



**Universidade Estadual Do Oeste do Paraná  
PGEAGRI**

**ESTATÍSTICA EXPERIMENTAL  
Prof.: Dra. Luciana Pagliosa**

**Resolução da Lista de Exercícios de Delineamento  
Inteiramente Casualizado - 2018**

Vanessa Mendes Pientosa  
Willyan Goergen de Souza

Cascavel - PR  
Maio de 2018

**Questão 1)** Pesquise e apresente um artigo de um experimento conduzido com um fator com um fator sob estudo, associado a um dos temas a seguir, que descreva: a escolha do delineamento experimental, quem são as unidades experimentais (parcelas), qual é o fator sob estudo, qual é a variável resposta (pode ser mais que uma), e das análises estatísticas realizadas. Os temas para pesquisa são: árvores frutíferas; café; cana-de-açúcar; e soja. Apresente as referências bibliográficas e anexe a essa lista a(s) página(s) do artigo que descreva as informações solicitadas, destacando estas informações nesta(s) página(s).

**Resposta:** Artigo sobre Cana-de-açúcar:

### **COMPOSIÇÃO QUÍMICA DA CANA-DE-AÇÚCAR (*Saccharum SPP.*) E DAS SILAGENS COM DIFERENTES ADITIVOS EM DUAS IDADES DE CORTE**

**Delineamento experimental:** O delineamento foi inteiramente Casualizado, em faixas, espaçadas de 50,0 x 1,50 m com o número de repetições definidos em função do tempo e do modo de uso das facas.

**Unidade experimental:** Para a unidade experimental foi utilizado um esquema fatorial 2 x 9, sendo testada duas idades de corte e nove formas de utilização.

**Fator sobre estudo:** Foram testadas as influencias de duas idades de corte de cana- de- açúcar (11 e 24 meses) e nove formas de utilização (cana picada acrescida de 1% da mistura ureia e sulfato de amônio na proporção 9:1, cana in natura e ensilada sem aditivo, cana in natura e ensilada com 1% de ureia, cana in natura e ensilada com 8% de MDPS e cana in natura e ensilada com 0,5% de sal mineral).

**Variável resposta:** Altura do corte, índice de danos e abalos as soqueiras. Teores de matéria seca (MS) e de proteína bruta (PB), teores de fibras em detergente neutro e ácido (FDN e FDA).

**Análises estatísticas realizadas:** As observações foram analisadas estatisticamente pelos procedimentos de análises de variância, por meio do programa Estatístico SISVAR (Sistema de análise de variância para dados balanceados). E para a comparação das médias entre os tratamentos, foi utilizado o teste de Scott-Knott, a 5% de probabilidade.

**Questão 2)** Uma pesquisa experimental foi conduzida com o propósito de comparar as seguintes variedades de Batata doce: Brazlândia (B), Jacaré (J), Paulistinha (P), Rainha (R) e Yellow Yam (Y). Considerando uma homogeneidade da área experimental, a área foi dividida em 20 parcelas de igual tamanho, atribuídas aleatoriamente às variedades. Uma área de 6m<sup>2</sup> dentro da parcela foi considerada como área útil (unidade de observação). Após o tempo necessário para a colheita, o pesquisador anotou a produção da batata doce, em kg por 6m<sup>2</sup> e obteve os resultados que seguem.

**Tabela 1:** Produção (em kg) de batata doce por variedade.

Variedade	Repetição	Produção
Brazlândia	1	47
Brazlândia	2	31,8
Brazlândia	3	42,8
Brazlândia	4	58,4
Jacaré	1	64,6
Jacaré	2	65
Jacaré	3	66
Jacaré	4	78
Paulistinha	1	47,7
Paulistinha	2	42
Paulistinha	3	49,7
Paulistinha	4	58,2
Rainha	1	82,2
Rainha	2	120,6
Rainha	3	100,2
Rainha	4	97
Yellow Yan	1	68,5
Yellow Yan	2	69,4
Yellow Yan	3	70,3
Yellow Yan	4	88,2

**a) - Qual a variável resposta de interesse nesse experimento?**

**R:** Comparar a produção das variedades de batata-doce.

**- Qual a unidade de medida da variável resposta?**

**R:** Unidade de medida da variável resposta é em quilogramas por 6 m<sup>2</sup>.

**- Qual o fator?**

**R:** Variedades de batata doce.

**- Quais são os tratamentos?**

**R:** Os tratamentos são a variedade de batata doce, sendo elas: Brazlândia, Jacaré, Paulistinha, Rainha e Yellow Yam.

**- O fator é de efeito fixo ou aleatório?**

**R:** O fator é de efeito fixo, uma vez que o conjunto de nível de fator (produção de batata doce), foi intencionalmente escolhido pelo pesquisador.

**- O fator é quantitativo ou qualitativo?**

R: É um fator qualitativo, pois está se fazendo uma comparação entre variedades de espécie de plantas.

**- Qual é o material experimental?**

R: Amostras das variedades de batata-doce.

**- Qual é a unidade experimental?**

R: Uma área experimental (considerando sua homogeneidade) que foi dividido em 20 parcelas de iguais tamanhos.

**- Qual é a unidade de observação?**

R: Uma área de 6m<sup>2</sup> dentro da parcela que foi considerada uma área útil, ou seja, a unidade de observação.

**- O experimento é balanceado ou desbalanceado?**

R: O experimento é balanceado, uma vez que apresenta o mesmo número de repetições para todos os tratamentos.

**- Cite alguns fatores (algumas condições) que o pesquisador possa ter controlado.**

R: No caso desse experimento ter sido conduzido em campo, visto sua área relativamente grande, o pesquisador pode ter utilizado o mesmo manejo de substrato para todos os tratamentos e irrigação, padronizado a profundidade do plantio e o espaçamento entre as plantas, escolha da área homogênea e dimensionamento igual das parcelas, porém ficando sujeito às intempéries do clima. Já se fosse o caso de ter sido conduzido em uma estufa, fatores como incidência de luz sobre os indivíduos, a quantidade de água, e temperatura poderiam ser mantidos mais controlados.

**- Segundo as recomendações de Pimentel Gomes, o número de parcelas é adequado?**

R: Para Pimentel Gomes (2000), o número de parcelas dos experimentos deve apresentar no mínimo 20 parcelas e possuir no mínimo 10 em graus de liberdade do resíduo. Sendo assim, o experimento atende as recomendações, possuindo 20 parcelas e 16 graus de liberdade do resíduo.

**b) Esse delineamento considerou na sua construção os princípios de experimentação? Se você fosse montar esse experimento, faça um exemplo de croqui para este delineamento.**

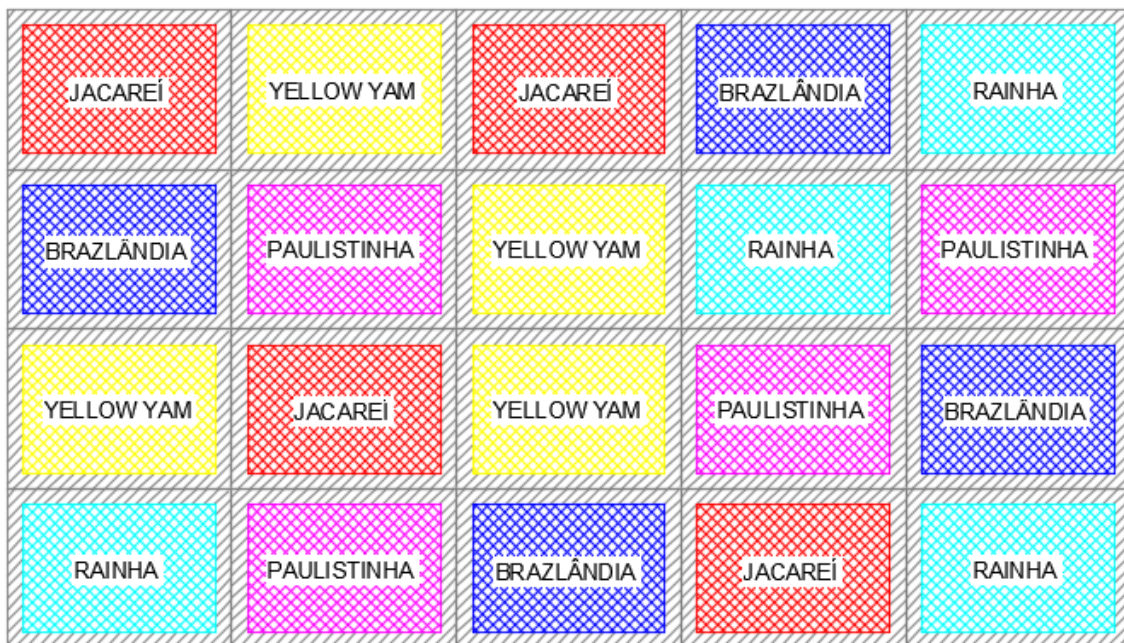
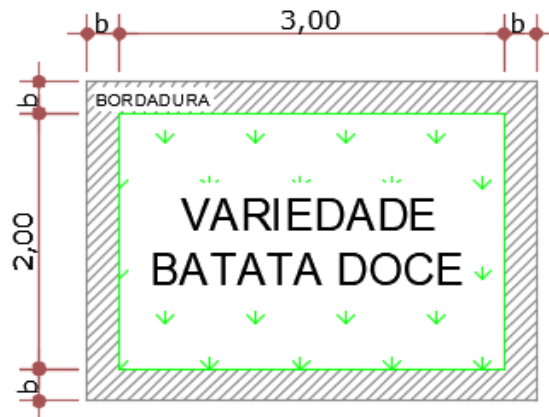
R: Para um experimento é necessário que se siga aos princípios básicos de experimentação, que foram atendidos:

- Repetição: é necessário que cada tratamento seja aplicado em mais do que uma parcela, e que se tenha no mínimo 20 parcelas, aumentando assim, a representatividade do material experimental na população. Utilizando 4 repetições para cada um dos 5 tratamentos, o experimento atinge as exigências da literatura;

- Casualização: as parcelas que receberam determinado tratamento foram atribuídas de forma aleatória;
- Controle local: é o controle dos fatores estranhos para minimizar o erro experimental, que foi atendido ao se estabelecer uma área homogênea para condução do experimento.

Resumindo, baseado nas informações apresentadas, pode-se dizer que o experimento possui os princípios da experimentação.

