

Exercício Prático
Universidade Federal do Ceará
Centro de Ciências
Departamento de Estatística e Matemática Aplicada
Bacharelado em Estatística
Planejamento de Experimentos

Antônio Arthur Silva de Lima
Romulo Barros de Freitas

09 de fevereiro de 2025

Problema

Nesse exercício proposto será realizada uma análise de variância (ANOVA) para um delineamento em blocos inteiramente casualizados (BIB). O objetivo é analisar as notas que 130 provadores (blocos) atribuíram sob a percepção de cor de uma determinada bebida. Cada provador teve acesso a quatro (repetições) de treze tratamentos, isso por conta de limitações impostas por especialistas da área de engenharia de alimentos. Dessa forma, por meio da análise de variância será possível inferir estatisticamente quais tratamentos refletem em maiores notas, segundo a percepção dos provadores. Toda a análise será realizada por meio do software R.

Ao todo, temos 130 provadores, 13 tratamentos e 4 repetições por bloco, sendo 40 repetições para cada tratamento.

Leitura da base

Primeiro, é realizada a leitura da base com o auxílio da biblioteca *readxl*, pois o conjunto de dados se encontra em uma planilha de excel, no formato *wide*, isto é, fora do padrão utilizado pela maioria dos pacotes que geram as análises a serem feitas. É necessário, portanto, converter o *dataset* do formato atual para um em que os blocos (provadores), os tratamentos e as notas (resposta) estejam em suas respectivas colunas. Este último formato é convencionalmente chamado de **tidy**.

```
df0 = read_excel('resultados_estatistica_Uva.xls', skip = 1)
blocos = rep(1:nrow(df0), each=4); blocos
tratamentos = c(t(df0[2:5])); tratamentos
tratamentos = as.numeric(tratamentos)
notas = c(t(df0[6:9]))
notas = as.numeric(notas); notas

df = as.data.frame(cbind(blocos, tratamentos, notas))
df$tratamentos = as.factor(df$tratamentos)
```

Quando analisamos a nota média de cada tratamento em relação à cor, observamos que o tratamento cinco possui uma menor nota média de avaliação sendo de 2,4. Por outro lado, os tratamentos onze e doze possuem as maiores notas médias, sendo 6,00 e 6,15, respectivamente.

```
tapply(df$notas, df$tratamentos, mean)
```

```
##   1    2    3    4    5    6    7    8    9    10   11   12   13  
## 3.175 3.300 5.875 5.125 2.400 5.750 3.925 4.175 5.825 5.875 6.000 6.150 5.725
```

Quando analisamos o coeficiente de variação das notas de cada tratamento, é possível perceber que o tratamento cinco possui o maior coeficiente de variação, com 75,96%. Por outro lado, os menores coeficientes de variação são referentes ao tratamento doze e ao tratamento dez, sendo de 25,39% e 31,83%, respectivamente.

```
round((tapply(df$notas, df$tratamentos, sd) / tapply(df$notas, df$tratamentos, mean)) * 100, 2) |>  
str_c("%")
```

```
## [1] "63.75%" "62.59%" "34.52%" "44.01%" "75.96%" "38.94%" "47.36%" "55.81%"  
## [9] "33.87%" "31.83%" "32.9%" "25.39%" "33.33%"
```

Partindo agora para a variabilidade nos tratamentos, esperamos que elas sejam estatisticamente iguais, ou seja, que os tratamentos sejam homogêneos em relação à variância. Podemos verificar esse pressuposto de várias formas. Primeiro, avaliando os desvios padrão de cada tratamento, vemos que não há diferenças significativas entre eles.

```
##      1      2      3      4      5      6      7      8  
## 2.024054 2.065591 2.027851 2.255336 1.822931 2.238933 1.858970 2.330264  
##      9     10     11     12     13  
## 1.972731 1.869972 1.974192 1.561557 1.907979
```

Também atestamos essa inferência através do boxplot logo abaixo, que revela um pouco mais a respeito da distribuição dos tratamentos.

```
df |> ggplot(aes(tratamentos, notas)) +  
  geom_boxplot()
```

