GET00182 – Estatística II
Aula de Exercícios – 1/12/2017
Profa. Ana Maria Farias – 2017/2

Exemplo1

Quatro catalisadores podem afetar a concentração de um componente em uma mistura líquida de três componentes. As seguintes concentrações foram obtidas:

Catalisador			
A	B	C	D
58,2	56,3	50,1	52,9
57,2	54,5	54,2	49,9
58,4	57,0	55,4	50,0
55,8	55,3		51,7
54,9			

Vamos fazer uma análise para ver se os quatro catalisadores têm o mesmo efeito sobre a concentração.

- Verificação das hipóteses do modelo

  Vamos assumir que temos amostras aleatórias simples independentes das quatro populações definidas pelo tipo de catalisador.

  - Normalidade

    Na Figura 1 temos o resultado do teste de Anderson-Darling para os quatro catalisadores; podemos assumir que as quatro amostras vêm de populações normais.

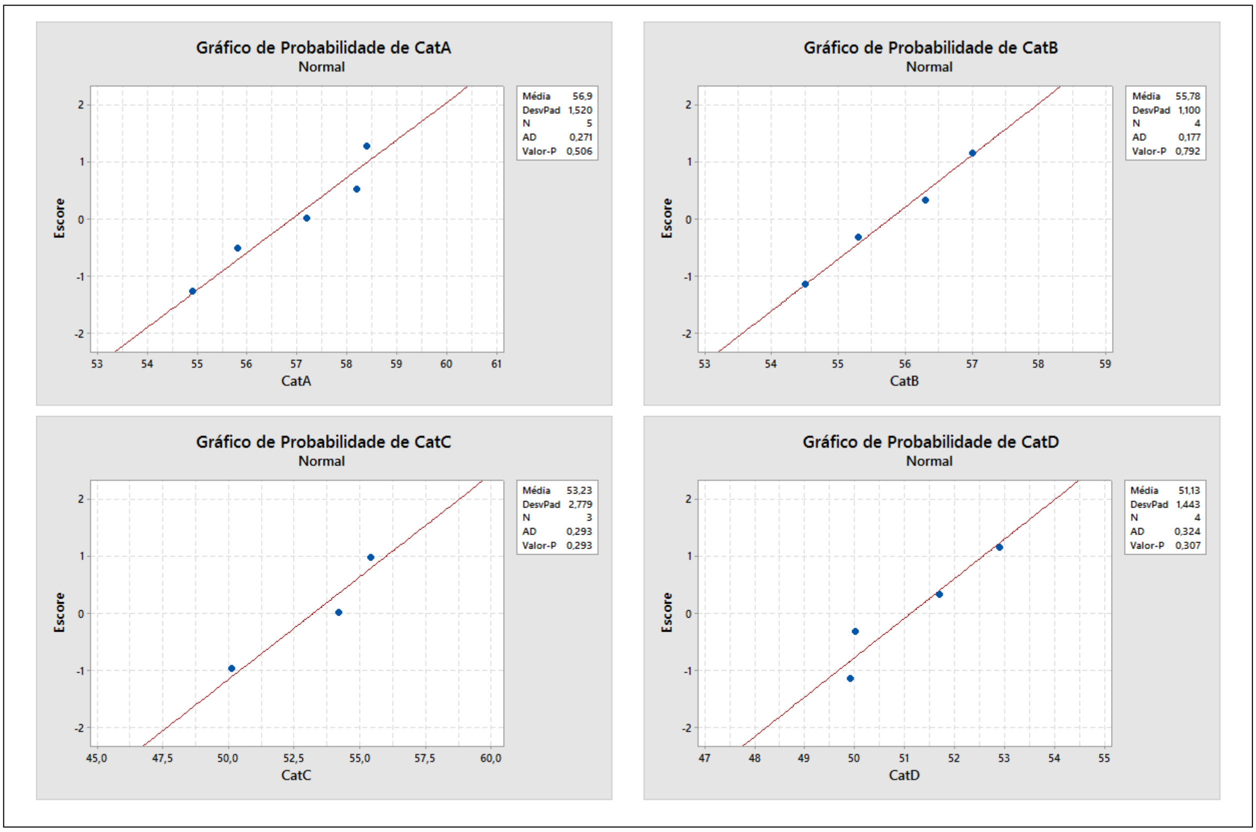


Figura 1 – Teste de normalidade

  - Homogeneidade das variâncias

    Na Figura 2 temos a saída do Minitab para o teste de igualdade de variâncias; como o valor P é grande, não se rejeita a hipótese de igualdade de variâncias

