

# Métodos Não Paramétricos

Profa. Teresa Cristina

01/2025

## Teste de Mann-Whitney

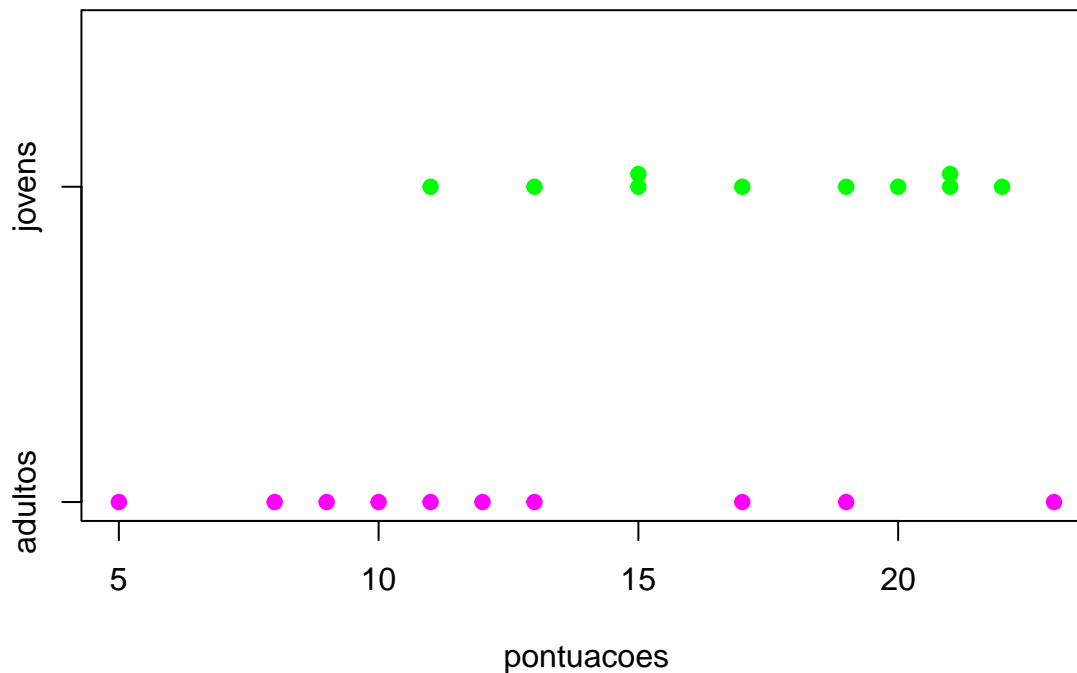
1. *Teste do relógio: pontuações para dois grupos: jovens (18 a 29 anos) e adultos (50 a 59 anos)*

```
pontuacoes <- matrix(c(11, 13, 15, 15, 17, 19, 20, 21, 21, 22,
                        8, 9, 10, 11, 12, 13, 5, 17, 19, 23),
                      ncol = 2, byrow=FALSE)

grupos <- matrix(c(rep("jovens", 10), rep("adultos", 10)), ncol = 2,
                 byrow = FALSE)
df <- data.frame(pontuacoes, grupos)
```

Pelo gráfico de pontos, podemos observar o comportamento das pontuações.

```
stripchart(pontuacoes ~ grupos, col = c("magenta", "green"),
           pch = c(19, 19), method = "stack", xlab = "pontuacoes")
```



Aplicando o teste de Mann-Whitney, obtemos:

```
jovens <- c(11, 13, 15, 15, 17, 19, 20, 21, 21, 22)
adultos <- c(8, 9, 10, 11, 12, 13, 5, 17, 19, 23)
```

```
wilcox.test(jovens, adultos, alternative = c("greater"),
            paired = FALSE, correct = FALSE)
```

```
## Warning in wilcox.test.default(jovens, adultos, alternative = c("greater"), :
## cannot compute exact p-value with ties
```

```
##
## Wilcoxon rank sum test
##
## data:  jovens and adultos
## W = 77, p-value = 0.0204
## alternative hypothesis: true location shift is greater than 0
```

Considerando o p-valor = 0,02234, podemos dizer que não existe evidências, nos dados, para aceitar  $H_0$  (ou seja, rejeitamos  $H_0$  com base nas evidências da amostra).