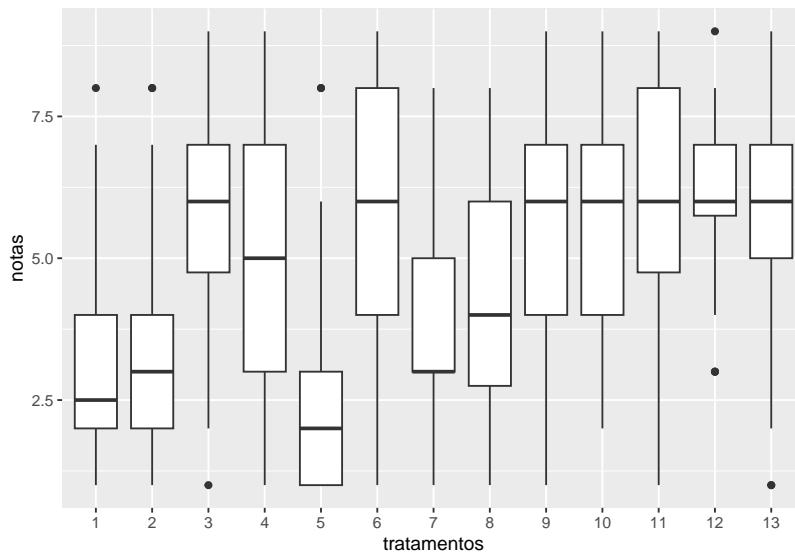


Figura 1: Boxplot das respostas por tratamento

Por fim, esperamos não rejeitar a hipótese nula no teste de Bartlett:

```
bartlett.test(df$notas, df$tratamentos)

##
##  Bartlett test of homogeneity of variances
##
## data: df$notas and df$tratamentos
## Bartlett's K-squared = 9.9705, df = 12, p-value = 0.6185
```

Com a hipótese de homogeneidade de variâncias satisfeita, queremos agora estudar o comportamento das médias de cada tratamento, a fim de sabermos se há notas semelhantes entre eles, ou seja, se podem ser “agrupadas”. Antes do quadro ANOVA, observando o gráfico de barras abaixo — que traz uma comparação das médias dos tratamentos — vemos que possivelmente há alguns tratamentos com médias estatisticamente iguais a outros.

```
df |>
  group_by(tratamentos) |>
  summarise(media = mean(notas), ep = sd(notas)/sqrt(n())) |>
  ggplot(aes(x = tratamentos, y = media)) +
  geom_bar(stat = 'identity', fill = 'white', color = 'black') +
  geom_errorbar(aes(ymin = media - ep, ymax = media + ep), width = 0.2) +
  geom_text(aes(label = round(media, 2)),
            vjust = -1.5,
            color = 'black',
            size = 3.5) +
  labs(y = 'notas médias',
       x = 'tratamentos')
```

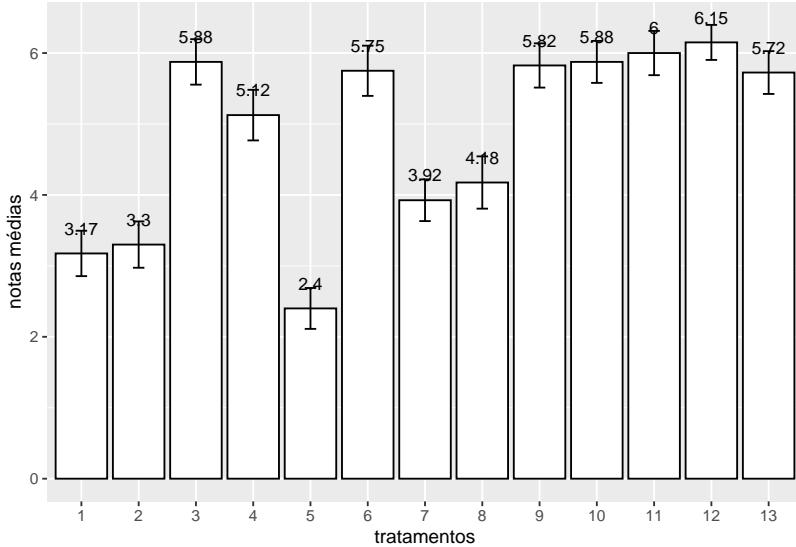


Figura 2: Comparação de médias entre tratamentos

Por meio da análise de variância é possível inferir que há o efeito dos blocos e dos tratamentos em relação à percepção de cor da bebida analisada, uma vez que o valor p para o efeito dos blocos e o efeito dos tratamentos

