

Planejamento e Pesquisa (ME 623A)
Segundo semestre de 2012
Prova I
Data: 11/09/2012

Nome: _____ RA: _____

Leia atentamente as instruções abaixo:

- Coloque seu nome completo e RA em todas as folhas que você recebeu, inclusive nesta.
- Utilize somente um dos lados de cada folha.
- Leia atentamente cada uma das questões.
- Enuncie, claramente, todos os resultados que você utilizar.
- Justifique, adequadamente, seus desenvolvimentos, sem, no entanto, escrever excessivamente.
- O(a) aluno(a) só poderá sair da sala após as 16h30, mesmo que já tenha finalizado a prova. Após a saída do(a) primeiro(a) aluno(a) não será permitido a entrada de nenhum(a) outro(a) aluno(a).
- Não é permitido empréstimo de material.
- Não serão dirimidas dúvidas de quaisquer natureza, após os 20 minutos iniciais.
- Resolva a prova, preferencialmente, à caneta, e procure ser organizado(a). Se fizer à lápis, destaque, à caneta, sua resposta.
- O(a) aluno(a) deverá portar sua carteira de estudante e apresentá-la, quando for solicitada sua assinatura.
- Contestações a respeito da nota, só serão consideradas se estiverem por escrito.
- A nota do aluno(a) será $\frac{NP}{NT} \times 10$, em que NP é o número de pontos obtidos na prova e NT é o número total de pontos da prova.
- Os resultados numéricos finais devem ser apresentados com duas casas decimais, apenas.
- A prova terá duração de 120 minutos, das 16h às 18h, improrrogáveis.

Faça uma excelente Prova!!

1. Considere um PCA com um único fator com k grupos e n_i observações por grupo, $i = 1, \dots, k$. Além disso, denote $n = \sum_{i=1}^k n_i$ (o número total de observações). Considere que $H_0 : \mu_1 = \mu_2 = \dots = \mu_k = \mu$ vs $H_1 : \text{há pelo menos uma diferença}$.

a) Escreva o modelo visto em classe, nas formas escalar e matricial, para analisar dados oriundos de tal experimento, sob a parametrização casela de referência (CR). Não se esqueça de escrever todas as suposições do modelo (200 pontos).

b) Prove que $\text{SQT} = \text{SQF} + \text{SQR}$, ou seja, prove que:

$$\sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y})^2 = \sum_{i=1}^k n_i (\bar{Y}_i - \bar{Y})^2 + \sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2.$$

Além disso, prove que, sob H_0 , $Q = \frac{1}{\hat{\sigma}^2} \sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - \mu)^2 \sim \chi^2_{(n)}$ (250 pontos).

c) Considere as seguintes hipóteses, $H_0 : \mathbf{C}_{(r \times p)} \boldsymbol{\beta}_{(p \times 1)} = \mathbf{0}_{(r \times 1)}$ vs $H_1 : \mathbf{C} \boldsymbol{\beta} \neq \mathbf{0}$, onde $\boldsymbol{\beta}$ é o vetor de parâmetros associado ao modelo normal linear sob a parametrização CR. Demonstre, **apenas apresentando justificativas para seus argumentos**, que se

$$Q = \frac{1}{r \hat{\sigma}^2} (\mathbf{C} \hat{\boldsymbol{\beta}})' (\mathbf{C} (\mathbf{X}' \mathbf{X})^{-1} \mathbf{C}')^{-1} (\mathbf{C} \hat{\boldsymbol{\beta}})$$

então, sob H_0 , $Q \sim F_{(q, n-p)}$, em que $\hat{\sigma}^2 = \text{QMR}$, QMR: quadrado médio residual (250 pontos).

