

Exercício 3

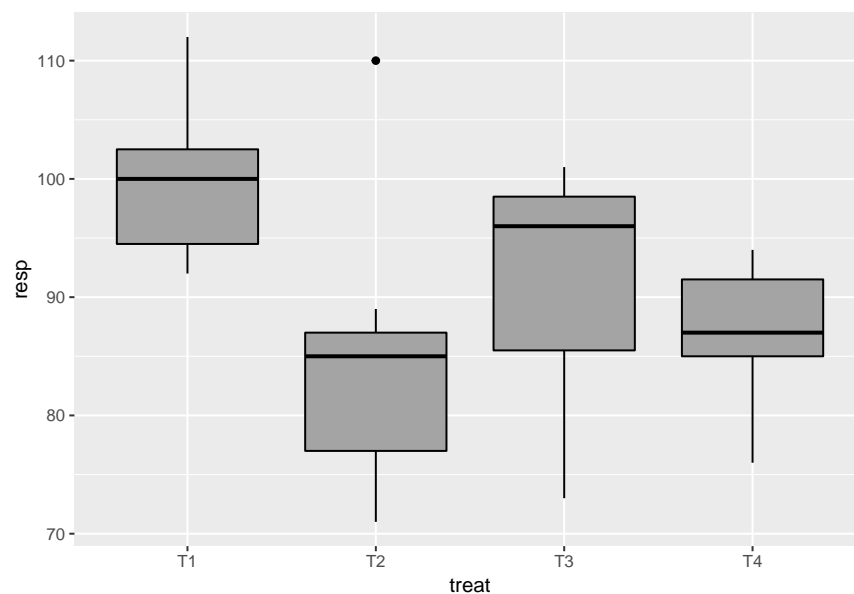
A hipertensão é um fator de risco para doenças cardiovasculares. Uma pesquisa foi conduzida para verificar o efeito de 4 tratamentos na normalização das medidas de pressão em pacientes hipertensos. Os dados estão apresentados a seguir

T1	T2	T3	T4
96	71	73	87
104	85	96	76
92	89	83	90
101	110	97	85
112	85	100	94
100	73	101	93
93	81	88	85

(a) Realize uma análise descritiva dos dados.

Resolução

```
##      resp      treat
##  Min.   : 71.00    T1:7
##  1st Qu.: 85.00    T2:7
##  Median : 91.00    T3:7
##  Mean   : 90.71    T4:7
##  3rd Qu.: 97.75
##  Max.   :112.00
```



Fazendo uma análise descritiva utilizando as estatísticas sumárias e um box-plot para cada tratamento podemos notar uma certa diferença na média entre os tratamentos de normalização das medidas de pressão, também podemos notar pelos intervalos interquartis de cada tratamento que as variâncias não são heterogêneas.

- (b) Proponha um modelo estatístico (estrutural e distribucional) apropriado para a análise destes dados. Represente o modelo proposto na forma matricial de um modelo linear. Interprete os parâmetros do modelo

Resolução

Um modelo estatístico estrutural de Delineamento Completamente Aleatorizado (DCA): 4 tratamentos, 7 réplicas (balanceado) com parametrização de casela de referência é da forma:

$$Y_{ij} = \begin{cases} \mu_1 + e_{ij}, & j = 1 \\ \mu_1 + \tau_j + e_{ij}, & j = 2, 3, 4 \end{cases}$$

Onde μ_1 é o valor esperado da resposta para unidades submetidas ao tratamento T1 (considerado como referência) e τ_j ($j = 2, 3, 4$) é o efeito do tratamento T_j em relação ao valor esperado da resposta ao tratamento T1 (é o desvio em relação a esta referência)

E $e_{ij} \sim N(0, \sigma^2)$ é o modelo distribucional para o DCA

E também pode ser representado na forma matricial:

