e a matriz de planejamento X .

Resolução

Para este problema o modelo de regressão equivalente ao modelo de covariância é:

$$y_{ij} = \mu + \alpha_2 I_2 + \alpha_3 I_3 + \gamma(X_{ij} - \bar{X}) + e_{ij}$$

Onde $i = 1, 2, 3, 4, 5, 6, 7$ e $j = 1, 2, 3$

Fazendo $x_{ij} = X_{ij} - \bar{X}_{..} = X_{ij} - 2,14$ (média do grupo 1), sendo que

I_2 é 1 quando a observação está sob o tratamento 2, 0 caso contrário;

I_3 é 1 quando a observação está sob o tratamento 3, 0 caso contrário.

O vetor $\beta = \begin{pmatrix} \mu \\ I_2 \\ I_3 \\ \gamma \end{pmatrix}$

O vetor $Y^T = (6, 8, 6, 4, 9, 7, 5, 7, 7, 3, 9, 7, 4, 8, 8, 3, 5, 5, 6, 7, 7)$

Matriz de planejamento de X é:

```
##      (Intercept) Fertilizante2 Fertilizante3 FertAntesM
## 1             1             0             0         0.86
## 2             1             1             0         1.86
## 3             1             0             1         0.86
## 4             1             0             0        -1.14
## 5             1             1             0         2.86
## 6             1             0             1        -0.14
## 7             1             0             0         0.86
## 8             1             1             0         2.86
## 9             1             0             1        -0.14
## 10            1             0             0        -1.14
## 11            1             1             0         1.86
## 12            1             0             1         0.86
## 13            1             0             0        -0.14
## 14            1             1             0         0.86
## 15            1             0             1         1.86
## 16            1             0             0        -1.14
## 17            1             1             0        -1.14
## 18            1             0             1        -1.14
## 19            1             0             0         1.86
## 20            1             1             0        -0.14
## 21            1             0             1         1.86
## attr(,"assign")
## [1] 0 1 1 2
## attr(,"contrasts")
## attr(,"contrasts")$Fertilizante
## [1] "contr.treatment"
```

O modelo de regressão ajustado (modelo completo) é dado por:

$$y_{ij} = 4.42 + 2.19I_2 + 1.86I_3 + 0.74x_{ij}$$

- (c) Escreva o modelo reduzido para testar a não existência de efeito de tratamento. Especifique a matriz de planejamento X correspondente a este modelo.

Resolução

Para testar $H_0 : \alpha_2 = \alpha_3 = 0$ contra H_1 : nem todos os α_i são iguais, ajustamos o seguinte modelo de regressão $y_{ij} = \mu + \gamma x_{ij} + e_{ij}$, (modelo reduzido).

Onde o modelo ajustado é:

$$y_{ij} = 5.6 + 1.01x_{ij}$$

E sua matriz de planejamento dada por:

```
##      (Intercept) FertAntesM
## 1             1         0.86
## 2             1         1.86
## 3             1         0.86
## 4             1        -1.14
## 5             1         2.86
## 6             1        -0.14
## 7             1         0.86
## 8             1         2.86
## 9             1        -0.14
## 10            1        -1.14
## 11            1         1.86
## 12            1         0.86
## 13            1        -0.14
## 14            1         0.86
## 15            1         1.86
## 16            1        -1.14
## 17            1        -1.14
## 18            1        -1.14
## 19            1         1.86
## 20            1        -0.14
## 21            1         1.86
## attr(,"assign")
## [1] 0 1
```

- (d) Teste a não existência de efeito de tratamento.

Resolução

Utilizando o modelo ajustado no item anterior, temos:

$$SQR_R = 27.234 \text{ e } gl_R = 19$$

Onde SQR_R é a soma de quadrado dos resíduos do modelo reduzido e SQR_C é a soma de quadrado dos resíduos do modelo completo, gl_R são os graus de liberdade do modelo reduzido e gl_C são os graus de liberdade do modelo completo.

