

# MAE0327 - Planejamento e Pesquisa II - Lista 1

Bruno de Castro Paul Schultze<sup>1</sup>

Guilherme Tamborra<sup>2</sup>

Gustavo de Oliveira Kanno<sup>3</sup>

Rodrigo Marcel Araujo Oliveira<sup>4</sup>

Rubens Santos Andrade Filho<sup>5</sup>

Setembro de 2020

## Sumário

<b>Questão 1</b>	<b>2</b>
a) T6	3
b) T12	5
c) T18	7
d) Simulação	8
d)	12
<b>Questão 2</b>	<b>16</b>
<b>Questão 3</b>	<b>16</b>
Tratamento dos Dados	16
Analisando	17
Conclusão	19
<b>Questão 4</b>	<b>19</b>
Organização dos Dados	19
Analisando	19
Conclusão	20

---

<sup>1</sup>Número USP: 10736862

<sup>2</sup>Número USP: 10736816

<sup>3</sup>Número USP: 9795810

<sup>4</sup>Número USP: 9299208

<sup>5</sup>Número USP: 10370336

## Questão 1

Consideramos os dados de consumo de oxigênio (O2cons) avaliado em três tipos de células (T6, T12 e T18) submetidas a um Tratamento (Placebo  $x$  Virum), na presença (=1) ou ausência (=0) de um Vírus.

```
# Carrega os dados
dados <- read.csv('O2cons.csv', stringsAsFactors = T)
dados <- dados %>%
  mutate(Virus = factor(Virus, 0:1, c("Ausente", "Presente")))
attach(dados)

# Tamanho dos grupos
table(Grup) %>% kable(caption='caption')
```

Tabela 1: caption

Grup	Freq
1	12
2	12
3	12
4	12

```
# Médias por grupo
dados %>% group_by(Grup) %>%
  summarise(across(T6:T18, mean, .names = 'Média {col}')) %>%
  kable(caption='caption')
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

Tabela 2: caption

Grup	Média T6	Média T12	Média T18
1	1,618	2,434	3,527
2	1,322	2,430	3,425
3	1,656	2,799	4,029
4	1,394	2,570	3,677

Os dados estão estruturados da seguinte forma:

- Dois fatores de efeito considerados de efeito fixo:
  - fator Tratamento com 2 níveis - Placebo, Virum;
  - fator Virus com 2 níveis - Ausente, Presente;
- 12 réplicas em cada grupo formado pela combinação dos fatores Tratamento e Virus.
- A resposta é multivariada contínua ( $p=3$ ) - consumo de oxigênio (O2cons) cada uma dos três tipos de células T6, T12 e T18. Entretanto faremos uma análise para cada resposta de forma separada.

Com o objetivo de avaliar se o consumo de oxigênio é afetado pelo Tratamento e pela exposição ao Virus, propomos uma Análise de Variância para os dados de cada tipo celular. Na análise consideramos o modelo estrutural e distribucional adotado, análise de diagnóstico das suposições do modelo, tabela da ANOVA com as correspondentes fontes de variação, possíveis comparações múltiplas entre médias e conclusão da análise.

Consideramos que foi feito um Delineamento Completamente Aleatorizado (DAC) com tratamentos compostos pela combinação dos níveis dos dois fatores de efeitos fixos com dois níveis cada. Dessa forma, supomos que os tratamentos foram aplicados às unidades experimentais de modo aleatório e independente. E assim, o modelo estrutural é (utilizando a parametrização Casela de Referência):

$$y_{ijk} = \begin{cases} \mu_{11} + \epsilon_{ijk}, & i = 1, \dots, 12, j = k = 1 \\ \mu_{11} + \tau_j + \beta_k + \gamma_{jk} + \epsilon_{ijk}, & i = 1, \dots, 12, j = k = 2 \end{cases}$$

Onde:

- $y_{ij}$ : é o consumo de oxigênio da unidade  $i$  submetida aos níveis  $j$  de Tratamento e  $k$  de Exposição ao vírus;
- $\mu_{11}$ : é o valor esperado da resposta para o nível Placebo do fator Tratamento e o nível Ausente do fator Vírus (casela de referência);
- $\tau_j$ : é o desvio em relação a  $\mu_{11}$  devido ao efeito principal do nível  $j$  do fator Tratamento;
- $\beta_k$ : é o desvio em relação a  $\mu_{11}$  devido ao efeito principal do nível  $k$  do fator Vírus;
- $\gamma_{jk}$ : é o efeito de interação entre os fatores. É o desvio do efeito aditivo dos fatores;
- $\epsilon_{ijk}$ : é um erro aleatório não controlado.

O erro não controlado é suposto suposto normal, independente e homocedástico. Dessa forma modelo distribucional é

$$\epsilon_{ijk} \sim N(0, \sigma^2) \Rightarrow y_{ij} \sim N(\mu_{ij}, \sigma^2)$$

E sob as suposições de normalidade, independência e homocedasticidade, queremos testar:

$$H_0 : \mu_{ij} = \mu_{kl}, \forall (i, j) \neq (k, l) \text{ vs } H_1 : \text{ existe pelo menos uma diferença entre as médias.}$$

## a) T6

```
o2 = read.csv("02cons.csv", header = TRUE)
o2$Trat = as.factor(o2$Trat)
o2$Virus = as.factor(o2$Virus)
```

Agora, vamos separar cada tipo de célula, e também retirar a variável “Grup”, já que não será utilizada.

