

Lista 01- Intervalo de Confiança e Teste de Hipoteses

Rogério O Souza

Friday, April 10, 2015

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

Disciplina: Analise de Dados

Professor: Washington S. Silva .

Lista de Exercicios - Tema: Intervalos de Confianca e Testes de Hipoteses

1. A autorização de despejo para uma industria requer que a concentração media mensal de COD seja inferior a 50 mg/L. A industria quer que isso seja interpretado como “50 mg/L esta dentro do intervalo de confianca da media, que vai ser estimada a partir de 20 observações por mes.” Para as 20 observações seguintes, estaria a industria em conformidade com esta interpretação do padrão?

Observações: 57 60 49 50 51 60 49 53 49 56 64 60 49 52 69 40 44 38 53 66

```
cod <- c(57, 60, 49, 50, 51, 60, 49, 53, 49, 56, 64, 60, 49, 52, 69, 40, 44, 38, 53, 66)
```

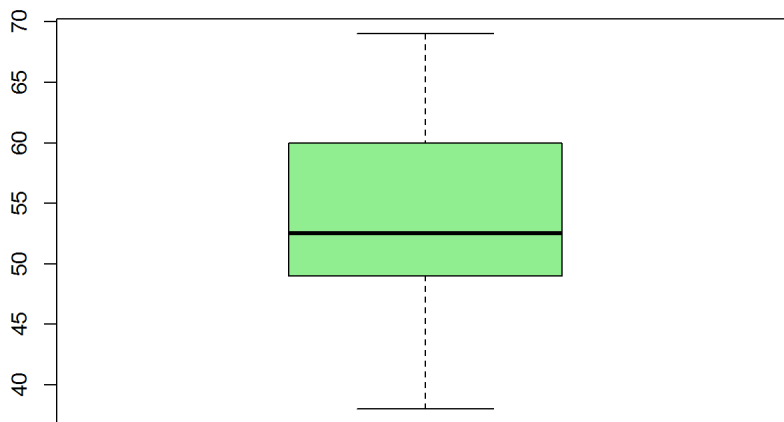
```
boxplot(cod, col="lightgreen", xlab="Observações COD") hist(cod)
```

```
cod <- c(57, 60, 49, 50, 51, 60, 49, 53, 49, 56, 64, 60, 49, 52, 69, 40, 44, 38, 53, 66)
```

```
summary(cod)
```

| ## | Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. |
|----|-------|---------|--------|-------|---------|-------|
| ## | 38.00 | 49.00 | 52.50 | 53.45 | 60.00 | 69.00 |

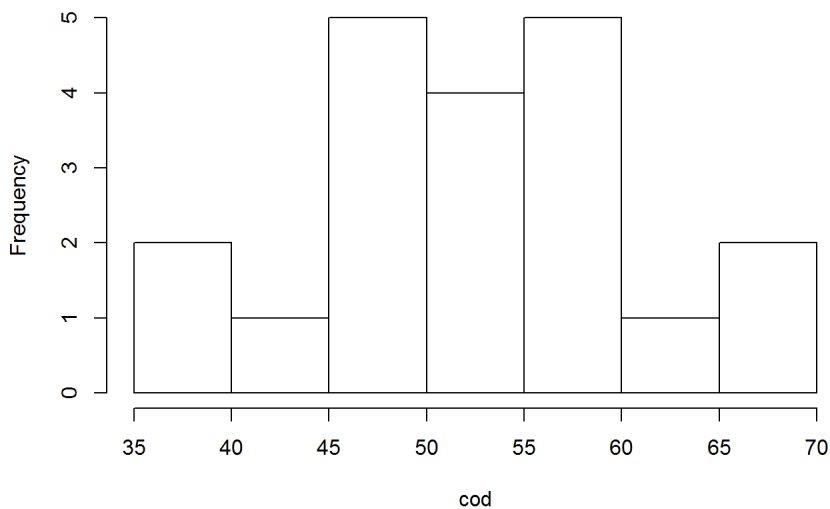
```
boxplot(cod, col="lightgreen", xlab="Observações COD")
```



Observações COD

```
hist(cod)
```

Histogram of cod



FUNÇÃO T- TESTE: Teste com $\alpha = 5\%$

```
t.test(cod, alternative=c("two.sided"),mu=50, conf.level = 0.95)
```

```
t.test(cod, alternative=c("two.sided"),mu=50, conf.level = 0.95)
```

```
##
## One Sample t-test
##
## data: cod
## t = 1.8894, df = 19, p-value = 0.07419
## alternative hypothesis: true mean is not equal to 50
## 95 percent confidence interval:
## 49.62825 57.27175
## sample estimates:
## mean of x
## 53.45
```

```
sort(cod)
```

```
sort(cod)
```

```
## [1] 38 40 44 49 49 49 49 50 51 52 53 53 56 57 60 60 60 64 66 69
```