**Figura 1:** Representação do croqui da parcela (em metros) e distribuição aleatorizada do experimento.



Obs: As variedades foram enumeradas de um a vinte e sorteados aleatoriamente sem reposição, representando a casualização.

- c) **Realize uma análise exploratória para os resultados apresentados da produção, para cada uma das variedades. Interprete e compare os resultados apresentados para cada um dos tratamentos.**

**Tabela 2:** Estatística descritiva geral dos dados.

<b>Produtividade geral (em kg/6m<sup>2</sup>) das 5 variedades de batata-doce</b>	
Soma	1347,50
Mín.	31,80
1º Quartil	49,20
Mediana	65,50
Média	67,38
3º Quartil	79,05
Máx.	120,50
Amplitude	88,70
Variância	492,93
Desvio Padrão	22,20
Coef. Variação (CV) (%)	32,95
Coef. de Assimetria (As)	0,62
Coef. de Curtose (CUR)	2,96

\* Os valores possuem unidade em quilograma/6m<sup>2</sup> e o C.V. em porcentagem

Observa-se a partir da tabela 2, que a soma geral da produtividade (em kg/6m<sup>2</sup>), para as 5 variedades de batata doce é de 1347,50 kg/6m<sup>2</sup>. A máxima dos valores é de 120,5 e a mínima é de 31,0, gerando uma amplitude de 88,70 kg/6m<sup>2</sup>, que mostra a dispersão dos valores de uma série. Se a amplitude for um número elevado, então os valores da série estão distribuídos e afastados; se a amplitude for um número baixo, então os valores na série estão próximos uns aos outros. Com até 49,20 kg/6m<sup>2</sup> de produtividade mostram-se 25% das observações, enquanto com 79,05 kg/6m<sup>2</sup>, apresentam-se 75%.

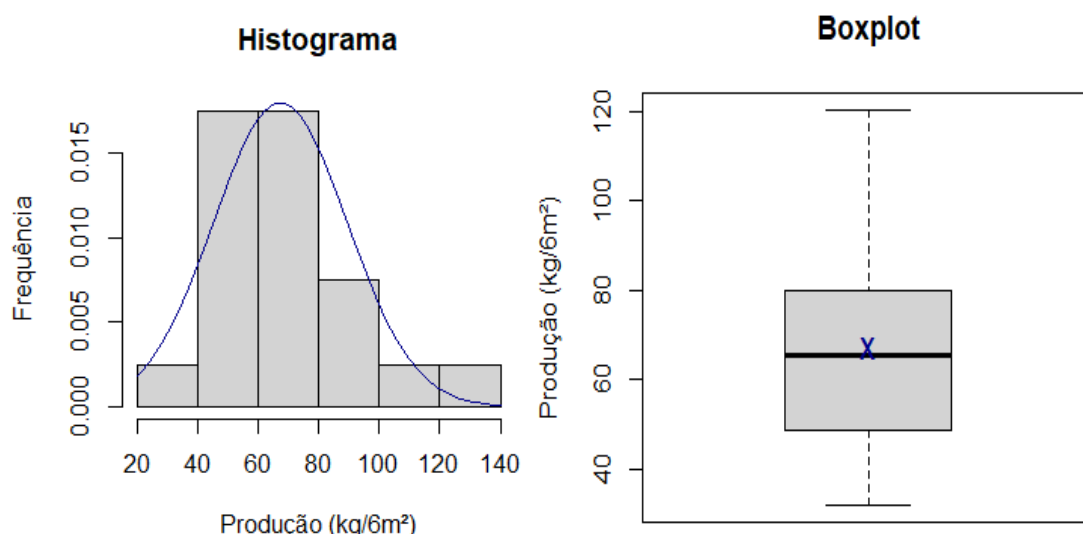
A Mediana apresentou valor de 65,50 kg/6m<sup>2</sup>, significando que 50 % das observações ficaram abaixo e acima desses valores.

A média de produção total das 5 variedades de batata doce foi de 67,38 kg/6m<sup>2</sup>. O Desvio padrão apresentou valor de 22,20, sendo esse um valor alto. O coeficiente de variação é de 32,95%, caracterizando assim, uma heterogeneidade dos dados.

O Coeficiente de assimetria (As) é um valor que quando igual à 0, diz-se que a distribuição é simétrica. Quando esse valor é maior que 0, como observado para esses dados, As de 0,62, diz-se que a distribuição é assimétrica positiva. Já se o valor fosse menor que 0, teria-se uma distribuição assimétrica negativa.

O coeficiente de curtose (Cur) é um valor que quando igual à 3, tem-se uma distribuição com caudas neutras (normais – mesocúrtica). Já se esse número for maior que 3, a distribuição é com caudas longas e pesadas (leptocúrtica). E se o valor calculado for menor que 3, tem-se distribuição com caudas curtas ou leves (platicúrtica). No caso do presente trabalho, o valor da curtose foi de 2,96, o que permite dizer que a curva se prolonga mais no eixo vertical em relação ao horizontal, se apresentando levemente platicúrtica.

**Figura 2:** Demonstra a distribuição de frequência de parcelas de rendimento da produção de batata doce.



Através do histograma (figura 2) tentou-se verificar se a forma da linha de distribuição normal está presente. No entanto, nota-se que isso não é possível uma vez que a curva está mais alongada para a direita do gráfico, evidenciando uma distribuição assimétrica positiva ( $As = 0,62 > 0$ ).

Com relação ao gráfico Boxplot (figura 2) verifica-se que a distribuição normal dos dados está em maior número no limite inferior da tabela, sendo essa distribuição abaixo da linha mediana representada pela linha central da caixa.

**Tabela 3:** Estatística descritiva dos dados de produtividade de batata doce, por tratamento.

Produtividade (em kg/6m <sup>2</sup> ) das por variedade de batata-doce					
Parâmetros estatísticos	Tratamentos				
	Brazlândia	Jacareí	Paulistinha	Rainha	Yellow Yam
Soma	180,00	273,60	197,60	399,90	296,40
Mín.	31,80	64,60	42,00	82,20	68,50
1º Quartil	40,05	64,90	46,27	93,30	69,50
Mediana	44,90	65,50	48,70	98,60	69,85
Média	45,00	68,40	49,40	99,97	74,10
3º Quartil	49,85	69,00	51,83	105,28	74,78
Máx.	58,40	78,00	58,20	120,50	88,20
Amplitude	26,60	13,40	16,20	38,30	19,70
Variância	120,88	41,31	45,06	248,71	88,90
Desvio Padrão	10,99	6,43	6,71	15,77	9,43
Coef. Variação (%)	24,43	9,40	13,59	15,77	12,72
Coef. Assimetria	0,03	1,13	0,35	0,29	1,13
Coef. Curtose	1,91	2,31	1,97	1,99	2,32

\* Os valores possuem unidade em quilogramas/6m<sup>2</sup> e o C.V. em porcentagem

**Variedade Brazlândia:** Observa-se uma soma de 180,00 em kg/6m<sup>2</sup> sendo o valor mínimo de 31,80 e máximo de 58,40, não sendo esse o máximo geral.

O valor do 1º quartil é de 40,05 kg/6m<sup>2</sup>, representando 25% das observações deste tratamento e 49,85 kg/6m<sup>2</sup> o 3º quartil correspondendo a um rendimento de 75 % das observações deste tratamento, sendo esse valor, abaixo do 3º quartil geral, que foi de 79,05 kg/6m<sup>2</sup>.

A mediana apresenta valor 44,90 kg/6m<sup>2</sup> um pouco abaixo da mediana geral que foi de 65,50 kg/6m<sup>2</sup>. A média desta variedade estudada foi de 45,00 kg/6m<sup>2</sup>, também menor que a média geral de 67,36 kg/6m<sup>2</sup>.

O Desvio padrão foi de 10,99 kg/6m<sup>2</sup> demonstrando uma dispersão da produtividade em torno da média. Já o coeficiente de variação (CV), apresentou um valor de 24,43% podendo esse ser considerado um alto valor, segundo Pimentel Gomes (2002) caracterizando esse tratamento com baixa precisão. Com relação a amplitude, o valor foi de 26,60, menor se comparado a amplitude geral, demonstrando que esses valores da série estão próximos um do outro.

Já o Coeficiente de assimetria apresentou um valor de 0,03, demonstrando uma assimetria positiva. O coeficiente de Curtose com resultado de 1,91, demonstrando uma distribuição de cauda curta (platicúrtica).

**Variedade Jacaré:** Apresentou um valor de soma de 273,60 kg/6m<sup>2</sup>, onde o valor máximo foi de 78,00 e o valor mínimo foi de 64,60.

O valor do 1º quartil foi de 64,90 kg/6m<sup>2</sup>, representando 25% das observações deste tratamento e o valor do 3º quartil, foi de 69,00 kg/6m<sup>2</sup> representando 75% das observações deste tratamento.

O valor da mediana foi de 65,50 kg/6m<sup>2</sup>, valor esse igual a mediada dos dados gerais (65,50 kg/6m<sup>2</sup>). A média deste tratamento foi de 68,40 kg/6m<sup>2</sup>, valor superior à média geral dos dados (67,38 kg/6m<sup>2</sup>).

O desvio padrão do presente tratamento foi de 6,43, representando assim, a dispersão da produtividade. Ainda para uma melhor compreensão e comparação dos dados, foi calculado o Coeficiente de variação de 9,40%, que segundo Pimentel Gomes (2002) é um valor considerado médio, podendo caracteriza-lo com média precisão. O coeficiente de assimetria apresentou o valor de 1,13, apresentando uma assimetria positiva. Já o coeficiente de curtose teve o valor de 2,3, resultando em uma distribuição de cauda curta (platicúrtica).

**Variedade Paulistinha:** observa-se que esta variedade apresentou a soma de 197,60 kg/6m<sup>2</sup>. O valor mínimo 42,00, sendo este acima do valor geral dos dados que foi de 31,80 kg/6m<sup>2</sup>. Já o valor máximo apresentou 50,20 kg/6m<sup>2</sup>, valor inferior ao máximo geral.

O valor do 1º quartil foi de 46,27 kg/6m<sup>2</sup> representando 25% das observações deste tratamento, enquanto o 3º quartil apresentou valor 51,83 kg/6m<sup>2</sup>, correspondendo a 75% das observações deste tratamento mesmo este valor sendo abaixo do valor da média geral que foi de 79,05 kg/6m<sup>2</sup>.