```
t6 = o2[c("Trat", "Virus", "T6")]
t12 = o2[c("Trat", "Virus", "T12")]
t18 = o2[c("Trat", "Virus", "T18")]
```

Faremos então a análise de variância ANOVA (vamos considerar que suas pressuposições são satisfeitas) para as células T6. Utilizaremos, para todos os tipos de célula, o seguinte modelo:

$$y_{ijk} = \mu + \tau_{tratamento} + \beta_{virus} + (\tau\beta)_{tratamento:virus} + e_{ij}$$

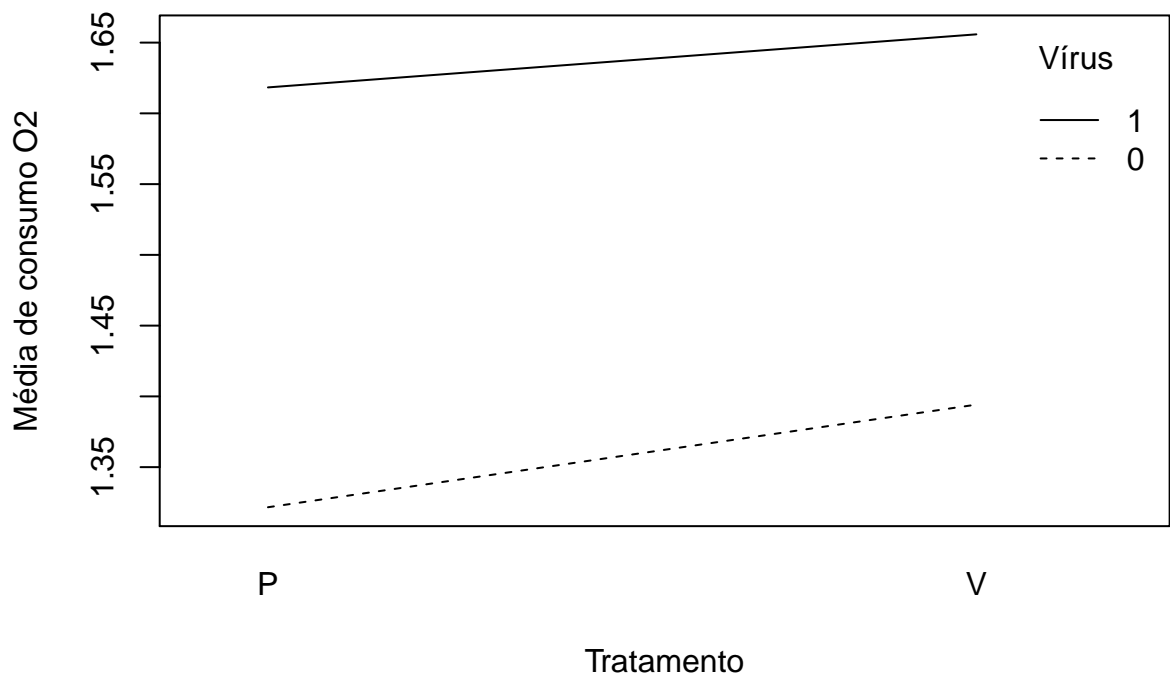
Assim, estamos considerando não somente os diferentes Tratamentos e a presença ou ausência do Vírus, mas também a interação entre tais fatores. Além disso,  $y_{ijk}$  é a i-ésima observação submetida ao j-ésimo tratamento ( $j \in \{Placebo, Virum\}$ ) e à k-ésima presença de vírus ( $k \in \{Presente, Ausente\}$ ). Utilizaremos, como de costume, caselas de referência.

```
t6.aov = aov(T6 ~ Trat * Virus, data = t6)
summary(t6.aov)
```

```
##              Df Sum Sq Mean Sq F value    Pr(>F)
## Trat          1  0.0363   0.0363    0.955    0.334
## Virus          1  0.9352   0.9352   24.597 1.1e-05 ***
## Trat:Virus     1  0.0037   0.0037    0.097    0.757
## Residuals     44  1.6729   0.0380
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Podemos observar, então, que a um nível de 5% somente a variável *Vírus* se mostra significativa, então a hipótese que mudanças no tratamento não interferem na média do consumo de oxigênio é mantida. Além disso, uma possível interação entre *Tratamento* e *Vírus* não é significativa para explicar variações no consumo de oxigênio. Como um adendo, o gráfico abaixo mostra que a interação entre *Vírus* e *Tratamento* é de fato muito baixa, visto que as retas são quase paralelas entre si.

```
interaction.plot(t6$Trat, t6$Virus, t6$T6, xlab = "Tratamento", ylab = "Média de consumo O2", trace.lab
```