A estatística para o teste de H_0 é dada por:

$$F^* = \frac{SQR_R - SQR_C}{gl_R - gl_C} : QMR_C = \frac{27.234 - 10.30}{19 - 17} = \frac{16.93}{2} = 8.46$$

Sob H_0 $F^* \sim F_{2,17}$ o que conduz a um $p\text{-value} < 0.001$. Assim, há evidências para rejeitarmos H_0 , ou seja, parece haver efeito do fertilizante sobre as produções de milho.

- (e) Estime por ponto e por intervalo as produções médias obtidas após a aplicação dos fertilizantes considerando a produção média inicial igual a 2.76. (Use $\gamma = 0.95$).

Resolução

Para o Fertilizante A_1 a produção média após a aplicação do mesmo é de 4.5, levando em consideração a produção média inicial igual a 2.76, obteve um ganho médio de $4.5 - 2.76 = 1.74$.

Para o Fertilizante A_2 a produção média após a aplicação do mesmo é de 7.57, levando em consideração a produção média inicial igual a 2.76, obteve um ganho médio de $7.57 - 2.76 = 4.81$.

Para o Fertilizante A_3 a produção média após a aplicação do mesmo é de 6.7, levando em consideração a produção média inicial igual a 2.76, obteve um ganho médio de $6.7 - 2.76 = 3.94$.

- (f) Você acha que a suposição de paralelismo é satisfeita neste problema? Justifique. Escreva a estatística adequada ao teste de paralelismo (especifique os modelos envolvidos no cálculo desta estatística).

Resolução

O modelo geral, com diferentes inclinações para as 3 promoções, é dado por

$$y_{ij} = \mu + \alpha_2 I_2 + \alpha_3 I_3 + \gamma x_{ij} + \beta_2 I_2 x_{ij} + \beta_3 I_3 x_{ij} + e_{ij}$$

Para testar $H_0 : \beta_2 = \beta_3 = 0$ contra H_1 : há pelo menos uma diferença, ajustamos o modelo geral, obtendo a seguinte equação:

$$y_{ij} = 4.42 + 2.27 I_2 + 1.94 I_3 + 0.97 x_{ij} - 0.29 I_2 x_{ij} - 0.37 I_3 x_{ij}$$

obtendo $SQR_G = 9.63$, $QMR_G = 0.333$ e $gl_G = 9$. A estatística para o teste de H_1 é dada por:

$$F^* = \frac{SQR_C - SQR_G}{gl_C - gl_G} : QMR_G = \frac{10.3 - 9.63}{17 - 15} = \frac{0.67}{2} = 0.335$$

Onde SQR_C é a soma de quadrado dos resíduos do modelo completo e SQR_R é a soma de quadrado dos resíduos do modelo geral, gl_G são os graus de liberdade do modelo geral e gl_C são os graus de liberdade do modelo completo.

Sob H_0 $F^* \sim F_{2,15}$ o que conduz a um $p\text{-value}=0.6$. Assim, não há evidências para rejeitarmos H_0 , ou seja, as 3 retas podem ser consideradas paralelas.

Códigos

```
# Lista 1 de planejamento II
setwd("/home/gui/Área de Trabalho/Lista de plan2")

library(car)
library(nortest)
library(onewaytests)
source("games.R")

# Exercício 1
```

```

# item a

dados <- read.csv("reab_Plan2", sep=";")
dados <- data.frame(dados$Dias,as.factor(dados$Media))
names(dados) <- c("Dias", "Medias")

fit.model<- lm(Dias~Medias, data=dados)
t.anova <- Anova(fit.model,type="III")

knitr::kable(caption = "Tabela de Anova", t.anova)

# item b

tapply(dados$Dias,dados$Medias,mean)

# item c

fit1.tk <- TukeyHSD(aov(fit.model), "Medias",conf.level = 0.99)
plot(fit1.tk)

# Exercício 2

# item c

dt <- read.csv("coracao.csv",sep=";",dec=',')
dt$Grupo <- as.factor(dt$Grupo)
attach(dt)

welch.test(Coracao~Grupo,data=dt,alpha = 0.1)

tb <- games.howell(Grupo,Coracao)

detach(dt)

# Exercício 3

# item a

dadosm <-read.csv('Mamoes.csv',sep=";",dec="," )
dadosm$tempo <- as.factor(dadosm$tempo)
dadosm$irradiacao <- as.factor(dadosm$irradiacao)
dadosm$tratamento <- as.factor(dadosm$tratamento)

attach(dadosm)

m1 <- tapply(Acidez, irradiacao, mean)
m2 <- tapply(Acidez, tempo, mean)
m3 <- tapply(Acidez, tratamento, mean)

d1 <- tapply(Acidez, irradiacao, sd)
d2 <- tapply(Acidez, tempo, sd)
d3 <- tapply(Acidez, tratamento, sd)

