

Métodos Não Paramétricos

Profa. Teresa Cristina

01/2025

Testes de associação:

1. Classe social e tipo de moradia

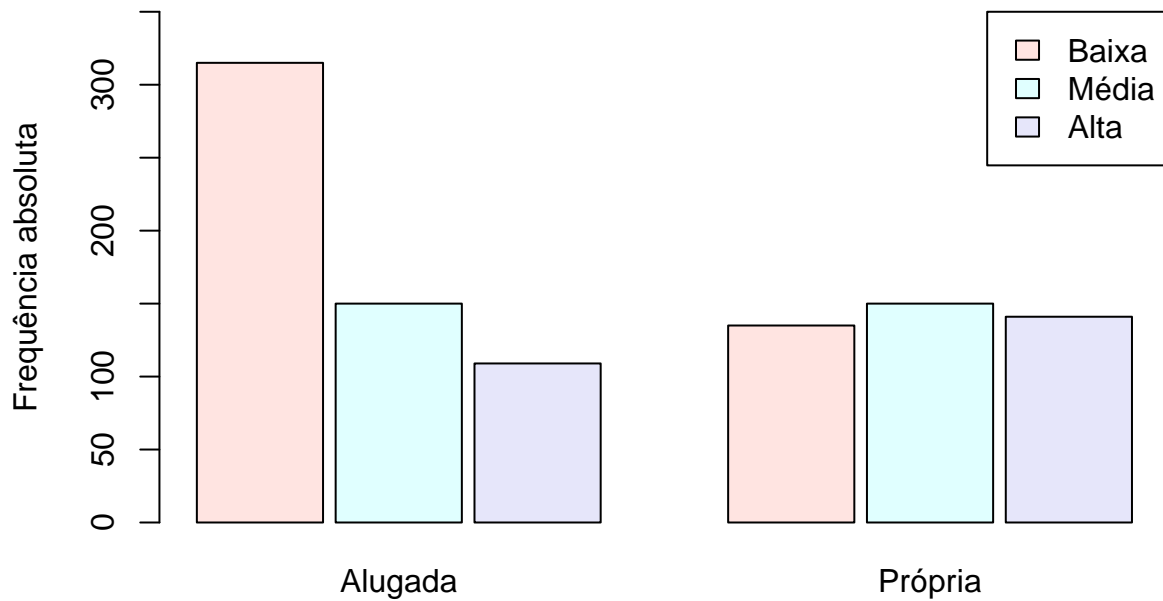
Em uma cidade foi realizada uma pesquisa a fim de verificar se existe associação entre a classe social e o tipo de moradia dos residentes. Na amostra, 1000 famílias responderam as perguntas: “Você reside em casa própria ou alugada? Qual sua classe social?” É possível afirmar que o tipo de moradia está associado à classe social?

```
c.social <- c("B", "M", "A")
moradia <- c("A", "P")
valores <- matrix(c(315,135, 150, 150, 109, 141), ncol = 2, byrow = T)
dimnames(valores) <- list(c("Baixa", "Média", "Alta"), c("Alugada",
                                                         "Própria"))
valores
```