2. A força de tensão do cimento “portland” está sendo estudada. Quatro diferentes técnicas de misturas podem ser usadas economicamente. Quanto maior a força de tensão, melhor a técnica de mistura. Os seguintes dados de um experimento inteiramente casualizado foram coletados:

Técnica de Mistura	Força de tensão (observação)			
	1	2	3	4
1	3129	3000	2865	2890
2	3200	3300	2975	3150
3	2800	2900	2985	3050
4	2600	2700	2600	2765

Abaixo encontram-se diversos resultados, incluindo: análises descritivas, testes de homogeneidade de variâncias, ajuste de um modelo normal linear completo (considerando os 4 grupos), sob parametrização casela de referência (técnica de mistura “1” é a referência), e análise de resíduos. Os grupos 1, 2, 3 e 4 correspondem, respectivamente, às técnicas de mistura 1, 2, 3 e 4.

Tabela 1: Análise descritiva dos dados da Questão 2

Técnica de mistura	Medida descritiva					
	Média	DP	Var.	CV%	Mínimo	Máximo
1	2971,00	120,56	14534,00	4,06	2865,00	3129,00
2	3156,25	135,98	18489,58	4,31	2975,00	3300,00
3	2933,75	108,27	11722,92	3,69	2800,00	3050,00
4	2666,25	80,97	6556,25	3,04	2600,00	2765,00

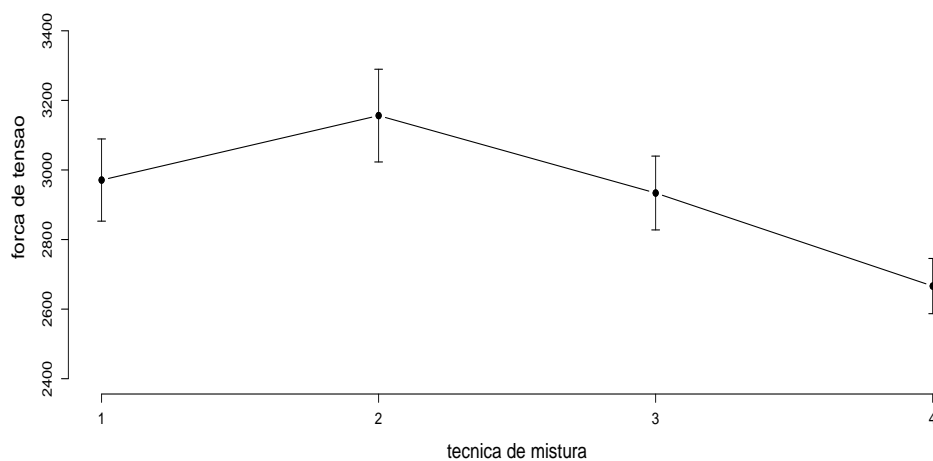


Figura 1: Gráficos de perfis médios para os dados da Questão 2

Teste de Bartlett = 3,772 (pvalor = 0,4378). Teste de Levene = 0,696 (pvalor=0,6033).

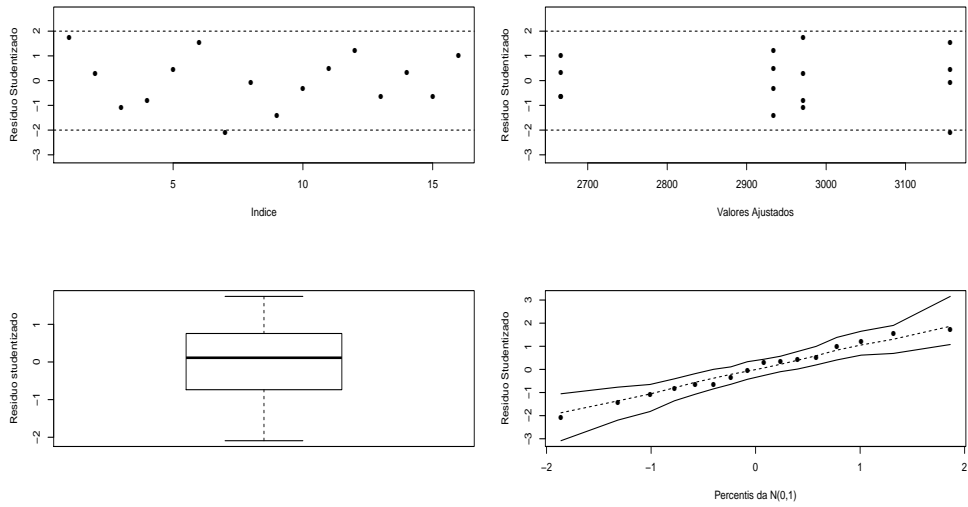


Figura 2: Análise de resíduos para o modelo completo para os dados da Questão 2

Tabela 2: Análise de variância para os dados da Questão 2

FV	SQ	GL	QM	Estatística F	pvalor
Técnica de Mistura	489740,00	3	163247,00	12,73	0,0004
Resíduo	153908,00	12	12826,00		
Total	643648,40	15			