O valor da mediana foi de 48,70 kg/6m<sup>2</sup>, valor esse menor que a mediana geral dos dados. A média deste tratamento foi de 49,40 kg/6m<sup>2</sup> valor esse inferior à média dos dados gerais.

O desvio padrão teve como resultado 6,71kg/6m<sup>2</sup>, representando a dispersão da produtividade. O coeficiente de variação apresentou 13,59 % resultado esse, segundo Pimentel Gomes, um resultado médio, caracterizando esse tratamento com precisão media.

O coeficiente de assimetria apresentou valor de 0,35, sendo esta superior a 0 podendo ser interpretado como uma assimetria positiva. O coeficiente de



curtose apresentou valor 1,97, resultando em uma distribuição de cauda curta (platicúrtica).

**Variedade Rainha:** nota-se que esta variedade apresentou soma igual a 399,9 kg/6m<sup>2</sup>. O valor mínimo 82,20 kg/6m<sup>2</sup>, sendo este valor bem acima do valor apresentado da média geral (31,80 kg/6m<sup>2</sup>) Com relação ao valor máximo foi de 120,50 kg/6m<sup>2</sup> sendo um valor igual a media geral dos dados (120,50 kg/6m<sup>2</sup>)

O valor do 1º quartil foi de 93,30 kg/6m<sup>2</sup>, representando 25% da observação deste tratamento, enquanto o 3º quartil apresentou um valor de 105,28 kg/6m<sup>2</sup>, representando 75% da observação deste tratamento sendo ambos os valores, superiores ao valor dos dados de 1º e 3º quartis dos dados gerais.

O valor médio deste tratamento foi de 99,97 kg/6m<sup>2</sup>, valor esse bem acima da média geral (67,38 kg/6m<sup>2</sup>), quanto a mediana foi de 98,60 kg/6m<sup>2</sup>, valor esse também acima da média geral dos dados (65,50 kg/6m<sup>2</sup>).

O desvio padrão obteve-se o resultado de 15,77 kg/6m<sup>2</sup>, representando a dispersão da produtividade. Já o coeficiente de variação apresentou resultado semelhante ao desvio padrão 15,77%, um resultado médio segundo Pimentel Gomes, o que caracteriza esse tratamento de precisão média.

O coeficiente de assimetria teve como resultado 0,29, maior que 0, resultando em uma interpretação de que esse tratamento tem simetria positiva. O coeficiente de curtose apresentou o valor de 1,99, resultando em uma distribuição cauda curta (platicúrtica).

**Variedade Yellow Yam:** O presente tratamento apresentou a soma igual a 296,40 kg/6m<sup>2</sup>. O valor mínimo foi de 68,50 kg/6m<sup>2</sup>, sendo esse valor acima dos dados gerais. Já o valor máximo foi de 88,20 kg/6m<sup>2</sup>, sendo esse valor inferior ao valor geral da máxima dos dados gerais (120,50 kg/6m<sup>2</sup>).

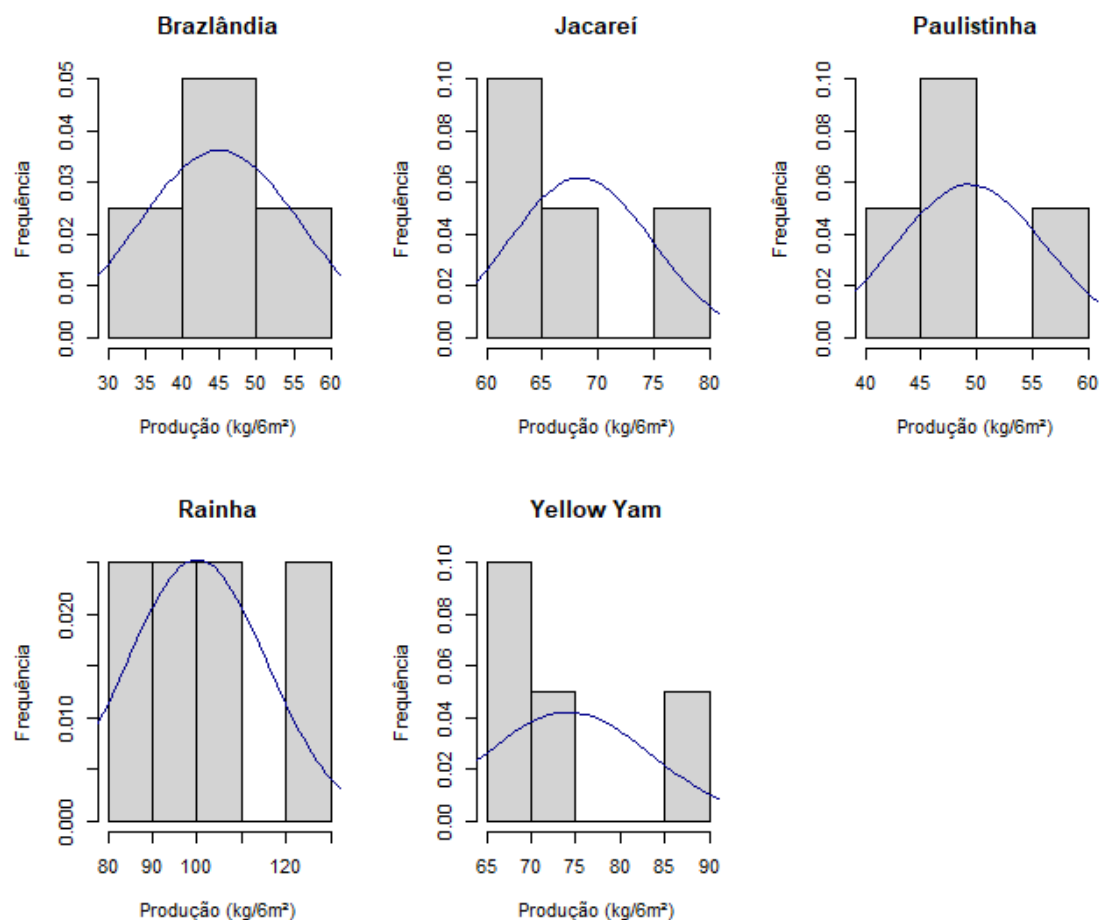
O valor do 1º quartil foi de 69,50 kg/6m<sup>2</sup> representando 25 % das observações deste tratamento, enquanto o 3º quartil teve o valor de 74,78 kg/6m<sup>2</sup> representando 75% das observações desta variável.

A mediana do tratamento em questão foi de 69,85 kg/6m<sup>2</sup> valor muito parecido com a mediana geral dos dados. Já a média foi de 74,10 kg/6m<sup>2</sup>, valor um pouco acima da média geral dos dados.

O resultado do desvio padrão foi de 9,43 kg/6m<sup>2</sup>, representando a dispersão da produtividade. Já o coeficiente de variação apresentou o valor de 12,72% um resultado que segundo Pimentel Gomes, é um valor médio, representando que esse tratamento é de precisão média.

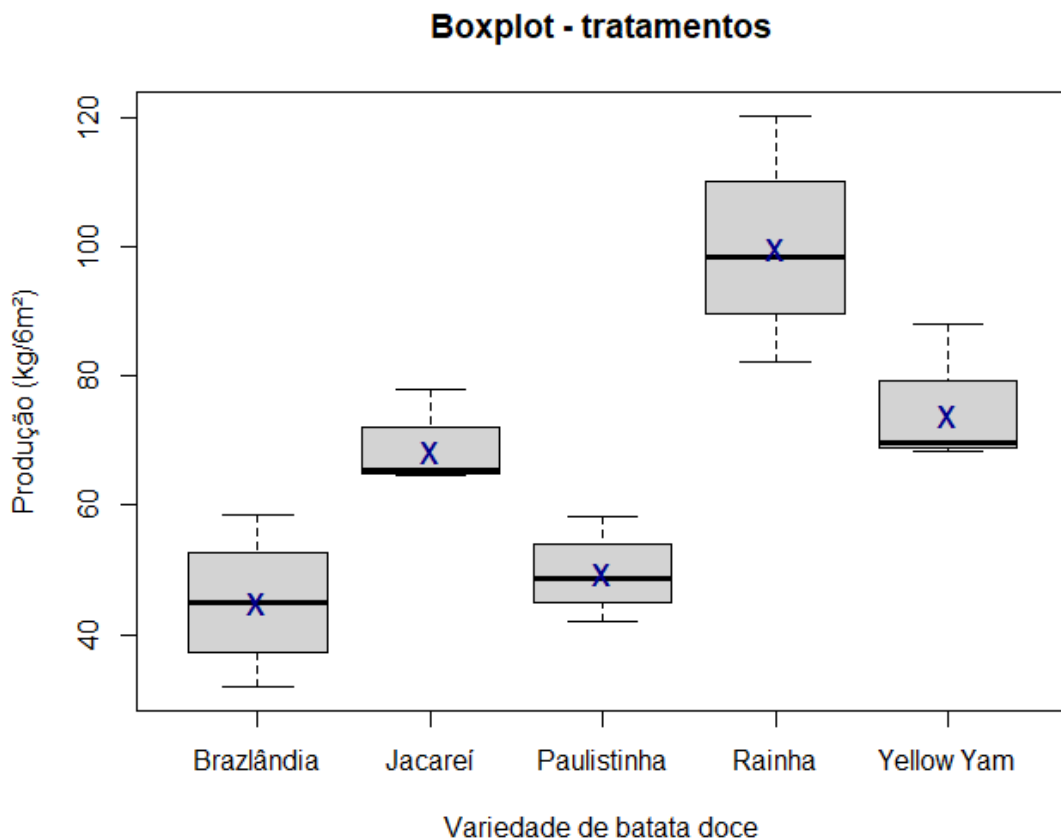
O coeficiente de assimetria teve o resultado de 1,13, sendo esse valor maior que 0 e representando uma assimetria positiva. Já o coeficiente de curtose teve o valor de 2,32 resultando em uma distribuição com cauda curta (platicúrtica).

