



PRO2271 ESTATÍSTICA

7. Análise de Variância

Escola Politécnica da Universidade de São Paulo | Departamento de Engenharia de Produção



Terminologia em A.V.

Experimento:

- procedimento para avaliar o efeito de um ou mais fatores em uma variável de resposta

Fator: variável controlada de um experimento

Variável de Resposta

- variável principal do estudo, considerada normal, com variância constante e média sujeita à influência dos diferentes fatores

Análise de Variância

- variância das médias entre os níveis de cada fator
- análise com um e dois fatores (com e sem repetição)
- múltiplos fatores: técnicas de Delineamento de Experimentos

Escola Politécnica da Universidade de São Paulo | Departamento de Engenharia de Produção

2



Análise de Variância - Um Fator



Objetivo:

- verificar se existem diferenças significativas entre as médias para os diferentes níveis do fator

Hipóteses:

- H_0 : não há diferença entre as médias
- H_1 : pelo menos uma média é diferente das demais

Dados:

	Nível 1	Nível 2	Nível 3
	X_{11}	X_{21}	X_{31}
	X_{12}	X_{22}	X_{32}
	X_{13}	X_{23}	X_{33}
	X_{14}	X_{24}	X_{34}
	X_{15}	X_{25}	X_{35}

Escola Politécnica da Universidade de São Paulo

Departamento de Engenharia de Produção

3



Exemplo 1

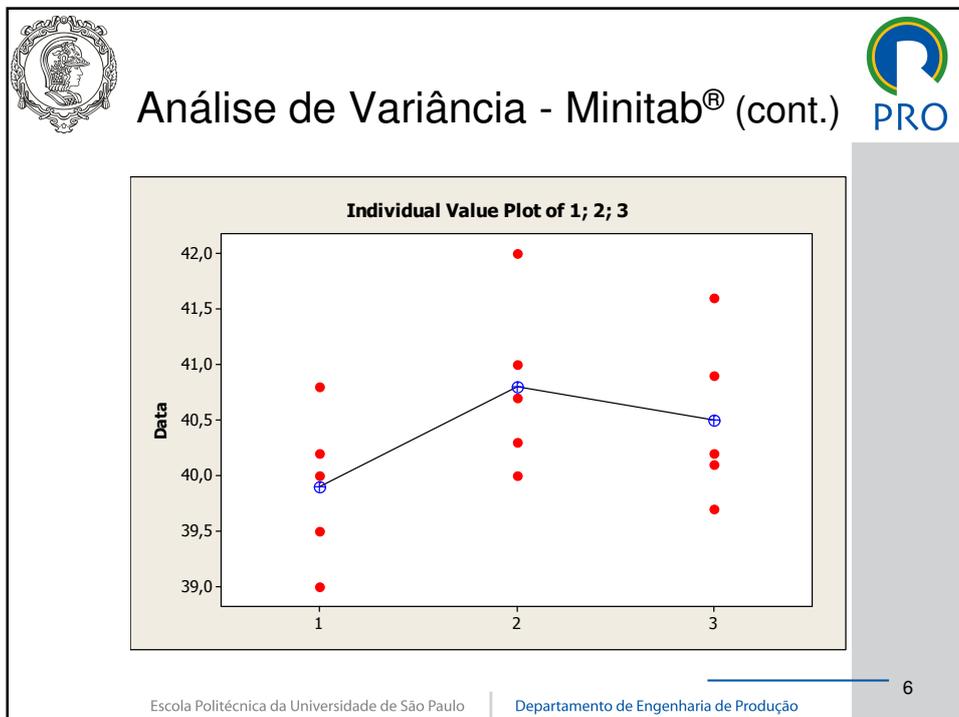
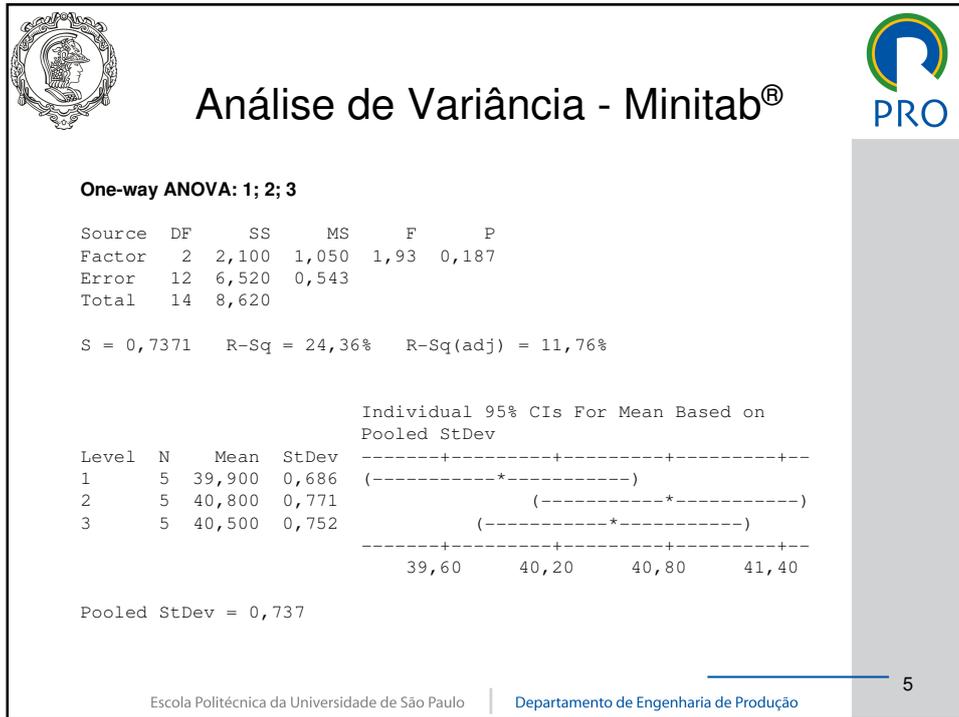