Tabela 3: Estimativas pontuais, erros-padrão e IC de 95% para os parâmetros do modelo relativo aos dados da Questão 2

Parâmetro	Estimativa	EP	IC(95%)	Estat. t	pvalor
μ (grupo 1)	2971,00	56,63	[2860,01; 3081,99]	52,468	< 0,0001
α_2 (grupo 2)	185,25	80,08	[28,29;342,21]	2,313	0,0392
α_3 (grupo 3)	-37,25	80,08	[-194,21;119,71]	-0,465	0,6501
α_4 (grupo 4)	-304,75	80,08	[-461,71;-147,79]	-3,806	0,0025

Estimativa da matriz de covariâncias dos parâmetros do modelo (Questão 2):

$$\tilde{\sigma}^2 (\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} 3206,42 & -3206,42 & -3206,42 & -3206,42 \\ -3206,42 & 6412,84 & 3206,42 & 3206,42 \\ -3206,42 & 3206,42 & 6412,84 & 3206,42 \\ -3206,42 & 3206,42 & 3206,42 & 6412,84 \end{bmatrix}$$

Responda os itens:

- Defina quem são: fator de interesse (quantos e quem são seus níveis), a variável resposta e o número de unidades experimentais por nível do fator (100 pontos).
- Escreva o modelo que foi ajustado ao conjunto de dados (com todas as suposições pertinentes), de acordo com o que foi dito acima, para comparar os tratamentos (níveis dos fatores) sob a parametrização casela de referência (CR) (150 pontos).
- Em relação à hipótese de homocedasticidade, usando os testes de Levene e Bartlett, qual sua conclusão ao nível de significância de $\alpha = 0,05$? Unindo sua conclusão à análise descritiva, você acha razoável a validade da hipótese de homocedasticidade? Justifique, adequadamente, sua resposta (150 pontos).
- O que você pode afirmar sobre a validade das suposições do modelo para o conjunto de dados em questão, utilizando os resultados da análise residual e dos testes de homocedasticidade? Justifique, adequadamente, sua resposta (200 pontos).
- Qual seria sua conclusão à respeito da hipótese de igualdade de médias, através da tabela ANOVA? E em relação à tabela com as estimativas dos parâmetros do modelo? Justifique, adequadamente, sua resposta (200 pontos).
- Com base em todos os resultados apresentados, seus comentários e conclusões relacionados ao item e), estabeleça e teste, ao menos uma hipótese, na estrutura $H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{0}$ vs $H_1 : \mathbf{C}\boldsymbol{\beta} \neq \mathbf{0}$. Quais foram suas conclusões em relação à esse(s) teste(s)? Justifique, adequadamente, suas afirmações (250 pontos).
- Com base em todos os resultados apresentados e obtidos, proponha, se for o caso, um modelo reduzido (com todas as suposições pertinentes). Até o momento, ou seja, antes de ajustar um modelo reduzido, quais suas conclusões com respeito ao desempenho das técnicas de mistura? Justifique, adequadamente, seus comentários (250 pontos).

3. O rendimento de um processo químico está sendo estudado. As duas variáveis mais importantes, acredita-se, são a pressão e a temperatura. Dois níveis de cada fator são selecionados, e um experimento fatorial completamente casualizado com duas repetições é realizado. Quanto maior o valor do rendimento, melhor o processo químico. Os dados de rendimento seguem:

Temperatura($^{\circ}C$)	Pressão (psig)	
	200	215
150	90,4	90,7
	90,2	90,6
160	90,1	90,5
	90,3	90,6

Abaixo encontram-se diversos resultados, incluindo análises descritivas, testes de homogeneidade de variâncias, ajuste de um modelo normal linear, sob parametrização casela de referência, incluindo interação entre os fatores (o tratamento pressão = 200 e temperatura = 150 C é o grupo de referência), e análise de resíduos. O fator A é a “pressão” e o fator B a “temperatura”.