Mediana

```
median(cod)
```

```
median(cod)
```

```
## [1] 52.5
```

WILCOX. TEST

```
wilcox.test(cod, mu=50, conf.int = TRUE)
```

```
wilcox.test(cod, mu=50, conf.int = TRUE)
```

```
##
## Wilcoxon signed rank test with continuity correction
##
## data: cod
## V = 139, p-value = 0.07905
## alternative hypothesis: true location is not equal to 50
## 95 percent confidence interval:
## 49.00001 58.00001
## sample estimates:
## (pseudo)median
## 53.66572
```

```
pwr.t.test(n=25, d=0.75, sig.level=.01, alternative="greater")
```

```
pwr.t.test(power=0.5, d=0.2, sig.level=.01, alternative="two.sided")
```

CONCLUSÃO

A hipótese nula deve ser rejeitada ao nível = 5%, é muito pouco provável observar uma concentração média menor que 50 mg/L baseados na análise das observações realizadas

O intervalo de confiança demonstrado pelo Wilcox.test, também indica que a análise das observações não são verdadeiras para o limite indicado de 50mg/L

Análise técnica da área ambiental: “há forte evidência de que as observações realizadas ultrapassem o limite de 50mg/L (valor de referência.)”

EXERCÍCIO 02

Um protocolo de garantia de qualidade laboratorial exige que as soluções padrão tenham 50 mg/l de COT, tais amostras são inseridas de forma aleatória no fluxo de trabalho. Os analistas são cegos a estas normas. Estime o viés e a precisão das 16

observações mais recentes deste padrão. O processo de medição de COT em conformidade com o padrão?

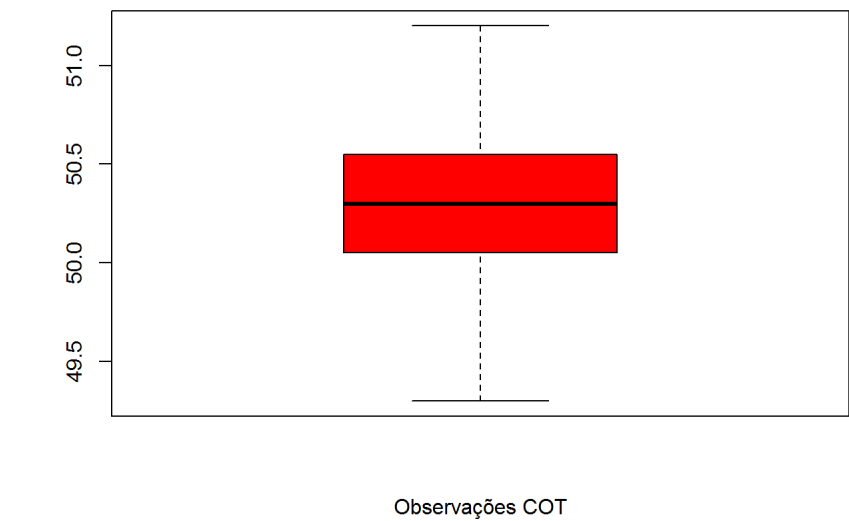
c(50.3, 51.2, 50.5, 50.2, 49.9, 50.2, 50.3, 50.5, 49.3, 50.0, 50.4, 50.1, 51.0, 49.8, 50.7, 50.6)

cot <- c(50.3, 51.2, 50.5, 50.2, 49.9, 50.2, 50.3, 50.5, 49.3, 50.0, 50.4, 50.1, 51.0, 49.8, 50.7, 50.6)