	1	2	3
1	39,5	40,7	41,6
2	39,0	41,0	40,1
3	40,2	42,0	40,9
4	40,8	40,0	39,7
5	40,0	40,3	40,2
\bar{x}_i	39,9	40,8	40,5
s_i^2	0,470	0,595	0,565

Escola Politécnica da Universidade de São Paulo

Departamento de Engenharia de Produção

4





Análise de Variância - Um Fator (cont.)



Estatísticas

- média e variância das amostras: \bar{x}_i e s_i^2
- variância do fator A: $s_A^2 = n \cdot s_{\bar{X}}^2$
- variância residual: $s_R^2 = \frac{s_1^2 + s_2^2 + \dots + s_m^2}{m}$
- razão das variâncias: $F_{m-1, m(n-1)} = \frac{s_A^2}{s_R^2}$

Rejeitar H_0 (médias iguais) se: $F_{m-1, m(n-1)} > F_{m-1, m(n-1), \alpha}$

Escola Politécnica da Universidade de São Paulo
Departamento de Engenharia de Produção

7



Exemplo 1 (cont.)



	1	2	3
1	39,5	40,7	41,6
2	39,0	41,0	40,1
3	40,2	42,0	40,9
4	40,8	40,0	39,7
5	40,0	40,3	40,2
\bar{x}_i	39,9	40,8	40,5
s_i^2	0,470	0,595	0,565

$\bar{\bar{x}} = 40,40$

$H_0: \mu_1 = \mu_2 = \mu_3$

$m = 3 \quad n = 5 \quad \alpha = 5\%$

$s_A^2 = 1,05$

$s_R^2 = 0,543$

$F_{2,12} = 1,93 \quad F_{2,12,5\%} = 3,89$

como $F < F_{5\%} \Rightarrow$ aceitar H_0

Escola Politécnica da Universidade de São Paulo
Departamento de Engenharia de Produção

8




Procedimento padrão de cálculo

Notação:

- m: qtd de amostras (níveis)
- n: tamanho das amostras
- STQ: soma total de quadrados
- SQA: soma de quadrados para o fator A
- SRQ: soma residual de quadrados

Cálculo das Variâncias:

$$s_A^2 = \frac{SQA}{m-1} \quad s_R^2 = \frac{SRQ}{m(n-1)} \quad s_T^2 = \frac{STQ}{mn-1}$$

Escola Politécnica da Universidade de São Paulo | Departamento de Engenharia de Produção

9




Procedimento de cálculo (cont.)

