

R APLICADO AOS DELINEAMENTOS EXPERIMENTAIS

Wander Matos de Aguiar¹, Elaine Borges Monteiro Cassiano²

¹Professor de Ensino Superior da Universidade Estadual de Mato Grosso do Sul - UEMS. Doutor em Ciências Ambientais e Sustentabilidade Agropecuária pela UCDB. wander.aguiar@outlook.com

²Professora de Ensino Superior do Instituto Federal de Mato Grosso do Sul - IFMS. Doutora em Ciências Ambientais e Sustentabilidade Agropecuária pela UCDB. elaine.cassiano@ifms.edu.br

RESUMO

Gratuito, o R é um sistema gratuito que conta com uma gama de elementos e pacotes capazes de realizar inúmeras análises estatísticas, dentre eles, a Análise de Variância (ANOVA), objetivo primordial desta trabalho, traduzindo-se, desta forma, em ferramenta indispensável à demonstração de similitudes e/ou divergências no confronto entre tratamentos dos mais diversos experimentos, razão pela qual a utilização de ferramentas e pacotes estatísticos, que normalmente são pagos, tem muito à contribuir com sua consecução, razão pela qual a possibilidade de sua execução em sistema estatístico gratuito tem o potencial de democratizar as ferramentas necessárias ao público em geral, sendo de grande valia para o ambiente acadêmico, principalmente para discentes. Diante disso, buscou-se a elaboração de uma série de rotinas (programação) voltadas à execução da ANOVA bem com da estatística descritiva básica e de análise gráfica em projetos de pesquisa em delineamentos inteiramente casualizados, prezando não só a automatização do processo, com a menor interferência possível nas linhas de comandos, como também, e sempre que possível, sua apresentação de forma didática.

PALAVRAS-CHAVE: ANOVA; R – Project; DIC.

1 INTRODUÇÃO

O R é um sistema gratuito de computação estatística e de gráficos, que consiste basicamente de uma linguagem de computação, capaz de gerar cálculos estatísticos, e também uma série de gráficos atrelados ou não à estas análises, com a capacidade de executar programas armazenados em arquivos de script (seu formato padrão).

Criado “originalmente por Ross Ihaka e por Robert Gentleman (por isso R), no departamento de Estatística da Universidade de Auckland, Nova Zelândia a partir da linguagem S” (SILVA, 2015), ele compila e roda em uma ampla variedade de plataformas, como por exemplo, o UNIX, o Mac OS e o Windows”, e que pode como a própria nomenclatura sugere, ser livremente redistribuído ou mesmo modificado (*The R Project for Statistical Computing*) e, talvez por este motivo, a inexistência de custo de aquisição, bem como na facilidade de sua obtenção, se apresente hoje como uma grande ferramenta de auxílio àqueles que demandam por análises estatísticas e gráficos das mais variadas.

R foi inicialmente escrito por Ross Ihaka e Robert Gentleman no Departamento de Estatística da Universidade de Auckland em Auckland, Nova Zelândia. Além disso, um grande grupo de indivíduos tem contribuído para R enviando relatórios de código e do erro.

Desde meados de 1997, houve um grupo central (o "R Core Team"), que pode modificar o arquivo R código fonte. O grupo é actualmente constituído por Doug Bates, John Chambers, Peter Dalgaard, Seth Falcon, Robert Gentleman, Kurt Hornik, Stefano Iacus, Ross Ihaka, Friedrich Leisch, Uwe Ligges, Thomas Lumley, Martin Maechler, Duncan Murdoch, Paul Murrell, Martyn Plummer, Brian Ripley, Deepayan Sarkar, Duncan Temple Lang, Luke Tierney, e Simon Urbanek. (The R Project for Statistical)

Partiremos do pressuposto, diante da relativa facilidade neste processo, bem como na diversidade de apostilas e manuais disponíveis na internet sobre o tema, inclusive no próprio *site* do sistema (<http://www.r-project.org/>), que o leitor possua algum

conhecimento, ainda que introdutório, na linguagem R e que esteja em dia com os conceitos dos delineamentos experimentais.

Por fim, duas observações são necessárias e vem ao encontro de algumas dificuldades encontradas no decorrer do desenvolvimento das rotinas objeto deste livro.

A primeira delas pelo fato de que todos os procedimentos aqui descritos são aplicáveis ao ambiente Windows, mais especificamente em sua versão 10.

A segunda, ante ao fato de que para a utilização das rotinas que a serem descritas, serão necessárias a instalação de alguns pacotes adicionais àqueles já componentes da biblioteca do R, à exemplo dos pacotes “xlsx”, “agricolae” e “multcomp”, cujos procedimentos de instalação encontram-se descritas na programação (Material e Métodos), mas que, desde início observamos, no ambiente acima especificado, a necessidade de que o Windows fosse iniciado em seu modo “administrador” bem como da existência de uma versão atualizada do Java que, especificamente para o desenvolvimento das rotinas, onde foram utilizados computadores com processador de 64 bits, se fez necessária a instalação de um arquivo específico para esta versão, mais especificamente pela instalação do arquivo “Windows Off-line (64 bits)”¹, disponível no site do Java.

Segundo Kabacoff (2015, p. 5), a maioria das plataformas de software de estatística comerciais podem custar até milhares, ou até dezenas de milhares, de dólares, o principal atrativo do R reside justamente na sua gratuidade, e alerta “Se você é um professor ou um estudante, os benefícios são óbvios”.

O presente artigo, por sua vez, tem por objetivo a apresentação de uma programação para uma série de rotinas capazes de realizar, com a utilização do R, a análise de variância (ANOVA), teste de médias e análise estatística descritiva básica e gráfica para experimentos em Delineamentos Inteiramente Casualizados.

Como a finalidade precípua das rotinas elaboradas se prendem à elaboração da Análise de Variância de Experimentos (ANOVA), optou-se pela sua realização logo após a importação dos dados, e que será acompanhada pela realização de três testes de médias, respectivamente TUKEY, DUNCAN e STUDENT-NEWMAN-KEULS.