$$Y_{7 \times 1} = X_{7 \times 4} \beta_{4 \times 1} + e_{7 \times 1}$$

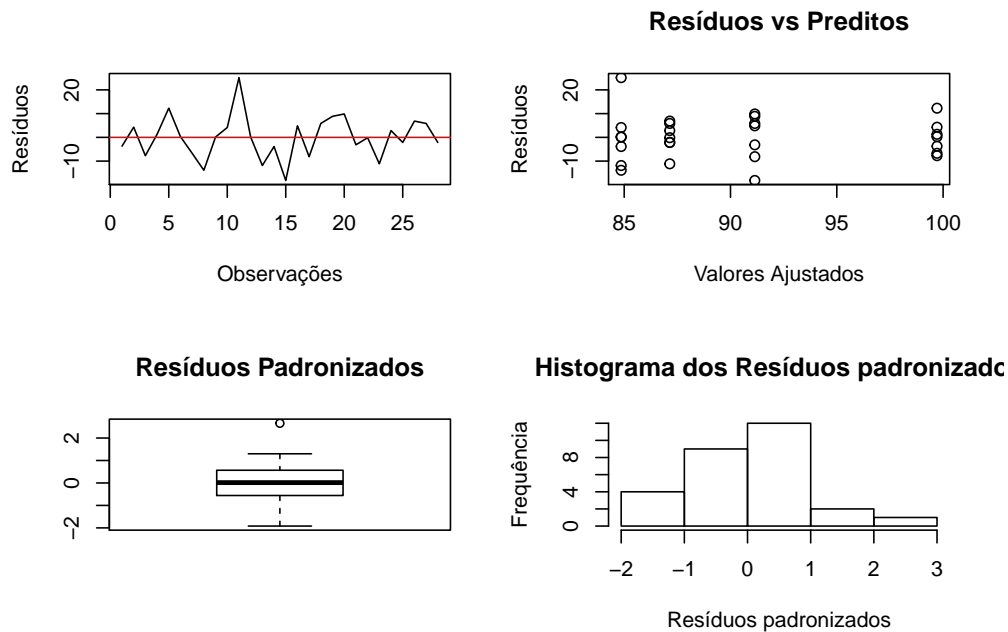
Onde $Y_{7 \times 1}$ é o vetor de observações, $X_{7 \times 4}$ é matriz de planejamento, $\beta_{4 \times 1}$ é o vetor de parâmetros do modelo e $e_{7 \times 1}$ é o vetor de resíduos.

- (c) Obtenha a tabela de ANOVA correspondente ao modelo proposto em b). Realize uma análise de diagnóstico. Há efeito significativo de “tratamento”?

Resolução

Table 1: Tabela de ANOVA

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
treat	3	897.7143	299.2381	3.343442	0.0359562
Residuals	24	2148.0000	89.5000	NA	NA



```
## Loading required package: carData

##
## Shapiro-Wilk normality test
##
## data:  res
## W = 0.9747, p-value = 0.7103

## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value Pr(>F)
## group 3  0.5875 0.6291
##      24

##
## Bartlett test of homogeneity of variances
##
## data:  resp by treat
## Bartlett's K-squared = 3.9568, df = 3, p-value = 0.2662
```

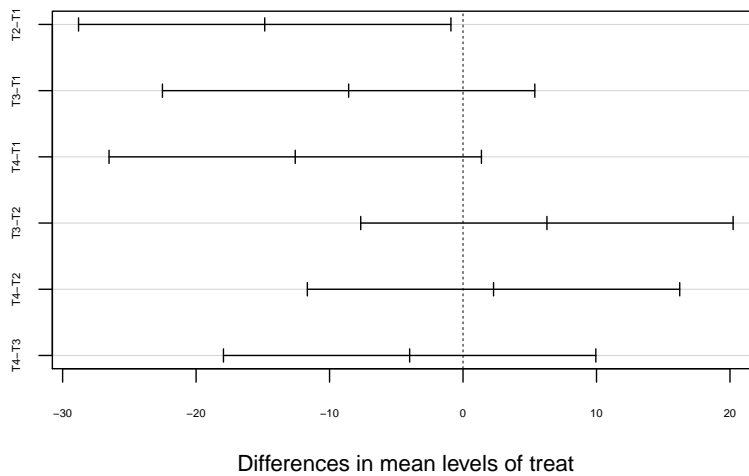
Observando o gráfico de resíduos x número de observação, podemos concluir que existe uma espécie de “zig-zag” o que como visto em aula nos garante a hipótese de independência, além disso observando o gráfico QQ-Plot e o teste shapiro obtemos um p-valor de 0.71, que a um nível de significância de 5% temos evidências estatísticas para aceitar que os dados seguem distribuição Normal, para avaliar a homocedasticidade podemos observar pelo gráfico de resíduos x valores ajustados que o mesmo está em um formato de “caixa”, além disso utilizando testes de levine e bartlett obtemos um p-valor de 0.794 e 0.266 respectivamente e a um nível de significância de 5% podemos aceitar a hipótese de homocedasticidade.

- (d) Obtenha intervalos de confiança para a diferença entre os pares de médias. Considere intervalos construídos sem correção para múltiplos testes, intervalos simultâneos com a correção de Bonferroni, Tukey e FDR. Compare os resultados e conclua sobre o possível efeito de tratamento.

Resolução

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = resp ~ treat, data = dados.df)
##
## $treat
##          diff          lwr          upr      p adj
## T2-T1 -14.857143 -28.806931 -0.9073548 0.0338694
## T3-T1 -8.571429 -22.521217  5.3783595 0.3480772
## T4-T1 -12.571429 -26.521217  1.3783595 0.0878504
## T3-T2  6.285714  -7.664074 20.2355024 0.6065419
## T4-T2  2.285714 -11.664074 16.2355024 0.9685950
## T4-T3 -4.000000 -17.949788  9.9497881 0.8578039
```

95% family-wise confidence level



Intervalos de confiança com

correção de bonferroni

```
##          results      adjustp
## treatT2 0.007186529 0.04311917
## treatT3 0.103008972 0.61805383
## treatT4 0.020274078 0.12164447
## treatT3 0.225868561 1.00000000
## treatT4 0.436685303 1.00000000
## treatT4 0.436685303 1.00000000
```

Intervalos de confiança FDR

