

# Exercícios da Lista 3

Aluno: Germano Andrade Brandão

19 Abril 2020

## Nota inicial

Para a resolução dos exercícios, foram utilizados os pacotes “ggplot2” (para construção de gráficos) e “ggpubr” (para alterar a organização os gráficos).

## Capítulo 3, Edição 2017: Ex 28 a 31

### Ex. 28

Para se estudar o desempenho de duas corretoras de ações, selecionou-se de cada uma delas, amostras aleatórias das ações negociadas. Para cada ação selecionada, computou-se a porcentagem de lucro apresentada durante um período fixado de tempo. Os dados estão a seguir:

```
corretora.A <- data.frame(periodo_1 = c(45, 62, 38, 55, 54, 65),
                          periodo_2 = c(60, 55, 48, 56, 59, 55),
                          periodo_3 = c(54, 70, 64, 55, 48, 60))
```

corretora.A

##	periodo_1	periodo_2	periodo_3
## 1	45	60	54
## 2	62	55	70
## 3	38	48	64
## 4	55	56	55
## 5	54	59	48
## 6	65	55	60

```
corretora.B <- data.frame(periodo_1 = c(57, 50, 59, 61, 57, 55, 59),
                          periodo_2 = c(55, 52, 55, 52, 57, 58, 51),
                          periodo_3 = c(58, 59, 56, 53, 50, 54, 56))
```

corretora.B

##	periodo_1	periodo_2	periodo_3
## 1	57	55	58
## 2	50	52	59
## 3	59	55	56
## 4	61	52	53
## 5	57	57	50
## 6	55	58	54
## 7	59	51	56

Que tipo de informação revelam esses dados? (Sugestão: Observar os quantis empíricos)

**Corretora A:** Percebemos, a partir de  $q_2 - x_1, x_n - q_2, q_2 - q_1$  e  $q_3 - q_2$ , que os dados das porcentagens das ações da corretora A não são totalmente distribuídos simetricamente.

```
Dados_Corret_A <-c(min(corretora.A),
  quantile(c(corretora.A$periodo_1,
    corretora.A$periodo_2,
    corretora.A$periodo_3),
    probs = c(0.25, 0.5, 0.75)),
  max(corretora.A))

Dados_Corret_A_M <- matrix(Dados_Corret_A,
  nrow = 1,
  ncol = 5,
  dimnames = list("A",
    c("x1", "q1", "q2", "q3", "xn")))
```

```
Dados_Corret_A_M
```

```
##   x1 q1 q2 q3 xn
## A 38 54 55 60 70
```

$q_2 - x_1 \cong x_n - q_2$

*#q2 - x1*

```
Dados_Corret_A_M[, "q2"] - Dados_Corret_A_M[, "x1"]
```

```
## [1] 17
```

*#xn - q2*

```
Dados_Corret_A_M[, "xn"] - Dados_Corret_A_M[, "q2"]
```

```
## [1] 15
```

$q_2 - q_1 \neq q_3 - q_2$

*#q2 - q1*

```
Dados_Corret_A_M[, "q2"] - Dados_Corret_A_M[, "q1"]
```

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

*#q3 - q2*

```
Dados_Corret_A_M[, "q3"] - Dados_Corret_A_M[, "q2"]
```

```
## [1] 5
```

**Corretora B:** Já na corretora B, percebemos, a partir de  $q_2 - x_1, x_n - q_2, q_2 - q_1$  e  $q_3 - q_2$ , que os dados das porcentagens das ações são distribuídos de uma forma mais homogênea.

```
Dados_Corret_B <-c(min(corretora.B),
  quantile(c(corretora.B$periodo_1,
    corretora.B$periodo_2,
    corretora.B$periodo_3),
    probs = c(0.25, 0.5, 0.75)),
  max(corretora.B))

Dados_Corret_B_M <- matrix(Dados_Corret_B,
  nrow = 1,
  ncol = 5,
  dimnames = list("A",
    c("x1", "q1", "q2", "q3", "xn")))
```

*#Esquema dos cinco números*

```
Dados_Corret_B_M
```

```
## x1 q1 q2 q3 xn
## A 50 53 56 58 61
```

$q_2 - x_1 \cong x_n - q_2$

```
#q2 - x1
Dados_Corret_B_M[, "q2"] - Dados_Corret_B_M[, "x1"]
```

```
## [1] 6
```

```
#xn - q2
Dados_Corret_B_M[, "xn"] - Dados_Corret_B_M[, "q2"]
```

```
## [1] 5
```

$q_2 - q_1 \cong q_3 - q_2$

```
#q2 - q1
Dados_Corret_B_M[, "q2"] - Dados_Corret_B_M[, "q1"]
```

```
## [1] 3
```

```
#q3 - q2
Dados_Corret_B_M[, "q3"] - Dados_Corret_B_M[, "q2"]
```

```
## [1] 2
```

## Ex. 29

Para verificar a homogeneidade das duas populações do problema anterior, um estatístico sugeriu que se usasse o quociente

$$F = \frac{\text{var}(X/A)}{\text{var}(X/B)}$$

sendo  $\text{var}(X/A)$  a variância de  $X$  para a corretora  $A$ ;  $X$  = porcentagem de lucro. No entanto, o profissional não disse qual decisão tomar baseado nesse valor. Que regra de decisão voce adotaria para dizer se são homogêneas ou não?

Faça sua análise aqui

```
# Resposta
```

- Caso  $\text{var}(X/A)$  seja *maior* que  $\text{var}(X/B)$ , o número  $F$  será maior que 1 e tão maior quanto a proporção da diferença entre as duas.
- Já se  $\text{var}(X/B)$  for *maior* que  $\text{var}(X/A)$ , o número  $F$  será menor que 1 e quanto maior a diferença entre as duas, mais perto de 0 será o quociente.
- E, por fim, caso as variâncias sejam homogêneas,  $F = 1$  ou  $F$  muito próximo a 1.