2 DESENVOLVENDO AS ROTINAS

Por aparentemente ser comum entre os usuários não possuidores de outros sistemas estatísticos, optou-se pelo lançamento dos dados a serem analisados via planilha do Excel, ante sua relativa facilidade de manuseio e ante a possibilidade de se realizar a importação dos dados diretamente daquela plataforma evitando assim ter que se realizar o re-lançamento de todos os dados no R, processo relativamente “entediante”, o que será feito através da utilização do pacote “xlsx” (DRAGULESCU, 2014), que poderá ser baixado diretamente do ambiente R (RGui 64-bit), um alternativa ideal para este fim, capaz de proporcionar a importação daquelas informações diretamente do Excel para o R.

Sendo a forma mais simples de delineamento experimental, o Delineamento Inteiramente Casualizado “é indicado quando as condições experimentais são homogêneas” (RIBEIRO JUNIOR, 2004, p. 178), onde se buscam as seguintes informações:

¹ Disponível em http://www.java.com/pt_BR/download/manual.jsp

Tabela 1: ANOVA em um experimento em Delineamento Inteiramente Casualizado (DIC)

FV	GL	SQ	QM	F
Tratamentos	t-1	SQTrat	SQTrat/(t-1)	
Resíduo	(tr -1) - t1	SQRes	SQRes/(tr -1) - t1	QMTrat / QMRes
Total	tr -1			

FV = Fontes de Variação; GL = Graus de liberdade; SQ = Soma dos quadrados; QM = Quadrado médio; SQTrat = Soma dos Quadrados do Tratamento; SQRes = Soma dos Quadrados do Resíduo; t=nº de tratamentos; tr=nº total de repetições.

Para Sonia Vieira (1999, p.67-68), “A comparação de médias só pode ser feita após a análise de variância. Isto porque todos os procedimentos para obter a d.m.s. exigem o cálculo do quadrado médio do resíduo” ao mesmo tempo em que ressalta que a própria análise de variância tem, como consequência, o conhecimento do valor de F, permitindo ao pesquisador verificar, estatisticamente, se as médias são ou não iguais a determinado nível de significância. Neste contexto, foram desenvolvidas também linhas de comando visando esta finalidade.

Muito embora o R possua função padrão instalada para o teste de média Tukey, optou-se pela busca de novos pacotes capazes de promoverem uma apresentação mais “intuitiva” dos resultados advindos das análises buscadas. Diante disso, foram localizados os pacotes “agricolae” (Mendiburu, 2014), capaz de apresentar o resultado do teste de médias, pelos métodos Tukey, Dunkan e SNK, e o pacote “multcomp”, que foram aliadas à visualização gráfica destes testes.

3 MATERIAL E MÉTODOS

3.1 PREPARAÇÃO PRECEDENTE À PROGRAMAÇÃO

Os dados utilizados, a fim de se demonstrar o funcionamento das rotinas aqui apresentadas, foram extraídos do livro “Estatística Experimental” de autoria de Sonia Vieira, que ora transcrevemos:

Para comparar a produtividade de quatro variedades de milho, um agrônomo tomou vinte parcelas similares e plantou a variedade A em cinco parcelas, a variedade B em outras cinco parcelas e assim por diante, até completar as vinte parcelas. É claro que o agrônomo sorteou uma variedade para cada parcela. (...) A produção de cada parcela foi anotada ao final do experimento e está representada na tabela a seguir, que também apresenta as médias de produção de cada variedade.(VIEIRA, 1999, p.45)

Com a utilização do Excel, os dados exemplificativamente apresentados pela autora, poderiam ser organizados da maneira como apresentada na Figura 1 abaixo, da mesma forma que alguns cálculos da estatística descritiva inicial, tais como médias, desvio padrão e coeficiente de variação, são facilmente executados, e que aqui serão apresentados por questões meramente elucidativas.

Produção de milho em Kg/100 m2 segundo a variedade

Repetições	Variedade 1	Variedade 2	Variedade 3	Variedade 4
1	25	31	22	33
2	26	25	26	29
3	21	28	28	31
4	22	27	25	34
5	21	24	29	28
Soma	115,0	135,0	130,0	155,0
Média nos Tratamentos	23,0	27,0	26,0	31,0
Desv.Padr.Tratamento	2,3	2,7	2,7	2,5
Coef.Var.Tratamento	10,2%	10,1%	10,5%	8,2%
Média Geral Experimento				26,8
Desv.Padr.Experimento				3,8
Coef.Var.Experimento				14,1%

Não obstante, muito embora seja possível a importação destes dados da forma como foram propostos acima, e visando uma maior facilitação do processo e, em consequência, a elaboração de um número menor de linhas de programação para as rotinas, uma nova disposição destes dados contribuirá em muito para esta finalidade (lado direito da figura).

	A	B	C
	Repet	Tratamento	Vlrs
1			
2	1	Var_1	25
3	2	Var_1	26
4	3	Var_1	21
5	4	Var_1	22
6	5	Var_1	21
7	6	Var_2	31
8	7	Var_2	25
9	8	Var_2	28
10	9	Var_2	27
11	10	Var_2	24
12	11	Var_3	22
13	12	Var_3	26
14	13	Var_3	28
15	14	Var_3	25
16	15	Var_3	29
17	16	Var_4	33
18	17	Var_4	29
19	18	Var_4	31
20	19	Var_4	34
21	20	Var_4	28

Figura 1: Montagem do Experimento no Excel

Os dados obtidos e descritos pela autora foram lançados em uma planilha do Excel, nomeada “Modelo_DIC em R” na guia “Plan1” (padrão na versão 10 do Excel).

Desde que seguidos esta padronização, os dados podem ser lançados em uma infinidade de combinações (em caso de outras modalidades de delineamentos), seja no número de tratamentos, blocos, parcelas, etc., e que serão prontamente reconhecidas pelas rotinas em R.