```

```

me1 <- tapply(Acidez, irradiacao, median)
me2 <- tapply(Acidez, tempo, median)
me3 <- tapply(Acidez, tratamento, median)

par(mfrow=c(1,2))
boxplot(Acidez~tempo, main="Acidez x Tempo")
boxplot(Acidez~irradiacao, main="Acidez x Irradiação")

par(mfrow=c(1,1))
boxplot(Acidez~tratamento, main="Acidez x Tratamento")

interaction.plot(tempo, irradiacao, Acidez, main="Interção tempo e irradiação")

# item b

fit.model1<- lm(Acidez~irradiacao*tempo)
Anova(fit.model1,type="III")

tb <- games.howell(tratamento,Acidez)

detach(dadosm)

# Exercício 4

# item b

horm <- read.csv("horm.txt",sep=";")
attach(horm)
sexo <- factor(horm$sexo)
desenvolvimentooosseo <- factor(horm$desenvolvimentooosseo)
tratamento <- factor(horm$tratamento)

fit.modelC<- lm(crescimento~sexo*desenvolvimentooosseo)
Anova(fit.modelC,type="III")

fit.modelI<- lm(crescimento~desenvolvimentooosseo)
Anova(fit.modelI,type="III")

# Diagnostico-----#

X <- model.matrix(fit.modelC)
n <- nrow(X)
p <- ncol(X)
H <- X%*%solve(t(X)%*%X)%*%t(X)
h <- diag(H)
lms <- summary(fit.modelC)
s <- lms$sigma
r <- resid(lms)
ts <- r/(s*sqrt(1-h))
di <- (1/p)*(h/(1-h))*(ts^2)
si <- lm.influence(fit.modelC)$sigma
tsi <- r/(si*sqrt(1-h))
a <- max(tsi)

```

```

b <- min(tsi)
par(mfrow=c(2,2))

plot(di,xlab="Indice", ylab="Distancia de Cook", pch=16)
abline(h=3*mean(di),lty=3)

plot(tsi,xlab="Indice", ylab="Residuo Padronizado",
      ylim=c(-5,5), pch=16)
abline(2,0,lty=2)
abline(-2,0,lty=2)

plot(fitted(fit.modeloC),tsi,xlab="Valor Ajustado",
      ylab="Residuo Padronizado", ylim=c(-5,5), pch=16)
abline(2,0,lty=2)
abline(-2,0,lty=2)

leveneTest(ts,tratamento)

# Testes de normalidade

shapiro.test(ts)

ad.test(ts)

# item d

df <- summary(fit.modeloI)$df[2]
alpha <- 0.05
n.comp <- 3
qt(1-alpha/(2*n.comp),df)

summary(fit.modeloI)
vcov(fit.modeloI)

detach(horm)

# Exercício 5

# item b

df <- read.csv("milho1.txt",sep=";")
df$Fertilizante <- as.factor(df$Fertilizante)
df$FertAntesM <- (df$FertAntes-2.14)
df$Fert2M <- (df$FertAntesM*df$Fert2)
df$Fert3M <- (df$FertAntesM*df$Fert3)
color = c("red", "blue", "black")[df$Fertilizante]

modeloancova <- lm(FertDepois~Fertilizante+FertAntesM,data=df)

plot(df$FertAntes,df$FertDepois,col=color,pch=16)

# Diagnostico-----#