## Ex. 30

Faça um gráfico *boxplot* para cada uma das corretoras e compare os dois conjuntos de dados a partir dos gráficos.

Escreva sua resposta aqui

```
C_A <- ggplot(data = data.frame(c(corretora.A$periodo_1,
                                corretora.A$periodo_2,
                                corretora.A$periodo_3)),
              aes(y = c(corretora.A$periodo_1,
                        corretora.A$periodo_2,
                        corretora.A$periodo_3))) +
```

```

geom_boxplot(fill = "seagreen1",
             outlier.shape = 8,
             outlier.colour = "red",
             outlier.size = 2) +
scale_x_continuous(b = NULL) +
theme_classic() +
labs(title = "Box Plot Corretora A",
     y = "% das ações") +
scale_y_continuous(breaks = c(38, 54, 55, 60, 70))

```

```

C_B <- ggplot(data = data.frame(c(corretora.B$periodo_1,
                                corretora.B$periodo_2,
                                corretora.B$periodo_3)),
             aes(y = c(corretora.B$periodo_1,
                        corretora.B$periodo_2,
                        corretora.B$periodo_3))) +
geom_boxplot(fill = "hotpink1") +
scale_x_continuous(b = NULL) +
theme_classic() +
labs(title = "Box Plot Corretora B",
     y = "% das ações") +
scale_y_continuous(breaks = c(50, 53, 56, 58, 61))

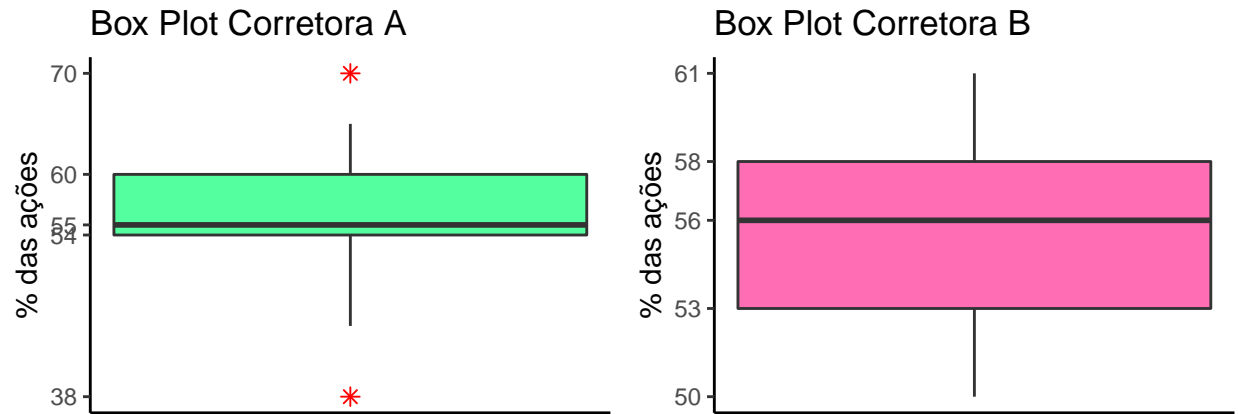
```

Podemos perceber novamente que os dados têm uma variação assimétrica na corretora A e, já na B, a distribuição em torno da mediana é mais normal.

```

ggarrange(C_A, C_B,
          ncol = 2, nrow = 2)

```



### Ex. 31

Para decidir se o desempenho das duas corretoras são semelhantes ou não, adotou-se o seguinte teste:

$$t = \frac{\bar{x}_A - \bar{x}_B}{S^2 * \sqrt{1/n_A + 1/n_B}}$$

$$S^2 = \frac{(n_A - 1)var(X/A) + (n_B - 1)var(X/B)}{n_A + n_B - 2}$$

Sendo  $n_A$  e  $n_B$  o número de ações selecionadas das corretoras  $A$  e  $B$  respectivamente.

Caso  $|t| < 2$ , os desempenhos são semelhantes, caso contrário, são diferentes. Qual seria a sua conclusão?

**Escreva sua resposta aqui**

- Como chegamos ao resultado de  $|t| = 0.1014008$ , concluímos de que o desempenho das duas corretoras são semelhantes.

*#Organizando os dados da corretora A em um único vetor:*

```
Corret_A <- c(corretora.A$periodo_1,
             corretora.A$periodo_2,
             corretora.A$periodo_3)
```

```
Corret_A
```

```
## [1] 45 62 38 55 54 65 60 55 48 56 59 55 54 70 64 55 48 60
```

```

#Organizando os dados da corretora B em um único vetor:
Corret_B <- c(corretora.B$periodo_1,
              corretora.B$periodo_2,
              corretora.B$periodo_3)
Corret_B

## [1] 57 50 59 61 57 55 59 55 52 55 52 57 58 51 58 59 56 53 50 54 56

n_A <- length(Corret_A)
n_B <- length(Corret_B)

media_A <- mean(Corret_A)
media_B <- mean(Corret_B)

var_A <- var(Corret_A)
var_B <- var(Corret_B)

#Cálculo do "S^2"
S_2 <- ((n_A - 1)*var_A + (n_B - 1)*var_B)/(n_A + n_B - 2)

#Cálculo do "t"
t <- (media_A - media_B)/(S_2*((1/n_A) - (1/n_B))^(1/2))

#módulo de "t" igual a:
abs(t)

## [1] 0.1014008

```