Somas de Quadrados:

$$SQA = n \cdot \sum_{i=1}^m (\bar{x}_i - \bar{\bar{x}})^2 \quad s_A^2 = \frac{SQA}{(m-1)}$$

$$SRQ = \sum_{i=1}^m \sum_{j=1}^n (x_{ij} - \bar{x}_i)^2 \quad s_R^2 = \frac{SRQ}{m(n-1)}$$

$$STQ = \sum_{i=1}^m \sum_{j=1}^n (x_{ij} - \bar{\bar{x}})^2$$

Escola Politécnica da Universidade de São Paulo | Departamento de Engenharia de Produção

10



Procedimento de cálculo (cont.)



Fórmulas:

$$STQ = Q - \frac{T^2}{mn} \qquad SQA = \frac{\sum_{i=1}^m T_i^2}{n} - \frac{T^2}{mn}$$

$$SRQ = STQ - SQA$$

onde:

- T_i : soma dos valores da amostra i
- T : soma total dos valores
- Q_i : soma dos quadrados da amostra i
- Q : soma total dos quadrados

Escola Politécnica da Universidade de São Paulo

Departamento de Engenharia de Produção

11



Procedimento de cálculo (cont.)



Tabela da A.V.

Fonte de Variação	Soma de Quadrados	Graus de Liberdade	Quadrados Médios	F	F α
Fator A	SQA	m-1	$s_A^2 = \frac{SQA}{(m-1)}$	s_A^2 / s_R^2	$F_{m-1, m(n-1), \alpha}$
Residual	SRQ	m(n-1)	$s_R^2 = \frac{SRQ}{m(n-1)}$		
Total	STQ	mn-1			

Uso de softwares para A.V. (Minitab e Excel)

Escola Politécnica da Universidade de São Paulo

Departamento de Engenharia de Produção

12




Exemplo 1 (cont.)

	1	2	3
1	39,5	40,7	41,6
2	39,0	41,0	40,1
3	40,2	42,0	40,9
4	40,8	40,0	39,7
5	40,0	40,3	40,2
T_i	199,50	204,00	202,50
Q_i	7961,93	8325,58	8203,51
T_i^2	39800,25	41616,0	41006,25

$H_0 : \mu_1 = \mu_2 = \mu_3$

$m = 3 \quad n = 5$

$T = 606,0 \quad Q = 24.491,02$

$\frac{\sum T_i^2}{n} = 24.484,50$

$\frac{T^2}{mn} = 24.482,40$

$STQ = 8,62 \quad SQA = 2,10$

$SRQ = 6,52$

Escola Politécnica da Universidade de São Paulo

Departamento de Engenharia de Produção

13




Exemplo 1 (cont.)

Fonte de Variação	Soma de Quadrados	Graus de Liberdade	Quadrados Médios	F	F _{5%}
Fator A	2,10	2	1,05	1,93	3,89
Residual	6,52	12	0,543		
Total	8,62	14			

Como $F < F_{5\%}$, aceita-se H_0 , ou seja, não há evidência estatística para afirmar que o Fator A influencia a média da variável de resposta.

Escola Politécnica da Universidade de São Paulo

Departamento de Engenharia de Produção

14



Exemplo 1 – Excel



Anova: fator único

RESUMO

Grupo	Contagem	Soma	Média	Variância
1	5	199,5	39,9	0,47
2	5	204	40,8	0,595
3	5	202,5	40,5	0,565

ANOVA

Fonte da variação	SQ	gl	MQ	F	valor-P	F crítico
Entre grupos	2,10	2	1,050	1,93	0,187	3,89
Dentro dos grupos	6,52	12	0,543			
Total	8,62	14				

Escola Politécnica da Universidade de São Paulo

Departamento de Engenharia de Produção

15



Método de Tukey



Quando se rejeita a Hipótese H_0 (igualdade das médias), deve-se identificar quais são as médias diferentes.

Há vários métodos de comparação das médias. Abaixo, apresenta-se o método de Tukey para A.V. com um único fator e amostras de mesmo tamanho.