```
##          results      adjustp
## treatT2 0.007186529 0.04311917
## treatT3 0.103008972 0.20601794
## treatT4 0.020274078 0.06082223
## treatT3 0.225868561 0.33880284
## treatT4 0.436685303 0.43668530
## treatT4 0.436685303 0.43668530
```

- (e) Considerando que o tratamento T1 corresponde a um Placebo, obtenha os intervalos de confiança com correção de Dunnett para estabelecer as comparações entre os demais tratamentos com T1. Comente os resultados.

Resolução

```
## Loading required package: tcltk

##
## 95% Dunnett confidence intervals
##
##              Diff      Lower      Upper Decision
## muT2-muT1 -14.865904 -27.545545 -2.186263 Reject H0
## muT3-muT1  -8.564392 -21.243108  4.114325   FTR H0
## muT4-muT1 -12.572897 -25.238401  0.092607   FTR H0
```

Logo podemos concluir que T1 e T2 tem diferença médias diferentes, enquanto o T1 e T3 e T1 e T4 tem médias iguais.

- (f) Proponha contrastes ortogonais entre os tratamentos. Qual é a SQ e o valor-p para cada contraste e para o conjunto desses contrastes.

Resolução

Propondo os contrastes ortogonais para testar as seguintes hipóteses, temos:

‘(T1,T4) vs (T2,T3)’ = (1, -1, -1, 1)

‘T1 vs T4’ = (1, 0, 0, -1)

‘T2 vs T3’ = (0, 1, -1, 0)

Assim o SQ e o valor-p para cada contraste e para o conjunto desses contrastes segue na tabela abaixo.

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## treat              3  897.7    299.2    3.343 0.0360 *
##   treat: (T1,T4) vs (T2,T3)  1   206.3    206.3    2.305 0.1420
##   treat: T1 vs T4           1   553.1    553.1    6.180 0.0203 *
##   treat: T2 vs T3           1   138.3    138.3    1.545 0.2259
## Residuals              24 2148.0     89.5
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Exercício 4

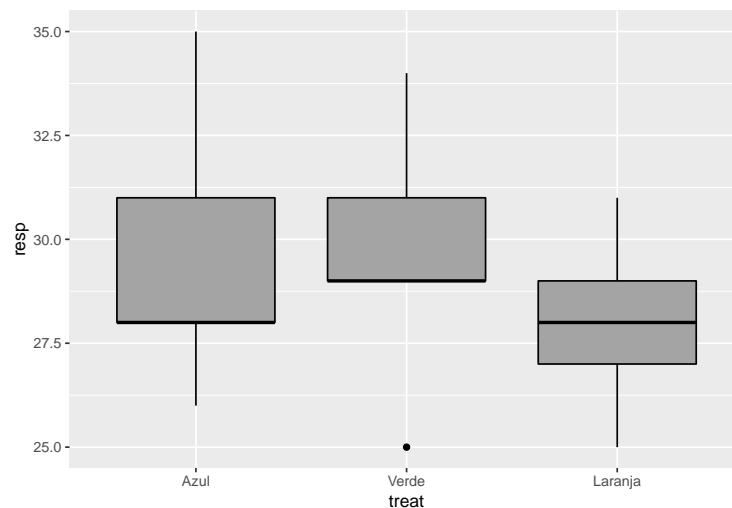
Uma pesquisa foi conduzida para investigar o efeito da cor do papel (no qual os questionários são impressos) na taxa de resposta. Para a pesquisa de opinião, 100 questionários foram distribuídos aleatoriamente em estacionamento de grandes supermercados. 15 supermercados foram escolhidos de uma área metropolitana e cada cor foi aleatoriamente atribuída a 5 deles. As taxas de respostas (em %) foram:

Supermercado / Cor	1	2	3	4	5
Azul	28	26	31	28	35
Verde	34	29	25	31	29
Laranja	53	25	27	29	28

(a) Realize uma análise descritiva dos dados. Comente os resultados.

Resolução

```
##      resp      treat
## Min.   :25.00 Azul   :5
## 1st Qu.:27.50 Verde  :5
## Median :29.00 Laranja:5
## Mean   :29.07
## 3rd Qu.:31.00
## Max.   :35.00
```



Fazendo uma análise descritiva, podemos notar que não existe muita diferença nas médias a partir da cor do panfleto, através de medidas resumo, as quais mostra aonde os dados estão comportados. Outro mecanismo para realizar a estatística descritiva foi o boxplot, este já diferenciando as cores do panfleto, que aparentemente retrata que as médias são bem próximas, o panfleto laranja possui uma menor variabilidade nos dados e o panfleto verde tem uma média maior que os demais.

(b) Assuma um modelo de ANOVA apropriado à análise dos dados. Que suposições foram adotadas? Obtenha os valores ajustados e os resíduos correspondentes ao ajuste do modelo.

Resolução

Obtendo os valores ajustados:

```
## [1] -0.36292738 -1.04341621  0.65780587 -0.36292738  2.01878353
## [6]  1.67853912 -0.02268296 -1.38366062  0.65780587 -0.02268296
## [11]  0.65780587 -1.38366062 -0.70317179 -0.02268296 -0.36292738
```

Os graus de liberdade do SQR, que são igual ao número de amostras menos o número de tratamentos ($n - K$):

```
## [1] 12
```

Os resíduos:

```
##          1          2          3          4          5
## -1.600000e+00 -3.600000e+00  1.400000e+00 -1.600000e+00  5.400000e+00
##          6          7          8          9         10
##  4.400000e+00 -6.000000e-01 -4.600000e+00  1.400000e+00 -6.000000e-01
##         11         12         13         14         15
##  3.000000e+00 -3.000000e+00 -1.000000e+00  1.000000e+00 -5.551115e-17
```

Os resíduos semi-studentizados:

```
##          1          2          3          4          5
## -5.227902e-01 -1.176278e+00  4.574414e-01 -5.227902e-01  1.764417e+00
##          6          7          8          9         10
##  1.437673e+00 -1.960463e-01 -1.503022e+00  4.574414e-01 -1.960463e-01
##         11         12         13         14         15
##  9.802317e-01 -9.802317e-01 -3.267439e-01  3.267439e-01 -1.813793e-17
```

Os resíduos studentizados:

```
##          1          2          3          4          5          6
## -0.5677538 -1.3610257  0.4950873 -0.5677538  2.2976360  1.7372797
##          7          8          9         10         11         12
## -0.2102764 -1.8398662  0.4950873 -0.2102764  1.1060879 -1.1060879
##         13         14         15
## -0.3517196  0.3517196  0.0000000
```

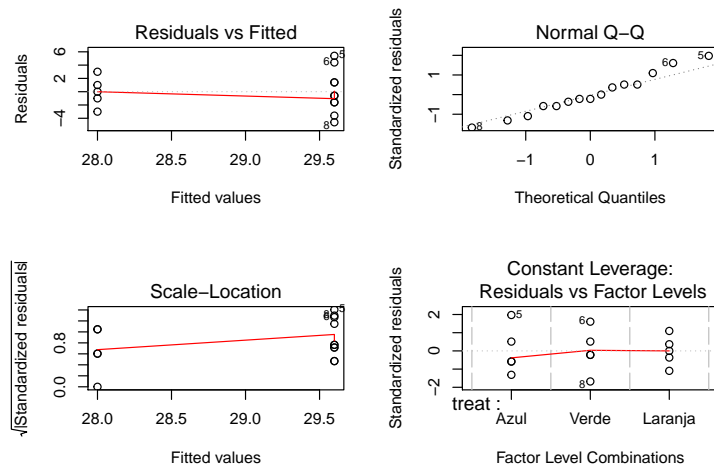
Para a realização de um modelo ANOVA é necessário as seguintes suposições estejam satisfeitas:

- Homocedasticidade (Erros com variâncias iguais)
- Normalidade nos dados
- Independencia (entre observações)

(c) Realize uma análise de diagnóstico dos resíduos. Comente se as suposições do modelo de ANOVA clássico estão satisfeitas a estes dados.

Resolução

Analisando os gráficos abaixo é possível ver que os resíduos estão se comportando como esperado e aparentemente há normalidade nos gráficos, como nas suposições.



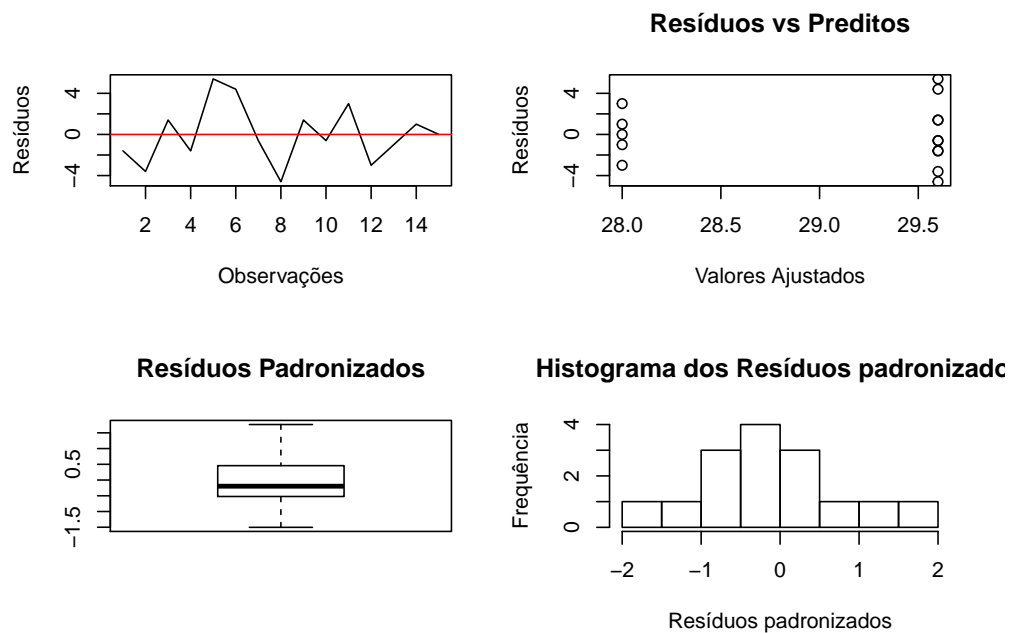
É possível perceber também a do cálculo das médias abaixo (Verde, Laranja e Azul respectivamente) que duas médias dos são exatamente iguais (Verde e Azul):