**Figura 3:** Distribuição da frequência de produtividade de batata doce para cada tratamento.



De acordo com a figura 3 é possível analisar a curva da distribuição normal em cada variedade de batata doce. Nota-se, visualmente, que a variedade Brazlândia apresenta uma distribuição praticamente simétrica, o que se comprova pelo seu Coeficiente de Assimetria ( $A_s$ ) com valor de 0,03. As variedades Paulistinha e Rainha, com valores de  $A_s$  de 0,35 e 0,29, respectivamente, apresentam distribuição ligeiramente positiva. Por fim, as variedades Jacaré e Yellow Yam possuem uma maior concentração dos dados a esquerda da média, causando assimetrias positivas mais acentuadas, com valores de  $A_s = 1,13$  para ambas.

**Figura 4:** Gráfico Boxplot para cada uma das variedades de batata doce.



Comparando os valores da produção das variedades entre si, através da figura 4, constata-se que a variedade Rainha apresenta a maior produção média e também o maior desvio padrão. As variedades Jacareí e Yellow Yam apresentam valores intermediários de produção, seguidas por Paulistinha e Brazlândia, esta apresentando o menor valor médio de quilogramas por 6m<sup>2</sup>.

**d) Construa a ANOVA e interprete os resultados e suas conclusões usando 5% de significância.**

**R:** A ANOVA, para que seja interpretada admite-se a hipótese de nulidade ( $H_0$ ), de efeitos nulos para os tratamentos, ou seja, que a diferença entre os tratamentos e a variável resposta não esteja exercendo efeito uma sobre a outra.

Para o trabalho em questão, se utilizou um nível de significância a 5% de probabilidade

*$H_0$ : Nenhuma variedade de batata doce está exercendo efeito sobre a variável resposta (produção de batata doce) ( $t_1 = t_2 = t_3 = t_4 = t_5 = 0$ ).*

*$H_1$ : Pelo menos uma variedade de batata doce está exercendo efeito sobre a variável resposta (produção de batata doce) ( $t_i \neq 0$ , sendo  $i = 1, \dots, 5$ ).*

**Tabela 4:** Análise de variância (ANOVA) para o delineamento de produção de batata doce.

Fonte de variação	Graus de liberdade	Soma dos quadrados (S.Q.)	Quadrado dos resíduos (Q.M.)	F calculado	F tab.	Pvalor
Tratamento	4	7731,11	1932,78	17,74	3,06 <sup>a</sup>	(1,482e-05) *** b
Resíduo	15	1634,57	108,97			
Total	19	9365,68				

<sup>a</sup> Ao nível de probabilidade de 5% da tabela unilateral de F para o caso de  $F > 1$ , Referências bibliográficas: Gomes, F.P. Curso de estatística experimental. 11 ed. Piracicaba: Livraria Nobel S.A., 1985, 466p.

<sup>b</sup> Obtido através do Software R (Anexo II), com grau de signif.: 0 '\*\*\*\*' 0.001 '\*\*\*' 0.01 '\*\*' 0.05 '.' 0.1 ' ' 1

De acordo com a tabela 1 o  $F_{\text{calculado}}$  (17,74) apresentou valor superior ao  $F_{\text{tab.}}$  (3,06) com o nível de significância a 5%, rejeitando a hipótese  $H_0$  de nulidade. Assim, é possível concluir que pelo menos um tratamento está influenciando a produtividade da batata doce.

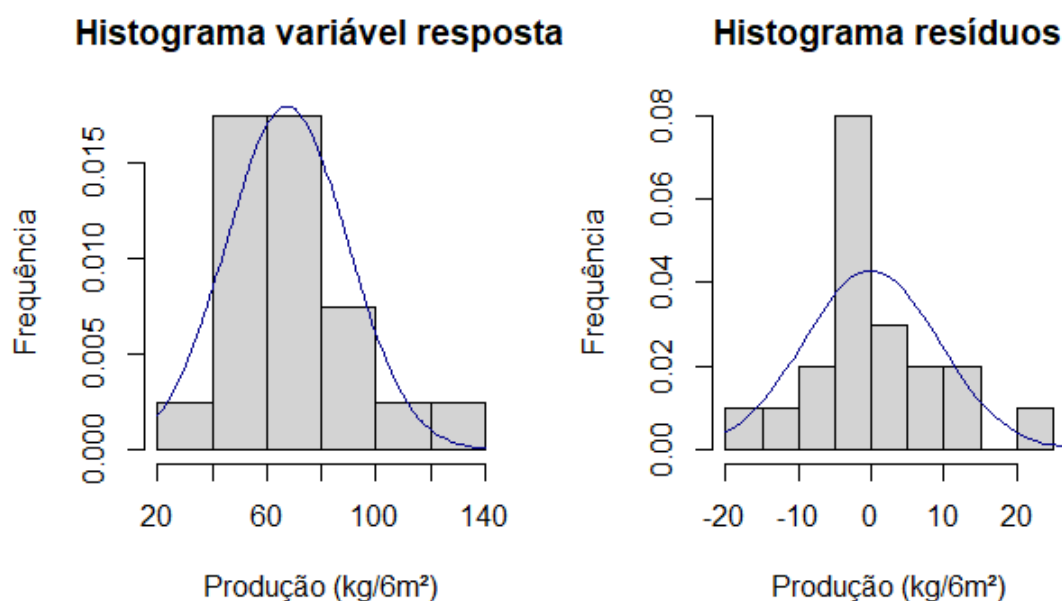
**e) Faça a análise das suposições do modelo (se houver necessidade de transformação dos dados, faça a análise de variância para o conjunto de dados transformados).**

**R:** As suposições associadas ao modelo são: normalidade, independência e homocedasticidade, além da análise dos resíduos padronizados.

**Normalidade**

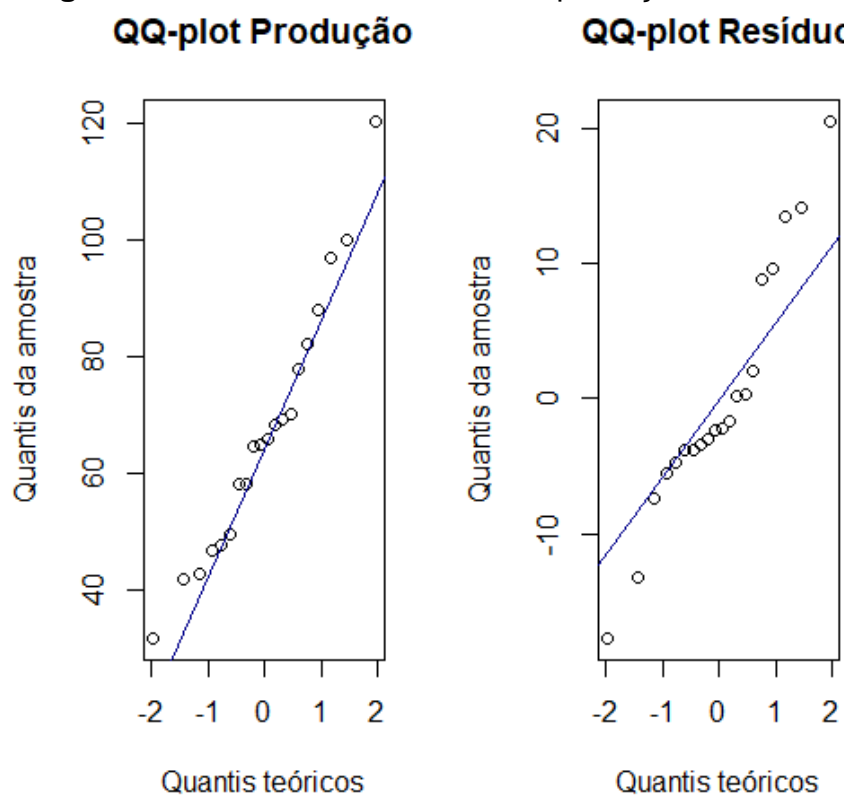
Para verificação da normalidade, primeiramente fez-se uma análise subjetiva dos gráficos de Histograma da variável resposta e dos resíduos, com as respectivas curvas da distribuição normal, e do gráfico QQ-plot.

**Figura 5:** Histogramas para verificação da normalidade



Busca-se nesses gráficos avaliar a similaridade da curva com o formato de sino da distribuição normal. Visualmente, tanto a variável resposta quanto os resíduos, apresentam uma distribuição praticamente simétrica.

**Figura 6:** Gráficos Quantil-Quantil da produção e do resíduo.



De acordo com a figura 6, os dados da produção, aparentemente, se mostraram com menor dispersão em torno da reta, em comparação com os dados do resíduo. Porém, pode-se dizer que ambas as nuvens de pontos se concentram em torno da reta, o que é um indício da normalidade.

Contudo, para uma análise mais objetiva, realizou-se o teste não paramétrico para distribuição normal de Shapiro-Wilk. Sendo as hipóteses testadas:

*H<sub>0</sub>: A produção de batata doce apresenta distribuição normal de probabilidade;*

*H<sub>1</sub>: A produção de batata doce não apresenta distribuição normal de probabilidade.*

**Tabela 5:** Os valores resultantes para o teste de Shapiro-Wilk foram:

<b>Shapiro-Wilk</b>	<b>W<sub>calc</sub></b>	<b>W(5%,20)</b>	<b>pvalor</b>
Produção	0,962	0,905 <sup>a</sup>	0,594
Resíduos da ANOVA	0,939		0,231

<sup>a</sup> Valor obtido na tabela de Shapiro-Wilk (PORTAL ACTION, *sd*).

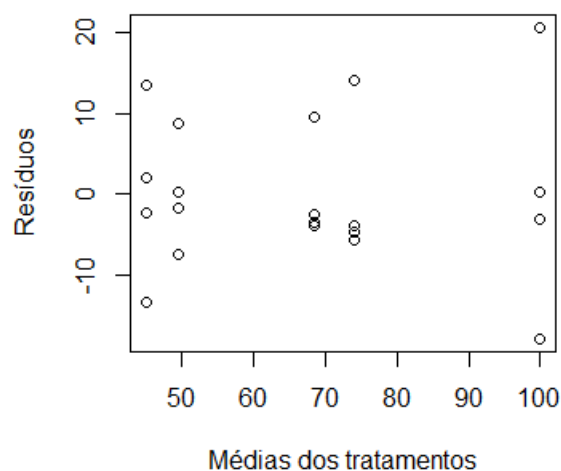
Com um grau de significância a 5%, como o W<sub>calculado</sub> (0,962) é maior que o W<sub>tabelado</sub> (0,905) para a variável produção, não se rejeita a hipótese nula, constatando-se a distribuição normal dos dados. Da mesma forma, com a mesma significância, o W<sub>calculado</sub> (0,939) para a variável dos resíduos permite concluir que os dados residuais seguem distribuição normal.