Procedimento:

1. Calcular $e = g_{m, m(n-1), \alpha} \sqrt{S_R^2 / n}$
2. Ordenar as médias e grifar diferenças menores que “e”.
3. As médias grifadas serão consideradas iguais.

t de student modificada (tabela)

Escola Politécnica da Universidade de São Paulo

Departamento de Engenharia de Produção

16



Análise de Variância: 2 fatores, sem repetição



Hipóteses:

- H_{0A} : não há diferenças entre as médias do fator A
- H_{1A} : o fator A influencia a média da variável de resposta
- H_{0B} : não há diferenças entre as médias do fator B
- H_{1B} : o fator B influencia a média da variável de resposta

Dados:

		B			
		1	2	3	4
A	1	X_{11}	X_{12}	X_{13}	X_{14}
	2	X_{21}	X_{22}	X_{23}	X_{24}
	3	X_{31}	X_{32}	X_{33}	X_{34}

Escola Politécnica da Universidade de São Paulo

Departamento de Engenharia de Produção

17



A.V. com 2 fatores, sem repetição



Fórmulas:

$$SQA = \frac{\sum_{i=1}^m T_i^2}{n} - \frac{T^2}{mn} \quad SQB = \frac{\sum_{j=1}^n T_j^2}{m} - \frac{T^2}{mn}$$

$$STQ = Q - \frac{T^2}{mn} \quad SRQ = STQ - SQA - SQB$$

$$s_A^2 = \frac{SQA}{m-1} \quad s_B^2 = \frac{SQB}{n-1} \quad s_R^2 = \frac{SRQ}{(m-1)(n-1)}$$

$$F_A = \frac{s_A^2}{s_R^2} \quad F_B = \frac{s_B^2}{s_R^2}$$

Escola Politécnica da Universidade de São Paulo

Departamento de Engenharia de Produção

18




A.V. com 2 fatores, sem repetição - Tabela

Tabela da A.V. com 2 fatores (sem interação)

Fonte de Variação	Soma de Quadrados	Graus de Liberdade	Quadrados Médios	F	F_α
Fator A	SQA	m-1	s_A^2	F_A	$F_{m-1, (m-1)(n-1), \alpha}$
Fator B	SQB	n-1	s_B^2	F_B	$F_{n-1, (m-1)(n-1), \alpha}$
Residual	SRQ	(m-1)(n-1)	s_R^2		
Total	STQ	mn-1			

Próximo passo: Análise de Variância com dois fatores e repetição, que permite avaliar, além da influência individual de cada fator, se há interação entre os fatores.

19

Escola Politécnica da Universidade de São Paulo | Departamento de Engenharia de Produção




Tabela F

v \ u	1	2	3	4	5	6	7	8	9
1	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59
·	·	·	·	·	·	·	·	·	·
·	·	·	·	·	·	·	·	·	·

20

Escola Politécnica da Universidade de São Paulo | Departamento de Engenharia de Produção

EXERCÍCIOS

- 1) Amostras de um lubrificante foram obtidas em quatro fornecedores e analisadas em laboratório. Os resultados apresentados na Tabela 7.1 constituem as viscosidades em cada amostra. Pode-se afirmar, ao nível de 5% de significância, que há diferenças entre os fornecedores quanto à viscosidade do lubrificante produzido?

Tabela 7.1 Resultados das análises de viscosidade nas amostras.

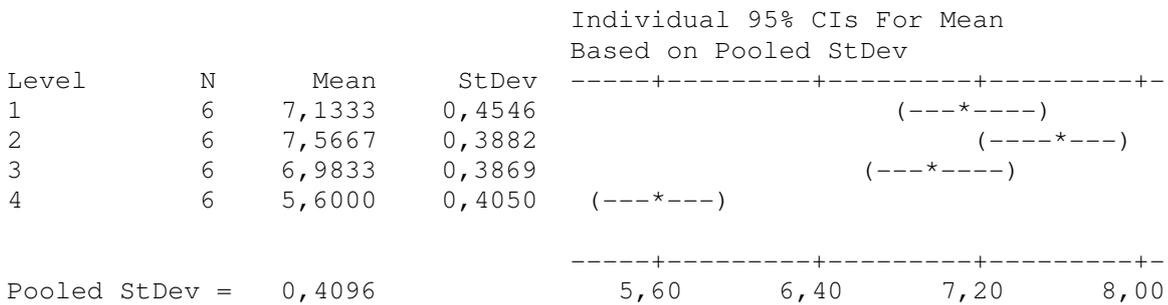
Fornecedor	Viscosidades					
1	6,6	7,9	7,3	7,2	6,8	7,0
2	7,9	7,7	7,8	6,9	7,3	7,8
3	7,4	6,4	6,9	6,7	7,2	7,3
4	5,3	6,3	5,4	5,6	5,8	5,2