foi quase nulo. Além disso, foi observada uma variancia residual (QMRES) de 1,738 e um confecienete de variação de 27,1%, indicando variabilidade moderada em relação à nota média geral observada de 4,87.

Em relação ao teste de Tukey, realizado ao nível de significância de 5%, foi observado que a média das notas dos tratamentos 6, 12, 9, 11, 13 e 10 não diferem estatisticamente entre si. Esses tratamentos foram os melhores avaliados no experimento e devem ser piorizados. Por outra perspectiva, os tratamentos 1, 2 e 5 não diferiram estatisticamente entre si e foram os que obtiveram as menores médias de avaliação no experimento.

```
##
## ANALYSIS BIB: df$notas
## Class level information
##
## Block: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34
## Trt : 1 2 4 10 3 5 11 6 12 7 13 8 9
##
## Number of observations: 520
##
## Analysis of Variance Table
##
## Response: df$notas
##              Df  Sum Sq Mean Sq F value    Pr(>F)
## block.unadj 129 1483.61 11.501  6.6182 < 2.2e-16 ***
## trt.adj      12  702.63 58.553 33.6945 < 2.2e-16 ***
## Residuals   378  656.87  1.738
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## coefficient of variation: 27.1 %
## df$notas Means: 4.869231
##
## df$tratamentos, statistics
##
##      df$notas mean.adj      SE r      std Min Max
## 1      3.175 3.015385 0.2295605 40 2.024054 1  8
## 2      3.300 3.069231 0.2295605 40 2.065591 2  9
## 3      5.875 5.361538 0.2295605 40 2.027851 1  9
## 4      5.125 5.492308 0.2295605 40 2.255336 3  9
## 5      2.400 2.230769 0.2295605 40 1.822931 1  9
## 6      5.750 6.084615 0.2295605 40 2.238933 1  8
## 7      3.925 4.415385 0.2295605 40 1.858970 1  9
## 8      4.175 4.215385 0.2295605 40 2.330264 1  9
## 9      5.825 5.992308 0.2295605 40 1.972731 1  8
## 10     5.875 5.707692 0.2295605 40 1.869972 1  9
## 11     6.000 5.938462 0.2295605 40 1.974192 1  8
## 12     6.150 6.053846 0.2295605 40 1.561557 1  8
## 13     5.725 5.723077 0.2295605 40 1.907979 1  9
##
## Tukey
## Alpha      : 0.05
## Std.err    : 0.2312341
## HSD        : 1.090329
## Parameters BIB
## Lambda     : 10
## treatmeans : 13
## Block size : 4
```

```

## Blocks      : 130
## Replication: 40
##
## Efficiency factor 0.8125
##
## <<< Book >>>
##
## Comparison between treatments means
##          Difference pvalue sig.
## 1 - 2    -0.05384615 1.0000
## 1 - 3    -2.34615385 0.0000 *** 
## 1 - 4    -2.47692308 0.0000 *** 
## 1 - 5     0.78461538 0.4467
## 1 - 6    -3.06923077 0.0000 *** 
## 1 - 7    -1.40000000 0.0016 ** 
## 1 - 8    -1.20000000 0.0168 * 
## 1 - 9    -2.97692308 0.0000 *** 
## 1 - 10   -2.69230769 0.0000 *** 
## 1 - 11   -2.92307692 0.0000 *** 
## 1 - 12   -3.03846154 0.0000 *** 
## 1 - 13   -2.70769231 0.0000 *** 
## 2 - 3    -2.29230769 0.0000 *** 
## 2 - 4    -2.42307692 0.0000 *** 
## 2 - 5     0.83846154 0.3362
## 2 - 6    -3.01538462 0.0000 *** 
## 2 - 7    -1.34615385 0.0032 ** 
## 2 - 8    -1.14615385 0.0292 * 
## 2 - 9    -2.92307692 0.0000 *** 
## 2 - 10   -2.63846154 0.0000 *** 
## 2 - 11   -2.86923077 0.0000 *** 
## 2 - 12   -2.98461538 0.0000 *** 
## 2 - 13   -2.65384615 0.0000 *** 
## 3 - 4    -0.13076923 1.0000
## 3 - 5     3.13076923 0.0000 *** 
## 3 - 6    -0.72307692 0.5837
## 3 - 7     0.94615385 0.1662
## 3 - 8     1.14615385 0.0292 * 
## 3 - 9    -0.63076923 0.7774
## 3 - 10   -0.34615385 0.9978
## 3 - 11   -0.57692308 0.8660
## 3 - 12   -0.69230769 0.6519
## 3 - 13   -0.36153846 0.9966
## 4 - 5     3.26153846 0.0000 *** 
## 4 - 6    -0.59230769 0.8431
## 4 - 7     1.07692308 0.0566 . 
## 4 - 8     1.27692308 0.0072 ** 
## 4 - 9    -0.50000000 0.9488
## 4 - 10   -0.21538462 1.0000
## 4 - 11   -0.44615385 0.9788
## 4 - 12   -0.56153846 0.8867
## 4 - 13   -0.23076923 1.0000
## 5 - 6    -3.85384615 0.0000 *** 
## 5 - 7    -2.18461538 0.0000 *** 
## 5 - 8    -1.98461538 0.0000 ***

