

Relatório Atividade Semana 7

Fernando Bispo

26/09/2021

Primeira Questão

Introdução

O conjunto de dados (**dados1**) contém as observações acerca do tempo de permanência na chamada telefônica dos clientes da companhia aérea e o tipo de áudio utilizado no momento em que os clientes aguardam atendimento, sendo estas variáveis quantitativa e qualitativa respectivamente.

O intuito é comparar a distribuição do tempo de espera dos clientes com os diferentes tipos de áudio, sendo estes: um anúncio sobre a companhia aérea e suas promoções atuais, música de elevador ou música clássica gravada.

Para tanto será utilizado o teste de hipótese de Análise de Variância de um fator, abreviadamente denominado **ANOVA** pois se faz necessário a comparação entre as variáveis, sendo estas compostas por mais de 2 classes.

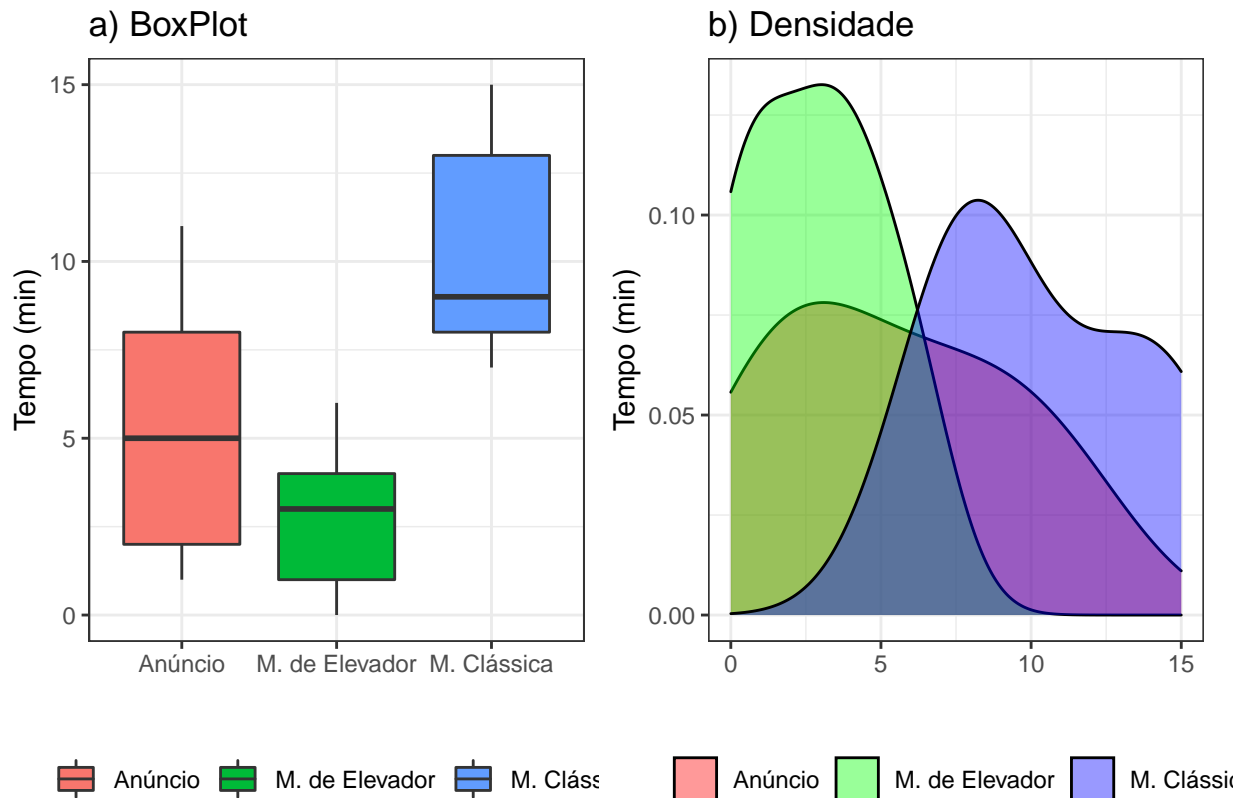
Análise Descritiva

Antes de se iniciar a análise inferencial é necessário fazer uma análise descritiva a fim de se gerar uma interpretação preliminar. Analisando o BoxPlot (Figura 1a) é possível identificar um comportamento assimétrico entre as categorias de tipos de áudio. Os clientes que escutaram Música Clássica aguardaram por mais tempo até o atendimento, quando comparado aos clientes que aguardaram ouvindo os outros tipos de áudio. É possível identificar que há uma maior intercessão entre os BoxPlots dos indivíduos que escutaram o Anúncio e os que escutaram Música de Elevador, indicando que não há diferenciação entre esses indivíduos quanto ao tempo esperado.

