

# Métodos Não Paramétricos

Profa. Teresa Cristina

01/2025

## Teste de Friedman

### Aplicações apresentadas na aula 08: teste de Friedman

#### 1. *Situações emocionais*

Em um estudo sobre hipnose, oito pessoas receberam uma tensão elétrica na superfície da pele (em milivolts), em quatro situações emocionais distintas: medo, alegria, tristeza e calma. Podemos afirmar que existe diferença na tensão entre os diferentes estados emocionais?

- dados: as oito das pessoas receberam o choque e as tensões, para cada emoção, foram registrados.
- análise descritiva:

```
sensacoes <- matrix(c(23.1, 57.6, 10.5, 23.6, 11.9, 54.6, 21.0, 20.3,
                      22.7, 53.2, 9.7, 19.6, 13.8, 47.1, 13.6, 23.6,
                      22.5, 53.7, 10.8, 21.1, 13.7, 39.2, 13.7, 16.3,
                      22.6, 53.1, 8.3, 21.6, 13.3, 37.0, 14.8, 14.8),
                    ncol = 4, byrow=FALSE, dimnames = list(1:8,
                    c("medo", "alegria", "tristeza", "calma")))
```

sensacoes

```
##   medo alegria tristeza calma
## 1 23.1    22.7    22.5  22.6
## 2 57.6    53.2    53.7  53.1
## 3 10.5     9.7    10.8   8.3
## 4 23.6    19.6    21.1  21.6
## 5 11.9    13.8    13.7  13.3
## 6 54.6    47.1    39.2  37.0
## 7 21.0    13.6    13.7  14.8
## 8 20.3    23.6    16.3  14.8
```

Obtemos as principais estatísticas descritivas:

```
summary(sensacoes)
```

```
##           medo           alegria           tristeza           calma
## Min.      :10.50   Min.       : 9.70   Min.      :10.80   Min.       : 8.30
## 1st Qu.:18.20   1st Qu.:13.75   1st Qu.:13.70   1st Qu.:14.43
## Median :22.05   Median :21.15   Median :18.70   Median :18.20
## Mean     :27.82   Mean     :25.41   Mean     :23.88   Mean     :23.19
## 3rd Qu.:31.35   3rd Qu.:29.48   3rd Qu.:26.68   3rd Qu.:26.20
## Max.     :57.60   Max.     :53.20   Max.     :53.70   Max.     :53.10
```

Para o cálculo do coeficiente de variação, por variedade, fazemos:

```
dp <- apply(sensacoes, 2, sd)
media <- apply(sensacoes, 2, mean)

cv <- round((dp/media) * 100, 2)
cv
```

```
##      medo  alegria tristeza   calma
##      65.15   63.24   62.71   64.14
```

% Ou ainda, para cada uma das sensações:

```
% cv_sensacoes_medo <- sd(sensacoes[, "medo"]) / mean(sensacoes[, % "medo"]) * 100 % cv_sensacoes_medo
```

O gráfico de pontos pode ser construído para observar o comportamento dos valores registrados. Para a construção do gráfico, considerando cada as diferentes sensações, rescrevemos as respostas da seguinte forma:

```
respostas <- matrix(c(23.1, 57.6, 10.5, 23.6, 11.9, 54.6, 21.0, 20.3,
                     22.7, 53.2, 9.7, 19.6, 13.8, 47.1, 13.6, 23.6,
                     22.5, 53.7, 10.8, 21.1, 13.7, 39.2, 13.7, 16.3,
                     22.6, 53.1, 8.3, 21.6, 13.3, 37.0, 14.8, 14.8),
                    ncol = 4, byrow = FALSE)
```

respostas

```
##      [,1] [,2] [,3] [,4]
## [1,] 23.1 22.7 22.5 22.6
## [2,] 57.6 53.2 53.7 53.1
## [3,] 10.5 9.7 10.8 8.3
## [4,] 23.6 19.6 21.1 21.6
## [5,] 11.9 13.8 13.7 13.3
## [6,] 54.6 47.1 39.2 37.0
## [7,] 21.0 13.6 13.7 14.8
## [8,] 20.3 23.6 16.3 14.8
```