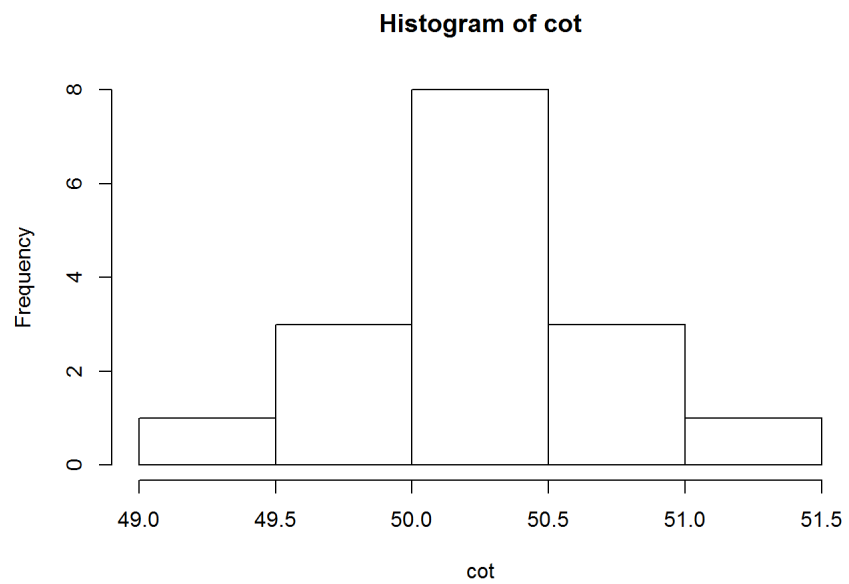
boxplot(cot, col="red",xlab="Observações COT") hist(cot)

```
cot <- c(50.3, 51.2, 50.5, 50.2, 49.9, 50.2, 50.3, 50.5, 49.3, 50.0, 50.4, 50.1, 51.0, 49.8, 50.7, 50.6)

boxplot(cot, col="red",xlab="Observações COT")
```



```
hist(cot)
```



FUNÇÃO T- TESTE: Teste com alpha = 5%

t.test(cot, alternative=c("two.sided"),mu=50, conf.level = 0.95)

```
t.test(cot, alternative=c("two.sided"),mu=50, conf.level = 0.95)
```

```
##
## One Sample t-test
##
## data: cot
## t = 2.7074, df = 15, p-value = 0.01622
## alternative hypothesis: true mean is not equal to 50
## 95 percent confidence interval:
##  50.06648 50.55852
## sample estimates:
## mean of x
##  50.3125
```

```
sort(cot)
```

```
sort(cot)
```

```
## [1] 49.3 49.8 49.9 50.0 50.1 50.2 50.2 50.3 50.3 50.4 50.5 50.5 50.6 50.7
## [15] 51.0 51.2
```

Mediana

```
median(cot)
```

```
median(cot)
```

```
## [1] 50.3
```

WILCOX. TEST

```
wilcox.test(cot, mu=50, conf.int = TRUE)
```

```
wilcox.test(cot, mu=50, conf.int = TRUE)
```

```
##
## Wilcoxon signed rank test with continuity correction
##
## data: cot
## V = 102, p-value = 0.01823
## alternative hypothesis: true location is not equal to 50
## 95 percent confidence interval:
##  50.09998 50.59994
## sample estimates:
## (pseudo)median
##      50.34994
```

```
pwr.t.test(n=25, d=0.75, sig.level=.01, alternative="greater")
```

```
pwr.t.test(power=0.5,d=0.2,sig.level=.01,alternative="two.sided")
```

CONCLUSÃO

A hipótese nula deve ser aceita ao nível = 5% de intervalo de confiança, é muito provável observar uma concentração média próxima de 50 mg/L COT baseados na análise das observações realizadas

O intervalo de confiança demonstrado pelo Wilcox.test, também indica que a análise das observações são verdadeiras para o limite indicado de 50mg/L

Análise técnica da área ambiental: “há forte evidência de que as observações realizadas estejam no limite de 50mg/L (valor de referência.)””

EXERCÍCIO 03

O gás produzido a partir da fermentação biológica é oferecido para a venda com o garantia de que o teor médio de metano de 72%. Uma amostra aleatória de $n = 7$ amostras de gás forneceu conteúdo de metano (em %) de 64, 65, 75, 67, 65, 74 e 75.

Efetue testes de hipóteses com níveis de significância de 0,10, 0,05, e 0,01 para determinar se é justo reivindicar uma média de 72

Calcular intervalos de confiança de 90%, 95% e 99% para avaliar a alegação de uma média de 72%.

```
tm <- c(64, 65, 75, 67, 65, 74, 75)
```

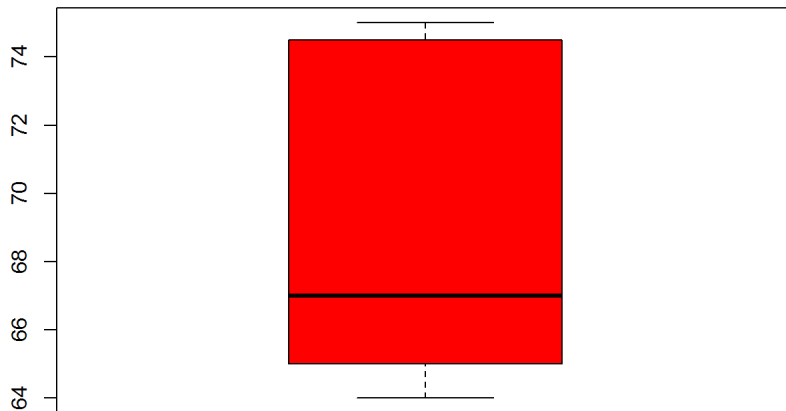
```
tm <- c(64, 65, 75, 67, 65, 74, 75)
```

```
# FUNÇÃO T-TESTE: Teste com alpha = 10%
```

```
t.test(tm, alternative=c("two.sided"),mu=72, conf.level = 0.90)
```

```
##
## One Sample t-test
##
## data:  tm
## t = -1.402, df = 6, p-value = 0.2105
## alternative hypothesis: true mean is not equal to 72
## 90 percent confidence interval:
##  65.52362 73.04781
## sample estimates:
## mean of x
##  69.28571
```

```
boxplot(tm, col="red",xlab="Observações Teor Metano")
```