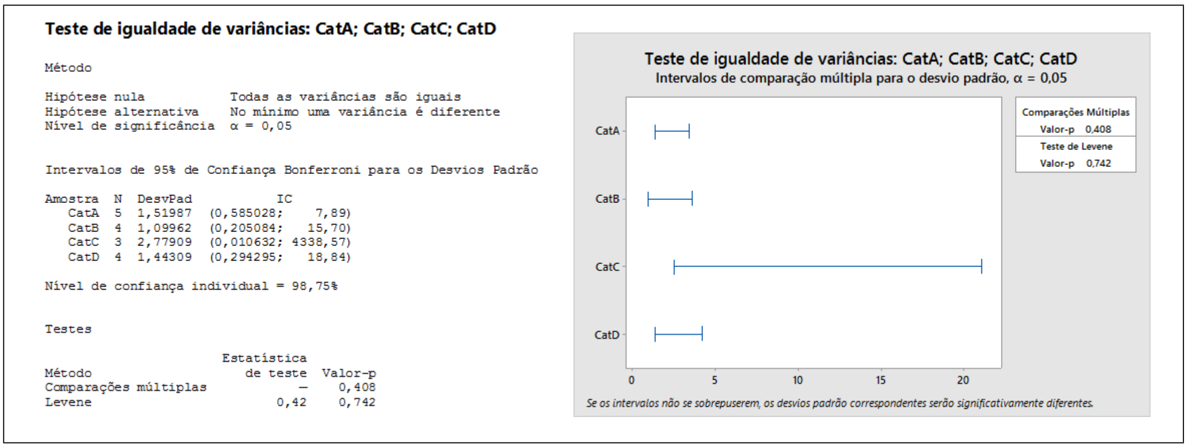


Figura 2 – Teste da igualdade de variâncias

Os intervalos de confiança de comparação múltipla para as variâncias se baseiam na hipótese de populações independentes com mesma curtose e os limites dos intervalos são construídos com base num estimador combinado da curtose com valores críticos da distribuição da amplitude de variáveis independentes e identicamente distribuídas como $N(0;1)$. Esses intervalos não podem ser considerados como intervalos de confiança para o desvio padrão populacional.

- ANOVA
  - Cálculos manuais

	Catalisador			
	A	B	C	D
	58,2	56,3	50,1	52,9
	57,2	54,5	54,2	49,9
	58,4	57,0	55,4	50,0
	55,8	55,3		51,7
	54,9			
$\sum x_i$	284,5	223,1	159,7	204,5
$\sum x_i^2$	16197,29	12447,03	8516,81	10461,31
n_i	5	4	3	4
Média	56,9	55,775	53,2333	51,125
Variância	2,31	1,209167	7,723333	2,0825

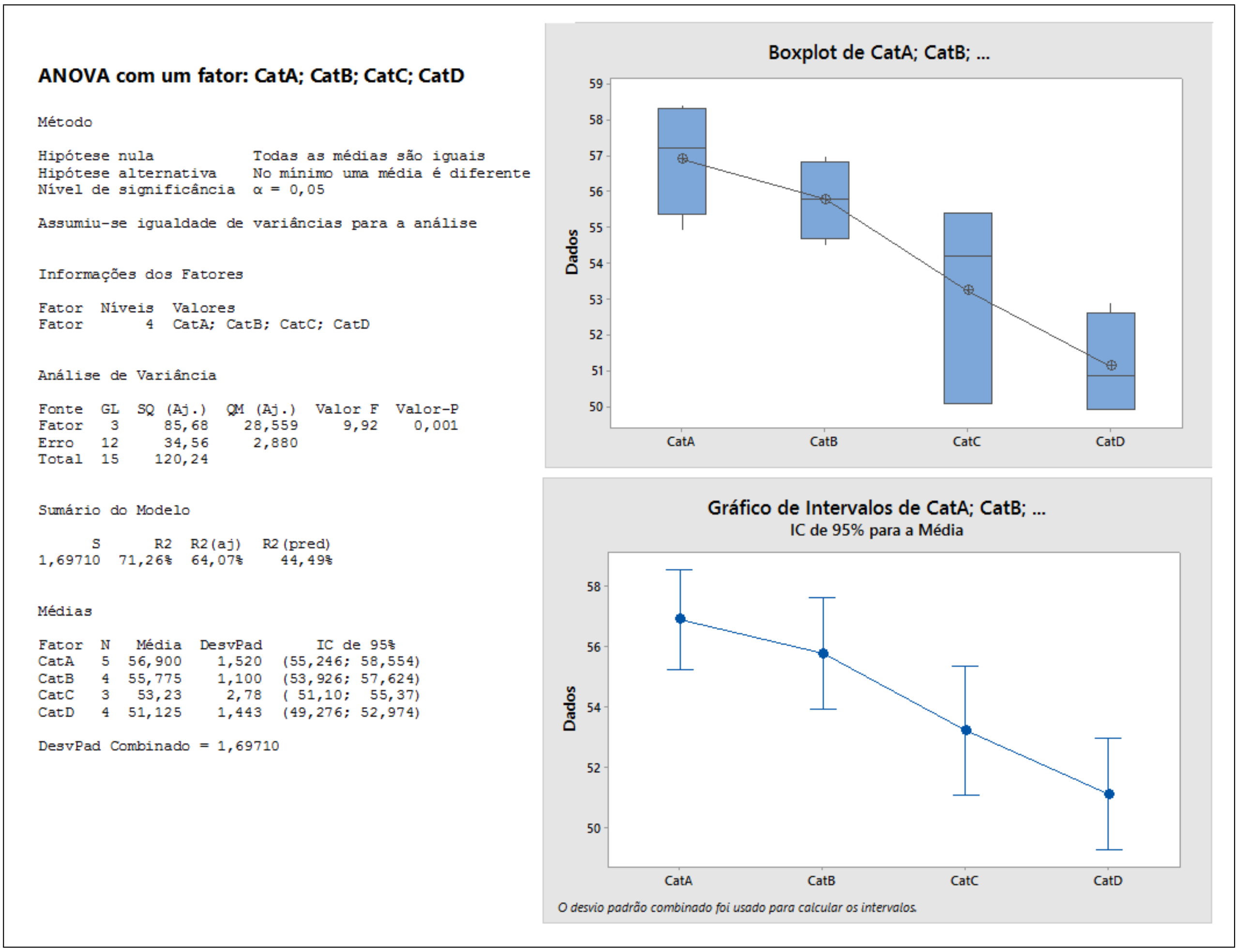
$$
\begin{aligned}
\text{SQT} &= (16197,29 + 12447,03 + 8516,81 + 10461,31) \\
&\quad - \frac{(284,5 + 223,1 + 159,7 + 204,5)^2}{(5+4+3+4)} = 120,2375 \\
\text{SQE} &= \left(16197,29 - \frac{284,5^2}{5}\right) + \left(12447,03 - \frac{223,1^2}{4}\right) + \left(8516,81 - \frac{159,7^2}{3}\right) \\
&\quad + \left(10461,31 - \frac{204,5^2}{4}\right) = 34,56167 \\
\text{SQA} &= \text{SQT} - \text{SQE} = 85,67583 \\
\text{MQA} &= \frac{85,67583}{4-1} = 28,558611 \\
\text{MQE} &= \widehat{\sigma}^2 = \frac{34,56167}{16-4} = 2,880139 \\
\text{desvio padrão combinado} &= \sqrt{2,880139} = 1,697097
\end{aligned}
$$