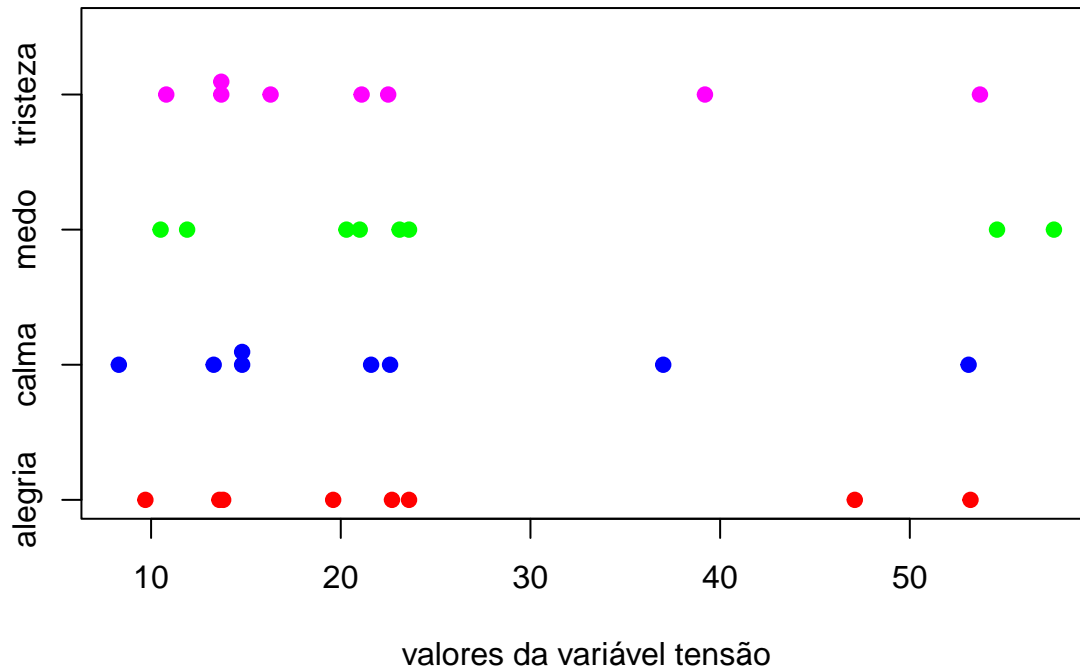
```
grupos <- matrix(c(rep("medo", 8), rep("alegria", 8),
                  rep("tristeza", 8), rep("calma", 8)), ncol = 4,
                  byrow = FALSE)
```

grupos

```
##      [,1] [,2] [,3] [,4]
## [1,] "medo" "alegria" "tristeza" "calma"
## [2,] "medo" "alegria" "tristeza" "calma"
## [3,] "medo" "alegria" "tristeza" "calma"
## [4,] "medo" "alegria" "tristeza" "calma"
## [5,] "medo" "alegria" "tristeza" "calma"
## [6,] "medo" "alegria" "tristeza" "calma"
## [7,] "medo" "alegria" "tristeza" "calma"
## [8,] "medo" "alegria" "tristeza" "calma"
```

Desta forma, o gráfico obtido é:

```
stripchart(respostas ~ grupos, col = c("red", "blue", "green", "magenta"),  
           pch = c(19, 19, 19, 19), method = "stack", xlab = "valores da variável tensão")
```



- aplicação do teste de Friedman:

```
friedman.test(sensacoes)
```

```
##  
## Friedman rank sum test  
##  
## data: sensacoes  
## Friedman chi-squared = 6.45, df = 3, p-value = 0.09166
```

Observando o resultado o gráfico e as medidas resumo, parece que não existe diferença entre os valores registrados. Pelo teste, dizemos que não existe evidência na amostra para rejeitar  $H_0$ , considerando p-valor = 0,09166.

## 2. Plantação de cana-de-açúcar

Em um estudo sobre variedades de cana-de-açúcar, foram registrados a produção (em toneladas por hectare), dos seis canteiros para as cinco variedades da cana. O que podemos concluir sobre as cinco variedades de cana?