3.2 LINHAS DE COMANDO OU PROGRAMAÇÃO

As linhas de comandos (ou programação) necessárias ao desenvolvimento de todas as análises estatísticas e gráficas, além do cálculo da ANOVA para Delineamentos Inteiramente Casualizados, que é o objetivo principal deste artigo, encontram-se no Quadro 1 e, sempre que possível, com algumas informações, normalmente precedidas do símbolo “#”, exemplificativas/explicativas dos comandos.

A função *aov* nos permite realizar análises de variância envolvendo dados balanceados. (...) Podemos utilizar também a função *lm*, acompanhada da *aov* ou da *anova*. A diferença básica entre a função *aov* e a *lm* é a forma tradicional dos resultados da análise

de variância que é apresentada pela *aov*, enquanto na *lm*, os resultados são apresentados em conformidade com aqueles da teoria dos modelos lineares. A função *aov* é na verdade um envelope da função *lm*, ou seja, a função em questão invoca a função *lm* em seus cálculos. (FERREIRA, 2013, p. 168)

O sistema R fornece um pacote básico instalado em seu sistema capaz de realizar a análise de variância (ANOVA).

Da mesma forma, para facilitar a compreensão, a programação foi dividida em Etapas.

Quadro 1: Programação do R

```

#Análise de Variância - ANOVA
#EM DIC - Delineamento Inteiramente ao Acaso
#####
#####

##Passo 1 -
#Instalar/Carregar pacotes necessários às análises e funções pré-estabelecidas
#Pacotes: xlsx, agricolae e multcomp

#Caso os pacotes abaixo ainda não estejam instalados, promover sua instalação
# através dos comandos:
install.packages("xlsx")
install.packages("agricolae")
install.packages("multcomp")

#Caso já tenham sido instalados, deve-se somente carregá-los no sistema:
require(xlsx)
require(agricolae)
require(multcomp)

#-----
#Função(ções) para Calcular Coeficiente de Variação
#-----
#Executar o comando abaixo para deixá-la ativa no sistema
cv<-function(x){coef<-sd(x)/mean(x)*100
return(coef)}

#-----
#Definindo cores, tipos de linhas e de pontos para Análises Gráficas
#-----
###Definindo as cores que serão utilizadas nos gráficos - vide tabela - help(points)
##Escolha uma das alternativas de Composição de Cores abaixo para os Gráficos:
#Alternativa de Vermelho à Amarelo
corDegrade <- colorRampPalette(c("red", "yellow"))

#Alternativa de Cinza à Branco
corDegrade <- colorRampPalette(c("darkgrey", "white"))

#Alternativa de Azul à Turquesa
corDegrade <- colorRampPalette(c("blue", "turquoise"))

#Alternativa de Azul à Ciano
corDegrade <- colorRampPalette(c("blue", "cyan"))

#Alternativa de Verde à Verde Limão
corDegrade <- colorRampPalette(c("green", "turquoise"))

###Definindo os tipos de points que serão utilizados nos gráficos -
#vide tabela - help(points)
CodsPoints<-c(1,2,6,0,8,17,19,21,23,25)

###Definindo os tipos de linhas que serão utilizados nos gráficos -
CodsLines<-c("solid", "dashed", "dotted", "dotdash", "longdash",
"solid", "dashed", "dotted", "dotdash", "longdash" )

#####
##Passo 2 -
#Importando a PLANILHA de TRATAMENTOS e os DADOS GERAIS - DIRETAMENTE do EXCEL
#e ESCOLHENDO a PASTA
#####

#Escolhendo o arquivo base para importação dos dados
NomeArq<-file.choose() # Atribuindo o local e o nome do
# à variável "NomeArq"
NomeArq # Mostra o nome do arquivo escolhido

#Importando os dados do arquivo escolhido
DadosImp = read.xlsx(NomeArq, sheetName = "Plan1", header=TRUE)
DadosImp #Exibindo os dados importados
names(DadosImp) #Atribuindo nomes aos dados importados

#####

```

```

##Passo 3 -
#PREPARANDO os dados e CALCULANDO a ANOVA em DIC
#####
#Transformando os valores relativos aos TRATAMENTOS em MATRIZ e FATORES
Tratamentos=as.factor(DadosImp[,2]) #Criando um fator para os nomes
Tratamentos #Mostra as etiquetas(tratamento)-cada
valor

NomeTrats<-levels(Tratamentos) #"Etiquetando" os nomes dos tratamentos
NomeTrats #Mostra o nome de cada grupo de
tratamentos

Valores = c(t(as.matrix(DadosImp[,3])))
Valores #Mostra os valores atribuídos à variável

Tratamentos;NomeTrats;Valores #Exibindo os dados criados

#-----
##Calculando e exibindo os valores da ANOVA
##COM demonstração das interações entre os tratamentos e blocos
ANOVA_DIC = aov(Valores ~ Tratamentos) #Calculo da ANOVA
summary(ANOVA_DIC) #Exibição dos valores da ANOVA
model.tables(ANOVA_DIC,type="means") #Visualizando as Médias dos Tratamentos
summary.lm(ANOVA_DIC) #Teste de Médias padrão do R

#####
##Passo 4 -
#Estatística DESCRITIVA dos EXPERIMENTOS
#####
###Demonstrando o RESUMO (Média, Mediana Mínimo e Máximo, etc. do conjunto de dados
summary(DadosImp) #Est.Descritiva de TODO o
EXPERIMENTO

##Estabelecendo as linhas da média GERAL e do DESVIO PADRÃO do Experimento TOTAL
MedExp=mean(Valores) #Média Geral do Experimento
lmin<-min(Valores) #MENOR valor das medições
lmax<-max(Valores) #MAIOR valor das medições
MedExp;lmin;lmax

DesvPad<-(sd(Valores)) #Desvio Padrão do EXPERIMENTO
LsSd<-MedExp+DesvPad #Limite Superior do Desvio Padrão
LiSd<-MedExp-DesvPad #Limite Inferior do Desvio Padrão
DesvPad;LsSd;LiSd

cv(Valores) #Coef. de Variação do Experimento

mT<-tapply(Valores,Tratamentos,mean) #Valor da Média em cada Tratamento
mT

qT=length(levels(as.factor(DadosImp[,2]))) #Determinando o número de Tratamentos

###Visualizando a Estatística Descritiva para cada TRATAMENTO
summary(Valores) #Média Geral - Todas as repetições
tapply(Valores,Tratamentos,sum) #Soma, por TRATAMENTO
tapply(Valores,Tratamentos,mean) #Média, por TRATAMENTO
tapply(Valores,Tratamentos,median) #Mediana, por TRATAMENTO
tapply(Valores,Tratamentos,sd) #Desvio Padrão, por TRATAMENTO
tapply(Valores,Tratamentos,var) #Variância, por TRATAMENTO
tapply(Valores,Tratamentos,range) #Alcance, por TRATAMENTO

#####
##Passo 5 -
#Testes de COMPARAÇÃO das MÉDIAS - TUKEY, DUNCAN e STUDENT-NEWMAN-KEULS
#####
Tukey_TRAT=HSD.test(ANOVA_DIC,'Tratamentos',alpha=0.05)
Tukey_TRAT

#Visualizando GRAFICAMENTE o teste de média dos TRATAMENTOS - TUKEY
plot(Tukey_TRAT,main="Comparação de Médias - TUKEY")

#Caso queira visualizar, no mesmo gráfico, a média geral do experimento e seu desvio Padrão, execute:
NPontos=(length(Valores)/qT)