## b) T12

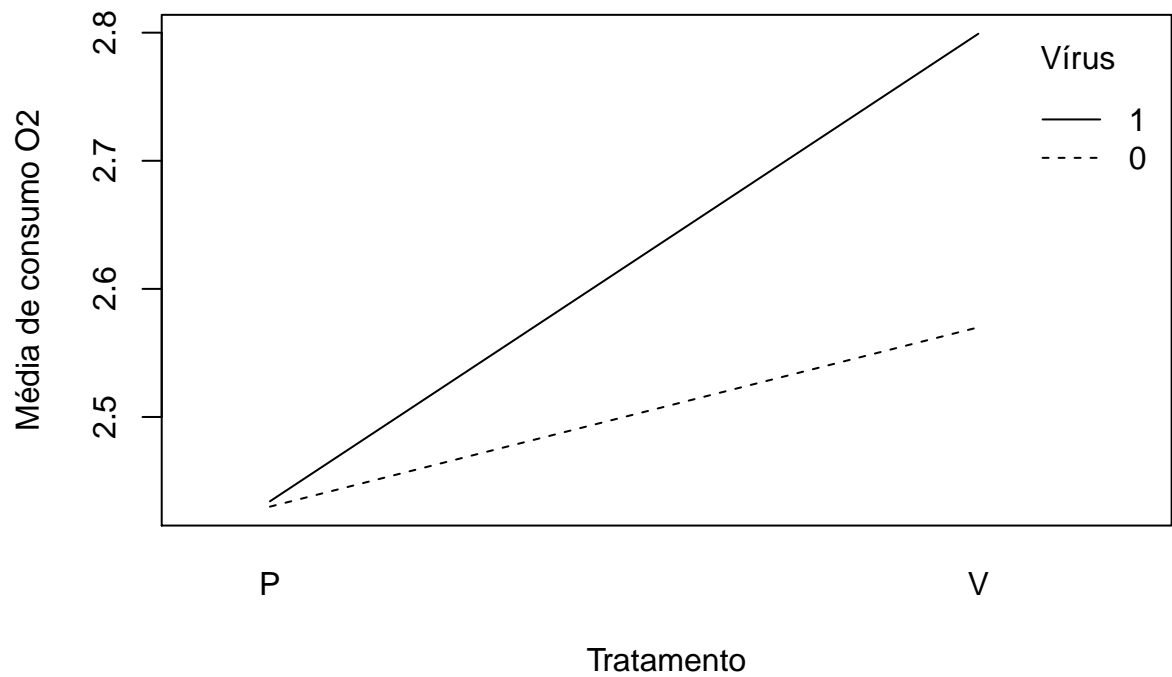
Agora vamos fazer a mesma coisa que na questão anterior, só que para células do tipo T12.

```
t12.aov = aov(T12 ~ Trat * Virus, data = t12)
summary(t12.aov)
```

```
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Trat       1  0.765   0.7651    9.196 0.00406 **
## Virus      1  0.163   0.1633     1.963 0.16819
## Trat:Virus  1  0.152   0.1519     1.825 0.18358
## Residuals 44  3.661   0.0832
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

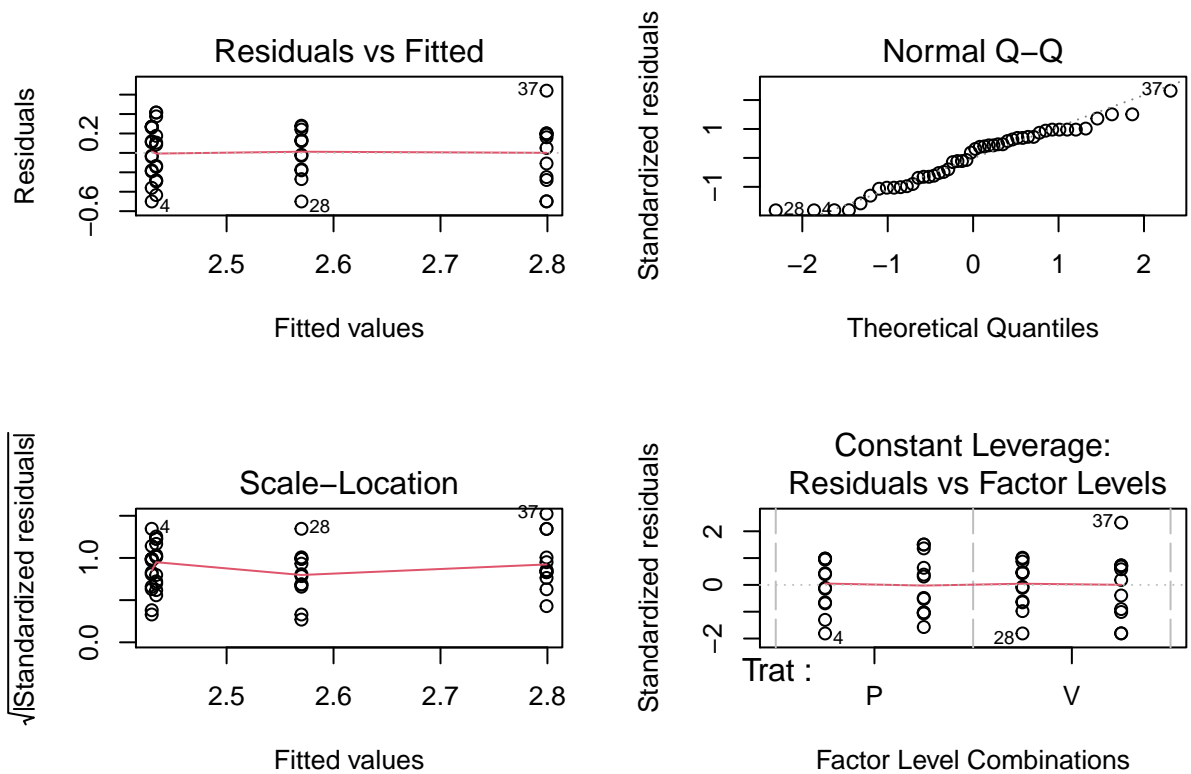
Podemos observar, então, que a um nível de 5% somente a variável *Tratamento* se mostra significativa, então a hipótese que mudanças no tratamento não interferem na média do consumo de oxigênio é rejeitada. É interessante notar que o algoritmo não mostra significância da interação entre *Tratamento* e *Vírus*, mas o gráfico de interação abaixo mostra que ela existe. Então, vamos investigar um pouco mais a fundo se tal interação é realmente não-significativa ou se seu efeito está sendo mascarado por algum ponto.

```
interaction.plot(t12$Trat, t12$Virus, t12$T12, xlab = "Tratamento", ylab = "Média de consumo O2", trace
```



Vamos então procurar pontos influentes.

```
par(mfrow = c(2,2))  
plot(t12.aov)
```



Observemos então que os pontos 28, 37 e 4 se destacam. Vamos retirá-los e refazer a análise de variância.