```
## [1] 29.6
```

```
## [1] 28
```

```
## [1] 29.6
```

No primeiro gráfico abaixo, os resíduos estão de maneira independente, cumprindo a última suposição, os demais gráficos é uma análise mais atenta sobre os resíduos:



Realindo os testes para a normalidade e homogeneidade das variâncias, aceitamos as hipóteses com seus respectivos p-valores:


```
##
## Shapiro-Wilk normality test
##
## data:  res
## W = 0.97404, p-value = 0.9127

## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value Pr(>F)
## group 2  0.1699 0.8457
##      12

##
## Bartlett test of homogeneity of variances
##
## data:  resp by treat
## Bartlett's K-squared = 0.7664, df = 2, p-value = 0.6817
```

(d) Construa a tabela de ANOVA e interprete os resultados.

Resolução

Primeiramente, na coluna Df, é obtido os graus de liberdade do tratamento (3 tratamentos menos 1) e do resíduos (15 dados amostrais menos 3 tratamentos). Em seguida, na coluna Sum Sq e Mean Sq, a soma de quadrados e esses divididos pelos seus respectivos graus de liberdade. Por fim, nas últimas duas colunas é obtido o F valor, dividindo o Mean Sq do tratamento pelo Mean Sq dos res?duos, e o P valor confirmando que as médias dos tratamentos são iguais.

Table 3: Tabela de ANOVA

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
treat	2	8.533333	4.266667	0.455516	0.6446466
Residuals	12	112.400000	9.366667	NA	NA

- e) Quando um executivo recebeu os resultados dessa pesquisa ele comentou: “Viu? Eu estava certo. Daqui para frente não podemos imprimir os questionários em papel branco, que é muito mais barato”. Esta conclusão reflete os resultados obtidos da análise dos dados?

Resolução

Primeiramente, diria que esta hipótese não está sendo testada na pesquisa, o real questionamento é se há diferença entre usar as cores azul, verde e laranja.

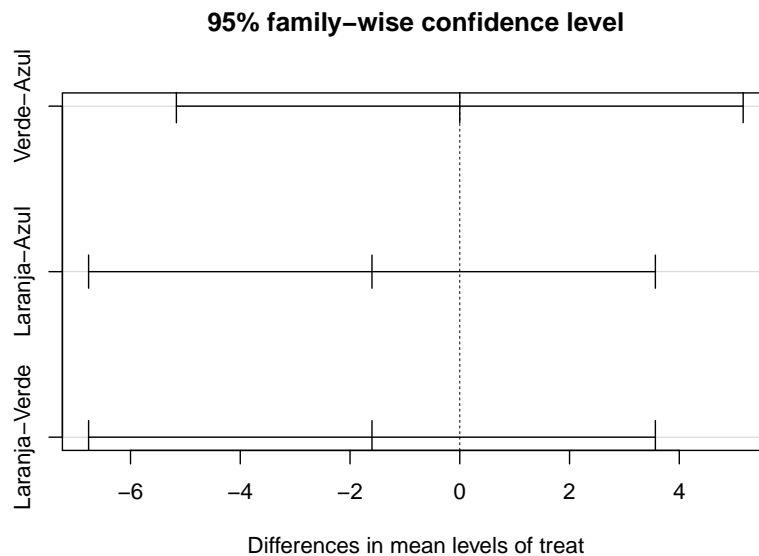
Partindo desse pressuposto, é realizado uma análise verificando se há diferença em pares entre duas cores.

```
## (Intercept)    treatVerde  treatLaranja
## 2.960000e+01  4.135041e-15 -1.600000e+00

## Tables of means
## Grand mean
##
## 29.06667
```

```
##
## treat
## treat
## Azul Verde Laranja
## 29.6 29.6 28.0
##
## Standard errors for differences of means
##      treat
##      1.936
## replic.    5

## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = resp ~ treat, data = base)
##
## $treat
##              diff          lwr          upr      p adj
## Verde-Azul      7.105427e-15 -5.163999  5.163999 1.0000000
## Laranja-Azul -1.600000e+00 -6.763999  3.563999 0.6943357
## Laranja-Verde -1.600000e+00 -6.763999  3.563999 0.6943357
```



Todos os intervalos tem o 0 contido, logo todas as cores são iguais entre si no quesito a ser testado.

Calculando o valor crítico do método de Tukey, e logo observando que as médias seriam consideradas diferentes com uma diferença de 3.77 da média do Azul.