  - Saída do Minitab

    Na Figura 3 temos a saída do resultado para a ANOVA.

    Os intervalos de confiança para as médias são construídos com base na distribuição t com 12 graus de liberdade, usando o desvio padrão combinado. Por exemplo, para CatA:

$$56,9 \pm 2,17881 \times \frac{1,69710}{\sqrt{5}} = (55,246\,;\,58,554)$$



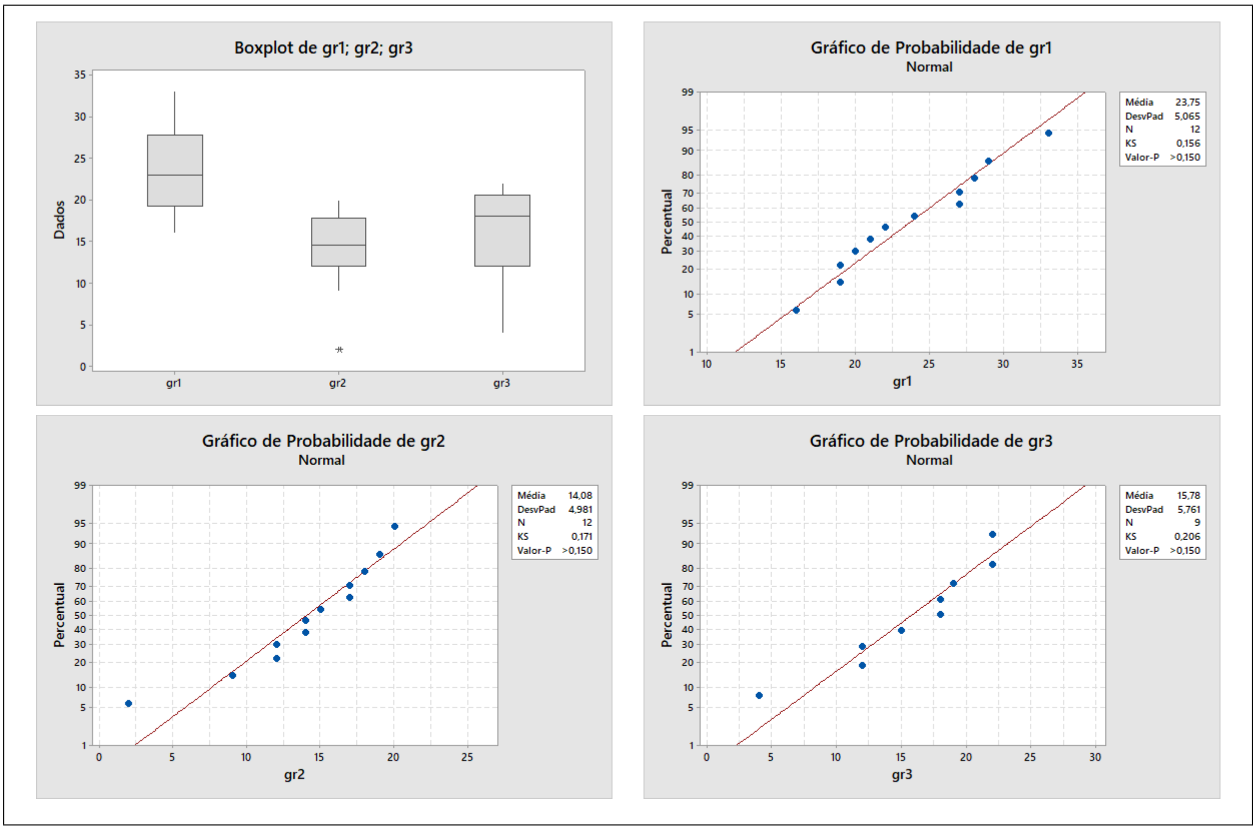
```
ANOVA com um fator: CatA; CatB; CatC; CatD

Método

Hipótese nula           Todas as médias são iguais
Hipótese alternativa    No mínimo uma média é diferente
Nível de significância  α = 0,05

Assumiu-se igualdade de variâncias para a análise


Informações dos Fatores

Fator  Níveis  Valores
Fator       4  CatA; CatB; CatC; CatD


Análise de Variância

Fonte  GL  SQ (Aj.)  QM (Aj.)  Valor F  Valor-P
Fator   3     85,68    28,559     9,92    0,001
Erro   12     34,56     2,880
Total  15    120,24


Sumário do Modelo

      S     R2  R2(aj)  R2(pred)
1,69710 71,26%  64,07%    44,49%


Médias

Fator  N   Média  DesvPad       IC de 95%
CatA   5  56,900    1,520  (55,246; 58,554)
CatB   4  55,775    1,100  (53,926; 57,624)
CatC   3   53,23     2,78  ( 51,10;  55,37)
CatD   4  51,125    1,443  (49,276; 52,974)

DesvPad Combinado = 1,69710
```