```
t12.aov2 = aov(T12 ~ Trat * Virus, data = t12[-c(37, 28, 4),])
summary(t12.aov2)
```

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## Trat       1  0.5656   0.5656   8.693 0.00525 **
## Virus      1  0.0183   0.0183   0.281 0.59910
## Trat:Virus  1  0.0781   0.0781   1.200 0.27966
## Residuals 41  2.6673   0.0651
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Vemos então que não, os pontos influentes não estavam mascarando um possível efeito significativo da interação entre *Vírus* e *Tratamento*, tal interação existe mas não é significativa.

### c) T18

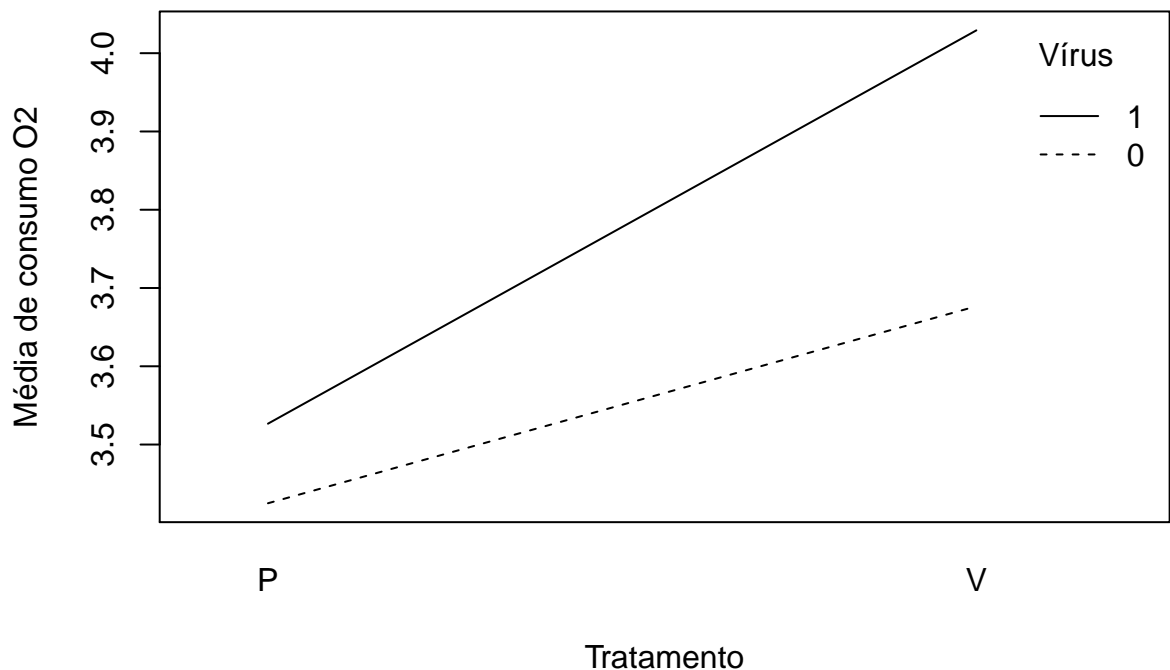
Faremos então o mesmo que nas questões anteriores, mas para células do tipo T18.

```
t18.aov = aov(T18 ~ Trat * Virus, data = t18)
summary(t18.aov)
```

```
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Trat       1  1.706   1.7063   18.083 0.000109 ***
## Virus      1  0.619   0.6188    6.558 0.013950 *
## Trat:Virus  1  0.189   0.1888    2.000 0.164306
## Residuals 44  4.152   0.0944
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Para esse tipo de célula tanto *Tratamento* quanto *Vírus* são significativas para explicar a variação do consumo médio de oxigênio a um nível de 5%. Novamente, a interação entre as duas não parece significativa.

```
interaction.plot(t18$Trat, t18$Virus, t18$T18, xlab = "Tratamento", ylab = "Média de consumo O2", trace
```



Notemos, novamente, que até existe certa interação, mas graficamente ela também não parece significativa.

#### d) Simulação

Agora o objetivo é construir um novo tipo de célula a partir de uma simulação. Construiremos então uma experimento DCA Fatorial 2x2 com Efeito de Interação, além disso o modelo estrutural é o



mesmo utilizado anteriormente, a saber:

$$y_{ijk} = \mu + \tau_{tratamento} + \beta_{virus} + (\tau\beta)_{tratamento:virus} + e_{ij}$$

Estamos considerando também que as observações tem distribuição Normal. Agora, vamos definir as variáveis arbitrariamente, para construir as simulações e conseguir o efeito desejado de interação.

```
set.seed(42)
mu = 10
tau = -5
beta = -5
inter11 = 11
inter01 = -3
inter10 = -3
n = 12
sigma = 2
```