```
## [1] 3.772929
```

I.C. para a diferença entre médias dos grupos diferença entre os grupos Azul e Verde

```
## treatVerde
## -5.163999
```

```
## treatVerde
##      5.163999
```

Diferença entre os grupos Azul e Laranja

```
## treatLaranja
##      -6.763999
```

```
## treatLaranja
##       3.563999
```

Diferença entre os grupos Verde e Laranja

```
## treatLaranja
##      -6.763999
```

```
## treatLaranja
##       3.563999
```

Exercício 5

Depois de ajustar um modelo ANOVA aos dados de um experimento completamente aleatorizado com um fator em 5 níveis, foram obtidos os seguintes resíduos:

Trat	T1	T2	T3	T4	T5
R	-2	26	-2	0	-4
E	0	-4	-6	NA	0
S	-4	-2	2	-2	-4
Í	NA	-2	2	-2	-4
D	12	-6	NA	0	NA
U	-2	NA	-2	-4	-6
O	2	-2	2	0	-2

Complete os NA's com valores de sua escolha e faça uma análise de diagnóstico do ajuste do modelo. Construa gráficos apropriados. Interprete.

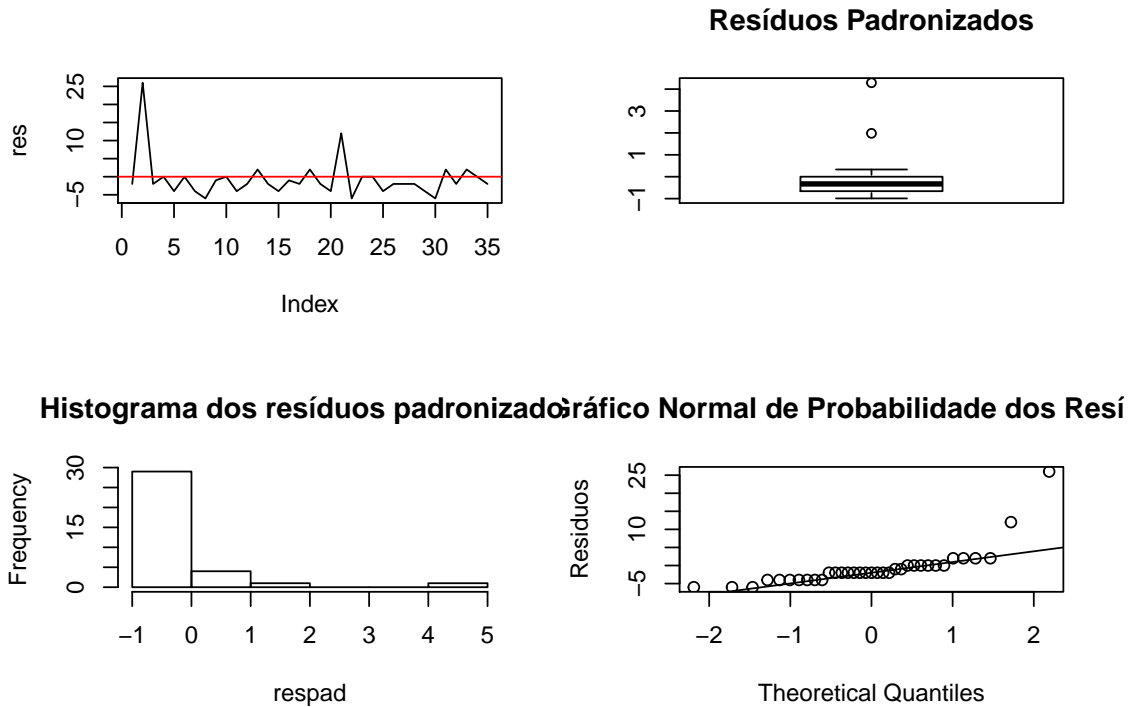
Resolução

Para substituição dos valores NA propomos a mediana por tratamento por ser uma medida robusta devido o fato de ter valores outliers nos tratamentos.

Trat	T1	T2	T3	T4	T5
R	-2	26	-2	0	-4
E	0	-4	-6	-1	0
S	-4	-2	2	-2	-4
Í	-1	-2	2	-2	-4
D	12	-6	0	0	-4
U	-2	-2	-2	-4	-6
O	2	-2	2	0	-2

Observação: Apresente em um Anexo os comandos (em R ou em outro aplicativo) usados na resolução de exercícios.