Tabela 4: Análise descritiva dos dados da Questão 3

Pressão	Temperatura	Medida descritiva					
		Média	DP	Var.	CV%	Mínimo	Máximo
200	150 C	90,30	9,50	0,02	0,16	90,20	90,40
	160 C	90,20	9,50	0,020	0,16	90,10	90,30
215	150 C	90,65	9,52	0,01	0,08	90,60	90,70
	160 C	90,55	9,52	0,01	0,08	90,50	90,60

Teste de Bartlett = 0,6301 (pvalor = 0,8895).

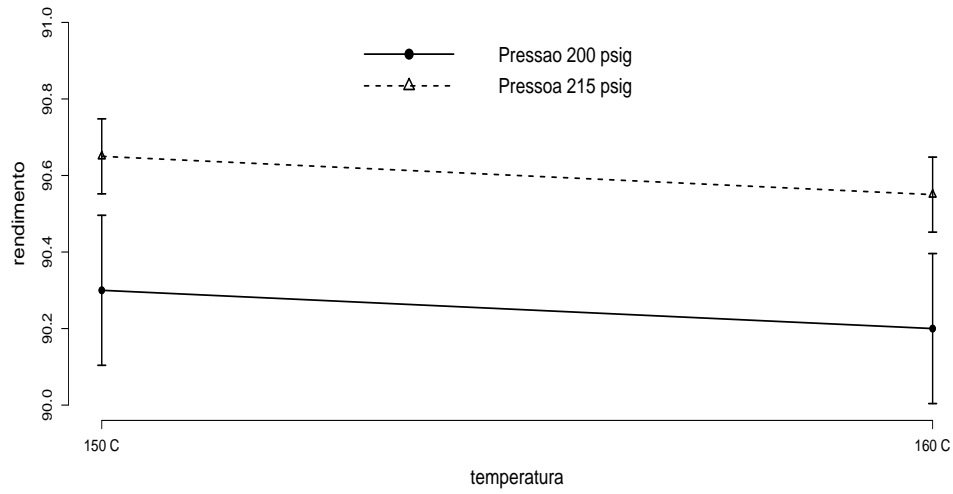


Figura 3: Gráficos de perfis médios para os dados da Questão 3

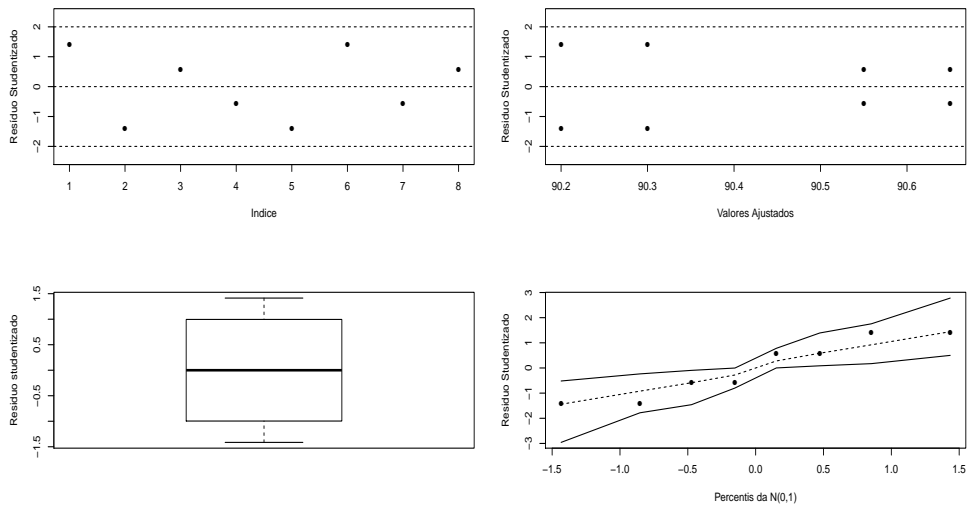


Figura 4: Análise de resíduos para o modelo completo para os dados da Questão 3

Tabela 5: Análise de variância para os dados da Questão 3

FV	SQ	GL	QM	Estatística F	pvalor
Pressão	0,25	1	0,25	19,6	0,01145
Temperatura	0,02	1	0,02	1,60	0,27458
Interação	< 0,01	1	<0,01	< 0,01	>0,9999
Resíduo	0,05	4	0,01		
Total	0,315	7			

Tabela 6: Estimativas pontuais, erros-padrão e IC de 95% para os parâmetros do modelo relativo aos dados da Questão 3

Parâmetro	Estimativa	EP	IC(95%)	Estat. t	pvalor
μ	90,30	0,08	[90,15;90,45]	1142,22	< 0,0001
α_2 (pressão)	0,35	0,11	[0,13 ;0,57]	3,13	0,0352
β_2 (temperatura)	-0,10	0,11	[-0,32;0,12]	-0,90	0,4216
$(\alpha\beta)_{22}$	$\approx 0,00$	0,16	[-0,31;0,31]	< 0,01	> 0,9999

Estimativa da matriz de covariâncias dos parâmetros do modelo (Questão 3):

$$\tilde{\sigma}^2 (\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} 0,00625 & -0,00625 & -0,00625 & 0,00625 \\ -0,00625 & 0,01250 & 0,00625 & -0,01250 \\ -0,00625 & 0,00625 & 0,01250 & -0,01250 \\ 0,00625 & -0,01250 & -0,01250 & 0,02500 \end{bmatrix}$$

Responda os itens:

- Defina quem são: fatores de interesse (quantos e quem são seus níveis), a variável resposta e o número de unidades experimentais por nível de tratamento (100 pontos).