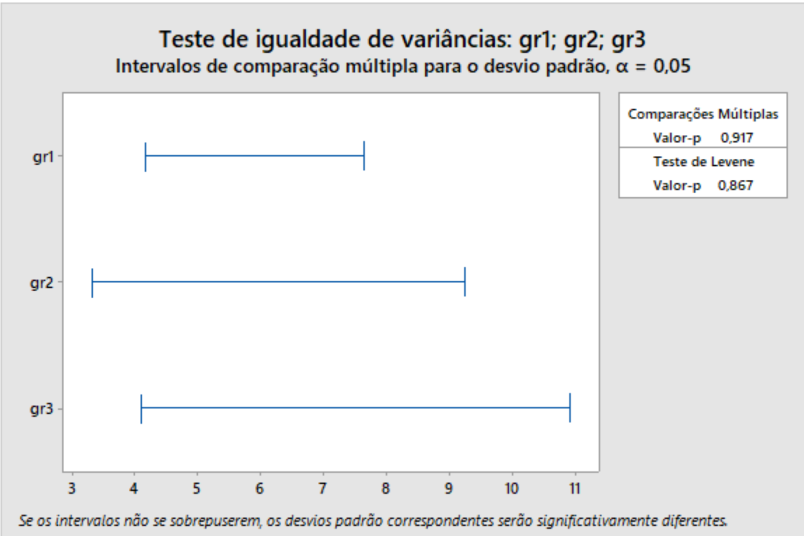

Figura 3 – Saída do Minitab

Exemplo 2

Como o desmatamento em uma floresta tropical afeta essa floresta muitos anos depois? Pesquisadores compararam lotes de florestas em Bornéu que nunca haviam sido desmatados (Grupo 1) e lotes semelhantes, próximos, que haviam sido desmatados 1 ano antes (Grupo 2) e 8 anos antes (Grupo 3). Embora o estudo não tenha sido um experimento, os autores explicam por que podemos considerar que os lotes tenham sido aleatoriamente selecionados. Os dados são mostrados a seguir. A variável de interesse é o número de árvores num lote.

Grupo 1	27	22	29	21	19	33	16	20	24	27	28	19
Grupo 2	12	12	15	9	20	18	17	14	14	2	17	19
Grupo 3	18	4	22	15	18	19	22	12	12			


Boxplot de gr1; gr2; gr3
Dados
gr1
gr2
gr3
Gráfico de Probabilidade de gr1
Normal
Percentual
Média 23,75
DesvPad 5,065
N 12
KS 0,156
Valor-P >0,150
Gráfico de Probabilidade de gr2
Normal
Média 14,08
DesvPad 4,981
N 12
KS 0,171
Valor-P >0,150
Gráfico de Probabilidade de gr3
Normal
Média 15,78
DesvPad 5,761
N 9
KS 0,206
Valor-P >0,150

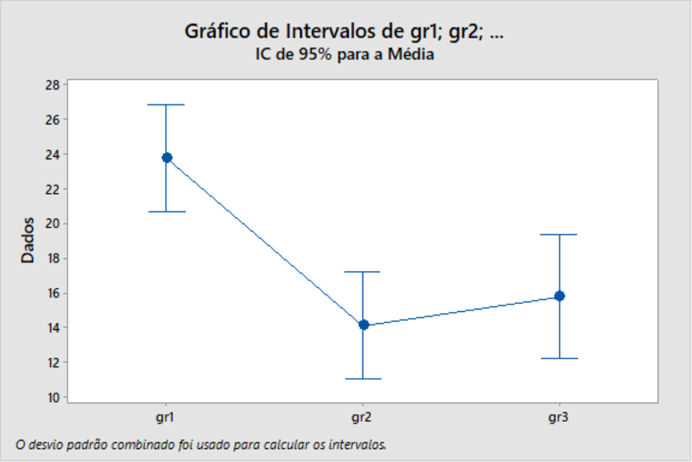

Teste de igualdade de variâncias: gr1; gr2; gr3
Intervalos de comparação múltipla para o desvio padrão, α = 0,05
Comparações Múltiplas
Valor-p 0,917
Teste de Levene
Valor-p 0,867
gr1
gr2
gr3
Se os intervalos não se sobrepuserem, os desvios padrão correspondentes serão significativamente diferentes.

ANOVA com um fator: gr1; gr2; gr3

Método

Hipótese nula	Todas as médias são iguais
Hipótese alternativa	Nem todas as médias são iguais
Nível de significância	α = 0,05

Assumiu-se igualdade de variâncias para a análise



Informações dos Fatores

Fator	Níveis	Valores
Fator	3	gr1; gr2; gr3

Análise de Variância

Fonte	GL	SQ (Aj.)	QM (Aj.)	Valor F	Valor-P
Fator	2	625,2	312,58	11,43	0,000
Erro	30	820,7	27,36		
Total	32	1445,9			

Sumário do Modelo

S	R2	R2(aj)	R2(pred)
5,23043	43,24%	39,45%	31,06%

Médias

Fator	N	Média	DesvPad	IC de 95%
gr1	12	23,75	5,07	(20,67; 26,83)
gr2	12	14,08	4,98	(11,00; 17,17)
gr3	9	15,78	5,76	(12,22; 19,34)