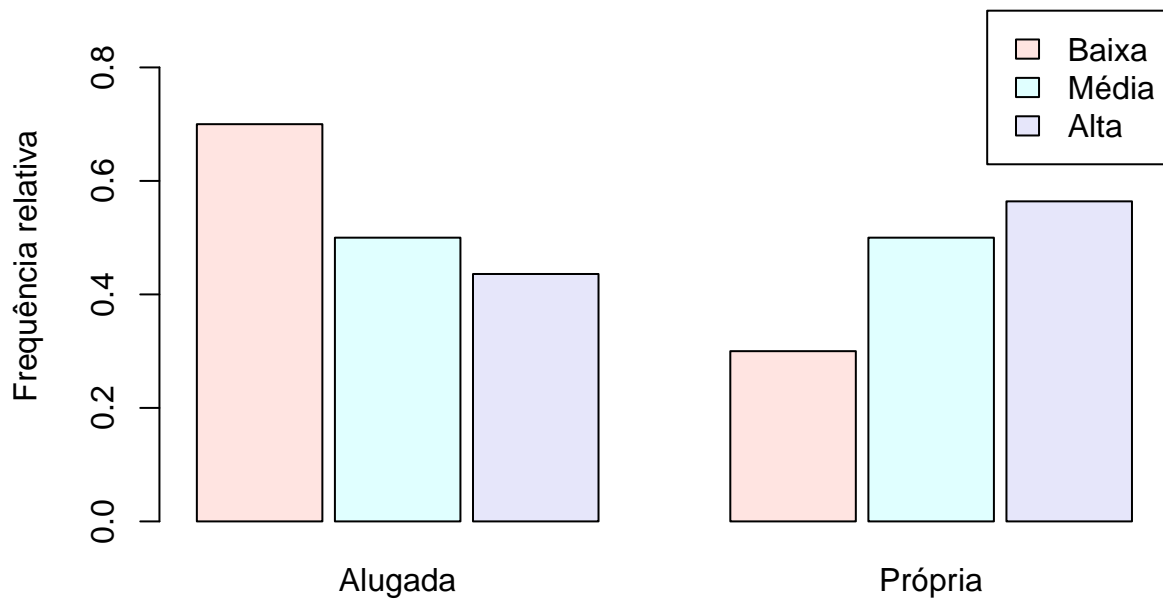
```
##           Alugada Própria
## Baixa      315      135
## Média      150      150
## Alta       109      141
```

- Observando o comportamento conjunto das variáveis:

```
barplot(valores, beside = T, space = c(.1, 1.01),
        col = c("mistyrose", "lightcyan", "lavender"),
        legend = rownames(valores),
        args.legend = list(x = "topright"),
        ylim = c(0, 350),
        ylab = "Frequência absoluta",
        main = "", font.main = 4)
```



```
barplot(prop.table(valores, margin = 1), beside = T, space = c(.1, 1.05),
        col = c("mistyrose", "lightcyan", "lavender"),
        legend = rownames(valores),
        args.legend = list(x = "topright"),
        ylim = c(0, 0.9),
        ylab = "Frequência relativa",
        main = "", font.main = 4)
```



- vamos obter o valor da estatística de Pearson:

```
n <- sum(valores)
n
```

```
## [1] 1000
```

```
valores      # tabela com nomes nas linhas e colunas
```

```
##      Alugada Própria
```

```
## Baixa      315      135
## Média      150      150
## Alta       109      141
```

```
margin.table(valores, margin = 1)
```

```
## Baixa Média Alta
##   450   300   250
```

```
margin.table(valores, margin = 2)
```

```
## Alugada Própria
##      574      426
```

```
prop.table(valores, margin = 1)
```

```
##      Alugada Própria
## Baixa  0.700  0.300
## Média  0.500  0.500
## Alta   0.436  0.564
```

```
teste.xsq <- chisq.test(valores) # estatística
teste.xsq
```

```
##
## Pearson's Chi-squared test
##
## data:  valores
## X-squared = 55.406, df = 2, p-value = 9.307e-13
```

```
teste.xsq$observed # valores observados
```

```
##      Alugada Própria
## Baixa  315      135
## Média  150      150
## Alta   109      141
```

```
teste.xsq$expected # valores esperados
```

```
##      Alugada Própria
## Baixa  258.3  191.7
## Média  172.2  127.8
## Alta   143.5  106.5
```

- os coeficientes de Cramér e, de McNemar-Siegel são dados por:

```
teste.xsq$statistic
```

```
## X-squared
##   55.4056
```

```
T_1 <- sqrt((teste.xsq$statistic)/(n*1))
T_1
```

```
## X-squared
##   0.235384
```

```
T_2 <- sqrt(teste.xsq$statistic/(n + teste.xsq$statistic))
T_2
```

```
## X-squared
```

```
## 0.2291222
```

2. Exercício

Existe associação entre consumo de energia e temperatura, considerando este conjunto de dados?

```
temperatura <- c("B", "M", "A")
consumo <- c("A", "B")
valores <- matrix(c(56, 24, 15, 25, 35, 20), ncol = 2, byrow = T)

dimnames(valores) <- list(c("Baixa", "Média", "Alta"), c("Alto", "Baixo"))
```