Tendo sido a hipótese nula aceita nas duas situações, observa-se que o pvalor de ambos é superior ao valor de 0,05 (correspondente a 5% de significância), confirmando a normalidade da distribuição dos dados do modelo.

### **Homocedasticidade**

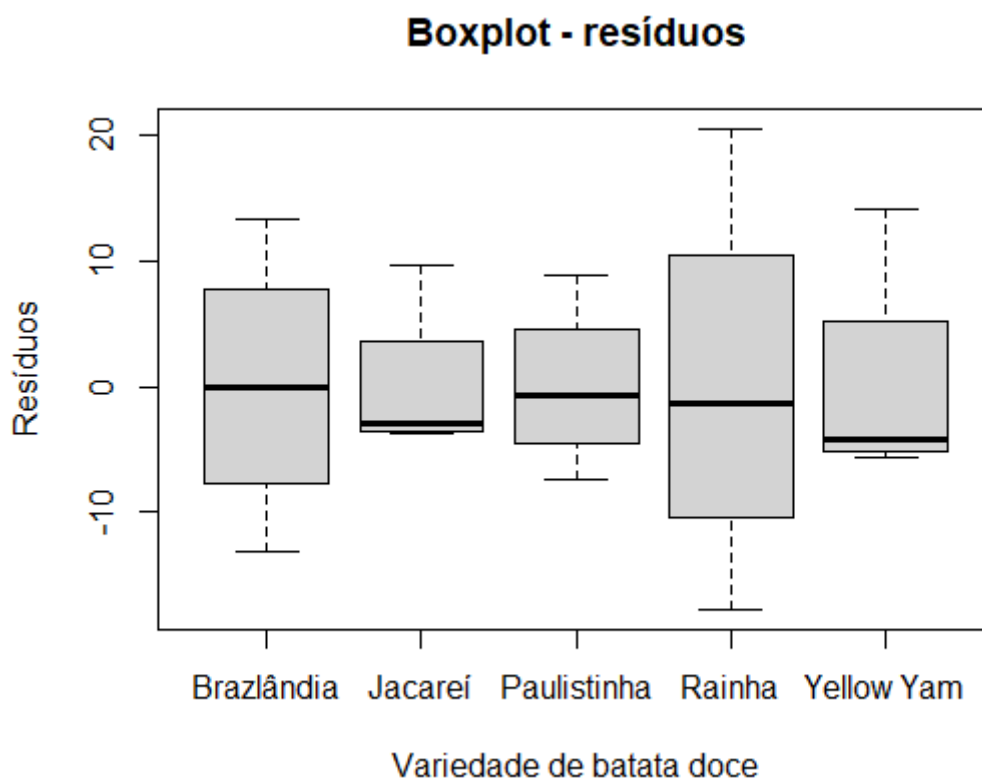
Nessa suposição, busca-se a homogeneidade das variâncias entre os tratamentos. Para isso, faz-se uma análise subjetiva, através do gráfico de dispersão dos resíduos pelas suas respectivas médias (Figura 7), e também o gráfico Boxplot dos resíduos para cada tratamento (Figura 8).

**Figura 7:** Gráfico de dispersão dos resíduos por suas respectivas médias dos tratamentos.



Aqui busca-se verificar a ausência de qualquer tendência na dispersão dos dados. Considerando os valores dos resíduos, é possível aferir uma distribuição com diferenças consideráveis quanto a amplitude, não sendo possível garantir, subjetivamente, a homocedasticidade.

**Figura 8:** Gráfico boxplot dos resíduos para cada tratamento.



Nessa análise, novamente não se nota nenhuma tendência visível na distribuição dos dados, mas constata-se as grandes diferenças de amplitude dos

valores dos resíduos. Sendo assim, para garantir, objetivamente, que os dados apresentam homocedasticidade, realizou-se o teste de Bartlett.

Para o teste, consideraram-se as hipóteses:

$H_0$ : Existe homogeneidade de variâncias entre as variedades de batata doce;

$H_1$ : Não existe homogeneidade de variâncias entre as variedades de batata doce.

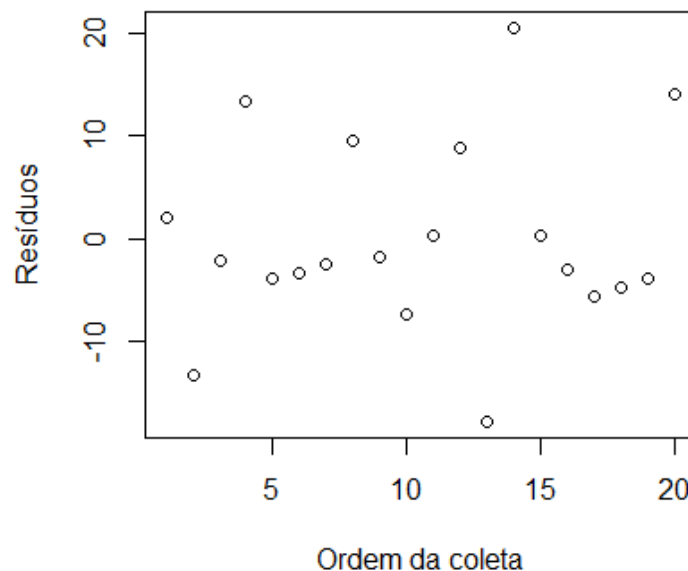
Os resultados do teste, a 5% de significância, foram um Qui-quadrado de 2,985 e um pvalor de 0,560, cujo é maior que 0,05 (referente ao grau de significância), permitindo não rejeitar estatisticamente a hipótese nula, garantindo a homocedasticidade dos dados da produção das variedades de batata doce analisadas.

### Independência

A independência dos erros geralmente é garantida no experimento ao trabalhar-se com o número de repetições adequado, além da aleatorização na etapa de distribuição dos tratamentos entre essas parcelas (Montgomery, 2001).

A fim de confirmar essa suposição, construiu-se o gráfico dos resíduos pela ordem de coleta dos dados.

**Figura 9:** Gráfico de dispersão entre os resíduos e sua respectiva ordem de coleta.



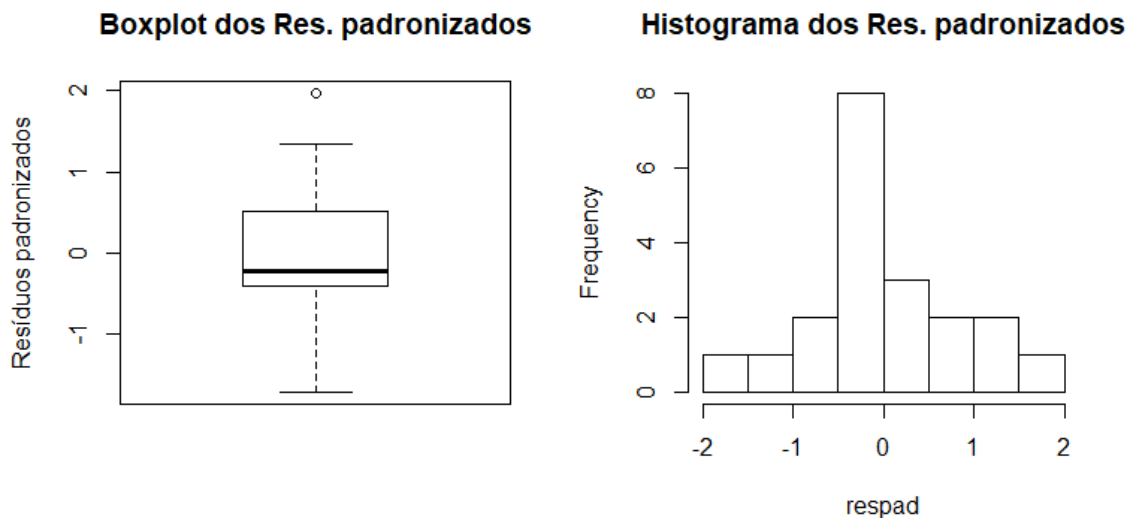
Como é observado no gráfico da Figura 9, ocorre uma distribuição aleatória da nuvem dos pontos, isto é, os erros não são dependentes, o que sugere que a suposição de independência está atendida no experimento.



### Análise dos resíduos padronizados

Esta análise foi realizada apenas para buscar outliers nos dados do conjunto. Através do software R (Anexo II), geraram-se os seguintes gráficos (Figura 10):

**Figura 10:** Gráficos de Boxplot e histograma para os resíduos padronizados.



Analisando os gráficos, nota-se que no Boxplot nenhum valor absoluto dos resíduos padronizados apresenta valor maior que 3, o que não indica nenhum potencial outlier, porém o mesmo gráfico assinala um ponto discrepante. Já no Histograma, verifica-se uma distribuição aparentemente normal.

**f) Faça o teste de Tukey ao nível de 5% de probabilidade. Tire conclusões.**

**R:** Tendo a ANOVA revelado que existe diferença significativa da produção de batata doce entre as variedades analisadas neste experimento, aplicou-se o teste Tukey, a 5% de significância, cujo realiza a comparação da produtividade entre todos os pares de tratamento. As hipóteses definidas para o teste foram:

$H_0$ : As médias de produção das variedades não diferem entre si ( $\mu_u = \mu_v$ );

$H_1$ : As médias de produção das variedades diferem entre si ( $\mu_u \neq \mu_v$ , com  $u \neq v$ ).

A Tabela 6 apresenta os resultados do teste Tukey a 5% de probabilidade, executado no software R (Anexo II).