Definimos as variáveis arbitrárias, agora vamos de fato simular as observações.

```
y.00 = rnorm(n, mu, sigma)
y.00 = as.data.frame(y.00)
y.00 = cbind(y.00, rep(0,12), rep(0,12))
colnames(y.00) = c("TSim", "Trat", "Virus")
y.01 = rnorm(n, mu + beta + inter01, sigma)
y.01 = as.data.frame(y.01)
y.01 = cbind(y.01, rep(0,12), rep(1,12))
colnames(y.01) = c("TSim", "Trat", "Virus")
y.10 = rnorm(n, mu + tau + inter10)
y.10 = as.data.frame(y.10)
y.10 = cbind(y.10, rep(1,12), rep(0,12))
colnames(y.10) = c("TSim", "Trat", "Virus")
y.11 = rnorm(n, mu + tau + beta + inter11, sigma)
y.11 = as.data.frame(y.11)
y.11 = cbind(y.11, rep(1,12), rep(1,12))
colnames(y.11) = c("TSim", "Trat", "Virus")

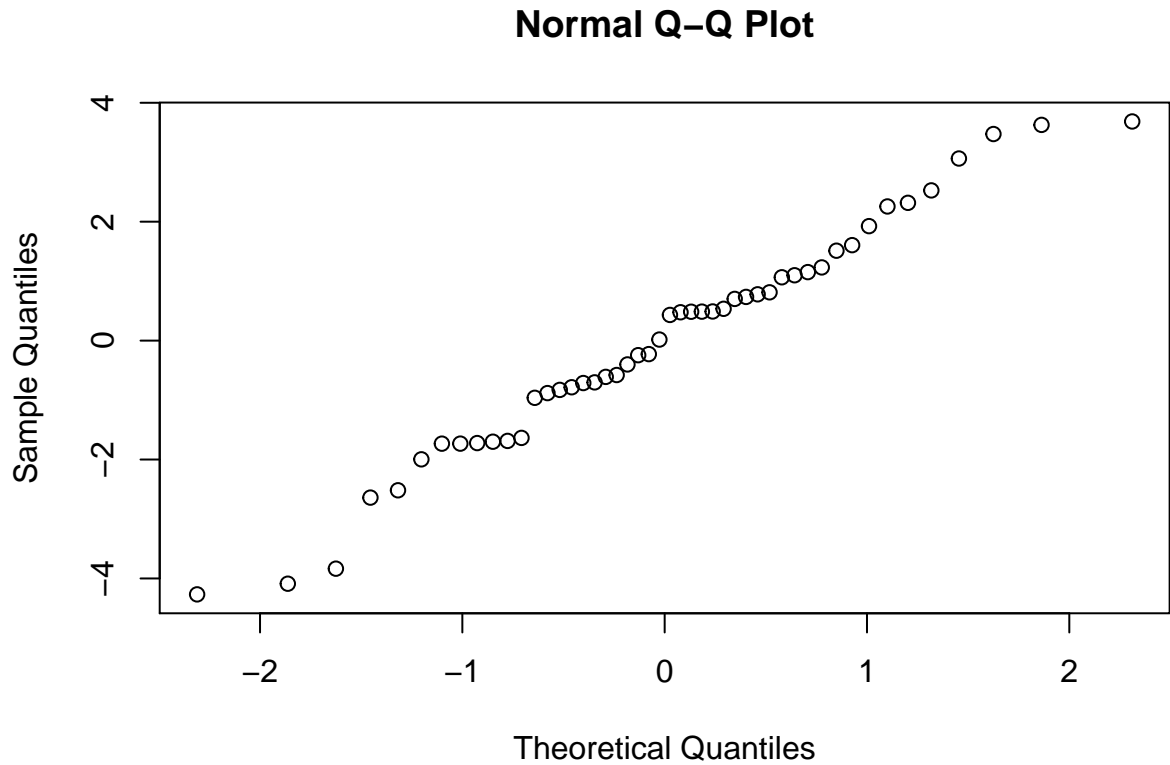
tsim = rbind(y.00, y.01, y.10, y.11)
tsim$Trat[tsim$Trat == 0] = "P"
tsim$Trat[tsim$Trat == 1] = "V"
tsim$Trat = as.factor(tsim$Trat)
tsim$Virus = as.factor(tsim$Virus)
rm(y.00, y.01, y.10, y.11)
```

Agora, com a simulação construída, vamos checar as suposições da ANOVA.

## Normalidade dos Resíduos

Para essa checagem, vamos usar duas abordagens. Primeiro um gráfico QQ comparando com a Distribuição Normal, e também o Teste de Shapiro Wilk.

```
tsim.aov = aov(TSim ~ Trat * Virus, data = tsim)
qqnorm(tsim.aov$residuals)
```



A distribuição parece de fato se adequar à normalidade, mas de qualquer forma utilizaremos o Teste de Shapiro Wilk para confirmar.

```
shapiro.test(tsim.aov$residuals)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  tsim.aov$residuals
## W = 0.9801, p-value = 0.5828
```

Como o p-valor é bem maior que 5%, não rejeitamos a hipótese de normalidade a um nível de 5%, então podemos continuar a análise.

## Homogeneidade da Variância

Para isso utilizaremos o Teste de Levene

```
leveneTest(TSim ~ Trat * Virus, data = tsim)
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value Pr(>F)
## group 3  1.4563 0.2395
##      44
```

Pelo p-valor não rejeitamos a hipótese de homogeneidade da variância.

## Independência entre Observações

Pela forma que a simulação foi construída, as observações são de fato independentes entre sim.

## ANOVA

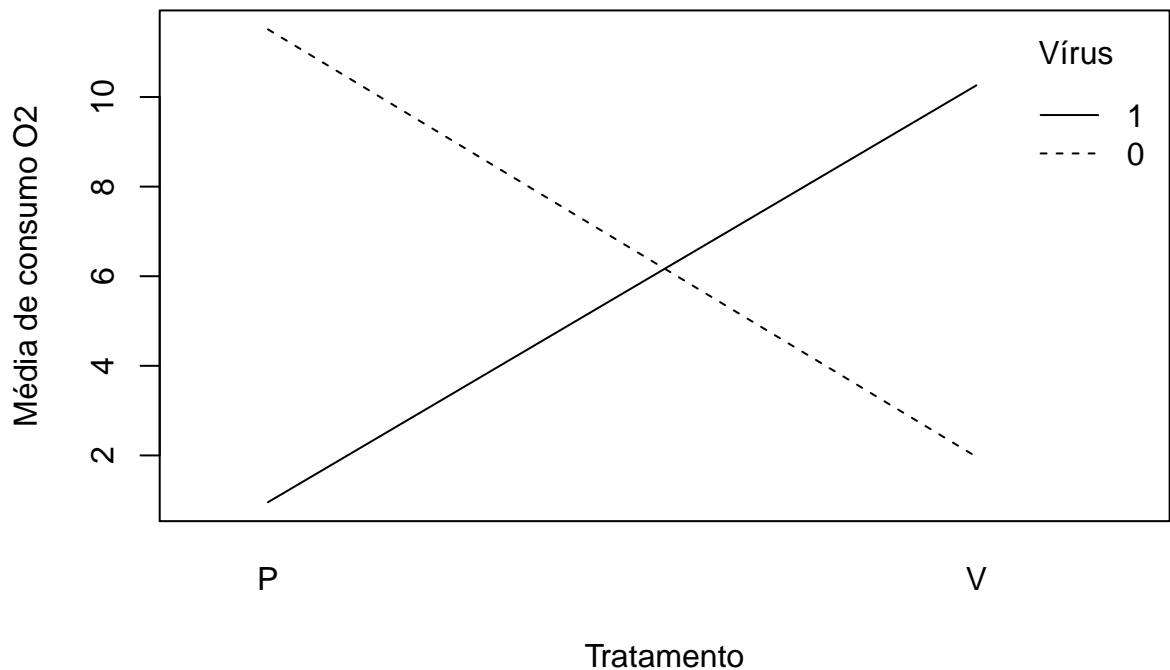
Agora que comprovamos que as suposições da ANOVA estão satisfeitas pela simulação, vamos aplicá-la.

```
tsim.aov = aov(TSim ~ Trat * Virus, data = tsim)
summary(tsim.aov)
```