3. Para realizar um teste de correlação para variáveis quantitativas, use:

```
# cor(x, y, method = c("pearson", "kendall", "spearman"))
# cor.test(x, y, method=c("pearson", "kendall", "spearman"))
```

a) Verifique se existe relação entre a diversidade de gafanhotos e o número de anos após a aplicação de um pesticida.

```
anos <- c(0, 1, 3, 5, 9, 12, 13, 15, 21, 25)
gafanhotos <- c(0.00, 0.19, 0.15, 1.49, 1.10, 1.12, 1.61, 1.42, 1.48, 1.92)
cor.test(anos,gafanhotos, method="spearman",alternative="two.sided")
```

```
##
## Spearman's rank correlation rho
##
## data: anos and gafanhotos
## S = 32, p-value = 0.008236
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##      rho
## 0.8060606
```

ou ainda,

```
cor.test(anos,gafanhotos, method="kendall",alternative="two.sided")
```

```
##
## Kendall's rank correlation tau
##
## data: anos and gafanhotos
## T = 38, p-value = 0.004687
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##      tau
## 0.6888889
```

- Para obter a correlação de Spearman parcial:

```
y.data <- data.frame(
  hl=c(7,15,19,15,21,22,57,15,20,18),
  disp=c(0.000,0.964,0.000,0.000,0.921,0.000,0.000,1.006,0.000,1.011),
  deg=c(9,2,3,4,1,3,1,3,6,1),
  BC=c(1.78e-02,1.05e-06,1.37e-05,7.18e-03,0.00e+00,0.00e+00,0.00e+00
    ,4.48e-03,2.10e-06,0.00e+00))
```

Exemplo da biblioteca.

correlação parcial entre as variáveis "hl" e "disp", eliminando o efeito das

```
# variaveis "deg" and "BC"
```

```
pcor.test(y.data$hl,y.data$disp,y.data[,c("deg","BC")])
```

```
##      estimate    p.value statistic  n gp Method  
## 1 -0.6720863 0.06789202 -2.223267 10 2 pearson
```

```
pcor.test(y.data[,1],y.data[,2],y.data[,c(3:4)])
```

```
##      estimate    p.value statistic  n gp Method  
## 1 -0.6720863 0.06789202 -2.223267 10 2 pearson
```

```
pcor.test(y.data[,1],y.data[,2],y.data[,c(1:2)])
```

```
##      estimate    p.value statistic  n gp Method  
## 1 -0.6720863 0.06789202 -2.223267 10 2 pearson
```

O que podemos concluir?

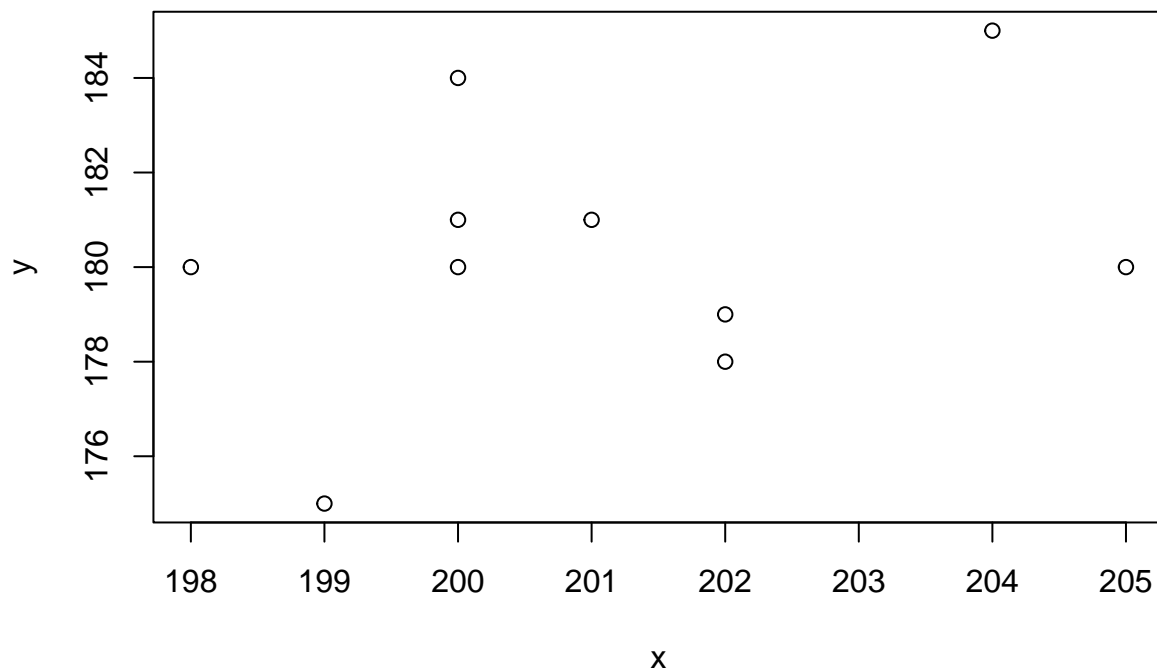
b) Considere os dados abaixo. Existe correlação parcial entre as variáveis?

```
x <- c(200, 204, 202, 205, 199, 200, 198, 200, 202, 201)
```