```



```

X <- model.matrix(modeloancova)
n <- nrow(X)
p <- ncol(X)
H <- X%*%solve(t(X)%*%X)%*%t(X)
h <- diag(H)
lms <- summary(modeloancova)
s <- lms$sigma
r <- resid(lms)
ts <- r/(s*sqrt(1-h))
di <- (1/p)*(h/(1-h))*(ts^2)
si <- lm.influence(modeloancova)$sigma
tsi <- r/(si*sqrt(1-h))
a <- max(tsi)
b <- min(tsi)
par(mfrow=c(2,2))
plot(h,xlab="Indice", ylab="Medida h", pch=16, ylim=c(0,1.0))
cut <- 3*p/n
abline(cut,0,lty=2)

plot(di,xlab="Indice", ylab="Distancia de Cook", pch=16)
abline(h=3*mean(di),lty=3)

plot(tsi,xlab="Indice", ylab="Residuo Padronizado",
      ylim=c(-5,5), pch=16)
abline(2,0,lty=2)
abline(-2,0,lty=2)

plot(fitted(modeloancova),tsi,xlab="Valor Ajustado",
      ylab="Residuo Padronizado", ylim=c(-5,5), pch=16)
abline(2,0,lty=2)
abline(-2,0,lty=2)

par(mfrow=c(1,1))

envelope <- function(fit0){
  Influence <- lm.influence(fit0)
  t.M <- resid(fit0)/(Influence$sigma*sqrt(1-Influence$hat))
  n <- length(t.M)
  X <- model.matrix(fit0)
  H <- X%*%solve(t(X)%*%X)%*%t(X)
  I <- diag(n)
  yy <- matrix(rnorm(n*100),n,100)
  A <- matrix(0,n,100)
  LI = LS = numeric(n)
  A <- (I - H)%*%yy/sqrt(diag(I - H))
  A <- apply(A,2,sort)
  B <- t(apply(A,1,sort))
  LI <- (B[,2]+B[,3])/2
  LS <- (B[,97]+B[,98])/2
  med <- apply(B,1,mean)
  aux <- range(t.M,LS,LI)
  #par(pty="s")
  qqnorm(t.M,xlab="Percentil da N(0,1)", ylab="Resíduo Studentizado modificado", ylim=aux,pch=16, main=

```

```

par(new=TRUE)
qqnorm(LI,axes=F,xlab="",ylab="",type="l",lty=1,ylim=aux)
par(new=TRUE)
qqnorm(LS,axes=F,xlab="",ylab="", type="l", lty=1,ylim=aux)
par(new=TRUE)
qqnorm(med,axes=F,xlab="",ylab="",type="l",ylim=aux,lty=2)
}

envelope(modeloancova)

summary(modeloancova)
Anova(modeloancova,type="III")
m.vcov.estparancova<- vcov(modeloancova)
m.vcov.estparancova

# item c

modres <- lm(FertDepois~FertAntesM,data=df)

model.matrix(modres)

# item d

modres.anova <- anova(modres)

modcom <- lm(FertDepois~Fert2+Fert3+FertAntesM,data=df)
modcom.anova <- anova(modcom)

SQP <- modres.anova$Sum[2]-modcom.anova$Sum[4]
SQP
GLP <- modres.anova$Df[2]-modcom.anova$Df[4]
GLP
QMP <-SQP/GLP
QMP
FcalcP <- QMP/modcom.anova$Mean[4]
FcalcP
pvalorP <- 1 - pf(QMP, GLP, modcom.anova$Df[4])
pvalorP

# item f

fit1 <- lm(FertDepois ~ Fert2 + Fert3 + FertAntesM + Fert2M + Fert3M, data=df)
summary(fit1)

fit1.anova <- anova(fit1)
fit1.anova

fit1.anova$Df
fit1.anova$Sum

SQParalel <- modcom.anova$Sum[4]-fit1.anova$Sum[6]
SQParalel
GLParalel <- modcom.anova$Df[4]-fit1.anova$Df[6]

```

```
GLParalel
QMParalel <-SQParalel/GLParalel
QMParalel
FcalcParalel <- QMParalel/fit1.anova$Mean[6]
FcalcParalel
pvalorParalel <- 1 - pf(FcalcParalel, GLParalel, fit1.anova$Df[6])
pvalorParalel
```