```

```

lines(c(0:(NPontos-1)),c(rep(MedExp,NPontos)),lty=CodsLines[1],col="darkred",lwd=1)
lines(c(0:(NPontos-1)),c(rep(LsSd,NPontos)),lty=CodsLines[2],col="red",lwd=1)
lines(c(0:(NPontos-1)),c(rep(LiSd,NPontos)),lty=CodsLines[2],col="red",lwd=1)
legend('topright', c("Média Experimento", "Desvio Padrão"),
      lty=c(CodsLines[1],CodsLines[2]), col=c("darkred","red"), bty='n', cex=0.5)

#Teste DUNCAN
Duncan_TRAT=duncan.test(ANOVA_DIC,"Tratamentos",alpha=0.05)
Duncan_TRAT

#Visualizando GRAFICAMENTE o teste de média dos TRATAMENTOS - DUNCAN
plot(Duncan_TRAT,main="Comparação de Médias - DUNCAN")

#Caso queira visualizar, no mesmo gráfico, a média geral do experimento e seu desvio Padrão, execute:

lines(c(0:(NPontos-1)),c(rep(MedExp,NPontos)),lty=CodsLines[1],col="darkred",lwd=1)
lines(c(0:(NPontos-1)),c(rep(LsSd,NPontos)),lty=CodsLines[2],col="red",lwd=1)
lines(c(0:(NPontos-1)),c(rep(LiSd,NPontos)),lty=CodsLines[2],col="red",lwd=1)
legend('topright', c("Média Experimento", "Desvio Padrão"),
      lty=c(CodsLines[1],CodsLines[2]), col=c("darkred","red"), bty='n', cex=0.5)

#-----
#Teste STUDENT-NEWMAN-KEULS
SNK_TRAT=SNK.test(ANOVA_DIC,"Tratamentos",alpha=0.05)
SNK_TRAT

#Visualizando GRAFICAMENTE o teste de média dos TRATAMENTOS - SNK
plot(SNK_TRAT,main="Comparação de Médias - SNK")

#Caso queira visualizar, no mesmo gráfico, a média geral do experimento e seu desvio Padrão, execute:

lines(c(0:(NPontos-1)),c(rep(MedExp,NPontos)),lty=CodsLines[1],col="darkred",lwd=1)
lines(c(0:(NPontos-1)),c(rep(LsSd,NPontos)),lty=CodsLines[2],col="red",lwd=1)
lines(c(0:(NPontos-1)),c(rep(LiSd,NPontos)),lty=CodsLines[2],col="red",lwd=1)
legend('topright', c("Média Experimento", "Desvio Padrão"),
      lty=c(CodsLines[1],CodsLines[2]), col=c("darkred","red"), bty='n', cex=0.5)

#-----
#Caso queira visualizar, no mesmo gráfico, todos os gráficos dos testes de médias
#-----
par(mfrow=c(2,2))
plot(Tukey_TRAT,main="Comparação de Médias - TUKEY")
#Caso queira visualizar, no mesmo gráfico, a média geral do experimento e seu desvio Padrão, execute:
lines(c(0:(NPontos-1)),c(rep(MedExp,NPontos)),lty=CodsLines[1],col="darkred",lwd=1)
lines(c(0:(NPontos-1)),c(rep(LsSd,NPontos)),lty=CodsLines[2],col="red",lwd=1)
lines(c(0:(NPontos-1)),c(rep(LiSd,NPontos)),lty=CodsLines[2],col="red",lwd=1)
legend('topright', c("Média Experimento", "Desvio Padrão"),
      lty=c(CodsLines[1],CodsLines[2]), col=c("darkred","red"), bty='n', cex=0.5)

plot(Duncan_TRAT,main="Comparação de Médias - DUNCAN")
#Caso queira visualizar, no mesmo gráfico, a média geral do experimento e seu desvio Padrão, execute:
lines(c(0:(NPontos-1)),c(rep(MedExp,NPontos)),lty=CodsLines[1],col="darkred",lwd=1)
lines(c(0:(NPontos-1)),c(rep(LsSd,NPontos)),lty=CodsLines[2],col="red",lwd=1)
lines(c(0:(NPontos-1)),c(rep(LiSd,NPontos)),lty=CodsLines[2],col="red",lwd=1)
legend('topright', c("Média Experimento", "Desvio Padrão"),
      lty=c(CodsLines[1],CodsLines[2]), col=c("darkred","red"), bty='n', cex=0.5)

plot(SNK_TRAT,main="Comparação de Médias - SNK")

#Caso queira visualizar, no mesmo gráfico, a média geral do experimento e seu desvio Padrão, execute:

lines(c(0:(NPontos-1)),c(rep(MedExp,NPontos)),lty=CodsLines[1],col="darkred",lwd=1)
lines(c(0:(NPontos-1)),c(rep(LsSd,NPontos)),lty=CodsLines[2],col="red",lwd=1)
lines(c(0:(NPontos-1)),c(rep(LiSd,NPontos)),lty=CodsLines[2],col="red",lwd=1)
legend('topright', c("Média Experimento", "Desvio Padrão"),
      lty=c(CodsLines[1],CodsLines[2]), col=c("darkred","red"), bty='n', cex=0.5)

#####
##Passo 6 -
##Análise GRÁFICA