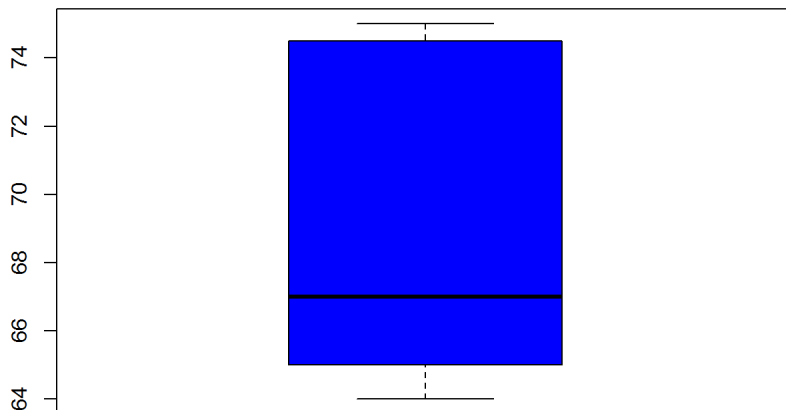
Observações Teor Metano

```
# FUNÇÃO T- TESTE: Teste com alpha = 5%
```

```
t.test(tm, alternative=c("two.sided"),mu=72, conf.level = 0.95)
```

```
##
## One Sample t-test
##
## data:  tm
## t = -1.402, df = 6, p-value = 0.2105
## alternative hypothesis: true mean is not equal to 72
## 95 percent confidence interval:
##  64.54836 74.02306
## sample estimates:
## mean of x
##  69.28571
```

```
boxplot(tm, col="blue",xlab="Observações Teor Metano")
```



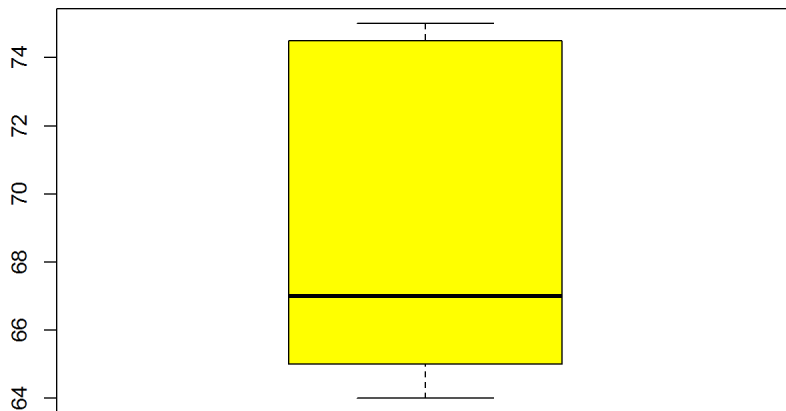
Observações Teor Metano

```
# FUNÇÃO T- TESTE: Teste com alpha = 1%
```

```
t.test(tm, alternative=c("two.sided"),mu=72, conf.level = 0.99)
```

```
##
## One Sample t-test
##
## data:  tm
## t = -1.402, df = 6, p-value = 0.2105
## alternative hypothesis: true mean is not equal to 72
## 99 percent confidence interval:
##  62.10794 76.46349
## sample estimates:
## mean of x
##  69.28571
```

```
boxplot(tm, col="yellow",xlab="Observações Teor Metano")
```



Observações Teor Metano

```
#WILCOX. TEST
```

```
wilcox.test(tm, mu=72, conf.int = TRUE)
```

```
##
## Wilcoxon signed rank test with continuity correction
##
## data:  tm
## V = 6, p-value = 0.2033
## alternative hypothesis: true location is not equal to 72
## 90 percent confidence interval:
##  64.99999 74.50002
## sample estimates:
## (pseudo)median
##  69.50002
```

CONCLUSÃO

A análise dos dados indica que a hipótese deve ser aceita de nível de Metano em 72%, com base na análise dos dados para os níveis de confiança, 10, 5 e 1%

O intervalo de confiança demonstrado pelo Wilcox.test, também indica que a análise das observações são verdadeiras para o limite indicado de 72% de Metano.