Analisando a densidade (Figura 1b) é constatado uma maior dispersão entre os indivíduos do grupo que ouviram o Anúncio, além de uma pequena assimetria positiva. Entre os indivíduos do grupo que ouviram Música de Elevador há uma assimetria positiva e uma assimetria negativa entre os indivíduos que ouviram Música Clássica.

Table 1: Tabela 1: Medidas Resumo

Tipo de Áudio	Min	Q1	Media	MD	Q3	Max	SD
Anúncio	1	2	5.4	5	8	11	4.16
Música de Elevador	0	1	2.8	3	4	6	2.39
Música Clássica	7	8	10.4	9	13	15	3.44



Na Tabela 1 há medidas resumo dos dados, em que é possível identificar maior dispersão (variabilidade) no grupo que ouviu o Anúncio e que em média o grupo que ouviu Música Clássica permaneceu por um maior tempo de espera.

Análise Inferencial

Antes da aplicação do teste ANOVA é necessário constatar algumas premissas:

1. Independência dos dados amostrais;
2. Normalidade de distribuição entre os dados;
3. Igualdade da Variância Populacional (Homogeneidade).

Análise dos pré requisitos

Verificação da Independência Para este item é assumido independência entre os dados do mesmo grupo e entre os dados dos grupos diferentes.

Verificação da normalidade dos dados. O método utilizado para verificação da normalidade entre os dados dos grupos é o método de *Shapiro-Wilk*, em que assume as seguintes hipóteses:

- **H0:** Os dados são provenientes de uma Distribuição Normal
- **H1:** Os dados **não são** provenientes de uma Distribuição Normal

```
##
## Shapiro-Wilk normality tests
##
## data: Tempo by Gravação
##
##              W p-value
## Anúncio      0.9468  0.7147
## Música de Elevador 0.9738  0.8989
## Música Clássica 0.9017  0.4195
```

Utilizando como base o *p-value* a um nível de significância de 5%, há indícios suficientes para aceitar a normalidade entre os dados, não sendo rejeitada assim a hipótese H0.

Verificação da homogeneidade das Variâncias Para a realização da verificação da homogeneidade das Variâncias é realizado o teste de Levene, que consiste num teste com as seguintes hipóteses:

- **H0:** As variâncias dos grupos são homogêneas.
- **H1:** As variâncias dos grupos **não são** homogêneas.

```
## Levene's Test for Homogeneity of Variance (center = mean)
##      Df F value Pr(>F)
## group  2  1.2275 0.3273
##      12
```

Utilizando como base o *p-value* há um nível de significância de 5%, há indícios suficientes para aceitar a homogeneidade entre os grupos, não sendo rejeitada assim a hipótese H0.

Após as análise das premissas há constatação da viabilidade da utilização do teste ANOVA nos dados coletados.

Teste ANOVA

O teste de hipótese de Análise Variância assume as seguintes hipóteses:

- **H0:** As médias dos grupos são todas iguais para um nível de significância de 5%.
- **H1:** Pelo menos uma média dos grupos **não é** igual para um nível de significância de 5%.

```
##      Df Sum Sq Mean Sq F value Pr(>F)
## gravacao    2  149.2    74.6   6.431 0.0126 *
## Residuals   12  139.2    11.6
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Conforme o nível de significância estabelecido, há indícios suficientes para rejeitar H0 pois o *p-value* menor que 5%, ou seja, há diferença entre as médias dos grupos e para identificar entre quais grupos há essa diferença é realizada a análise de PostHoc pelo método de Bonferroni, sendo este o de maior precisão.

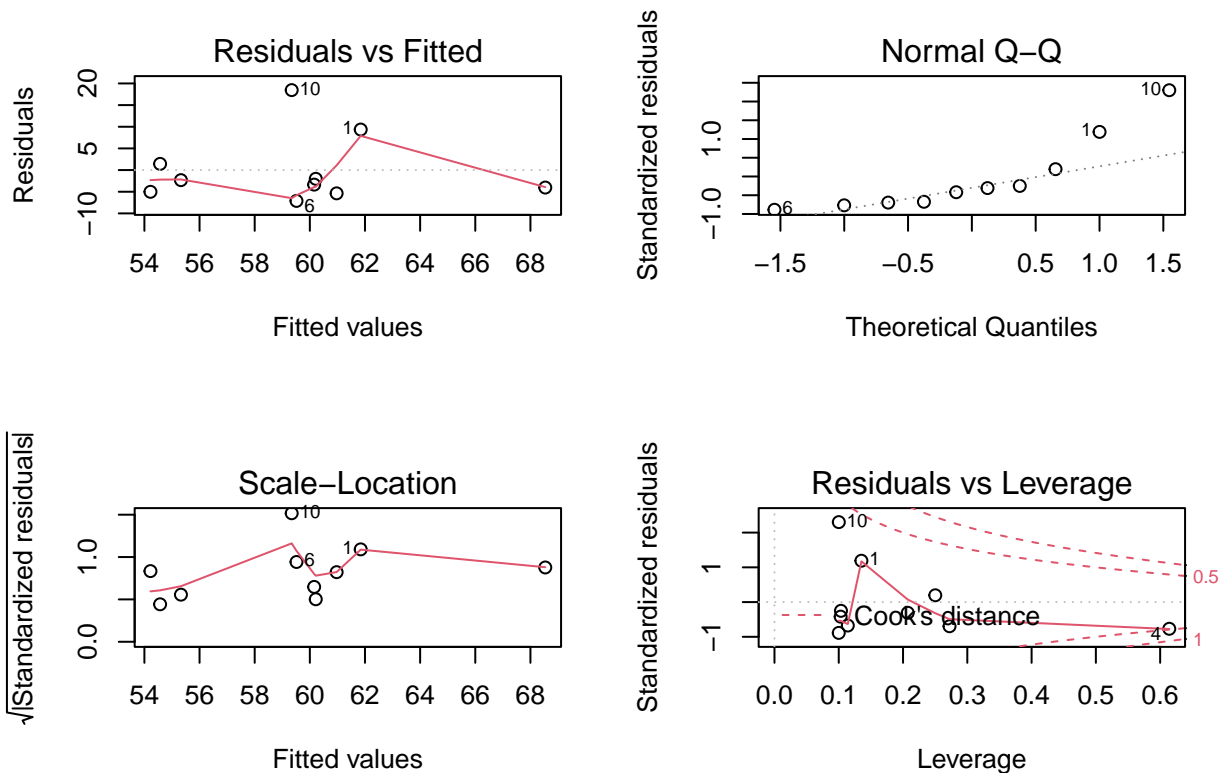
Análise de PostHoc

```
##
## Posthoc multiple comparisons of means : Bonferroni
## 95% family-wise confidence level
```

```
##
## $gravacao
##
##          diff      lwr.ci      upr.ci      pval
## Música de Elevador-Anúncio    -2.6 -8.5871683   3.387168 0.7520
## Música Clássica-Anúncio        5.0 -0.9871683  10.987168 0.1161
## Música Clássica-Música de Elevador  7.6  1.6128317  13.587168 0.0125 *
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Conclui-se que não existe diferença entre os grupos que ouviram Música de Elevador e os que ouviram o Anúncio, nem entre os grupos que ouviram Música de Clássica e os que ouviram o Anúncio, mas existe diferença entre os grupos que ouviram Música de Clássica e os que ouviram Música de Elevador.

Segunda Questão



```
##
## Call:
## lm(formula = Wt ~ Ht, data = dados2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -7.1166 -4.7744 -2.8412  0.5696 18.4581
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept) -36.8759    64.4728   -0.572    0.583
## Ht           0.5821     0.3892    1.496     0.173
##
## Residual standard error: 8.456 on 8 degrees of freedom
## Multiple R-squared:  0.2185, Adjusted R-squared:  0.1208
## F-statistic: 2.237 on 1 and 8 DF,  p-value: 0.1731
```