



Bioestatística

Atividade no R

Prof. Leticia Raposo

2019

Doença Cardíaca

O banco de dados disponível no arquivo **ExercicioDoencacardiaca.txt** contém 303 observações e 9 variáveis. As descrições de cada variável são:

- **age**: idade em anos
- **sex**: sexo (1 = masculino; 0 = feminino)
- **cp**: tipo de dor no peito (1 = angina típica; 2 = angina atípica; 3 = sem dor angínica; 4 = assintomático)
- **trestbps**: pressão sanguínea em repouso (em mm/Hg ao ser admitido no hospital)
- **chol**: colesterol em mg/dl
- **fbs**: açúcar no sangue em jejum > 120 mg/dl (1 = V; 0 = F)
- **thalach**: frequência cardíaca máxima alcançada
- **exang**: exercício induziu angina (1 = sim; 0 = não)
- **num**: diagnóstico de doença cardíaca (0 = $< 50\%$ estreitamento de diâmetro; 1 = $> 50\%$ estreitamento de diâmetro)

Antes de iniciar a análise, verifique e responda:

- (0,3 ponto) As variáveis foram lidas (codificadas) corretamente pelo R? Se não, codifique-as corretamente.
- (0,2 ponto) Há dados ausentes?

1. Para as variáveis trestbps e chol, pede-se:

- (0,3 ponto) Calcule a média aritmética, a mediana e a moda.
- (0,3 ponto) Calcule o primeiro e o terceiro quartis e também o IQR.
- (0,8 ponto) Calcule as medidas de dispersão (amplitude, variância, desvio-padrão e coeficiente de variação). O que pode afirmar sobre os coeficiente de variação?
- (0,4 ponto) Verifique se a distribuição é simétrica, assimétrica positiva ou assimétrica negativa.
- (0,4 ponto) Construa o histograma e o boxplot para as variáveis em estudo.
- (0,3 ponto) É possível observar outliers para as duas variáveis analisadas?

2. Para a variável cp, pede-se:

- (1,0 ponto) Construa uma tabela de distribuição de frequências. Comente o resultado.
- (0,5 ponto) Elabore um gráfico de barras.

3. Para as variáveis fbs e num, pede-se:

- (1,0 ponto) Construa tabelas de contingência com os perfis linha e coluna.
- (0,5 ponto) Calcule o valor da estatística qui-quadrado. O que podemos interpretar com esse resultado?
- (0,5 ponto) Construa um gráfico com barras empilhadas.

4. Para as variáveis chol e thalach:

- (1,0 ponto) Construa um gráfico de dispersão e avalie se há indícios de relação entre as variáveis.
- (1,0 ponto) Calcule também o coeficiente de correlação de Pearson e interprete-o.



5. (1,5 pontos) Construa boxplots para a variável `trestbps`, segundo a variável `num`. Descreva o que é observado no gráfico.

DESAFIO! (Bônus de 1,0 ponto) Construa um gráfico não ensinado em sala de aula utilizando alguma(s) variável(is) do banco de dados. Explique o que o gráfico mostra.