**Tabela 6:** Teste Tukey (5% de significância).

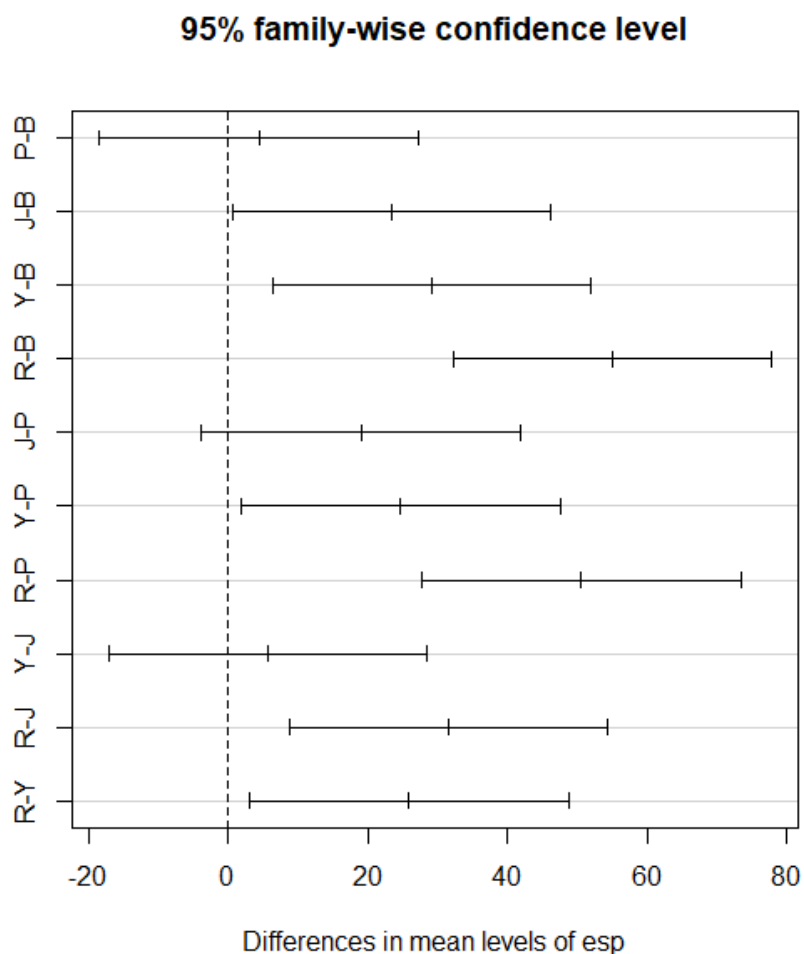
Par de variedades	Diferença das médias	d.m.s.	Valor mín. intervalo confiança	Valor máx. intervalo confiança	pvalor
J – B	23,40	22,79	0,61	46,19	0,043 *
P – B	4,40	22,79	-18,39	27,19	0,974 <sup>n.s.</sup>
R – B	54,98	22,79	32,18	77,77	<0,001 *
Y – B	29,10	22,79	6,31	51,89	0,010 *
P – J	19,00	22,79	-41,79	3,79	0,126 <sup>n.s.</sup>
R – J	31,58	22,79	8,78	54,37	0,005 *
Y – J	5,70	22,79	-17,09	28,49	0,935 <sup>n.s.</sup>
R – P	50,58	22,79	27,78	73,37	<0,001 *
Y – P	24,70	22,79	1,91	47,49	0,031 *
Y-R	25,88	22,79	-48,67	-3,08	0,023 *

\* significativo a 5%; <sup>n.s.</sup> não significativo a 5%.

A decisão de rejeição ou aceitação da hipótese nula, para cada par de tratamento avaliado, pode ser tomada de duas formas: através da comparação da diferença das médias com o d.m.s., sendo aceito  $H_0$  quando  $|\mu - mv| > d.m.s.$ ; ou analisando o pvalor com o grau de significância (nesse caso 5%), aceitando-se  $H_0$ , quando  $pvalor < 0,05$ .

Uma terceira forma de verificar a aceitação da hipótese nula é observando os intervalos de confiança (Figura 6) entre as médias de Tukey (95% de confiabilidade), sendo que os valores que contém o 0 (zero), significa que a diferença estatística não foi significativa.

**Figura 11:** Intervalos de confiança para a diferença entre médias do Teste Tukey com 95% de confiabilidade.



Analisando a figura, confirma-se as decisões tomadas na Tabela 7. Isto é, os pares de tratamento Brazlândia-Paulistinha, Jacareí-Paulistinha e Jacareí-YellowYam não foram significativamente diferentes entre si, ao nível de 5% de confiança.

A tabela 7 apresenta as diferenças entre os pares de tratamento do Teste Tukey a 5% de significância de forma resumida. A figura T, de forma análoga, apresenta o gráfico de barras para a produtividade dos tratamentos, onde tem-se uma melhor observação das variedades que são estatisticamente (5%) iguais.

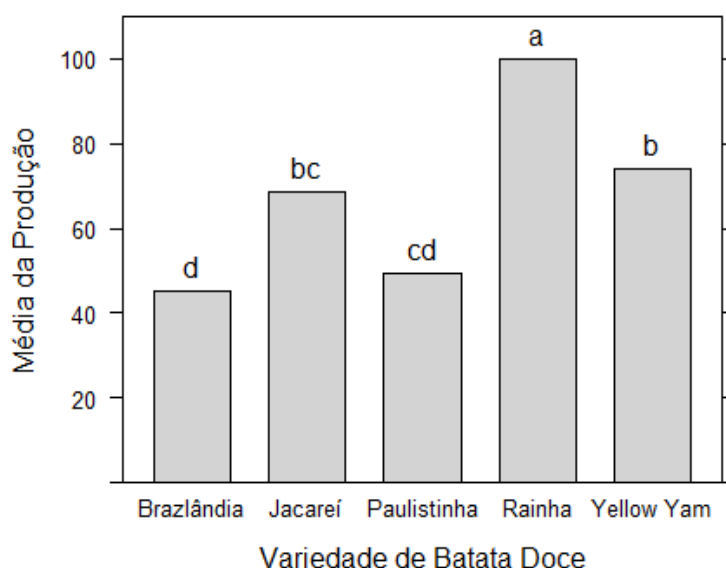
**Tabela 7:** Resultado do Teste Tukey, 5% de significância, Coeficiente de variação: 15,49%.

Variedade de batata doce (Tratamentos)	Média de produção (kg/6m <sup>2</sup> )	Teste Tukey 5% *
Rainha	99,98	a
Yellow Yam	74,10	b
Jacareí	68,40	bc
Paulistinha	49,40	cd
Brazlândia	45,00	d

\*Letras diferentes significam que os tratamentos são, em média, estatisticamente diferentes, com 5% de significância

O teste permite afirmar que a variedade Rainha, cuja apresenta a maior média de produção (99,98 kg/6m<sup>2</sup>), difere estatisticamente das demais variedades, com um grau de significância de 5%. Em seguida, as variedades com a segunda maior média foram Yellow Yam e Jacareí (médias de produção de 74,10 e 68,40 kg/6m<sup>2</sup>, respectivamente), consideradas estatisticamente iguais (5%). A produtividade da variedade Paulistinha, com valor médio de 49,40 kg/6m<sup>2</sup>, não pode ser considerada estatisticamente diferente (5%) da variedade Jacareí, e nem da Brazlândia, que foi a variedade com a menor média de produção (45,00 kg/6m<sup>2</sup>).

**Figura 12:** Gráfico de barras das média de produção de batata doce (kg/6m<sup>2</sup>) com a comparação do Teste Tukey (95% de confiança).



**g) Supondo que a variedade Brazlândia seja considerada como grupo controle, realize o teste de hipótese adequado para a comparação das demais variedades com o controle, com 5% de significância.**

**R:** Para realização de teste de hipóteses com um grupo controle o indicado é o teste Dunnet. Sendo a variedade Brazlândia considerada o controle, tem-se as hipóteses:

$H_0$ : A média de produção da variedade de batata doce é igual a média do tratamento controle ( $\mu_i = \mu_{\text{controle}}$ );

$H_1$ : A média de produção da variedade de batata doce é diferente da média do tratamento controle ( $\mu_i \neq \mu_{\text{controle}}$ );

O teste Dunnett foi realizado no software R (Anexo II), e os resultados estão na Tabela 8.

**Tabela 8:** Teste Dunnett, Variedade Controle: Brazlândia, 95% de confiança (5% de significância).

Hipóteses Lineares	Estimativa	Erro Padrão	Tvalor(calc)	Ttab(4,15),5% <sup>a</sup>	Pvalor <sup>b</sup>
Jacareí-Brazlândia	23,40	7,38	3,17	2,36	0,0213*
Paulistinha-Brazlândia	4,40	7,38	0,60	2,36	0,9329
Rainha-Brazlândia	54,98	7,38	7,45	2,36	<0,001***
YellowYam-Brazlândia	29,10	7,38	3,94	2,36	0,0044**

<sup>a</sup> Valor de t obtido em “Tabela A.8 – Valores de t para comparações unilaterais para o teste de Dunnett entre o número de médias do tratamento, excluindo o controle, para os níveis de probabilidade de  $p=0,95$  e  $p=0,99$ ”;

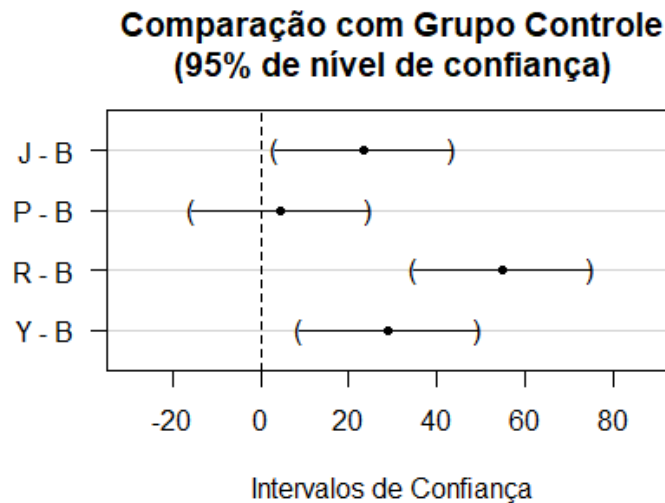
<sup>b</sup> Pvalor obtido no software R (Anexo II), com grau de signif.: 0 ‘\*\*\*\*’ 0.001 ‘\*\*\*’ 0.01 ‘\*\*’ 0.05 ‘.’ 0.1 ‘.’ 1.

Para a tomada de decisão quanto as hipóteses estatísticas, compara-se o Tvalor calculado e o Tvalor tabelado (com 5 % de significância), sendo que quando  $T_{\text{valor}}(\text{calc}) > T_{\text{tab}}$ , rejeita-se  $H_0$ , concluindo que a média de produção do tratamento é significativamente (5%) diferente da média do controle. A mesma decisão pode ser obtida comparando o pvalor com o grau de significância do teste Dunnett (5%), sendo que  $p_{\text{valor}} < 0,05$ , também se rejeita a hipótese nula.