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## Trat       1    0.2      0.2    0.043 0.8373
## Virus      1   15.4     15.4    3.944 0.0533 .
## Trat:Virus  1 1065.5   1065.5  272.864 <2e-16 ***
## Residuals  44   171.8      3.9
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Podemos ver que, a um nível de 5%, somente a interação entre *Tratamento* e *Vírus* mostrou-se significativa para explicar a variância do consumo médio de oxigênio. Agora vamos checar através do gráfico de interação se essa satisfaz a interação sugerida no enunciado da questão.

```
interaction.plot(tsim$Trat, tsim$Virus, tsim$TSim, xlab = "Tratamento", ylab = "Média de consumo O2", t
```



De fato satisfaz.

d)

Simulamos os dados de um novo tipo celular (digamos TSim) em que há efeito de interação entre Tratamento e exposição ao Virus. Propomos uma análise dos dados gerados.

```
#Simulando dados
#Experimento DCA Fatorial 2x2 com efeito de interação
#Modelo estrutural:  $y = \mu_i + \tau_{1j} + \tau_{2k} + \tau_{12jk} + e$ 
set.seed(1)
mi <- 4
tau1.1 <- (-1.8)
tau2.1 <- (-2)
tau12.11 <- 6
sigma <- 1
mi00 <- mi
mi01 <- mi + tau2.1
mi10 <- mi + tau1.1
mi11 <- mi + tau1.1 + tau2.1 + tau12.11
r <- 12
TSim <- round( c(mi00, mi01, mi10, mi11) + rnorm(4*r, 0, sigma) , 2)
f1 <- rep(0:1, each=2*r)
f2 <- rep(rep(0:1, each=r), 2)
```

```
f12<-rep(1:4,each=r)
datSim<-data.frame(cbind(factor(f1),factor(f2),f12,TSim))
# datSim %>% kable(caption='caption')
str(datSim)
```

```
## 'data.frame':  48 obs. of  4 variables:
## $ V1 : num  1 1 1 1 1 1 1 1 1 1 ...
## $ V2 : num  1 1 1 1 1 1 1 1 1 1 ...
## $ f12 : num  1 1 1 1 1 1 1 1 1 1 ...
## $ TSim: num  3.37 2.18 1.36 7.8 4.33 1.18 2.69 6.94 4.58 1.69 ...
```

```
interaction.plot(f1, f2, TSim)
```

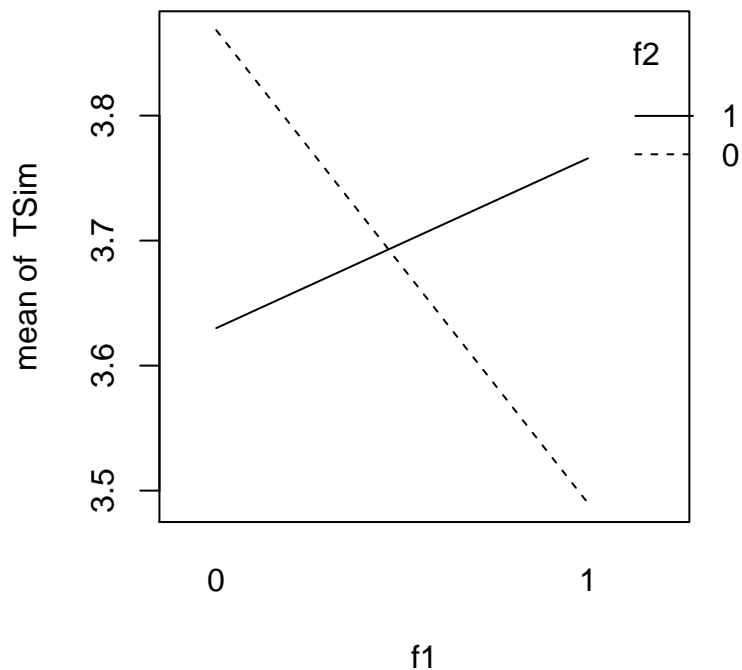


Figura 1: legenda

```
tapply(TSim,f12,mean) %>% t() %>% kable(caption="alguma legenda")
```

Tabela 3: alguma legenda

1	2	3	4
3,868	3,63	3,49	3,766

```
fitSim <- aov(TSim ~ f1*f2, datSim)
#Para fatorial 2x2 é equivalente usar "aov(TSim~factor(f1)*factor(f2))"
anova(fitSim) %>% kable(caption="fasdf")
```

Tabela 4: fasdf

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
f1	1	0,176	0,176	0,046	0,831
f2	1	0,004	0,004	0,001	0,974
f1:f2	1	0,793	0,793	0,206	0,652
Residuals	44	169,133	3,844	NA	NA

