# Atividade 2 - Estatística Básica 1

Victor Lima

04/06/2023

# Códigos e conclusões da Atividade 2 de Estatísitca básica 1

### Carregando bibliotecas e dados de imoveiscwbav

```
library(carData)
library(datasets)
library(BSDA)
load("imoveiscwbav.RData")
imoveis <- imoveiscwbav
```

### a) Estimar o intervalo de confiança para a média da variável “price” com 95% de confiança

Aplicado Teste de Z

### Definindo as variáveis do desvio padrão e vetor de preços

```
sd <- sd(imoveis$price) # Desvio padrão do preço
x <- imoveis$price # vetor de preços
```

### Iniciando z-test

```
options(scipen = 999)
z.test(x, y = NULL, alternative = "two.sided", mu = 0, sigma.x = sd, sigma.y = NULL, conf.lev
el = 0.95)
```

Saída

```
One-sample z-Test

data:  x
z = 42.9, p-value < 0.00000000000000022
alternative hypothesis: true mean is not equal to 0

95 percent confidence interval:
909638.3 996735.0

sample estimates:
mean of x
953186.7
```

Conclusão

Assim, podemos definir a partir do Z-test que o intervalo de confiança é entre 909638,30 e 996735,00. E a média é de 953186,70.

## b) Fazer o teste de diferença entre médias para as variáveis “parea” e “tarea”.

Aplicado Teste de Z

### Definindo os desvios padrões de parea e tarea

```
sdx <- sd(imoveis$tarea) # Desvio padrão de parea
sdy <- sd(imoveis$parea) # Desvio padrão de tarea
```

### Definindo os vetores de parea e tarea

```
x <- imoveis$tarea # Vetor de parea
y <- imoveis$parea # vetor de tarea
```

### Iniciando z-test

```
z.test(x, y, alternative = "two.sided", mu = 0, sigma.x = sdx, sigma.y = sdy, conf.level = 0.95)
```

Saída

```
Two-sample z-Test

data:  x and y
z = 22.195, p-value < 0.00000000000000022
alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:
57.18056 68.25752

sample estimates:
mean of x mean of y
183.9224  121.2033
```

Conclusão

Assim, podemos definir a partir do Z-test que a diferença entre as medias está entre 57,18 m2 e 68,25 m2, para as médias de 183,92 m2 e 121,20 m2 (diferença = 62,7191 m2), com 95% de confiança.

## c) Fazer o teste de diferença entre variâncias para as variáveis “parea” e “tarea”.

Aplicado Teste de F

### Definindo os vetores de parea e tarea

```
x <- imoveis$tarea # Vetor de parea
y <- imoveis$parea # vetor de tarea
```

### Iniciando o teste de F

```
var.test(x, y, alternative = "two.sided", conf.level = 0.95)
```

Saída

```
F test to compare two variances

data:  x and y
F = 3.1296, num df = 540, denom df = 540, p-value < 0.00000000000000022
alternative hypothesis: true ratio of variances is not equal to 1

95 percent confidence interval:
2.643346 3.705343

sample estimates:
ratio of variances
3.129617
```

Conclusão

A razão da diferenca eh de 3.1296 e a estatistica F = 3.129617 com 540 graus de liberdade no numerador e denominador. Devemos confrontar esse valor da estatistica F com o valor da tabela F, conforme funcao abaixo:

```
qf(0.95, 540, 540) # 1.152221
```

Como a estatística F (3.1296) é superior ao valor tabelado (1.152221), consideramos que as variâncias não são estatisticamente iguais.

## d) Fazer o Teste de Wilcoxon-Mann-Whitney para amostras independentes para as variáveis “parea” e “tarea”.

Aplicado Teste de Wilcoxon-Mann-Whitney

### Definindo os vetores de parea e tarea

```
x <- imoveis$tarea # Vetor de parea
y <- imoveis$parea # vetor de tarea
```

### Aplicando o Teste de Wilcoxon-Mann-Whitney

```
options(scipen = 999)
wilcox.test(x, y, alternative = "two.sided")
```

Saída

```
Wilcoxon rank sum test with continuity correction

data:  x and y
W = 239897, p-value < 0.00000000000000022
alternative hypothesis: true location shift is not equal to 0
```

Conclusão

Como o resultado do p-value é menor que 0.05, as amostras são independentes (rejeitamos H0 de que as amostras sao identicas)