EXERCÍCIO 04

Os seguintes dados foram obtidos a partir de medidas pareadas de nitrito em água e em águas residuais por eletrodo direto de íon-seletivo e um método colorimétrico. Os dois métodos forneceram resultados consistentes?

Método Medidas de Nitrito

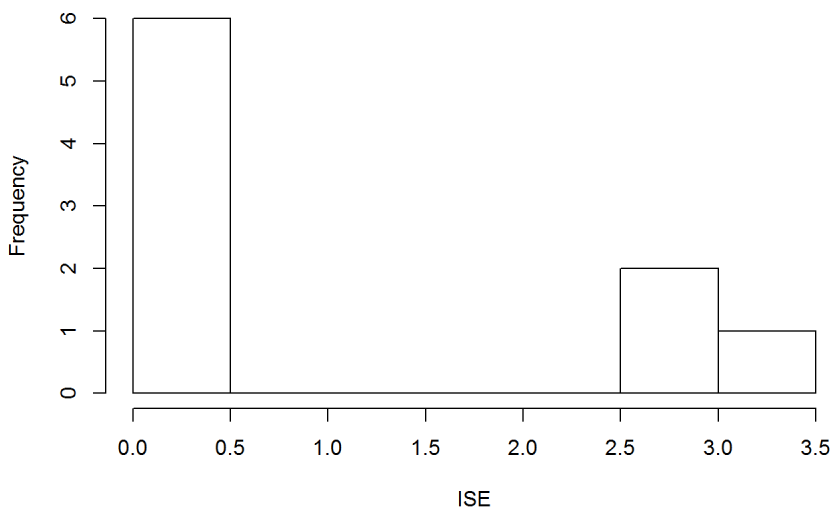
ISE 0.32 0.36 0.24 0.11 0.11 0.44 2.79 2.99 3.47

Colorimetric 0.36 0.37 0.21 0.09 0.11 0.42 2.77 2.91 3.52

Metodo <- c(ISE, Colorimetric) ISE <- c(0.32, 0.36, 0.24, 0.11, 0.11, 0.44, 2.79, 2.99, 3.47) Colorimetric <- c(0.36, 0.37, 0.21, 0.09, 0.11, 0.42, 2.77, 2.91, 3.52) Metodo
t.test(ISE,Colorimetric, var.equal = TRUE)

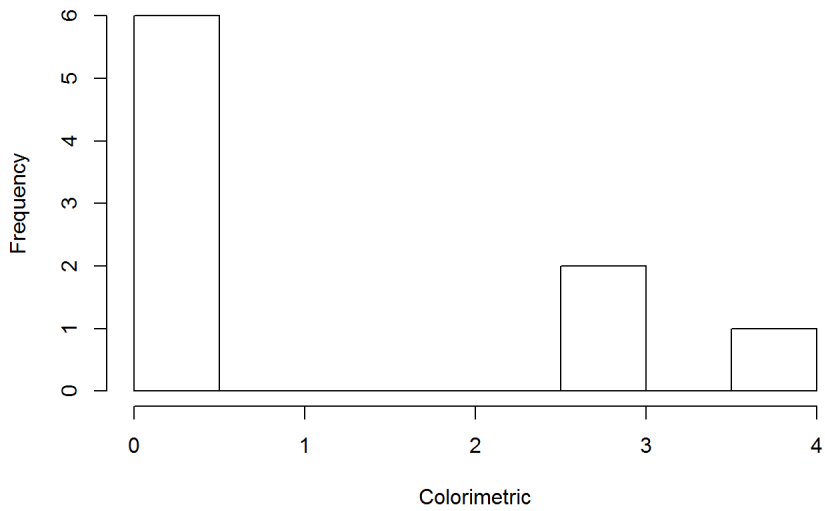
```
ISE <- c(0.32, 0.36, 0.24, 0.11, 0.11, 0.44, 2.79, 2.99, 3.47)
Colorimetric <- c(0.36, 0.37, 0.21, 0.09, 0.11, 0.42, 2.77, 2.91, 3.52)
hist(ISE)
```