### Vamos conferir a estatística teste do exemplo

- obtendo os postos para a amostra combinada

```
pontuacoes <- matrix(c(11, 13, 15, 15, 17, 19, 20, 21, 21, 22,
                        8, 9, 10, 11, 12, 13, 5, 17, 19, 23),
                      ncol = 1, byrow=FALSE)

grupos <- matrix(c(rep("jovens", 10), rep("adultos", 10)), ncol = 1,
                 byrow = FALSE)
df <- data.frame(pontuacoes, grupos, rank(pontuacoes))
df
```

```
##      pontuacoes  grupos rank.pontuacoes.
## 1          11  jovens           5.5
## 2          13  jovens           8.5
## 3          15  jovens          10.5
## 4          15  jovens          10.5
## 5          17  jovens          12.5
## 6          19  jovens          14.5
## 7          20  jovens          16.0
## 8          21  jovens          17.5
## 9          21  jovens          17.5
## 10         22  jovens          19.0
## 11           8 adultos           2.0
## 12           9 adultos           3.0
## 13          10 adultos           4.0
## 14          11 adultos           5.5
## 15          12 adultos           7.0
## 16          13 adultos           8.5
## 17           5 adultos           1.0
## 18          17 adultos          12.5
## 19          19 adultos          14.5
## 20         23 adultos          20.0
```

```
attach(df)
```

```
## The following objects are masked _by_ .GlobalEnv:
##
##      grupos, pontuacoes
```

- ordenando os postos

```
df%>%arrange(pontuacoes)
```

```
##   pontuacoes grupos rank.pontuacoes.
## 1          5 adultos          1.0
## 2          8 adultos          2.0
## 3          9 adultos          3.0
## 4         10 adultos          4.0
## 5         11 jovens          5.5
## 6         11 adultos          5.5
## 7         12 adultos          7.0
## 8         13 jovens          8.5
## 9         13 adultos          8.5
## 10        15 jovens         10.5
## 11        15 jovens         10.5
## 12        17 jovens         12.5
## 13        17 adultos         12.5
## 14        19 jovens         14.5
## 15        19 adultos         14.5
## 16        20 jovens         16.0
## 17        21 jovens         17.5
## 18        21 jovens         17.5
## 19        22 jovens         19.0
## 20        23 adultos         20.0
```

- extraindo postos para o grupo jovens e obtendo a estatística teste.

```
posto.j <-df[df$grupos == "jovens",]
posto.j
```

```
##   pontuacoes grupos rank.pontuacoes.
## 1         11 jovens          5.5
## 2         13 jovens          8.5
## 3         15 jovens         10.5
## 4         15 jovens         10.5
## 5         17 jovens         12.5
## 6         19 jovens         14.5
## 7         20 jovens         16.0
## 8         21 jovens         17.5
## 9         21 jovens         17.5
## 10        22 jovens         19.0
```

```
R.j <- sum(posto.j$rank.pontuacoes)
T <- R.j - (length(jovens))*(length(jovens)+1)/2
T
```

```
## [1] 77
```

Observe que  $T$  é a estatística  $W$  obtida acima.

### Exercício: *Dieta de porcos*

(Gibbons, J. D.; Subhabrata C. Nonparametric statistical inference: revised and expanded. CRC press, 2014.)

Um dos problemas para pequenos fazendeiros que comprem porcos jovens para engordar e vender para abate é se há alguma diferença no ganho de peso para porcos machos e fêmeas, quando os dois gêneros são submetidos a tratamentos de alimentação idênticos. Se houver diferença, o fazendeiro pode otimizar a produção comprando apenas um gênero de porcos para engorda. Como um serviço público, uma estação experimental agrícola decidiu executar um experimento controlado para determinar se o gênero é um fator importante no ganho de peso. Eles colocaram oito porcos machos jovens em um curral e oito fêmeas jovens em outro curral e no curral, os tratamentos de alimentação eram idênticos por um período fixo de tempo. Os pesos iniciais estavam todos entre 35 e 50 lb (1 libra = 0,453592 Kg), e as quantidades de ganho de peso em libras para os dois gêneros são registradas abaixo. Infelizmente, uma das porcas morreu e, apenas sete observações naquele grupo. Os pesos registrados foram:

fêmeas: (9.31, 9.57; 10.21, 8.86, 8.52, 10.53, 9.21)

machos: (9.14, 9.98, 8.46, 8.93, 10.14, 10.17, 11.04, 9.43)

O que podemos concluir sobre a dieta dos porcos machos e fêmeas?