```

```
#####  
#-----  
#Criando um BARPLOT para os TRATAMENTOS  
#-----  
###Criando o BARPLOT  
  
Trats <- data.frame(tapply(Valores,Tratamentos,list))  
NomeTrats=colnames(Trats)  
  
barplot(mT,ylim=c(lmin,lmax),col=corDegrade(qT),main="Tratamentos")  
  
###Delimitando a Média Geral e o Desvio Padrão do Experimento  
lines(c(0:(NPontos-1)),c(rep(MedExp,NPontos)),lty=CodsLines[1],col="black",lwd=1)  
lines(c(0:(NPontos-1)),c(rep(LsSd,NPontos)),lty=CodsLines[2],col="darkgrey",lwd=2)  
lines(c(0:(NPontos-1)),c(rep(LiSd,NPontos)),lty=CodsLines[2],col="darkgrey",lwd=2)  
legend('topright', c("Média Experimento", "Desvio Padrão"),  
lty=c(CodsLines[1],CodsLines[2]), col=c("black", "darkgrey"), bty='n', cex=0.5)  
  
#-----  
#Criando um BOXPLOT para os TRATAMENTOS  
#-----  
###BOXPLOT com Título e na HORIZONTAL  
boxplot(tapply(Valores,Tratamentos,list),horizontal=FALSE,main="Boxplot dos Tratamentos",  
xlab="Tratamentos",ylab="Desempenho",col=corDegrade(qT))  
  
###Delimitando a Média Geral e o Desvio Padrão do Experimento  
lines(c(0:(NPontos-1)),c(rep(MedExp,NPontos)),lty=CodsLines[1],col="black",lwd=1)  
lines(c(0:(NPontos-1)),c(rep(LsSd,NPontos)),lty=CodsLines[2],col="darkgrey",lwd=2)  
lines(c(0:(NPontos-1)),c(rep(LiSd,NPontos)),lty=CodsLines[2],col="darkgrey",lwd=2)  
legend('topleft', c("Média Geral do Experimento", "Desvio Padrão"),  
lty=c(CodsLines[1],CodsLines[2]), col=c("black", "darkgrey"), bty='n', cex=0.5)  
  
#-----  
#PLOT CONFRONTANDO TODOS os TRATAMENTOS ao REDOR da MÉDIA GERAL do EXPERIMENTO  
#Finalidade - Visualizar as médias graficamente  
#-----  
  
Trats <- data.frame(tapply(Valores,Tratamentos,list))  
  
##Gráfico com titulos  
plot(c(1,(length(Valores)/qT)),c(min(Valores),max(Valores)),main="Todos os tratamentos através da média geral",  
xlab="Repetições", ylab="Valores", type="n")  
  
###Delimitando a Média Geral e o Desvio Padrão do Experimento  
lines(c(1:NPontos),c(rep(MedExp,NPontos)),col="black",lwd=1)  
lines(c(1:NPontos),c(rep(LsSd,NPontos)),lty=CodsLines[2],col="darkgrey",lwd=2)  
lines(c(1:NPontos),c(rep(LiSd,NPontos)),lty=CodsLines[2],col="darkgrey",lwd=2)  
  
###Inserindo LEGENDAS no gráfico  
legend('topright', NomeTrats, pch=CodsPoints[1:length(NomeTrats)],  
col=corDegrade(length(NomeTrats)),bty='n', cex=.5)  
legend('bottomright', c("Média Geral do Experimento", "Desvio Padrão"),  
lty=c(CodsLines[1],CodsLines[2]), col=c("black", "darkgrey"), bty='n', cex=0.5)  
  
#Por PONTOS  
###Executar somente até o número total de Tratamentos do Experimento  
points(c(1:length(Trats[,1])),Trats[,1],pch=CodsPoints[1],lwd=1,col=corDegrade(qT)[1])  
points(c(1:length(Trats[,2])),Trats[,2],pch=CodsPoints[2],lwd=2,col=corDegrade(qT)[2])  
points(c(1:length(Trats[,3])),Trats[,3],pch=CodsPoints[3],lwd=1,col=corDegrade(qT)[3])  
points(c(1:length(Trats[,4])),Trats[,4],pch=CodsPoints[4],lwd=2,col=corDegrade(qT)[4])  
points(c(1:length(Trats[,5])),Trats[,5],pch=CodsPoints[5],lwd=1,col=corDegrade(qT)[5])  
points(c(1:length(Trats[,6])),Trats[,6],pch=CodsPoints[6],lwd=2,col=corDegrade(qT)[6])  
points(c(1:length(Trats[,7])),Trats[,7],pch=CodsPoints[7],lwd=1,col=corDegrade(qT)[7])  
points(c(1:length(Trats[,8])),Trats[,8],pch=CodsPoints[8],lwd=2,col=corDegrade(qT)[8])  
points(c(1:length(Trats[,9])),Trats[,9],pch=CodsPoints[9],lwd=1,col=corDegrade(qT)[9])  
points(c(1:length(Trats[,10])),Trats[,10],pch=CodsPoints[10],lwd=2,col=corDegrade(qT)[10])  
  
#PLOT (Dispersão) por LINHAS  
##Gráfico com titulos  
Trats <- data.frame(tapply(Valores,Tratamentos,list))
```

```
plot(c(1,(length(Valores)/qT)),c(min(Valores),max(Valores)),main="Todos os tratamentos através da média geral",
      xlab="Repetições", ylab="Valores", type="n")

###Delimitando a Média Geral e o Desvio Padrão do Experimento
lines(c(1:NPontos),c(rep(MedExp,NPontos)),col="black",lwd=1)
lines(c(1:NPontos),c(rep(LsSd,NPontos)),lty=CodsLines[2],col="darkgrey",lwd=2)
lines(c(1:NPontos),c(rep(LiSd,NPontos)),lty=CodsLines[2],col="darkgrey",lwd=2)

###Inserindo LEGENDAS no gráfico
legend('topright', NomeTrats, pch=CodsPoints[1:length(NomeTrats)],
      col=corDegrade(length(NomeTrats)),bty='n', cex=.5)
legend('bottomright', c("Média Geral do Experimento","Desvio Padrão"),
      lty=c(CodsLines[1],CodsLines[2]), col=c("black","darkgrey"), bty='n', cex=0.5)

lines(c(1:length(Trats[,1])),Trats[,1],lty=CodsLines[1],lwd=1,col=corDegrade(qT)[1])
lines(c(1:length(Trats[,2])),Trats[,2],lty=CodsLines[2],lwd=2,col=corDegrade(qT)[2])
lines(c(1:length(Trats[,3])),Trats[,3],lty=CodsLines[3],lwd=3,col=corDegrade(qT)[3])
lines(c(1:length(Trats[,4])),Trats[,4],lty=CodsLines[4],lwd=4,col=corDegrade(qT)[4])
lines(c(1:length(Trats[,5])),Trats[,5],lty=CodsLines[5],lwd=5,col=corDegrade(qT)[5])
lines(c(1:length(Trats[,6])),Trats[,6],lty=CodsLines[6],lwd=6,col=corDegrade(qT)[6])
lines(c(1:length(Trats[,7])),Trats[,7],lty=CodsLines[7],lwd=7,col=corDegrade(qT)[7])
lines(c(1:length(Trats[,8])),Trats[,8],lty=CodsLines[8],lwd=8,col=corDegrade(qT)[8])
lines(c(1:length(Trats[,9])),Trats[,9],lty=CodsLines[9],lwd=9,col=corDegrade(qT)[9])
lines(c(1:length(Trats[,10])),Trats[,10],lty=CodsLines[10],lwd=10,col=corDegrade(qT)[10])
```

