

## Exercício 2 - Metodologia de Pesquisa em Computação

Gabriel Luciano Gomes (265673)      Gabriel Oliveira Campos (265146)  
Paulo Junio Reis Rodrigues (265674)

October 16, 2020

### 1 Interseção dos intervalos de confiança

Para este exercício, o seguinte código foi utilizado para análise dos dados.

```
ex1 <- read.csv("D:/dbs/MO430/ex1.csv")

ex1$type <- sapply(ex1$type, as.character)
ex1[which(ex1$type == 'No'), "type"] <- 0
ex1[which(ex1$type == 'Yes'), "type"] <- 1
ex1$type <- sapply(ex1$type, as.numeric)

bpNo <- array()
bpYes <- array()
y <- 1
z <- 1

for(i in 1:length(ex1$type)){
  if(ex1$type[i] == 0){
    bpNo[y] <- ex1$bp[i]
    y <- y+1
  }else{
    bpYes[z] <- ex1$bp[i]
    z <- z+1
  }
}

# ----- CALCULO DO P VALOR ----- #
t.test(bpNo, bpYes)$p.value
wilcox.test(bpNo, bpYes)$p.value

# ----- T TESTE ----- #
t.test(bpNo)$conf.int
t.test(bpYes)$conf.int
```

```

# ----- WILCOXON RANK SUM ----- #

wilcox.test(bpNo, conf.int=T)$conf.int
wilcox.test(bpYes, conf.int=T)$conf.int


# ----- BOOTSTRAP ----- #
library(boot)
auxf <- function(dado, indice){
  return(mean(dado[indice]))
}

bbNo = boot(bpNo, R=5000, statistic=auxf)
boot.ci(bbNo, type="bca")

bbYes = boot(bpYes, R=5000, statistic=auxf)
boot.ci(bbYes, type="bca")

```

### 1.1 Há interseção ente os intervalos de confiança dos 2 conjuntos de dados? Isso está de acordo com o teste de diferença significativa entre eles?

Ao observar os resultados, é possível visualizar que o intervalo de confiança dos conjuntos não fazem interseção. Isso porque em nenhum dos três testes houve a interseção do limites dos conjuntos. Ainda assim, tem-se que o p-valor em ambos os testes (T-test e Wilcoxon rank sum) foram inferior a 0.05, fato que justifica a diferença significativa entre as bases e, por isso, não há interseção entre os limites.

```

> # ----- CALCULO DO P VALOR ----- #
> t.test(bpNo, bpYes)$p.value
[1] 0.003664884
> wilcox.test(bpNo, bpYes)$p.value
[1] 0.002294375
>
>
> # ----- T TESTE REVERSO ----- #
> t.test(bpNo)$conf.int
[1] 67.63705 71.45386
attr(,"conf.level")
[1] 0.95
> t.test(bpYes)$conf.int
[1] 71.78434 77.39213
attr(,"conf.level")
[1] 0.95
>
> # ----- WILCOXON RANK SUM REVERSO ----- #
>
> wilcox.test(bpNo, conf.int=T)$conf.int
[1] 67.50002 71.00003
attr(,"conf.level")
[1] 0.95
> wilcox.test(bpYes, conf.int=T)$conf.int
[1] 71.99996 77.00002
attr(,"conf.level")
[1] 0.95
>
>
> # ----- BOOTSTRAP ----- #
> library(boot)
> auxf <- function(dado, indice){
+   return(mean(dado[indice]))
+ }
>
> bbNo = boot(bpNo, R=5000, statistic=auxf)
> boot.ci(bbNo, type="bca")
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 5000 bootstrap replicates

CALL :
boot.ci(boot.out = bbNo, type = "bca")

Intervals :
Level      BCa
95%      (67.73, 71.44 )
calculations and Intervals on Original Scale
>
> bbyes = boot(bpYes, R=5000, statistic=auxf)
> boot.ci(bbyes, type="bca")
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 5000 bootstrap replicates

CALL :
boot.ci(boot.out = bbyes, type = "bca")

Intervals :
Level      BCa
95%      (71.74, 77.25 )
calculations and Intervals on Original Scale

```

Figure 1: Imagem mostra todos os resultados da questão 1

## 2 Tamanho de efeito

Para cálculo do Cohen D, a função **cohen.d** da biblioteca **effsize** foi utilizada, conforme o código a seguir. Além disso, o parâmetro **pooled=TRUE** foi utilizado para o uso do *pooled standard deviation* nos cálculos.

```
# ----- COHEN D ----- #  
library(effsize)  
cohen.d(bpYes, bpNo, pooled=TRUE)  
  
> # ----- COHEN D ----- #  
> library(effsize)  
> cohen.d(bpYes, bpNo, pooled=TRUE)  
  
Cohen's d  
  
d estimate: 0.4480343 (small)  
95 percent confidence interval:  
  lower      upper  
0.1503741 0.7456945
```

Figure 2: Imagem mostra todos os resultados da questão 2

### 2.1 Usando o intervalo de confiança do Cohen D, diga se a existe uma diferença significativa entre os dois conjuntos de dados.

Após obter o resultado do Cohen D, é possível visualizar que o intervalo não cruza o valor 0. Com isso, pode-se dizer que a diferença entre os conjuntos é significativa.