DesvPad Combinado = 5,23043

Parte 2: Comparações múltiplas

Exemplo 1 revisitado

- Intervalos de confiança de Bonferroni

  $t_{12;0,025/4} = t_{12;0,00625} = 2,93446$

  $$\overline{x}_{A\cdot} - \overline{x}_{B\cdot} : (56,9 - 55,775) \pm 2,93446\sqrt{2,880139\left(\frac{1}{5}+\frac{1}{4}\right)} = (-2,2157\,;\,4,4657)$$

  $$\overline{x}_{A\cdot} - \overline{x}_{C\cdot} : (56,9 - 53,230) \pm 2,93446\sqrt{2,880139\left(\frac{1}{5}+\frac{1}{3}\right)} = (0,0297\,;\,7,3036)$$

  $$\overline{x}_{A\cdot} - \overline{x}_{D\cdot} : (56,9 - 51,125) \pm 2,93446\sqrt{2,880139\left(\frac{1}{5}+\frac{1}{4}\right)} = (2,4343\,;\,9,1158)$$

  $$\overline{x}_{B\cdot} - \overline{x}_{C\cdot} : (55,775 - 53,230) \pm 2,93446\sqrt{2,880139\left(\frac{1}{4}+\frac{1}{3}\right)} = (-1,2619\,;\,6,3453)$$

  $$\overline{x}_{B\cdot} - \overline{x}_{D\cdot} : (55,775 - 51,125) \pm 2,93446\sqrt{2,880139\left(\frac{1}{4}+\frac{1}{4}\right)} = (1,1286\,;\,8,1714)$$

  $$\overline{x}_{C\cdot} - \overline{x}_{D\cdot} : (53,230 - 51,125) \pm 2,93446\sqrt{2,880139\left(\frac{1}{3}+\frac{1}{4}\right)} = (-1,6953\,;\,5,9119)$$

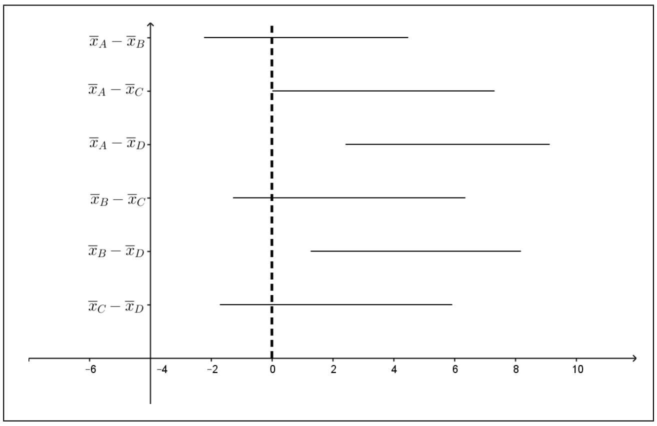


Figura 4 – Intervalos de confiança de Bonferroni

O esquema das comparações resulta em

Fator	N	Média	Agrupamento		
Cat1	5	56,900	A		
Cat2	4	55,775	A	B	
Cat3	3	53,23		B	C
Cat4	4	51,125			C

- Teste da DMS de Fisher

$t_{12;0,025/4} = t_{12;0,025} = 2,17881$

$$\overline{x}_{B\cdot} - \overline{x}_{A\cdot} : (55,775 - 56,9) \pm 2,17881\sqrt{2,880139\left(\frac{1}{5}+\frac{1}{4}\right)} = (-3,61\,;\,1,36)$$

$$\overline{x}_{C\cdot} - \overline{x}_{A\cdot} : (53,230 - 56,9) \pm 2,934462,17881\sqrt{2,880139\left(\frac{1}{5}+\frac{1}{3}\right)} = (-6,37\,;\,-0,97)$$

$$\overline{x}_{D\cdot} - \overline{x}_{A\cdot} : (51,125 - 56,9) \pm 2,17881\sqrt{2,880139\left(\frac{1}{5}+\frac{1}{4}\right)} = (-8,26\,;\,-3,29)$$

$$\overline{x}_{C\cdot} - \overline{x}_{B\cdot} : (53,230 - 55,775) \pm 2,17881\sqrt{2,880139\left(\frac{1}{4}+\frac{1}{3}\right)} = (-5,37\,;\,0,28)$$

$$\overline{x}_{D\cdot} - \overline{x}_{B\cdot} : (51,125 - 55,775) \pm 2,17881\sqrt{2,880139\left(\frac{1}{4}+\frac{1}{4}\right)} = (-7,27\,;\,-2,03)$$

$$\overline{x}_{D\cdot} - \overline{x}_{C\cdot} : (51,125 - 53,230) \pm 2,17881\sqrt{2,880139\left(\frac{1}{3}+\frac{1}{4}\right)} = (-4,93\,;\,0,72)$$