Na programação, foram utilizadas linhas de comando que, dentre outras, tiveram a função de:

- Instalação/Carregamento de “pacotes”² visando a realização das análises;
- Definição de algumas “configurações” que utilizadas na execução do R;
- Comandos destinados à realização e visualização da estatística descritiva básica, não só do experimento como um todo, mas também em cada tratamento, que proporcionaram informações relativas às médias, valores mínimos e máximos, medianas e quartis, bem como soma, desvio padrão e variância, também por tratamentos, e ainda o coeficiente de variação total do experimento sob análise;
- Por fim, alguma programação fora efetivada visando a análise (visualização) gráfica (BARPLOT, BOXPLOT, GRÁFICO DE DISPERSÃO, entre outros) dos resultados apresentados, cujo escopo versou em se contribuir com interpretação dos resultados obtidos;
- A fim de se facilitar esta visualização, foram atribuídos valores-padrão às cores e tipos de “pontos” a serem utilizados nos gráficos ora apresentados, e que encontram-se descritas nas primeiras linhas de comandos, com a atribuição das variáveis “NomesCores” e “CodsPoints”. Assim, o gráfico de barras (das médias nos tratamentos e nos blocos) pode ser visto na figura 20;
- Outro expediente bastante interessante utilizado consistiu na elaboração de um Boxplot, “ou diagrama de caixa, gráfico que capta importantes aspectos de um conjunto de dados através do seu resumo dos cinco números, formado pelos seguintes valores: valor mínimo, primeiro quartil, segundo quartil, terceiro quartil e valor máximo.” (UFF, 2015).

Segundo Nieuwenhuis (2007, p. 53, tradução nossa), a produção de gráficos pode ser uma forma de se familiarizar com os seus dados ou para reforçar os resultados apresentados o que, felizmente, pode ser feito facilmente pelo R, que já vem com

² Pacotes nada mais são do que convenções utilizadas para se organizar e/ou padronizar a distribuição de funções extras do R.

algumas funções gráficas padrão em sua versão de instalação, que transitam desde um simples gráfico de barras, até um gráfico de dispersão. Contudo, diversos outros pacotes gráficos encontram-se à disposição para uma gama de aplicações e resultados dos mais variados.

3 RESULTADOS

Após o cálculo da ANOVA, para o exemplo utilizado, torna-se possível a verificação, para o exemplo apresentado, da ocorrência de efeitos ($p < 0,05$) nos tratamentos.