Histogram of ISE

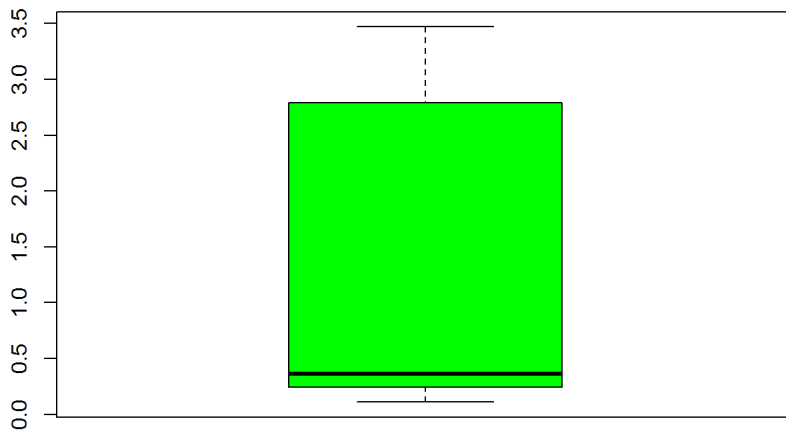


```
hist(Colorimetric)
```

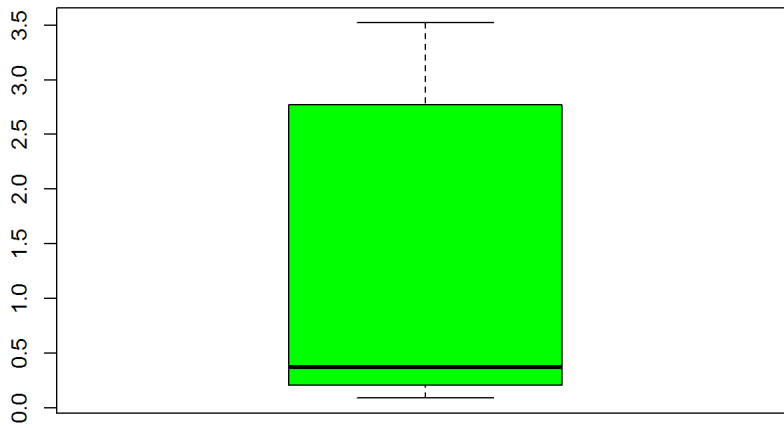
Histogram of Colorimetric



```
boxplot(ISE, col="green")
```



```
boxplot(Colorimetric, col="green")
```



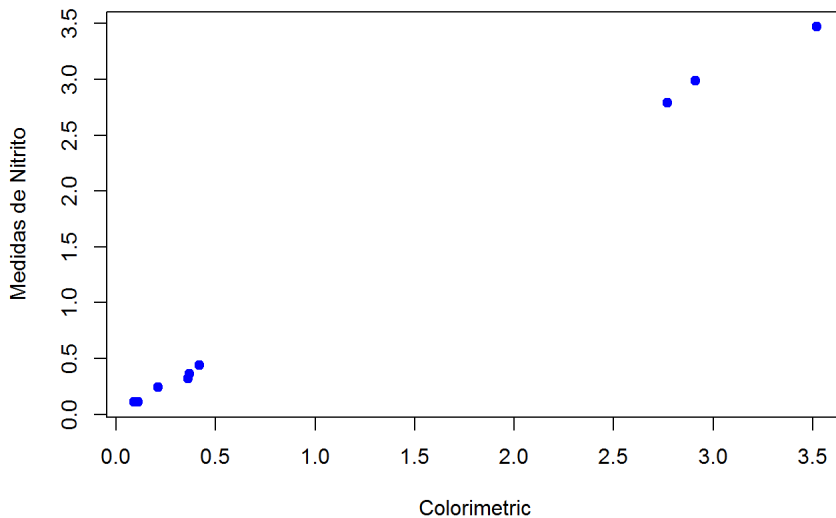
```
t.test(ISE,Colorimetric, var.equal = TRUE)
```

```
##
## Two Sample t-test
##
## data: ISE and Colorimetric
## t = 0.0116, df = 16, p-value = 0.9909
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -1.414623 1.430179
## sample estimates:
## mean of x mean of y
## 1.203333 1.195556
```

```
cor.test(ISE,Colorimetric, method="pearson")
```

```
##
## Pearson's product-moment correlation
##
## data: ISE and Colorimetric
## t = 96.8499, df = 7, p-value = 3.297e-12
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.9981537 0.9999247
## sample estimates:
## cor
## 0.9996271
```

```
plot(ISE~Colorimetric, ylab="Medidas de Nitrito", pch=19, col="blue")
```



CONCLUSÃO

A análise dos dados indica que a hipótese deve ser aceita de que os resultados foram consistentes a análise t.test para duas amostras, indicam intervalos de confiança e médias próximas entre as amostras

o teste de correlação de Pearson, também indica que a hipótese deve ser aceita, visto o intervalo de confiança entre 0.9981537 0.9999247;