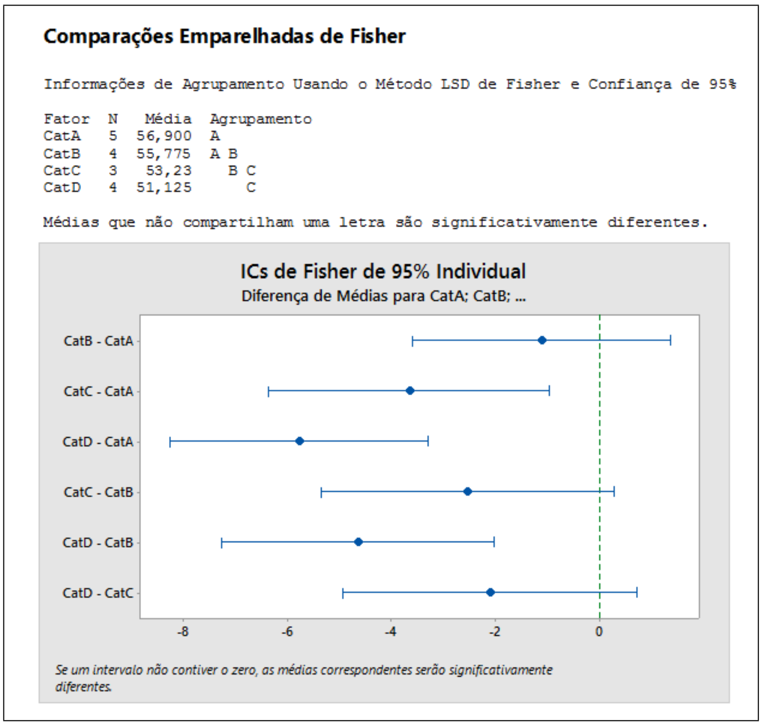


Figura 5 – Teste da DMS de Fisher - Minitab

- Teste da DHS de Tukey

  $q_{4,12;0,05} = 4,199$ pela Tabela 1.

$$\overline{x}_{B\cdot} - \overline{x}_{A\cdot} : (55,775 - 56,9) \pm 4,199\sqrt{\frac{2,880139}{2}\left(\frac{1}{5}+\frac{1}{4}\right)} = (-4,51\,;\,2,26)$$

$$\overline{x}_{C\cdot} - \overline{x}_{A\cdot} : (53,230 - 56,9) \pm 4,199\sqrt{\frac{2,880139}{2}\left(\frac{1}{5}+\frac{1}{3}\right)} = (-7,35\,;\,0,01)$$

$$\overline{x}_{D\cdot} - \overline{x}_{A\cdot} : (51,125 - 56,9) \pm 4,199\sqrt{\frac{2,880139}{2}\left(\frac{1}{5}+\frac{1}{4}\right)} = (-9,16\,;\,-2,39)$$

$$\overline{x}_{C\cdot} - \overline{x}_{B\cdot} : (53,230 - 55,775) \pm 4,199\sqrt{\frac{2,880139}{2}\left(\frac{1}{4}+\frac{1}{3}\right)} = (-6,39\,;\,1,31)$$

$$\overline{x}_{D\cdot} - \overline{x}_{B\cdot} : (51,125 - 55,775) \pm 4,199\sqrt{\frac{2,880139}{2}\left(\frac{1}{4}+\frac{1}{4}\right)} = (-8,214\,;\,-1,09)$$

$$\overline{x}_{D\cdot} - \overline{x}_{C\cdot} : (51,125 - 53,230) \pm 4,199\sqrt{2,880139\left(\frac{1}{3}+\frac{1}{4}\right)} = (-5,96\,;\,1,74)$$

Comparações Emparelhadas de Tukey

Informações de Agrupamento Usando Método de Tukey e Confiança de 95%

Fator	N	Média	Agrupamento
CatA	5	56,900	A
CatB	4	55,775	A
CatC	3	53,23	A B
CatD	4	51,125	B

Médias que não compartilham uma letra são significativamente diferentes.

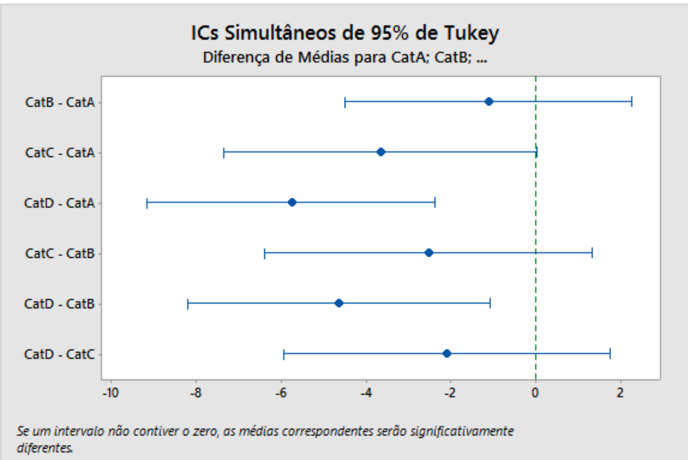


Figura 6 – Teste da DHS de Tukey - Minitab

- Teste de Duncan

$$n_H = \frac{4}{\frac{1}{5}+\frac{1}{4}+\frac{1}{3}+\frac{1}{4}} = 3,870968$$

A estimativa do erro padrão da média é

$$\sqrt{s_{\overline{x}}} = \sqrt{\frac{2,880139}{3,870968}} = 0,862575$$

Da tabela das diferenças significantes de Duncan com $\alpha = 0,05$ e $\nu = 16 - 4 = 12$ obtemos

- $p = 2$ $\quad r_2 = 3,082 \Rightarrow R_2 = \sqrt{s_{\overline{x}}} \times r_2 = 0,862575 \times 3,082 = 2,658456$
- $p = 3$ $\quad r_3 = 3,225 \Rightarrow R_3 = \sqrt{s_{\overline{x}}} \times r_3 = 0,862575 \times 3,225 = 2,781804$
- $p = 4$ $\quad r_4 = 3,313 \Rightarrow R_4 = \sqrt{s_{\overline{x}}} \times r_4 = 0,862575 \times 3,313 = 2,857711$

As médias ordenadas são

(D)	(C)	(B)	(A)
51,1250	53,2333	55,7750	56,9000

Comparações

- $(A)-(D)$ $\quad 56,900 - 51,1250 = 5,775 > 2,857711 = R_4$
- $(A)-(C)$ $\quad 56,900 - 53,2333 = 3,6667 > 2,781804 = R_3$
- $(A)-(B)$ $\quad 56,900 - 55,775 = 1,125 < 2,658456 = R_2$
- $(B)-(D)$ $\quad 55,775 - 51,125 = 4,65 > 2,781804 = R_3$
- $(B)-(C)$ $\quad 55,775 - 53,2333 = 2,5417 < 2,658456 = R_2$
- $(C)-(D)$ $\quad 53,2333 - 51,125 = 2,1083 < 2,658456 = R_2$