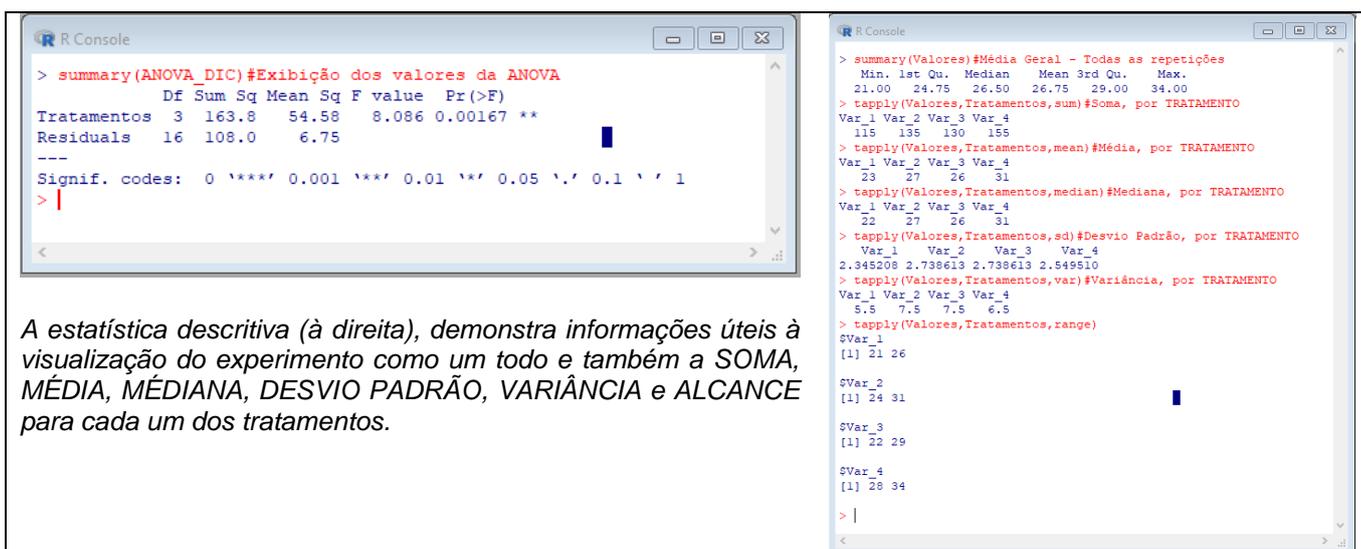


Figura 2: Cálculo da ANOVA e Visualização da Estatística descritiva

A título meramente exemplificativo, as figuras 3 e 4 demonstram os resultados desta programação no que se refere ao teste Tukey. Os demais testes (Duncan e SNK) seguem o mesmo padrão gráfico de resposta.

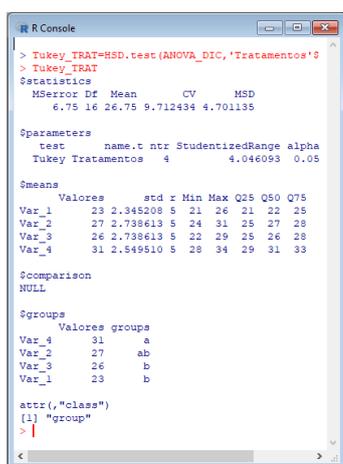


Figura 3: Teste Tukey de Médias

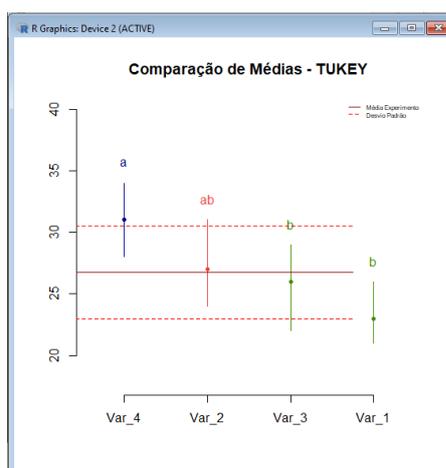


Figura 4: Visualização gráfica do Teste Tukey de Médias

Note-se que os resultados trazidos pelos testes advindos dos pacotes instalados, são bem mais intuitivos (através da atribuição de letras) e facilitam, ainda mais quando aliado ao elemento gráfico (Figura 4), em muito sua interpretação.

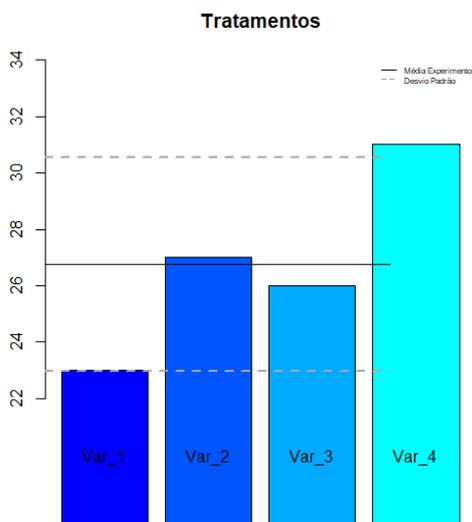


Figura 5: BARPLOT dos Tratamentos

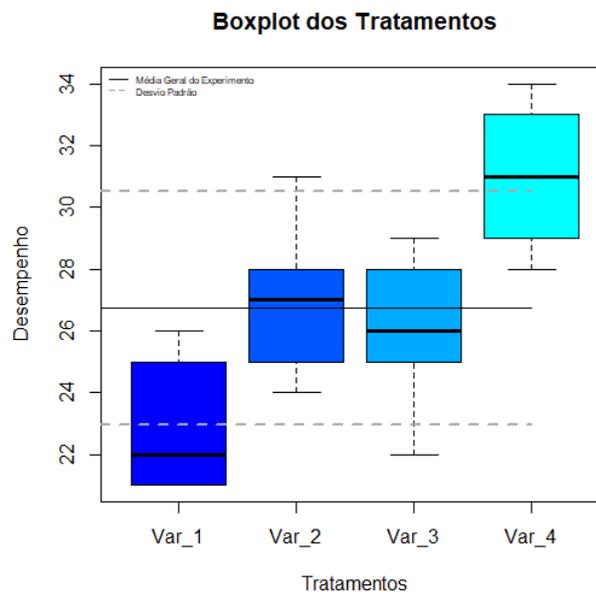


Figura 6: BOXPLOT dos Tratamentos

O BARPLOT (utilizado normalmente para variáveis qualitativas e quantitativas discretas, indica a contagem com que cada valor ocorre para a variável analisada) e o BOXPLOT que, segundo Tukey (1977) é um gráfico para visualizar uma única variável que consiste em apresentar visualmente 5 medidas resumo.