- 2) (Cont.1) No exercício anterior, foi rejeitada a hipótese de igualdade das médias. Identifique, a partir do teste de Tukey com $\alpha = 5\%$, quais são as médias diferentes. (g 4, 20, 5% = 3,96)

One-way Analysis of Variance

Analysis of Variance for Viscosidade

Source	DF	SS	MS	F	P
Fornecedor	3	13,025	4,342	25,88	0,000
Error	20	3,355	0,168		
Total	23	16,380			



- 3) A Tabela 7.2 apresenta os resultados de uma análise interlaboratorial, envolvendo três laboratórios e quatro amostras de um material de referência. Que conclusões podem ser tiradas a respeito da equivalência dos laboratórios? ($\alpha = 5\%$)

Tabela 7.2 Resultados das análises por laboratório.

Laboratório	Material				T _i	Q _i
	1	2	3	4		
1	100,4	99,0	101,5	103,5		
2	96,7	102,1	102,2	101,8		
3	103,5	103,7	105	104,6		
T _j						
Q _j						

$$T^2 / mn =$$

$$\sum T_i^2 =$$

$$\sum T_j^2 =$$

$$STQ =$$

$$SQA =$$

$$SQB =$$

$$SRQ =$$

Tabela 7.3 Análise de variância para os fatores laboratório e material.

Fonte de Variação	Soma de Quadrados	Graus de Liberdade	Quadrados Médios	F	F _{5%}
Laboratório					
Material					
Resíduo					
Total					

- 4) (Cont.3) Resolva o exercício anterior, considerando apenas o fator laboratório. Interprete os resultados obtidos.

- 5) A área de qualidade deseja avaliar a capacidade de um sistema de medição. Para isto, cinco itens foram retirados aleatoriamente da produção e três operadores escolhidos. Cada operador realizou duas medições em cada item. Verifique, a partir da análise de variância, se há diferenças significativas entre os operadores.

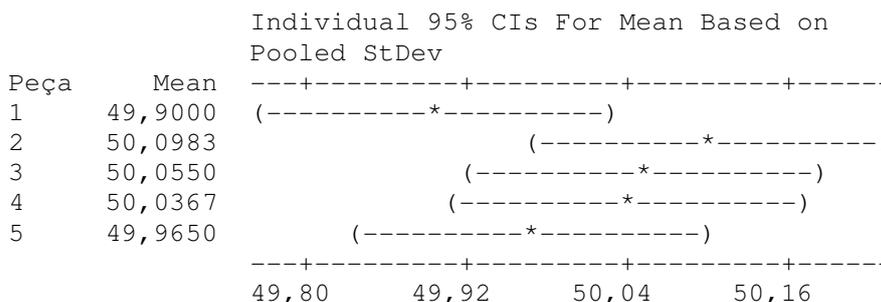
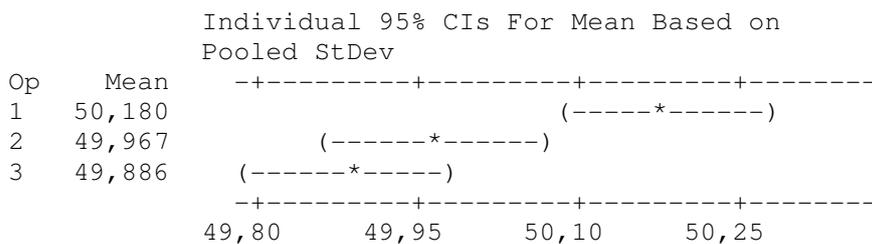
Tabela 7.4 Resultados das medições.