O esquema das comparações resulta em

Fator	N	Média	Agrupamento		
Cat1	5	56,900	A		
Cat2	4	55,775	A	B	
Cat3	3	53,23		B	C
Cat4	4	51,125			C

Exemplo 2: Saída do Minitab

Comparações Emparelhadas de Fisher

Informações de Agrupamento Usando o Método LSD de Fisher e Confiança de 95%

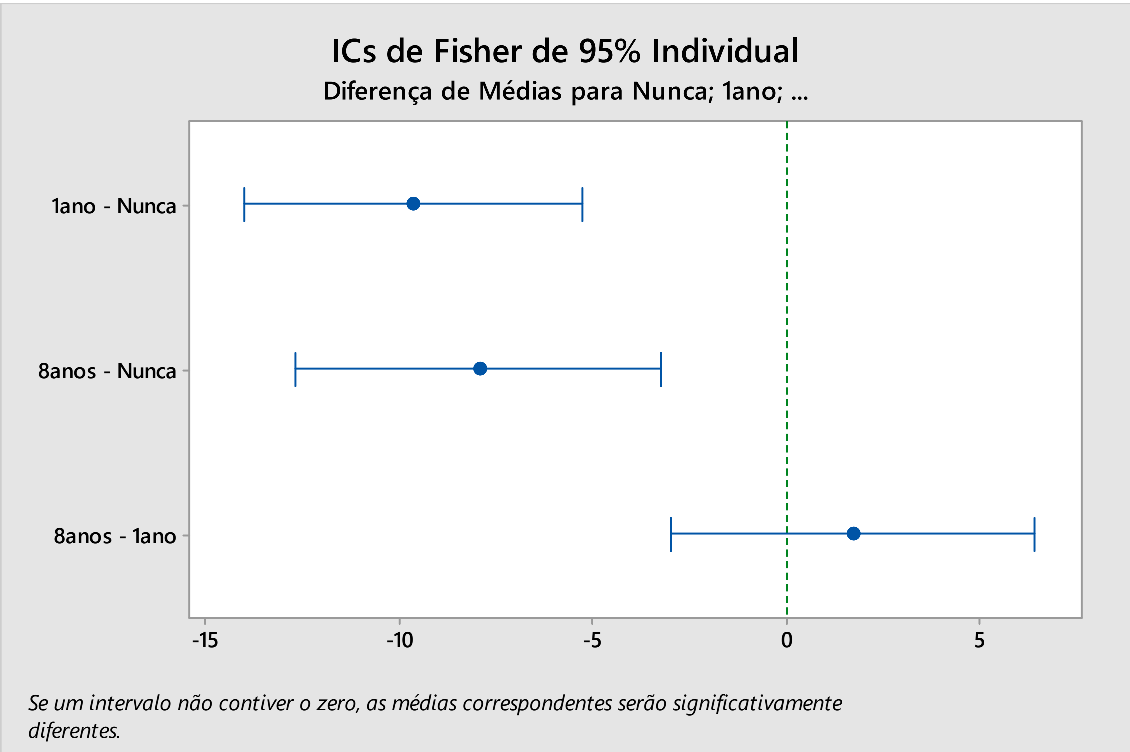
Fator	N	Média	Agrupamento	
Nunca	12	23,75	A	
8anos	9	15,78		B
1ano	12	14,08		B

Médias que não compartilham uma letra são significativamente diferentes.

Testes Individuais de Fisher para as Diferenças de Médias

Diferença de Níveis	Diferença de Médias	EP da Diferença	IC de 95%	Valor-T	Valor-P Ajustado
1ano - Nunca	-9,67	2,14	(-14,03; -5,31)	-4,53	0,000
8anos - Nunca	-7,97	2,31	(-12,68; -3,26)	-3,46	0,002
8anos - 1ano	1,69	2,31	(-3,02; 6,40)	0,73	0,468