Desta forma, observou-se que apenas a variedade Paulistinha não diferiu estatisticamente do grupo controle (Brazlândia). As outras variedades, Jacareí, Yellow Yam e Rainha apresentaram diferença significativa (5%) quanto às médias de produção ao serem comparadas com o tratamento controle.

A figura 13 apresenta os intervalos de confiança para as diferenças médias entre os tratamentos e o grupo controle, sendo que o intervalo que contem o 0 (zero) indica que não há diferença estatística (a 5% de significância) entre a produção média da variedade comparada e o tratamento controle.

**Figura 13:** Gráfico do intervalo de confiança do teste Dunnett para a diferença média entre os tratamentos e o controle (Brazlândia).



- h) Supondo que (1) as variedades Brazlândia, Jacareí e Paulistinha têm aspectos quanto a formato muito similares, comparando com as variedades Rainha e Yellow Yam, e (2) as variedades Brazlândia e Jacareí possuem características genéticas muito similares, comparando com a variedade Paulistinha; desdobre a análise de variância em contrastes ortogonais que considerem as características acima.

**R:** O número de contrastes a serem empregados se dá pelo  $C = (\text{número de tratamentos} - 1)$ . Neste caso, um total de 4 contrastes podem ser elaborados.

Os contrastes formulados, e suas respectivas comparações, estão enumerados a seguir (o valor a ser comparado é a média de produtividade de cada tratamento e os termos dos contrastes estão representados pelas iniciais do nome de cada variedade):

$C1 = (2B + 2J + 2P - 3R - 3Y)$  atendendo a condição (1) da questão, isto é, comparando as variedades B, J e P com R e Y;

$C2 = (1B + 1J - 2P)$  realizando a comparação da condição (2) da questão, ou seja, comparando B e J com P;

$C3 = (1B - 1J)$  comparando as variedades B e J entre si;

$C4 = (1R - 1Y)$  aqui sendo comparado as variedades R e Y entre si.

Verifica-se então se todos são contrastes, somando todos os coeficientes adotados dentro de cada contraste, sendo necessário resultar 0 (zero):

$$C1 = 2 + 2 + 2 - 3 - 3 = 0$$

$$C2 = 1 + 1 - 2 + 0 + 0 = 0$$

$$C3 = 1 - 1 + 0 + 0 + 0 = 0$$

$$C4 = 0 + 0 + 0 + 1 - 1 = 0$$

Para verificar se os contrastes apresentam variação independente entre si, calcula-se o somatório da multiplicação dos coeficientes, par a par, e observa-se novamente se o resultado é nulo (zero):

$$C1 * C2 = 2*1 + 2*1 + 2*(-2) + (-3)*0 + (-3)*0 = 0$$

$$C1 * C3 = 2*1 + 2*(-1) + 2*0 + (-3)*0 + (-3)*0 = 0$$

$$C1 * C4 = 2*0 + 2*0 + 2*0 + (-3)*1 + (-3)*(-1) = 0$$

$$C2 * C3 = 1*1 + 1*(-1) + (-2)*0 + 0*0 + 0*0 = 0$$

$$C2 * C4 = 1*0 + 1*0 + (-2)*0 + 0*1 + 0*(-1) = 0$$

$$C3 * C4 = 1*0 + (-1)*0 + 0*0 + 0*1 + 0*(-1) = 0$$

Com essa análise, afere-se que os contrastes são ortogonais. Parte-se então para o desdobramento dos contrastes na ANOVA, com o uso do software R. Os resultados estão apresentados na Tabela 9 a seguir. As hipóteses de teste são:

$H_0: C = 0$  os grupos comparados são iguais;

$H_1: C \neq 0$  os grupos comparados são diferentes.

**Tabela 9:** Desdobramento da análise de variância (ANOVA) para os contrastes.

Fonte de variação	Graus de liberdade	Soma dos quadrados (S.Q.)	Quadrado dos resíduos (Q.M.)	Fcalc.	F <sub>tab.</sub> <sup>a</sup>	Pvalor <sup>b</sup>
Tratamento	4	7731	1933	17,74	3,06	1,48e-05 ***
C1	1	5155	5155	47,31	4,54	5,25e-06 ***
C2	1	142	142	1,30	4,54	0,271
C3	1	1095	1095	10,05	4,54	0,006 **
C4	1	1339	1339	12,29	4,54	0,003 **
Resíduo	15	1635	109			
Total	19	9366				

<sup>a</sup>Ao nível de probabilidade de 5% da tabela unilateral de F para o caso de  $F > 1$ , Referências bibliográficas: Gomes, F.P. Curso de estatística experimental. 11 ed. Piracicaba: Livraria Nobel S.A., 1985, 466p;

<sup>b</sup>Obtido através do Software R (Anexo II), com grau de signif.: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

A tomada de decisão, em relação as hipóteses de teste, ocorre comparando os valores do  $F_{\text{calculado}}$  e do  $F_{\text{tabelado}}$ , de modo que se  $F_{\text{calc}} > F_{\text{tab}}$ , rejeita-se  $H_0$ , com um grau de 5% de significância, podendo se afirmar que a comparação existente no contraste é estatisticamente diferente. Em caso de não haver evidências para se rejeitar  $H_0$ , assume-se que os grupos comparados no contraste são iguais entre si (5%).

Assim sendo, o contraste C1, que teve a hipótese nula rejeitada ( $F_{\text{calc}} > F_{\text{tab}}$ ), mostra que as médias de produção das variedades Brazlândia, Jacaré e Paulistinha não são significativamente iguais (ao nível de 5% de confiança) às médias das variedades Rainha e Yellow Yam. Da mesma forma, os contrastes C3 e C4 mostram, respectivamente, que a produção média da variedade Brazlândia é significativamente diferente da Jacaré, e a variedade Rainha diferente da Yellow Yam. Já o contraste C2 teve a hipótese nula mantida ( $F_{\text{calc}} < F_{\text{tab}}$ ), de onde pode-se concluir com 5% de significância que a produção média das variedades Brazlândia e Jacaré são iguais, estatisticamente, à Paulistinha.



## REFERÊNCIAS BIBLIOGRÁFICAS

GOMES, F.P. **Curso de estatística experimental**. 11 ed. Piracicaba: Livraria Nobel S.A., 1985, 466p.

MONTGOMERY, D. C. **Introdução ao controle estatístico da qualidade**. 4. ed. [S.l.]: E Fourth Edition, 2001.

PIMENTEL GOMES, F. **Curso de estatística experimental**. 13 ed. São Paulo: Nobel, 2000. 477p.

PORTAL ACTION. **6.4 – Teste de Shapiro-Wilk**. Disponível em <<http://www.portalaction.com.br/inferencia/64-teste-de-shapiro-wilk>> Acesso em 02 mai 2018.

SANTOS, R.V. EVANGELISTA, A.R. PINTO, J.C. FILHO, C.C.C.C. SOUZA, R.M. Composição química da cana-de-açúcar (*saccharum spp.*) e das silagens com diferentes aditivos em duas idades de corte. **Ciência agrotecnologia**., Lavras, v.30, n.6, p. 1181-1189, nov./dec. 2006.

STELL, R.G.D e TORRIE, J.H. **Principles and procedures of statistics: with special reference to the biological science**. New York: Mc Graw-Hill Book Company, Inc., 1960. 481 p.

## **Anexo I** – Artigo analisado na questão 1

**Título:** Composição química da cana-de-açúcar (*saccharum spp.*) e das silagens com diferentes aditivos em duas idades de corte.

**Autores:** Roberto Valadares Santos, Antônio Ricardo Evangelista, José Cardoso Pinto, Cristovão Colombo de Carvalho Couto Filho, Ronan Magalhães de Souza.

**Ano:** 2006

## Anexo II – Scrip usado no software R para resolução da questão 2

```
## _____ QUESTÃO 2 _____ ##
```

```
##-----##
```

```
## ____ ALTERNATIVA C ____ ##
```

```
##-----##
```

```
dados <- read.table("dados_lista1_produção_batata_doce.txt", header =  
T) ## lendo um conjunto de dados em txt, header = T tem nome das  
colunas na 1ª linha ##
```

```
dados
```

```
attach(dados)
```

```
## __ ANÁLISE DESCRITIVA GERAL __ ##
```

```
summary(prod) ## resumo descritivo geral ##
```

```
var(prod) ## variância geral ##
```

```
sd(prod) ## desvio padrão geral ##
```

```
((sd(prod))/(mean(prod)))*100 ## coeficiente de variação ##
```

```
sum(prod) ## soma geral ##
```

```
max(prod)-min(prod) ## amplitude dos dados ##
```

```
require(moments)
```

```
## Assimetria (AS) ##
```

```
## AS = 0 distribuição simétrica;
```

```
## AS > 0 distribuição assimétrica positiva;
```

```
## AS < 0 distribuição assimétrica negativa.
```

```
skewness(prod)
```

```
## Curtose (CUR) ##
```

```
## CUR = 3 distribuição com caudas neutras (normais - mesocúrtica);
```

```
## CUR > 3 distribuição com caudas longas ou pesadas (leptocúrtica);
```

```
## CUR < 3 distribuição com caudas curtas ou leves (platicúrtica).
```

```
kurtosis(prod) #curtose#
```

```
## histograma ##
```

```
x <- seq(20, 140, length.out = 120)
d <- dnorm(x, mean=mean(prod), sd=sd(prod))
hist(prod, main = "Histograma", xlab = "Produção (kg/6m²)", ylab =
"Frequência",
col=c("lightgray"), prob=T)
lines (x, d, col = "darkblue")
```

```
## boxplot ##
```

```
(média = mean(prod))
boxplot(prod, main = "Boxplot", ylab = "Produção (kg/6m²)",
col=c("lightgray"))
points(média, pch='x', cex=1.5, col='darkblue')
```

```
## __ ANÁLISE DESCRITIVA POR TRATAMENTO __ ##
```

```
tapply(prod, esp, sum)
tapply(prod, esp, summary)
tapply(prod, esp, max)-tapply(prod, esp, min) ## amplitude(esp) =
(max-min) ##
tapply(prod, esp, mean)
tapply(prod, esp, var)
tapply(prod, esp, sd)
(tapply(prod, esp, sd)/tapply(prod, esp, mean))*100 ## coeficiente de
variação = sd/mean*100 ##
tapply(prod, esp, skewness)
tapply(prod, esp, kurtosis)
```

```
## histograma de cada tratamento ##
```

```
dados2 <-
read.table("dados_lista1_produção_batata_doce_histograma_especies.txt")
```

```
, header = T) ## lendo um conjunto de dados em txt, header = T tem  
nome das colunas na 1ª linha ##
```

```
dados2
```

```
attach(dados2)
```

```
par(mfrow=c(2,3))
```

```
xB <- seq(20, 140, length.out=120)  
dB <- dnorm(x, mean=mean(dados2$B), sd=sd(dados2$B))  
hist(dados2$B, main = "Brazlândia", xlab = "Produção (kg/6m²)",  
ylab = "Frequência",  
col=c("lightgray"), prob=T)  
lines (xB, dB, col = "darkblue")
```

```
xJ <- seq(20, 140, length.out=120)  
dJ <- dnorm(x, mean=mean(dados2$J), sd=sd(dados2$J))  
hist(dados2$J, main = "Jacareí", xlab = "Produção (kg/6m²)",  
ylab = "Frequência",  
col=c("lightgray"), prob=T)  
lines (xJ, dJ, col = "darkblue")
```

```
xP <- seq(20, 140, length.out=120)  
dP <- dnorm(x, mean=mean(dados2$P), sd=sd(dados2$P))  
hist(dados2$P, main = "Paulistinha", xlab = "Produção (kg/6m²)",  
ylab = "Frequência",  
col=c("lightgray"), prob=T)  
lines (xP, dP, col = "darkblue")
```

```
xR <- seq(20, 140, length.out=120)  
dR <- dnorm(x, mean=mean(dados2$R), sd=sd(dados2$R))  
hist(dados2$R, main = "Rainha", xlab = "Produção (kg/6m²)",  
ylab = "Frequência",  
col=c("lightgray"), prob=T)  
lines (xR, dR, col = "darkblue")
```

```
xy <- seq(20, 140, length.out=120)
```

```

dY <- dnorm(x, mean=mean(dados2$Y), sd=sd(dados2$Y))
hist(dados2$Y, main = "Yellow Yam", xlab = "Produção (kg/6m²)",
ylab = "Frequência",
col=c("lightgray"), prob=T)
lines (xY, dY, col = "darkblue")