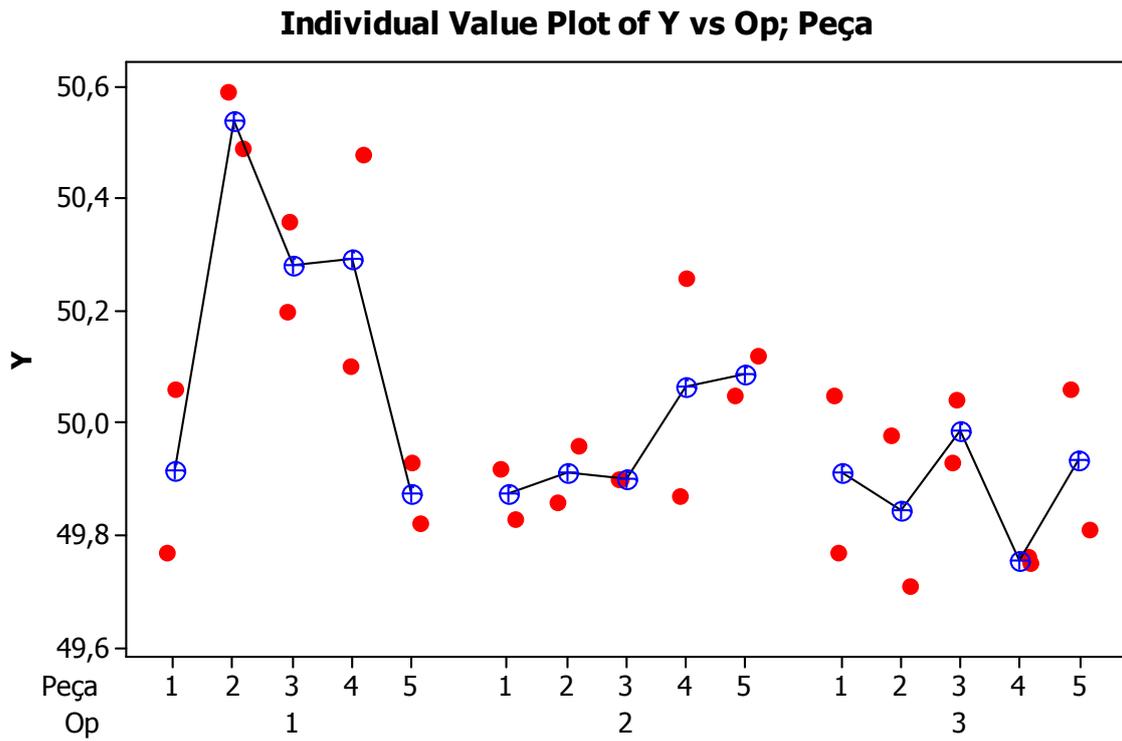
Material	Operador 1		Operador 2		Operador 3	
	1	2	1	2	1	2
1	49,77	50,06	49,92	49,83	50,05	49,77
2	50,49	50,59	49,86	49,96	49,98	49,71
3	50,20	50,36	49,90	49,90	49,93	50,04
4	50,10	50,48	50,26	49,87	49,76	49,75
5	49,93	49,82	50,12	50,05	49,81	50,06

Two-way ANOVA: Y versus Op; Peça

Source	DF	SS	MS	F	P
Op	2	0,46122	0,230610	10,21	0,002
Peça	4	0,14795	0,036988	1,64	0,216
Interaction	8	0,62465	0,078081	3,46	0,018
Error	15	0,33865	0,022577		
Total	29	1,57247			

S = 0,1503 R-Sq = 78,46% R-Sq(adj) = 58,36%





EXERCÍCIOS PROPOSTOS

- 1) Três marcas de veículos com motor 1.0 foram avaliados por uma revista especializada. Verifique, a partir dos dados abaixo, se há diferenças significativas de consumo de combustível entre os três modelos testados.

A	B	C
7,4	7,1	7,7
7,0	6,7	7,1
6,9	7,3	8,2
6,8	7,7	7,4
7,6	7,7	7,2
6,7	7,9	7,5
7,0	6,3	7,2
6,6	7,1	7,0
7,2	7,6	7,4
6,8	6,8	8,3

- 2) Os resultados a seguir representam as gramaturas de amostras de papel obtidas em quatro configurações do processo produtivo. Verifique, a partir da análise de variância, que há diferença significativa na gramatura e, a seguir, aplique o método de Tukey para identificar as configurações com médias diferentes.

Dados: $\alpha = 5\%$ e $g_{4, 20, 5\%} = 3,96$.

Configuração do Processo	Observações					
	1	2	3	4	5	6
1	77	78	79	80	81	85
2	82	83	85	86	86	89
3	84	86	87	88	88	92
4	82	84	86	86	87	89

- 3) Resolva, por A.V, o exercício 9 do capítulo 6, página 15. Compare os resultados da A.V. com os resultados anteriores.
- 4) Idem, para o exercício 10 do capítulo 6, página 15.
- 5) Quais são os objetivos e premissas da Análise de Variância. Explique a lógica e cite três possíveis exemplos de aplicação da técnica em sua área de formação.