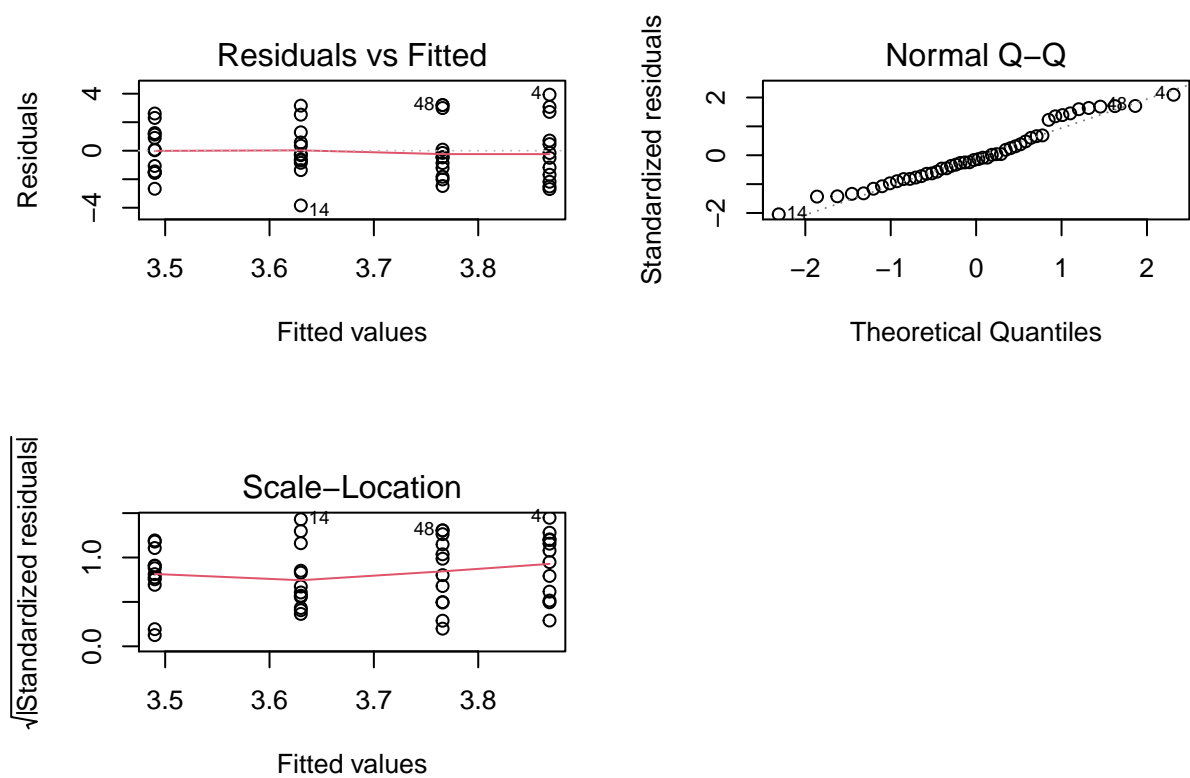
```
fitSim$coefficients %>% t() %>% kable(caption="Coeficientes")
```

Tabela 5: Coeficientes

(Intercept)	f1	f2	f1:f2
3,868	-0,378	-0,238	0,514

```
par(mfrow=c(2,2))
plot(fitSim)
```

```
## hat values (leverages) are all = 0.08333333
## and there are no factor predictors; no plot no. 5
```



Na simulação consideramos o seguinte padrão de variação das médias de O2cons:

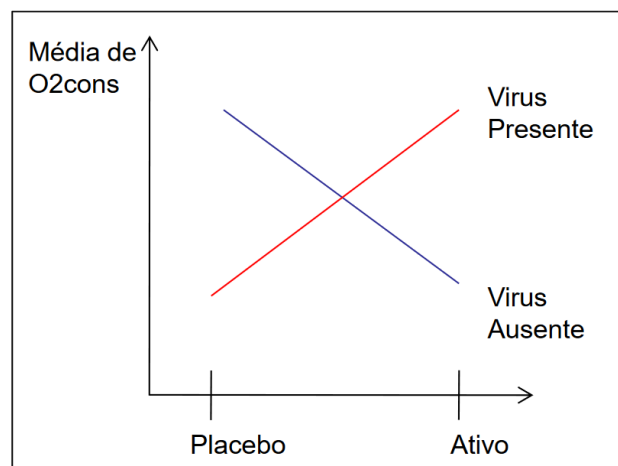


Figura 2: Padrão de variação das médias de O2cons.

## Questão 2

## Questão 3

### Tratamento dos Dados

```
library(MANOVA.RM)
dat.long<-o2cons
#dat.long
names(dat.long)
```

```
## [1] "02"          "Staphylococci" "Time"          "Group"
## [5] "Subject"
```

```
names(dat.long)<-c("02","Staph","Time","Grup","Subj")
#dat.long

library(reshape2)
dat.wide <- dcast(dat.long, Subj + Grup + Staph~ Time, value.var="02")
#dat.wide

dat<-dat.wide
attach(dat)
```

```
## The following object is masked from dados:
##
##      Grup
```

```
names(dat)
```

```
## [1] "Subj"  "Grup"  "Staph" "6"     "12"    "18"
```

```
dat<-cbind(dat,c(rep(1:2,12),rep(3:4,12)))
names(dat)<-c("Subj","Trat","Virus","T6","T12","T18","Grup")
dat<-dat[,2:7]
attach(dat)
```

```
## The following object is masked from dat (pos = 3):
##
##      Grup
```

```
## The following objects are masked from dados:
##
##      Grup, T12, T18, T6, Trat, Virus
```



```
str(dat)
```

```
## 'data.frame': 48 obs. of 6 variables:
## $ Trat : Factor w/ 2 levels "P","V": 1 1 1 1 1 1 1 1 1 1 ...
## $ Virus: int 1 0 1 0 1 0 1 0 1 0 ...
## $ T6 : num 1.48 1.04 1.48 1.04 1.8 1.5 1.65 1.35 1.49 1.19 ...
## $ T12 : num 2.81 2.07 2.52 1.93 2.15 2.7 2.85 2.4 2.54 2.24 ...
## $ T18 : num 3.56 2.81 3.41 2.89 3.2 3.75 3.75 3.3 3.58 3.84 ...
## $ Grup : int 1 2 1 2 1 2 1 2 1 2 ...
```

