

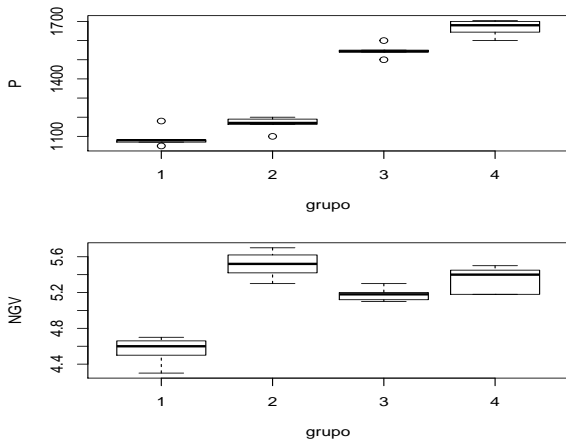
Exemplo sobre produção de feijões

Prof. Caio Azevedo

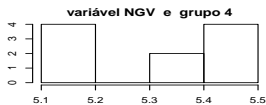
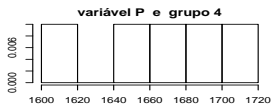
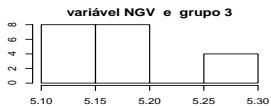
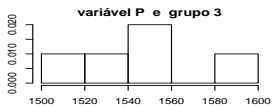
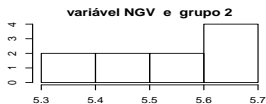
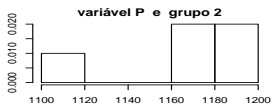
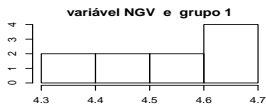
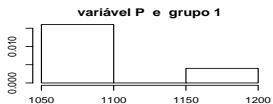
- Os dados são provenientes do exemplo 6.1 do livro [Ferreira 1996, p. 226] e correspondem a um estudo sobre quatro variedades de feijão, denotadas por grupos 1, 2, 3 e 4.
- Durante a seca foram medidas a produtividade (P) em kg/ha e o número de grãos por vagem (NGV).
- Cinco observações foram coletadas para cada variedade.

- Seja Y_{ijk} : o valor da k -ésima variável ($k = 1, 2$), para o j -ésimo indivíduo ($j = 1, \dots, 5$) do i -ésimo grupo ($i = 1, 2, 3, 4$). Os dados são balanceados (em relação aos grupos).
- Utilizamos a suposição que $\mathbf{Y}_{ij} = (Y_{ij1}, Y_{ij2})' \stackrel{ind.}{\sim} N_2(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$, independentes entre si.

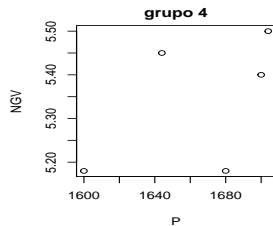
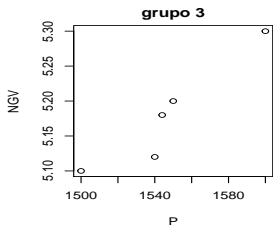
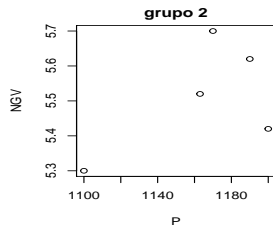
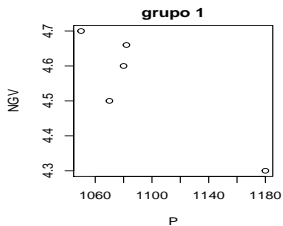
Boxplots



Histogramas



Gráficos de dispersão



Medidas descritivas da variável P

Grupo	Média	DP	Mínimo	Mediana	Máximo	n
1	1092.4	50.58458	1050	1080	1180	5
2	1164.6	39.06149	1100	1170	1200	5
3	1546.8	35.65389	1500	1544	1600	5
4	1665.6	43.68982	1600	1680	1704	5

Medidas descritivas da variável NGV

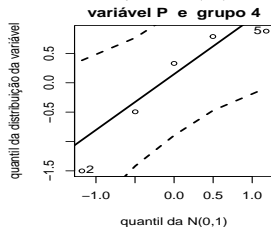
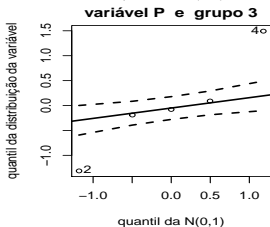
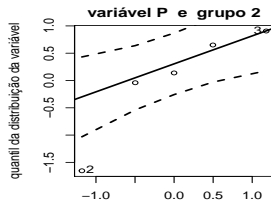
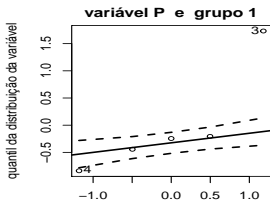
Grupo	Média	DP	Mínimo	Mediana	Máximo	n
1	4.552	0.1597498	4.3	4.6	4.7	5
2	5.512	0.1584929	5.3	5.52	5.7	5
3	5.18	0.07874008	5.1	5.18	5.3	5
4	5.342	0.1520526	5.18	5.4	5.5	5

- Variâncias na diagonal principal, correlações abaixo e covariâncias acima (da diagonal principal).