- Escreva o modelo que foi ajustado ao conjunto de dados (com todas as suposições pertinentes), de acordo com o que foi dito acima, para comparar os tratamentos sob a parametrização casela de referência (CR). O que você deduz, através somente do gráfico de perfis, sobre a existência de interação e dos efeitos principais? (200 pontos)
- Em relação à hipótese de homocedasticidade, usando o teste de Bartlett, qual sua conclusão ao nível de significância de $\alpha = 0,05$? Unindo sua conclusão à análise descritiva, você acha razoável a validade da hipótese de homocedasticidade? Justifique, adequadamente, sua resposta (150 pontos).
- O que você pode afirmar sobre a verificação das suposições do modelo para o conjunto de dados em questão, utilizando os resultados da análise residual e do teste de homocedasticidade? Justifique, adequadamente, sua resposta (200 pontos).
- Qual seria sua conclusão à respeito da existência de interação e dos efeitos principais de temperatura e pressão, através da tabela ANOVA? E em relação à tabela com as estimativas dos parâmetros do modelo? Justifique, adequadamente, sua resposta (250 pontos).
- Com base em todos os resultados apresentados, seus comentários e conclusões relacionados ao item e), estabeleça e teste, ao menos uma hipótese, na estrutura $H_0 : \mathbf{C}\beta = \mathbf{0}$ vs $H_1 : \mathbf{C}\beta \neq \mathbf{0}$. Quais

foram suas conclusões em relação à esse(s) teste(s)? Justifique, adequadamente, suas afirmações (250 pontos).

- g) Com base em todos os resultados apresentados e obtidos, proponha, se for o caso, um modelo reduzido (com todas as suposições pertinentes). Até o momento, ou seja, antes de ajustar um modelo reduzido, quais suas conclusões com respeito ao rendimento do processo químico em questão, em função da temperatura e pressão? Justifique, adequadamente, seus comentários (250 pontos).

Siglas

- PCA: planejamento completamente casualizado.
- SQT: Soma de quadrados total.
- SQF: Soma de quadrados dos fatores.
- SQR: Soma de quadrados dos resíduos.