## Analizando

Teste F global sob premissas clássicas e Teste de aleatorização:

```
library(lmPerm)
anova(aov(T6~Trat*factor(Virus), data = dat))
```

```
## Analysis of Variance Table
##
## Response: T6
##
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
## Trat	1	0.03630	0.03630	0.9547	0.3339
## factor(Virus)	1	0.93521	0.93521	24.5973	1.102e-05 ***
## Trat:factor(Virus)	1	0.00368	0.00368	0.0967	0.7573
## Residuals	44	1.67292	0.03802		

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova(lmp(T6~Trat*factor(Virus), data = dat))
```

```
## [1] "Settings: unique SS "
## Analysis of Variance Table
##
## Response: T6
##
```

	Df	R Sum Sq	R Mean Sq	Iter	Pr(Prob)
## Trat	1	0.03630	0.03630	51	0.6667
## factor(Virus)	1	0.93521	0.93521	5000	<2e-16 ***
## Trat:factor(Virus)	1	0.00368	0.00368	74	0.5811
## Residuals	44	1.67292	0.03802		

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

No T6, ambas as respostas foram satisfatórias. tanto a anova quanto o teste de aleatorização rejeitam as variáveis menos significativas retornando um p-valor elevado e aceitariam a variável significativa, retornando um p-valor pequeno.

```
anova(aov(T12~Trat*factor(Virus), data = dat))
```

```
## Analysis of Variance Table
##
## Response: T12
##              Df Sum Sq Mean Sq F value    Pr(>F)
## Trat          1  0.7651  0.76507   9.1957 0.004058 **
## factor(Virus)  1  0.1633  0.16333   1.9631 0.168191
## Trat:factor(Virus) 1  0.1519  0.15187   1.8254 0.183575
## Residuals     44  3.6608  0.08320
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova(lmp(T12~Trat*factor(Virus), data = dat))
```

```
## [1] "Settings: unique SS "
## Analysis of Variance Table
##
## Response: T12
##              Df R Sum Sq R Mean Sq Iter Pr(Prob)
## Trat          1  0.7651  0.76508 5000  0.0088 **
## factor(Virus)  1  0.1633  0.16333  860  0.1047
## Trat:factor(Virus) 1  0.1519  0.15187  167  0.3772
## Residuals     44  3.6608  0.08320
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Para T12, apesar do teste de aleatorização ter obtido resultados distantes da Anova, ele tbm retornou p-valores altos para as variáveis menos significativas e bem baixo para a mais, ou seja, ambas tiveram o mesmo desempenho.

```
anova(aov(T18~Trat*factor(Virus), data = dat))
```

```
## Analysis of Variance Table
##
## Response: T18
##              Df Sum Sq Mean Sq F value    Pr(>F)
## Trat          1  1.7063  1.70630 18.0825 0.0001085 ***
## factor(Virus)  1  0.6188  0.61880   6.5578 0.0139496 *
## Trat:factor(Virus) 1  0.1888  0.18875   2.0003 0.1643057
## Residuals     44  4.1519  0.09436
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova(lmp(T18~Trat*factor(Virus), data = dat))
```

```
## [1] "Settings: unique SS "
## Analysis of Variance Table
##
## Response: T18
##              Df R Sum Sq R Mean Sq Iter Pr(Prob)
## Trat          1  1.7063  1.70630 5000  <2e-16 ***
```

```
## factor(Virus)      1  0.6188  0.61880 5000  0.0080 **
## Trat:factor(Virus) 1  0.1888  0.18875 752  0.1184
## Residuals         44  4.1519  0.09436
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

No T18, ambos os testes retornaram p-valores muito próximos, sendo provavelmente os mais aproximados dos dois testes

## Conclusão

Tanto o Teste F global sob premissas clássicas quanto o Teste de aleatorização obtiveram desempenhos muito próximos, sendo que, para os valores de significância mais utilizados (0.05, 0.01, etc), ambos os testes levariam as mesmas conclusões.

## Questão 4

### Organização dos Dados

```
df = data.frame(taxa_cresc = c(0.02, 0.27, 0.11, 0.48, 0.15, 0.24, 0.35, 0.04, 0.45, -0.01, 0.14, 0.18,
faixa_etaria = c(rep(c(1, 2, 3, 4), times=4)),
medicamento = c("A", "B", "C", "D", "B", "C", "D", "A", "D", "A", "B", "C", "C", "D", "A", "B")
)

df$faixa_etaria <- as.factor(df$faixa_etaria)
df$medicamento <- as.factor(df$medicamento)
```

### Analisando

```
anova(aov(taxa_cresc ~ medicamento + faixa_etaria, data = df))

## Analysis of Variance Table
##
## Response: taxa_cresc
##           Df Sum Sq Mean Sq F value    Pr(>F)
## medicamento  3 0.43332  0.144440   86.917 5.775e-07 ***
## faixa_etaria  3 0.03462  0.011540    6.944  0.01021 *
## Residuals    9 0.01496  0.001662
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova(lmp(taxa_cresc ~ medicamento + faixa_etaria, data = df))
```

```
## [1] "Settings:  unique SS "  
## Analysis of Variance Table  
##  
## Response: taxa_cresc  
##           Df R Sum Sq R Mean Sq Iter Pr(Prob)  
## medicamento    3  0.43332  0.144440 5000  <2e-16 ***  
## faixa_etaria    3  0.03462  0.011540 5000  0.0086 **  
## Residuals       9  0.01496  0.001662  
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

## Conclusão

como é possível ver, ambos os testes tiveram resultados próximos e retornaram um baixo p-valor para a variável mais significativa, ou seja, os dois testes levariam as mesmas conclusões.