EXERCÍCIO 05

Limites extremamente baixos existem agora para metais pesados em limites de águas residuais de efluentes. Pensa-se frequentemente que, sempre que a concentração de metais pesados é demasiado elevada, o problema pode ser corrigido forçando as indústrias a interromper a descarga da substância nociva. É possível, no entanto, que a concentração alvo do efluente ser tão baixa que elas podem ser ultrapassadas pela concentração em esgoto doméstico. Amostras de água potável foram coletadas de dois bairros residenciais, um servido pelo abastecimento de água da cidade e outro servido por poços particulares. As concentrações de mercúrio observados estão listados na Tabela abaixo. Para estudos futuros sobre concentrações de mercúrio em

áreas residenciais, seria conveniente ser capaz de coletar dados em qualquer bairro, sem ter que se preocupar que o abastecimento de água pode afetar o resultado. Existe alguma diferença no teor de mercúrio das duas áreas residenciais?

Concentração de Mercurio

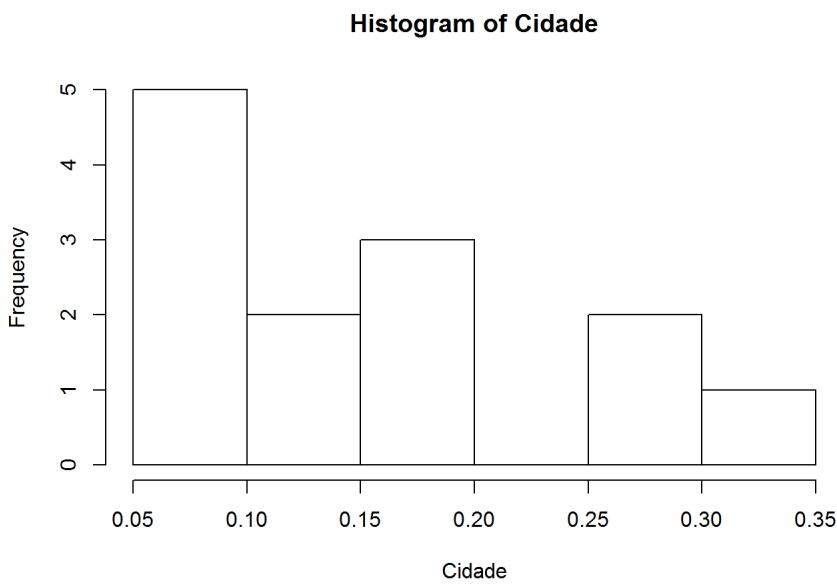
Cidade (n = 13)0.34 0.18 0.13 0.09 0.16 0.09 0.16 0.10 0.14 0.26 0.06 0.26 0.07

Part. (n=10) 0.26 0.06 0.16 0.19 0.32 0.16 0.08 0.05 0.10 0.13

Cidade <- c(0.34, 0.18, 0.13, 0.09, 0.16, 0.09, 0.16, 0.10, 0.14, 0.26, 0.06, 0.26, 0.07) Part <- c(0.26, 0.06, 0.16, 0.19, 0.32, 0.16, 0.08, 0.05, 0.10, 0.13)

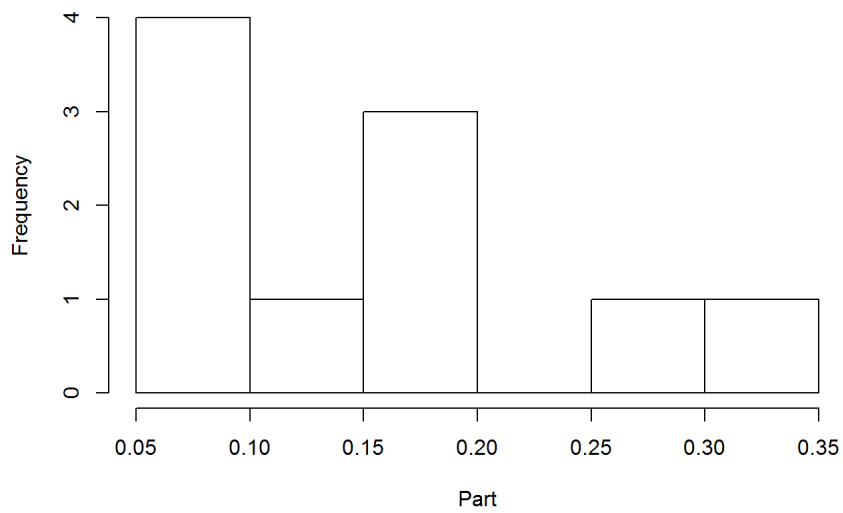
t.test(Cidade,Part, var.equal = TRUE)

```
Cidade <- c(0.34, 0.18, 0.13, 0.09, 0.16, 0.09, 0.16, 0.10, 0.14, 0.26, 0.06, 0.26, 0.07)
Part <- c(0.26, 0.06, 0.16, 0.19, 0.32, 0.16, 0.08, 0.05, 0.10, 0.13)
hist(Cidade)
```

