

# Apostila de estatística básica

## R Commander

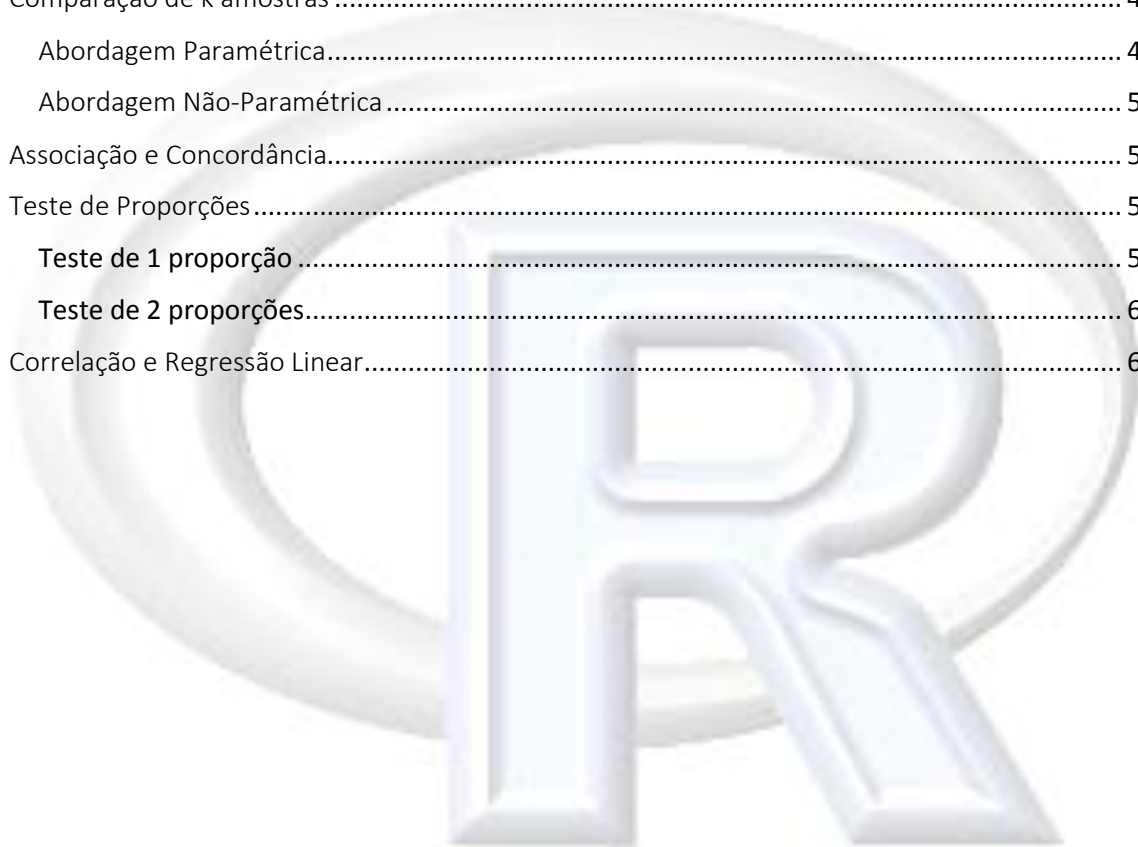
Organizador: Daniel Magalhães Lima

Autores:



## Sumário

|  |   |
|--|---|
| Importando dados .....   | 3 |
| Explorando dados – Tendência central, dispersão e gráficos ..... | 3 |
| Teste de normalidade.....  | 4 |
| Teste de hipóteses (1 amostra) .....                             | 4 |
| Comparação de duas amostras .....                                | 4 |
| Abordagem Paramétrica.....                                       | 4 |
| Abordagem Não-Paramétrica.....                                   | 4 |
| Comparação de k amostras .....                                   | 4 |
| Abordagem Paramétrica.....                                       | 4 |
| Abordagem Não-Paramétrica.....                                   | 5 |
| Associação e Concordância.....                                   | 5 |
| Teste de Proporções.....   | 5 |
| Teste de 1 proporção .....                                       | 5 |
| Teste de 2 proporções.....                                       | 6 |
| Correlação e Regressão Linear.....                               | 6 |



## Importando dados

- Acessar o Rcmdr através do R: comando `require("Rcmdr")` ou `library("Rcmdr")`
  - Clique em Dados -> Importar arquivos de dados -> from Excel file
  - Defina o nome do conjunto de dados
  - Encontre e selecione o arquivo que contenha os dados
  - Clique em "OK"
- Para visualizar os dados clique em "Ver conjunto de dados"
- Para mudar os nomes das variáveis: Dados -> Modificação de variáveis no conjunto de dados -> Renomear variáveis...
  - Após selecionar na variável desejada, mudar o nome para por exemplo nomes desejados.

## Explorando dados – Tendência central, dispersão e gráficos

- Clique em Estatísticas-> Resumos-> Conjunto de dados ativo
  - Na Saída ou "Output" aparece o "summary" (resumo) com os resultados dos principais valores das "Variable(s):" como média, mediana, ...
  - Clique em Estatísticas -> Resumos -> Resumos numéricos
    - No botão "Resuma por grupos", com a seleção de uma variável de agrupamento, há a possibilidade de se obter uma análise descritiva estratificada por grupos.
  - Selecione as variáveis de interesse e depois, na aba "Estatísticas", selecione as medidas que desejar (ex.: Média, Desvio padrão, Erro Padrão da média, Intervalo-Interquartil, etc), clique em "OK" para efetuar os cálculos.
  - Na Saída ou "Output" aparecerão os valores solicitados

A aba "Gráficos" contém uma série de gráficos fáceis de serem confeccionados, a seguir damos dois exemplos de gráficos a serem feitos.