```

```
## boxplot separado por tratamento ##
```

```

(médias = tapply(prod, esp, mean))
boxplot(prod ~ esp, main = "Boxplot - tratamentos",
xlab = "Variedade de batata doce", ylab = "Produção (kg/6m²)", names =
c("Brazlândia", "Jacareí", "Paulistinha", "Rainha", "Yellow Yam"),
col=c("lightgray"))
points(médias, pch='x', cex=1.5, col='darkblue')

```

```

##-----##
## ____ ALTERNATIVA D ____ ##
##-----##

```

```
## ANÁLISE DE VARIÂNCIA ##
```

```

dados <- read.table("dados_lista1_produção_batata_doce.txt", header =
T) ## lendo um conjunto de dados em txt, header = T tem nome das
colunas na 1ª linha ##

```

```
dados
```

```
attach(dados)
```

```

esp <- as.factor(esp) ## informação de que a esp é um fator ##
anv <- aov(prod ~ esp) ## cálculo da análise de variância ##
anova(anv) ## demonstrativo da Tabela ANOVA ##

```

```

##-----##
## ____ ALTERNATIVA E ____ ##
##-----##

anv$residuals          ## resíduos ##
cbind(dados,anv$residuals)

## ANÁLISE DAS PRESSUPOSIÇÕES DO MODELO ##

## ____ normalidade ____ ##

# histogramas com a curva normal #
par(mfrow = c(1,2))
x <- seq(20, 140, length.out = 120)          ## o intervalo de
valores foi verificado no histograma ##
d <- dnorm(x, mean = mean(prod), sd = sd(prod))
hist(prod, main = "Histograma variável resposta", xlab = "Produção
(kg/6m²)", ylab = "Frequência", prob = T, col=c('lightgrey'))
lines(x,d, col = "darkblue")

xres <- seq(-20, 30, length.out = 50)          ## o intervalo de
valores foi verificado no histograma ##
dres <- dnorm(xres, mean = mean(anv$residuals), sd =
sd(anv$residuals))
hist(anv$residuals, main = "Histograma resíduos", xlab = "Produção
(kg/6m²)", ylab = "Frequência", prob = T, col=c('lightgrey'))
lines(xres,dres, col = "darkblue")

# Gráficos QQ-plot #

par(mfrow = c(1,2))      ## abre uma janela gráfica com espaço para 2
gráficos ##

qqnorm(prod, main = "QQ-plot Produção",
ylab="Quantis da amostra", xlab="Quantis teóricos")
qqline(prod, col = "darkblue")

qqnorm(anv$residuals, main = "QQ-plot Resíduo",
ylab="Quantis da amostra", xlab="Quantis teóricos")
qqline(anv$residuals, col = "darkblue")

```

```

# teste de Shapiro-wilk #

shapiro.test(prod)    ## teste de normalidade de Shapiro wilk ##
shapiro.test(anv$residuals) ## teste de normalidade de Shapiro wilk ##

## __ homogeneidade das variâncias __ ##

# gráfico de dispersão resíduos x médias #

n <- length(prod)          ## nº total de parcelas
I <- length(levels(esp))    ## nº de tratamentos
r <- n/I                   ## nº de repetições

repmean <- rep(tapply(prod, esp, mean), each = r)
plot(repmean, anv$residuals, xlab = "Médias dos tratamentos", ylab =
"Resíduos")

# boxplot para resíduos/tratamento #

boxplot(anv$residuals ~ esp, main = "Boxplot - resíduos",
xlab = "Variedade de batata doce", ylab = "Resíduos", names =
c("Brazlândia", "Jacareí", "Paulistinha", "Rainha", "Yellow Yam"),
col=c("lightgray"))

# testes estatísticos #

bartlett.test(prod ~ esp)  ## teste de Bartlett: homogeneidade das
variâncias ##

## __ independência __ ##

plot(1:n, anv$residuals, xlab = "Ordem da coleta", ylab = "Resíduos")

## análise dos resíduos padronizados: avaliação de outliers ##

```



```

res <- resid(anv)
res

respad <- res/sqrt(sum(res^2)/anv$df.res)  ## resíduos padronizados
##
respad

par(mfrow = c(1,2))
boxplot(respad, ylab = "Resíduos padronizados", main = "Boxplot dos
Res. padronizados")
hist(respad, main = "Histograma dos Res. padronizados")

##-----##
## ____ ALTERNATIVA F ____ ##
##-----##

## Teste Tukey ##

qtukey(0.95,5,15)                ## valor da tabela de
Tukey ##
qtabtukey <- qtukey(0.95,5,15)

dms <- (qtabtukey*(sqrt(108.97/4)))  ## valor do dms ##
dms

TukeyHSD(anv)                    ## teste Tukey, sem pacote
"laercio" ##

## gráfico do intervalo de confiança para as diferenças médias ##
plot(TukeyHSD(anv,wich='esp',ordered=TRUE,conf.level=0.95))

require(laercio)                  ## pacote laercio ##
LTukey(anv,"esp")                 ## teste Tukey, default =
5%, com pacote "laercio" ##

## gráfico em colunas com a comparação de médias ##

```

```

espe <- c("Brazlândia","Jacareí","Paulistinha","Rainha","Yellow Yam")
espe <- as.factor(espe)
medi <- c(45,68.4,49.4,99.975,74.1)
medi

```

```

let <- c("d","bc","cd","a","b")
let <- as.vector(let)
max(medi)

```

```

die <- data.frame(espe, medi, let)

```

```

require(lattice)

```

```

barchart(medi~espe, data=die, horiz=FALSE,
          ylab="Média da Produção", ylim=c(0,110),
          xlab="Variedade de Batata Doce", col = "lightgray",
          panel=function(x, y, subscripts, ...){
            panel.barchart(x, y, subscripts=subscripts, ...)
            panel.text(x, y, label=die[subscripts,"let"], pos=3, cex =
1)
            })

```

```

##-----##

```

```

## ____ ALTERNATIVA G ____ ##

```

```

##-----##

```

```

# TESTE DUNNETT: COMPARAÇÃO DOS TRATAMENTOS COM O CONTROLE #

```

```

# SEMPRE CONSIDERA QUE O 1º TRATAMENTO DO BANCO DE DADOS É O CONTROLE
#

```

```

require(multcomp)

```

```

esp <- as.factor(esp)
anv <- aov(prod ~ esp)

```

```
ht <- glht(anv, linfct = mcp(esp = "Dunnett")) ## o trat 1 no banco
de dados sempre será a testemunha ##
```

```
summary(ht)
```

```
confint(ht, level = 0.95)
```

```
plot(ht, xlim=c(-30,90), xlab="Intervalos de Confiança",
```

```
main="Comparação com Grupo Controle
```

```
(95% de nível de confiança)) ## gráfico do
intervalo de confiança para a diferença média ##
```

```
# obs.: intervalo que contém o zero, então o tratamento e o controle
são iguais #
```

```
##-----##
```

```
## ____ ALTERNATIVA H ____ ##
```

```
##-----##
```

```
## Contrastes Ortogonais ##
```

```
# C1 = 2B + 2J + 2P - 3R - 3Y
```

```
# C2 = 1B + 1J - 2P + 0R + 0Y
```

```
# C3 = 1B - 1J + 0J + 0R + 0Y
```

```
# C4 = 0B + 0J + 0P + 1R - 1Y
```

```
dados <- read.table("dados_lista1_produção_batata_doce.txt", header =
T) ## lendo um conjunto de dados em txt, header = T tem nome das
colunas na 1ª linha ##
```

```
dados
```

```
attach(dados)
```

```
con1 <- matrix(c(2,2,2,-3,-3,1,1,-2,0,0,1,-1,0,0,0,0,0,0,1,-1),
```

```
nrow = 5, ncol = 4)
```

```
con1
```

```
esp <- as.factor(esp) ## informação de que a espécie é um fator
#
```

```
contrasts(esp) <- con1
```

```
contrasts(esp)
```

```
anv <- aov(prod ~ esp) ## cálculo da análise de variância ##
```

```
anova(anv)
```

```
summary.aov(anv, split=list("esp"=list("c1" = 1, "c2" = 2, "c3" = 3,  
"c4" = 4)))
```

```
##-----##
```

```
##-----##
```

```
##-----Fim-----##
```