```
## [1] -2 26 -2 0 -4 0 -4 -6 -1 0 -4 -2 2 -2 -4 -1 -2 2 -2 -4 12 -6 0
## [24] 0 -4 -2 -2 -2 -4 -6 2 -2 2 0 -2
```



```
##
## Shapiro-Wilk normality test
##
## data:  res
## W = 0.61228, p-value = 2.071e-08
```

Podemos notar com os gráficos que nenhuma das suposições de para realizar a ANOVA utilizando DCA está satisfeita, pois o de independência parece ter uma linha de tendência, o teste para homocedasticidade retornou um p-valor < 0.001 que a um nível de significância de 5% rejeitamos a hipótese que os dados tem variâncias homogêneas, além disso analisando o gráfico QQ e o teste de Shapiro que retornou um p-valor < 0.001 que a um nível de significância de 5% rejeitamos a hipótese que os dados seguem distribuição normal, ou seja, em hipótese alguma o modelo ANOVA com DCA poderia ser utilizado com tais dados, neste caso seria melhor procurar algum outro modelo estatístico que tenha suposições mais “relaxadas” para se adequar a este problema.

Códigos

```
library(ggplot2)
library(car)
```

```

library(asbio)
library(gmodels)

#Exercício 3

# item a
dados <- c(96 , 71, 73 , 87 , 104, 85 , 96 , 76,
           92 , 89, 83 , 90 , 101, 110, 97 , 85,
           112, 85, 100, 94 , 100, 73 , 101, 93,
           93 , 81, 88 , 85)

dados.df <- as.data.frame(t(matrix(dados, 4, 7)))
colnames(dados.df) <- c("T1", "T2", "T3", "T4")
dados.df <- stack(dados.df )
colnames(dados.df) <- c("resp", "treat")

summary(dados.df)

ggplot(dados.df, aes(x=treat, y=resp)) +
  geom_boxplot(fill='#A4A4A4', color="black")

# item c
model <- lm(resp~treat, dados.df)
model1 <- aov(resp~treat, dados.df)

anov <- anova(model)

knitr::kable(caption = "Tabela de ANOVA", anov)

# Verificação de independência
s2 <- anov$"Mean Sq"[2]
res <- model1$res
par(mfrow=c(2,2))
plot(res, type="l",xlab="Observações", ylab="Resíduos")
abline(h=0, col="red")

plot(model$fit, model1$res, xlab="Valores Ajustados", ylab="Resíduos")
title("Resíduos vs Preditos")

respad <- (res/sqrt(s2))
boxplot(respad)
title("Resíduos Padronizados")

hist(respad, main=NULL, xlab="Resíduos padronizados",ylab="Frequência")
title("Histograma dos Resíduos padronizados")

# Teste de normalidade de Shapiro-Wilk
shapiro.test(res)

# Teste de homocedasticidade
leveneTest(model)
bartlett.test(resp~treat,data = dados.df)

```

```

# item d

fit.tu <- TukeyHSD(model1)
fit.tu
plot(fit.tu,cex.axis = 0.5)

# Calculando o valor crítico do método de Tukey
n <- dim(dados.df)[1]
K <- length(levels(dados.df$treat))

qtu <- qtkey(0.95,nmeans=K,df=n-K)

# I.C. para a diferença entre médias dos grupos
# Diferença entre os grupos 1 e 4
r <- 7 # número de réplicas (balanceado)
dif41 <- model$coefficients[4]
deltatu <- (qtu/sqrt(2))*sqrt(s2)*sqrt(1/r+1/r)
ic.li41 <- dif41 - deltatatu
ic.ls41 <- dif41 + deltatatu

# Diferença entre os grupos 2 e 4
dif42 <- model$coefficients[4] - model$coefficients[2]
ic.li42 <- dif42 - deltatatu
ic.ls42 <- dif42 + deltatatu

deno <- sqrt(s2)*sqrt(1/r+1/r)
dif21 <- model$coefficients[2]
dif31 <- model$coefficients[3]
dif41 <- model$coefficients[4]
dif32 <- model$coefficients[3] - model$coefficients[2]
dif42 <- model$coefficients[4] - model$coefficients[2]
dif43 <- model$coefficients[4] - model$coefficients[3]
t21 <- dif21/deno
t31 <- dif31/deno
t41 <- dif41/deno
t32 <- dif32/deno
t42 <- dif42/deno
t43 <- dif43/deno

pt21<- pt(c(abs(t21)), df=model$df.residual, lower.tail=FALSE)
pt31<- pt(c(abs(t31)), df=model$df.residual, lower.tail=FALSE)
pt41<- pt(c(abs(t41)), df=model$df.residual, lower.tail=FALSE)
pt32<- pt(c(abs(t32)), df=model$df.residual, lower.tail=FALSE)
pt42<- pt(c(abs(t42)), df=model$df.residual, lower.tail=FALSE)
pt43<- pt(c(abs(t43)), df=model$df.residual, lower.tail=FALSE)

results <- 2*c(pt21,pt31,pt41,pt32,pt42,pt43)

adjustp <- p.adjust(results,method="bonferroni")
cbind(results,adjustp)

```

```

adjustp <- p.adjust(results,method="fdr")
cbind(results,adjustp)

# item e

with(dados.df, pairw.anova(y=resp, x=treat, control="T1", method="dunnett"))