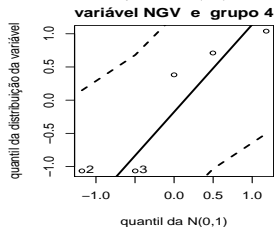
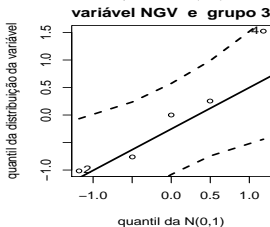
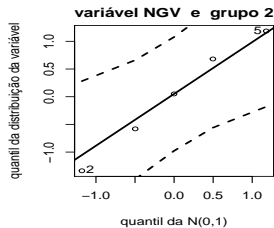
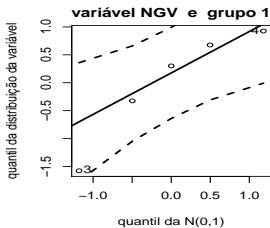
$$\tilde{\Psi}_1 = \begin{bmatrix} P & 2558.80 & -7.2260 \\ NVG & -0.8942 & 0.0255 \end{bmatrix} \quad \tilde{\Psi}_2 = \begin{bmatrix} P & 1525.800 & 3.5460 \\ NVG & 0.5727 & 0.0251 \end{bmatrix}$$

$$\tilde{\Psi}_3 = \begin{bmatrix} P & 1271.20 & 2.6500 \\ NVG & 0.9439 & 0.0062 \end{bmatrix} \quad \tilde{\Psi}_4 = \begin{bmatrix} P & 1908.800 & 3.5060 \\ NVG & 0.5277 & 0.0231 \end{bmatrix}$$

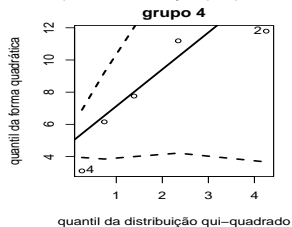
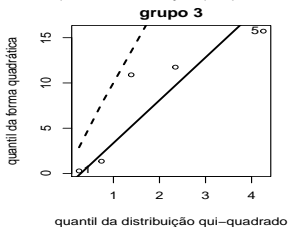
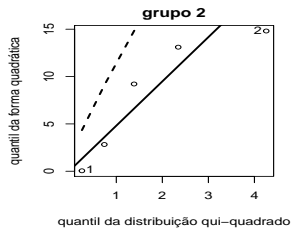
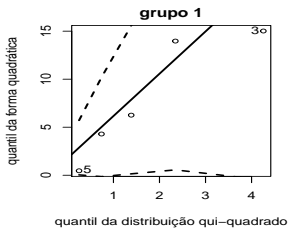
Gráficos de quantis-quantis da variável P



Gráficos de quantis-quantis da variável NGV



Gráficos de quantis-quantis das formas quadráticas



- Estamos interessados em testar inicialmente

$H_0 : \Sigma_1 = \Sigma_2 = \Sigma_3 = \Sigma_4$ vs. H_1 : alguma igualdade não vale.

- Utilizando o teste de Box, obtemos uma estatística de teste 16.09509 e valor-p 0.0649, o que nos leva a não rejeitar a hipótese nula de igualdade das matrizes de covariâncias ao nível de 5% de significância.

- Dessa forma, vamos aplicar a MANOVA para avaliar as médias dos quatro grupos. Nesse caso testamos as hipóteses

$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$ vs. H_1 : alguma igualdade não vale.

Estatísticas

Estatística	Valor	Aproxim. pela dist. F.	p-valor
Wilks	0.0442	183.43	< 0.001
Pillai	0.9557	183.43	< 0.001
Hotelling-Lawley	21.58	183.43	< 0.001
Roy	21.58	183.43	< 0.001

Com esses testes, rejeitamos a hipótese nula de igualdade dos vetores de médias μ_1 , μ_2 , μ_3 e μ_4 ao nível de 5% de significância.

Comparações múltiplas

- Comparações referentes à variável P:

$$H_0 : \mu_{i1} = \mu_{k1} \quad \text{vs} \quad H_1 : \mu_{i1} \neq \mu_{k1} \quad i \neq k, i, k = 1, 2, 3, 4$$

- Comparações referentes à variável NGV:

$$H_0 : \mu_{i2} = \mu_{k2} \quad \text{vs} \quad H_1 : \mu_{i2} \neq \mu_{k2} \quad i \neq k, i, k = 1, 2, 3, 4$$

- Note que o objeto manova usa a parametrização por casela de referência no R.

Resultados

- Estatística do teste e p-valor entre parênteses.
- Variável P:
 - grupo 1 x grupo 2: 7.18 (0.0074)
 - grupo 1 x grupo 3: 284.23 (0)
 - grupo 1 x grupo 4: 452.27 (0)
 - grupo 2 x grupo 3: 201.08 (0)
 - grupo 2 x grupo 4: 345.51 (0)
 - grupo 3 x grupo 4: 19.43 (0)

Resultados

- Variável NGV:
 - grupo 1 x grupo 2: 115.26 (0)
 - grupo 1 x grupo 3: 49.32 (0)
 - grupo 1 x grupo 4: 78.05 (0)
 - grupo 2 x grupo 3: 13.78 (< 0.001)
 - grupo 2 x grupo 4: 3.61 (0.0573)
 - grupo 3 x grupo 4: 3.28 (0.07)
- Notamos que, com exceção das médias referentes a NGV nos grupos 2, 3 e 4, todas as demais médias são significativamente diferentes ao nível de 5% de significância.

Referências I

 FERREIRA, D. F. *Análise Multivariada*. Lavras, MG: Universidade Federal de Lavras, 1996.