- dados: nos seis canteiro foram plantadas as cinco variedades de cana e, a produção, para cada canteiro foi registrada.
- análise descritiva:

```
producao <- matrix(c(110.6, 119.5, 120.1, 105.3, 130.8, 138.1,
                    116.7, 128.4, 131.5, 114.8, 146.8, 155.5,
                    128.7, 140.2, 130.3, 138.7, 146.0, 149.8,
                    140.3, 150.0, 150.9, 144.7, 153.9, 156.0,
                    143.4, 153.8, 151.5, 144.1, 154.6, 159.3),
                  ncol = 5, byrow=FALSE, dimnames = list(1:6,
                  c("V.1", "V.2", "V.3", "V.4", "V.5")))
```

```
producao
```

```
##      V.1  V.2  V.3  V.4  V.5
## 1 110.6 116.7 128.7 140.3 143.4
## 2 119.5 128.4 140.2 150.0 153.8
## 3 120.1 131.5 130.3 150.9 151.5
## 4 105.3 114.8 138.7 144.7 144.1
## 5 130.8 146.8 146.0 153.9 154.6
## 6 138.1 155.5 149.8 156.0 159.3
```

```
summary(producao)
```

```
##      V.1      V.2      V.3      V.4
## Min.   :105.3 Min.   :114.8 Min.   :128.7 Min.   :140.3
## 1st Qu.:112.8 1st Qu.:119.6 1st Qu.:132.4 1st Qu.:146.0
## Median :119.8 Median :129.9 Median :139.4 Median :150.4
## Mean   :120.7 Mean   :132.3 Mean   :138.9 Mean   :149.3
## 3rd Qu.:128.1 3rd Qu.:143.0 3rd Qu.:144.6 3rd Qu.:153.2
## Max.   :138.1 Max.   :155.5 Max.   :149.8 Max.   :156.0
##      V.5
## Min.   :143.4
## 1st Qu.:145.9
## Median :152.7
## Mean   :151.1
## 3rd Qu.:154.4
## Max.   :159.3
```

Para o cálculo do coeficiente de variação, por variedade, fazemos:

```
dp <- apply(producao, 2, sd)
media <- apply(producao, 2, mean)

cv <- round((dp/media) * 100, 2)
cv
```

```
##      V.1  V.2  V.3  V.4  V.5
## 10.12 12.25 6.01 3.92 4.13
```

O gráfico de pontos pode ser construído para observar o comportamento dos valores registrados. Para a construção do gráfico, considerando cada tipo de variedade de cana-de-açúcar, rescrevemos as respostas da

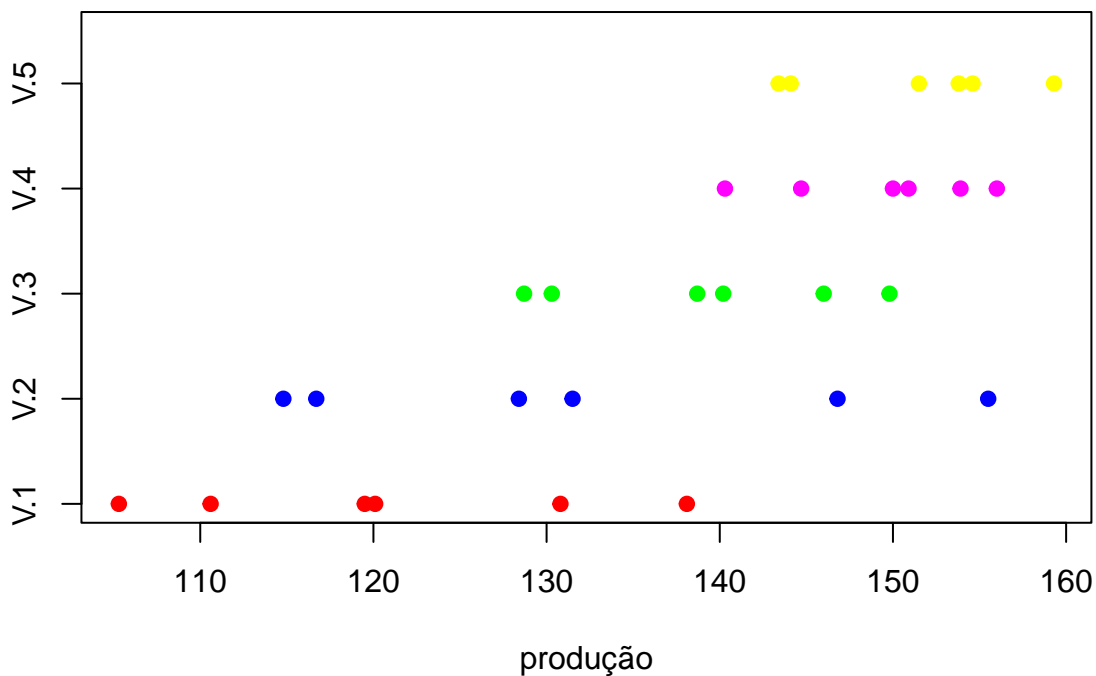
seguinte forma:

```
respostas <- matrix(c(110.6, 119.5, 120.1, 105.3, 130.8, 138.1,
                      116.7, 128.4, 131.5, 114.8, 146.8, 155.5,
                      128.7, 140.2, 130.3, 138.7, 146.0, 149.8,
                      140.3, 150.0, 150.9, 144.7, 153.9, 156.0,
                      143.4, 153.8, 151.5, 144.1, 154.6, 159.3),
                    ncol = 5, byrow=FALSE)

grupos <- matrix(c(rep("V.1", 6), rep("V.2", 6), rep("V.3", 6),
                    rep("V.4", 6), rep("V.5", 6)), ncol = 5,
                  byrow = FALSE)
```