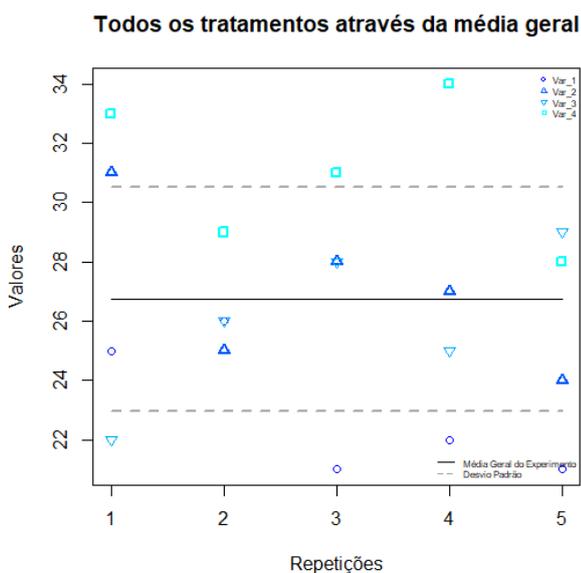


Figura 7: Gráfico de Dispersão em Pontos

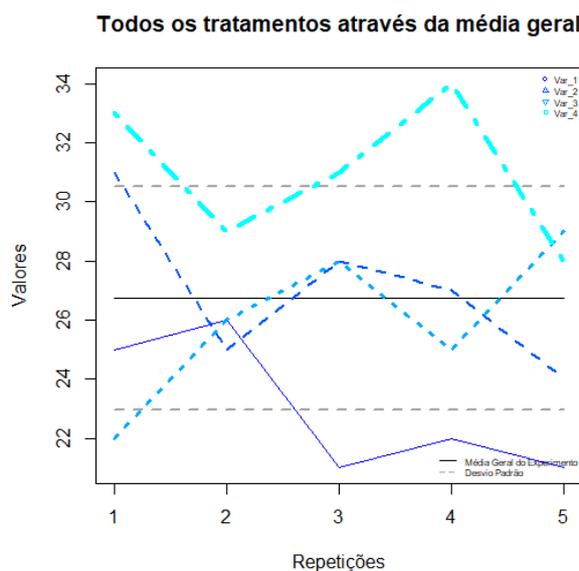


Figura 8: Gráfico de Dispersão em Linhas

Por fim, a visualização, através de gráficos de dispersão que, igualmente, nos permitem evidenciar a posição, dentre os limites estabelecidos para o experimento, de cada uma das unidades experimentais. Por sua vez, o gráfico de Dispersão em Linhas,

tem aplicabilidade em momentos que sejam necessários a visualização, por exemplo, da evolução das medidas obtidas.

4 CONSIDERAÇÕES FINAIS

Como inicialmente salientado, a Análise de Variância exerce papel fundamental nos experimentos.

A existência de um sistema estatístico gratuito tem muito a contribuir com aqueles que apresentam em sua vivência profissional a necessidade de se realizar experimentos e análises estatísticas, dentre as quais a ANOVA, com seus testes de médias e análises gráficas, capazes de abreviar a elaboração de seus resultados, razão pela qual optamos pela geração presente artigo, cujo intuito principal foi o de lançar mão da elaboração de linhas de comando da maneira mais simples possível, visando a possibilidade de se realizarem as referidas análises e testes com a mínima interferência do usuário, ao menos nas linhas de programação, buscando com isso a facilitação de todo o processo.

Diante disso, após a apresentação do R e seus benefícios, buscamos abordar a programação necessária a esse fim e, sempre que possível, dos resultados desta programação, com análises que vão desde a estatística descritiva básica até a ANOVA com seus testes de médias (métodos TUKEY, DUNCAN e STUDENT-NEWMAN-KEULS), considerados por alguns como os principais nesta categoria de análise de experimentos e, especificamente para o presente caso, relativos a Experimentos em Delineamentos Inteiramente Casualizados.

A programação aqui referida teve por objetivo, única e exclusivamente, de se demonstrar uma metodologia, dentre as várias possíveis, para que referidas análises fossem possíveis, utilizando-se de funções já prontas, testadas e facilmente adaptadas ao sistema R, capazes de dar maior celeridade e confiabilidade nos resultados dela advindos e que podem ser, com relativa facilidade, remanejadas para outras hipóteses e necessidades de experimentos e/ou delineamentos (blocos casualizados, quadrado latino, esquemas fatoriais, etc.)

REFERÊNCIAS

DRAGULESCU, Adrian A. **xlsx**: read, write, format Excel 2007 and Excel 97/2000/XP/2003 files. R package version 0.5.7. Url = <http://CRAN.R-project.org/package=xlsx>.

KABACOFF, Robert I. **R in action**: Data analysis and graphics with R. 2nd ed. U.S.A.: Manning Publications, 2015.

HOTHORN, Torsten; BRETZ, Frank e WESTFALL, Peter. **Simultaneous Inference in General**

MENDIBURU, Felipe de. **agricolae**: Statistical Procedures for Agricultural Research. R package version 1.2-1. Url = <http://CRAN.R-project.org/package=agricolae>.
Microsoft Store. Disponível em http://www.microsoftstore.com/store/msbr/pt_BR/cat/Office/categoryID.66256800?tduid=b3be4b824978894d77e918b2a9996823. Acesso em: 24 maio 2015.

NIEUWENHUIS, Rense. **Applied R for the quantitative social scientist**. Breda: 2007.
Parametric Models. Biometrical Journal 50(3), 346-363, 2008.

R Core Team. **A language and environment for statistical computing**. Disponível em http://cran.r-project.org/doc/FAQ/R-FAQ.html#What-is-R_003f. Acesso em: 24 maio 2015.

RIBEIRO JÚNIOR, José Ivo. **Análises estatísticas no Excel**. Guia prático. Viçosa-MG: 2004, p. 178.

SILVA, Wilson Martins da. **A superferramenta estatística para realizar trabalhos acadêmicos de forma livre, rápida e gratuita**. Belém: UEPA, 2015.

TUKEY, John. 1977. **Exploratory Data Analysis**. Addison-Wesley.

VIEIRA, Sonia. **Estatística Experimental**. 2. ed. São Paulo: Atlas, 1999.