```

```

## 5 - 9   -3.76153846 0.0000 ***  

## 5 - 10  -3.47692308 0.0000 ***  

## 5 - 11  -3.70769231 0.0000 ***  

## 5 - 12  -3.82307692 0.0000 ***  

## 5 - 13  -3.49230769 0.0000 ***  

## 6 - 7   1.66923077 0.0000 ***  

## 6 - 8   1.86923077 0.0000 ***  

## 6 - 9   0.09230769 1.0000  

## 6 - 10  0.37692308 0.9951  

## 6 - 11  0.14615385 1.0000  

## 6 - 12  0.03076923 1.0000  

## 6 - 13  0.36153846 0.9966  

## 7 - 8   0.20000000 1.0000  

## 7 - 9   -1.57692308 0.0002 ***  

## 7 - 10  -1.29230769 0.0060 **  

## 7 - 11  -1.52307692 0.0003 ***  

## 7 - 12  -1.63846154 0.0001 ***  

## 7 - 13  -1.30769231 0.0051 **  

## 8 - 9   -1.77692308 0.0000 ***  

## 8 - 10  -1.49230769 0.0005 ***  

## 8 - 11  -1.72307692 0.0000 ***  

## 8 - 12  -1.83846154 0.0000 ***  

## 8 - 13  -1.50769231 0.0004 ***  

## 9 - 10  0.28461538 0.9997  

## 9 - 11  0.05384615 1.0000  

## 9 - 12  -0.06153846 1.0000  

## 9 - 13  0.26923077 0.9998  

## 10 - 11 -0.23076923 1.0000  

## 10 - 12 -0.34615385 0.9978  

## 10 - 13 -0.01538462 1.0000  

## 11 - 12 -0.11538462 1.0000  

## 11 - 13  0.21538462 1.0000  

## 12 - 13  0.33076923 0.9986  

##  

## Treatments with the same letter are not significantly different.  

##  

##      df$notas groups  

## 6  6.084615      a  

## 12 6.053846     a  

## 9  5.992308     a  

## 11 5.938462     a  

## 13 5.723077     a  

## 10 5.707692     a  

## 4  5.492308     ab  

## 3  5.361538     ab  

## 7  4.415385     bc  

## 8  4.215385     c  

## 2  3.069231     d  

## 1  3.015385     d  

## 5  2.230769     d

```