Desta forma, o gráfico obtido é:

```
stripchart(respostas ~ grupos, col = c("red", "blue", "green", "magenta", "yellow"),
            pch = c(19, 19, 19, 19, 19), method = "stack", xlab = "produção")
```



- aplicação do teste de Friedman:

```
friedman.test(producao)
```

```
##
## Friedman rank sum test
##
## data:  producao
## Friedman chi-squared = 22.133, df = 4, p-value = 0.0001885
```

Observando o gráfico de pontos, notamos diferença na variabilidade e no intervalo que dos valores registrados e, observando também as medidas resumo, parece existir diferença entre as variedades (produção). Pelo teste, podemos dizer que existe evidência na amostra para rejeitar a hipótese  $H_0$ , isto seja, de que as variedades têm a mesma produção, considerando p-valor  $< 0,001$ .

*Comentário:* Vamos confirmar o valor da estatística, aplicando a fórmula (apresentada na aula 08: teste de Friedman). Neste exemplo temos:

```
respostas <- matrix(c(110.6, 119.5, 120.1, 105.3, 130.8, 138.1,
                      116.7, 128.4, 131.5, 114.8, 146.8, 155.5,
                      128.7, 140.2, 130.3, 138.7, 146.0, 149.8,
                      140.3, 150.0, 150.9, 144.7, 153.9, 156.0,
                      143.4, 153.8, 151.5, 144.1, 154.6, 159.3),
                    ncol = 5, byrow=FALSE)
```

- atribuindo postos para cada um dos blocos, obtemos

```
b.1 <- rank(respostas[1,]) # atribuicao de postos na linha 1
b.2 <- rank(respostas[2,])
b.3 <- rank(respostas[3,])
b.4 <- rank(respostas[4,])
b.5 <- rank(respostas[5,])
b.6 <- rank(respostas[6,])
todos.postos <- rbind(b.1, b.2, b.3, b.4, b.5, b.6)
todos.postos
```

```
##      [,1] [,2] [,3] [,4] [,5]
## b.1    1    2    3    4    5
## b.2    1    2    3    4    5
## b.3    1    3    2    4    5
## b.4    1    2    3    5    4
## b.5    1    3    2    4    5
## b.6    1    3    2    4    5
```

- somando os postos por tipo de variedade

```
postos <- cbind(sum(todos.postos[,1]), sum(todos.postos[,2]),
                sum(todos.postos[,3]), sum(todos.postos[,4]),
                sum(todos.postos[,5]))
```

```
postos
```

```
##      [,1] [,2] [,3] [,4] [,5]
## [1,]    6   15   15   25   29
```

```
soma.postos <- sum(postos)
soma.postos
```

```
## [1] 90
```

```
postos^2
```

```
##      [,1] [,2] [,3] [,4] [,5]
## [1,]   36  225  225  625  841
```

```
soma.postos.2 <- sum(postos^2)
soma.postos.2
```

```
## [1] 1952
```

- verificando o valor do teste:

```
teste.Friedman <- (12/(6*5*6))*soma.postos.2 - 3*6*6
teste.Friedman
```

```
## [1] 22.13333
```