- Clique em Gráficos -> Histograma
  - Selecione na aba "Dados" a variável ser analisada
    - No botão "Gráfico por grupos", com a seleção de uma variável de agrupamento, há a possibilidade de se obter histogramas para cada um dos grupos.
  - Na aba "Opções", há a possibilidade de alterar o título do gráfico, dos eixos e a escala do eixo.
  - Clique em "OK"
  - O gráfico é gerado no GUI (Interface Gráfica) do R
- Clique em Gráficos -> Boxplot
  - Selecione em "Dados" uma variável
  - Na aba "Opções", nomeie os rótulos dos eixos X e Y, e o nome do gráfico
  - Clique em "OK"
  - O gráfico é gerado no GUI (Interface Gráfica) do R

## Teste de normalidade

- Clique em Estatísticas -> Resumos -> Teste de normalidade de Shapiro-Wilk.
  - Selecione a variável a ser analisada e clique em OK
  - Avalie no "Output" o valor de P.

## Teste de hipóteses (1 amostra)

- Clique em Estatísticas -> Médias -> Teste t para uma amostra
  - Selecione a variável a ser analisada
  - Defina a hipótese a ser testada: média da população ser maior, menor ou diferente de  $\mu$ .
  - Defina  $\mu$  em "Hipótese nula: um =" e o "Nível de Confiança".

## Comparação de duas amostras

### Abordagem Paramétrica

- Clique em Estatísticas -> Médias -> Teste t para amostras independentes
  - Na aba "Dados" selecione a variável de interesse em "Variável resposta" e a variável de agrupamento em "Grupos".
  - Na aba "Opções" defina a hipótese alternativa e o nível de confiança.
  - Clique em "Ok" e avalie os resultados do teste em "Output"

### Abordagem Não-Paramétrica

- Clique em Estatísticas -> Testes Não-Paramétricos -> Teste de Wilcoxon (2 amostras)
  - Na aba "Dados" selecione a variável de interesse em "Variável resposta" e a variável de agrupamento em "Grupos".
  - Na aba "Opções" defina a hipótese alternativa e tipo de teste.
  - Clique em "Ok" e avalie os resultados do teste em "Output"

## Comparação de k amostras

### Abordagem Paramétrica

- Após verificar a normalidade devemos testar a igualdade de variâncias, para isto clique em Estatísticas -> Variâncias -> XXXX

- Selecione a variável a ser analisada em “Variável Resposta” e a variável de agrupamento em “Fatores”.
- Clique em “Ok” e avalie os resultados do teste em “Output”
- Clique em Estatística -> Médias -> ANOVA para um fator
  - Selecione a variável a ser analisada em “Variável Resposta” e a variável de agrupamento em “Fatores”.
  - Selecione a opção “Comparação de médias 1 a 1”.
  - Clique em “Ok” e avalie os resultados do teste em “Output”

## Abordagem Não-Paramétrica

- Clique em Estatísticas -> Testes Não-Paramétricos -> Teste de Kruskal-Wallis
  - Selecione a variável a ser analisada em “Variável Resposta” e a variável de agrupamento em “Fatores”.
  - Clique em “Ok” e avalie os resultados do teste em “Output”

## Associação e Concordância

- Clique em Estatísticas -> Tabelas de contingência -> Tabela de dupla entrada...
  - Selecione as variáveis que você deseja analisar.
  - Na aba “Estatísticas” selecione em “Computar percentagens” a opção “Percentual nas linhas”
  - Na seção “Teste de Hipótese” selecione o método a ser realizado
  - Clique em “Ok” e avalie os resultados do teste em “Output”
- Através do Rcmdr não é possível calcular o coeficiente de concordância Kappa, calculo que tem ser feito através do R
  - Digitar no R console ou através do R script: `install.packages("psych")`
  - Digitar `library(psych)`
  - Crie uma tabela através do comando abaixo substituindo os termos sublinhados pelos da sua tabela:
    - `objeto <- table(nome da planilha$variável 1, nome da planilha$variável 2)`
    - Aperte enter
  - Digite a função:
    - `wkappa(objeto)`
    - Aperte enter
  - Interprete o valor de Kappa

## Teste de Proporções

### Teste de 1 proporção

- Clique em Estatísticas -> Frequências/Proporções -> Teste de frequência/proporção (1 amostra)

- Na aba “Dados” selecione a variável a ser testada, na aba “Opções” defina a hipótese alternativa, a proporção hipotética da hipótese nula, o nível de confiança e o tipo de teste.
- Clique em “Ok” e avalie os resultados do teste em “Output”

### Teste de 2 proporções

- Clique em Estatísticas -> Frequências/Proporções -> Teste de frequência/proporção (2 amostras)
  - Selecione a variável a ser analisada em “Variável Resposta” e a variável de agrupamento em “Grupos”.
  - Na aba “Opções” defina a hipótese alternativa, o nível de confiança e o tipo de teste.
  - Clique em “Ok” e avalie os resultados do teste em “Output”

### Correlação e Regressão Linear

- Clique em Gráficos->Diagrama de dispersão
  - No campo "variável-x" selecione a variável independente e em "variável-y" selecione a variável dependente.
  - Clique em “Ok” e avalie o gráfico na interface do R
- Clique em Estatísticas ->Resumos ->Teste de correlação
  - Selecione as duas variáveis a serem analisadas.
  - Defina a hipótese e o teste de correlação a ser efetuado.
  - Clique em “Ok” e avalie os resultados do teste em “Output”