# item f

cmat <- rbind('(T1,T4) vs (T2,T3)' = c( 1, -1, -1,  1),  # Define a matriz dos contrastes ortogonais
              'T1 vs T4'          = c( 1,  0,  0, -1),
              'T2 vs T3'          = c( 0,  1, -1,  0))

model2 <- aov(resp ~ treat,
              data=dados.df,
              contrasts=list(treat=make.contrasts(cmat))) # make.contrasts (gmodels): gera matriz dos con

summary(model2,                                     # ANOVA com a SQDtra e GLtra desdobrados em con
        split=list(treat=list('(T1,T4) vs (T2,T3)'=1,
                              'T1 vs T4'   =2,
                              'T2 vs T3'   =3)))

# Exercício 4

#item a

base.r <- c(28,34,31,
            26,29,25,
            31,25,27,
            28,31,29,
            35,29,28)

base.r <- as.data.frame(t(matrix(base.r,3,5)))
rownames(base.r) <- c("1","2","3","4","5")
colnames(base.r) <- c("Azul","Verde","Laranja")

base <- stack(base.r)
colnames(base) <- c("resp", "treat")

n <- dim(base)[1]
K <- length(levels(base$treat))

summary(base)

ggplot(base, aes(x=treat, y=resp)) +
  geom_boxplot(fill='#A4A4A4', color="black")

# item b
(base$resp-mean(base$resp))/sqrt(var(base$resp))

fit <- aov(resp~treat, base)
anov <- anova(fit)

fit$df.residual

```

```

fit$residuals

fit$residuals/sqrt(anov[3]$`Mean Sq`[2])

rstudent(lm(base$resp~base$treat))

# item c

par(mfrow=c(2,2))
plot(fit)

mean(base.r$Verde)
mean(base.r$Laranja)
mean(base.r$Azul)

s2 <- anov$"Mean Sq"[2]
res <- fit$res
par(mfrow=c(2,2))
plot(res, type="l", xlab="Observações", ylab="Resíduos")
abline(h=0, col="red")

plot(fit$fit, fit$res, xlab="Valores Ajustados", ylab="Resíduos")
title("Resíduos vs Preditos")

respad <- (res/sqrt(s2))
boxplot(respad)
title("Resíduos Padronizados")

hist(respad, main=NULL, xlab="Resíduos padronizados", ylab="Frequência")
title("Histograma dos Resíduos padronizados")

shapiro.test(res)

leveneTest(fit)
bartlett.test(res~treat, base)

# item d

knitr::kable(caption = "Tabela de ANOVA", anov)

# item e

fit$coefficients
model.tables(fit, "means", se = TRUE)

fit.tu <- TukeyHSD(fit)
fit.tu
plot(fit.tu)

qtu <- qtkey(0.95, nmeans=K, df=n-K)
qtu

```



```

r <- 5
dif.av <- fit$coefficients[2]
deltatu <- (qtu/sqrt(2))*sqrt(s2)*sqrt(1/r+1/r)
ic.liav <- dif.av - deltatatu
ic.lsav <- dif.av + deltatatu
ic.liav
ic.lsav

dif.al <- fit$coefficients[3]
ic.lial <- dif.al - deltatatu
ic.lsal <- dif.al + deltatatu
ic.lial
ic.lsal

dif.vl <- fit$coefficients[3] - fit$coefficients[2]
ic.livl <- dif.vl - deltatatu
ic.lsvl <- dif.vl + deltatatu
ic.livl
ic.lsvl

# Exercício 5

res.NA <- c(-2,26,-2,0 ,-4,
            0 ,-4,-6,NA, 0,
            -4,-2,2 ,-2,-4,
            NA,-2,2 ,-2,-4,
            12,-6,NA,0 ,NA,
            -2,NA,-2,-4,-6,
            2 ,-2,2 ,0 ,-2)

res <- c(-2,26 ,-2,0 ,-4,
         0 ,-4,-6,-1, 0,
         -4,-2,2 ,-2,-4,
         -1,-2,2 ,-2,-4,
         12,-6,0 ,0 ,-4,
         -2,-2,-2,-4,-6,
         2 ,-2,2 ,0 ,-2)

res.df <- as.data.frame(t(matrix(res, 5, 7)))
colnames(res.df) <- c("T1", "T2", "T3", "T4","T5")
res.df <- stack(res.df)
colnames(res.df) <- c("resp", "treat")

gl.res <- length(res)-5 # Graus de liberdade n-k
res

s2 <- sum(res^2)/gl.res # estimativa da variância residual

# Verificação de independência
plot(res, type="l") # res x índice das observações
abline(h=0, col="red")

```

```

respad <- (res/sqrt(s2)) # resíduos padronizados
boxplot(respad)
title("Resíduos Padronizados")

hist(respad, main=NULL)
title("Histograma dos resíduos padronizados")

qqnorm(res,ylab="Resíduos", main=NULL)
qqline(res)
title("Gráfico Normal de Probabilidade dos Resíduos")

# Teste de normalidade de Shapiro-Wilk
# Qual é a hipótese H0
shapiro.test(res)

# Teste de homogeneidade de variâncias
bartlett.test(resp~treat,data = res.df)

```