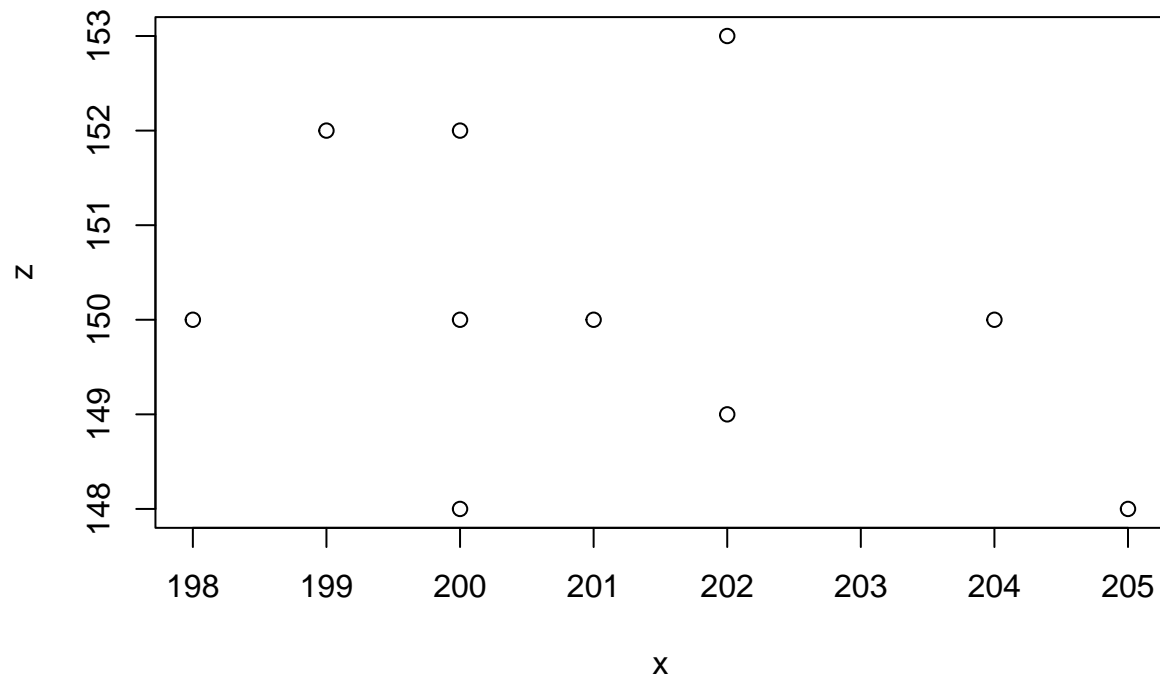
```
y <- c(180, 185, 179, 180, 175, 184, 180, 181, 178, 181)
```

```
z <- c(152, 150, 149, 148, 152, 150, 150, 148, 153, 150)
```

```
plot(x,y)
```



```
plot(x,z)
```



```
plot(y,z)
```