Nível de confiança simultâneo = 88,04%



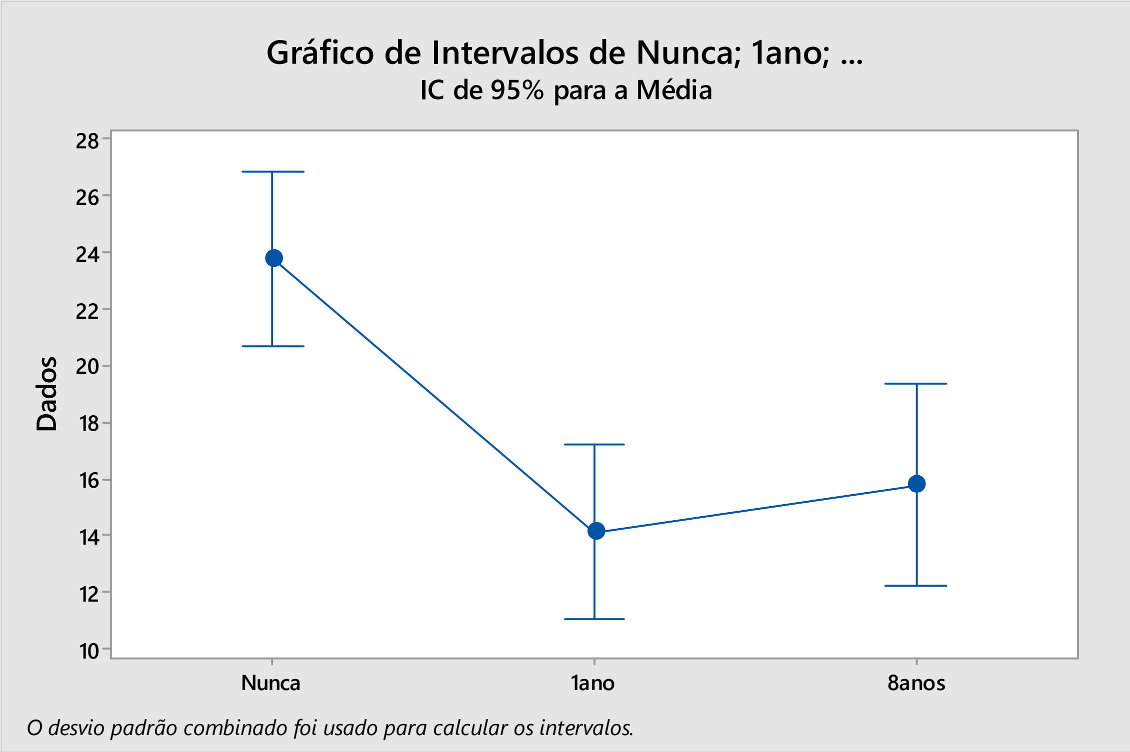
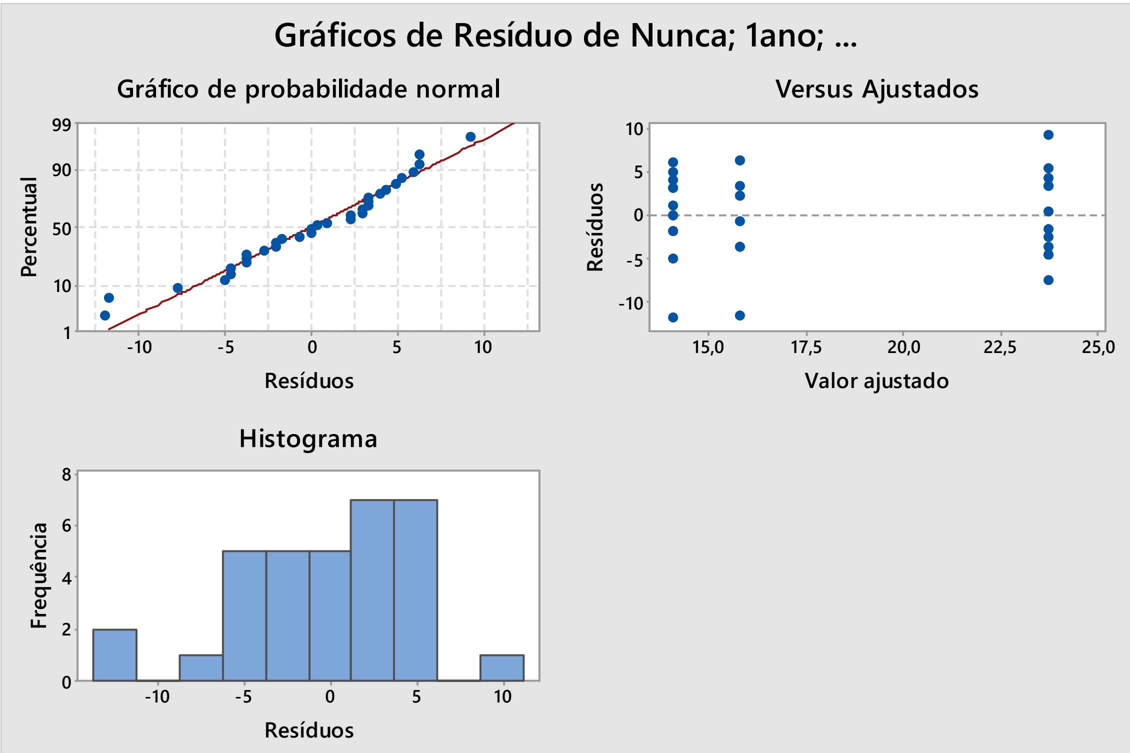


Gráfico de Intervalos de Nunca; 1ano; ...
IC de 95% para a Média
Dados
28
26
24
22
20
18
16
14
12
10
Nunca
1ano
8anos
O desvio padrão combinado foi usado para calcular os intervalos.


Gráficos de Resíduo de Nunca; 1ano; ...
Gráfico de probabilidade normal
Percentual
99
90
50
10
1
-10
-5
0
5
10
Resíduos
Versus Ajustados
Resíduos
10
5
0
-5
-10
15,0
17,5
20,0
22,5
25,0
Valor ajustado
Histograma
Frequência
8
6
4
2
0
-10
-5
0
5
10
Resíduos

Comparações Emparelhadas de Tukey

Informações de Agrupamento Usando Método de Tukey e Confiança de 95%

Fator	N	Média	Agrupamento	
Nunca	12	23,75	A	
8anos	9	15,78		B
1ano	12	14,08		B

Médias que não compartilham uma letra são significativamente diferentes.

Testes Simultâneos de Tukey para as Diferenças de Médias

Diferença de Níveis	Diferença de Médias	EP da Diferença	IC de 95%	Valor-T	Valor-P Ajustado
1ano - Nunca	-9,67	2,14	(-14,94; -4,40)	-4,53	0,000
8anos - Nunca	-7,97	2,31	(-13,66; -2,28)	-3,46	0,005
8anos - 1ano	1,69	2,31	(-4,00; 7,39)	0,73	0,745

Nível de confiança individual = 98,05%