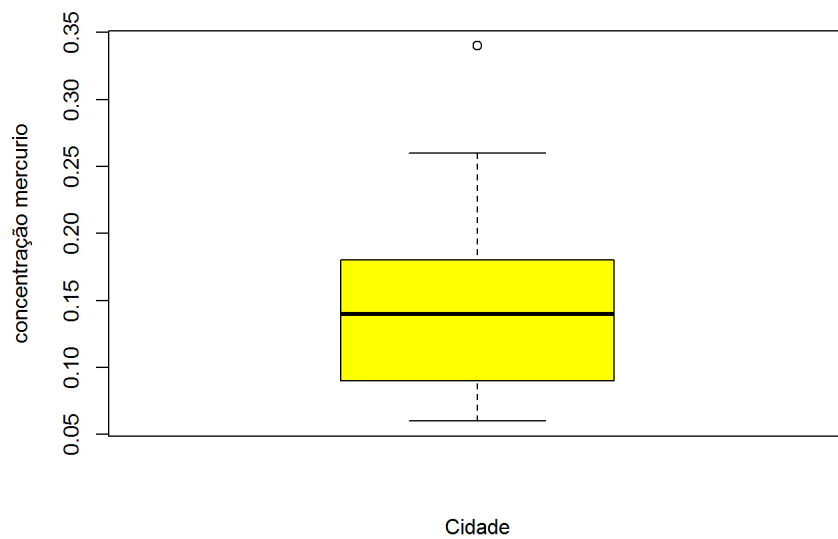


```
hist(Part)
```

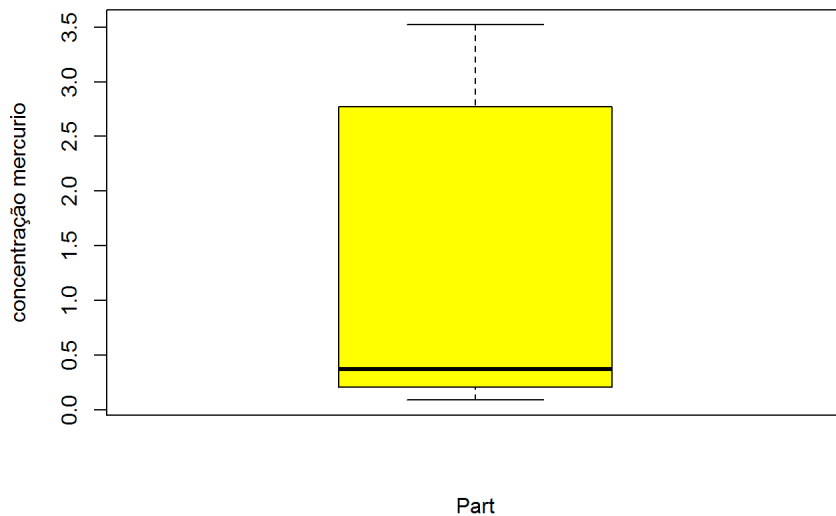
Histogram of Part



```
boxplot(Cidade, col="yellow", xlab= "Cidade",ylab="concentração mercurio")
```



```
boxplot(Colorimetric, col="yellow",xlab="Part", ylab="concentração mercurio")
```



```
t.test(Cidade,Part, var.equal = TRUE)
```

```
##
## Two Sample t-test
##
## data: Cidade and Part
## t = 0.1643, df = 21, p-value = 0.871
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.06902507 0.08087122
## sample estimates:
## mean of x mean of y
## 0.1569231 0.1510000
```

```
t.test(Cidade)
```

```
##
## One Sample t-test
##
## data: Cidade
## t = 6.7039, df = 12, p-value = 2.184e-05
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
## 0.1059217 0.2079245
## sample estimates:
## mean of x
## 0.1569231
```

```
t.test(Part)
```

```
##
## One Sample t-test
##
## data: Part
## t = 5.4658, df = 9, p-value = 0.0003974
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
## 0.08850457 0.21349543
## sample estimates:
## mean of x
## 0.151
```

CONCLUSÃO

A análise dos dados indica que a hipótese deve ser aceita de que há resultados diferentes entre as duas observações,

enquanto o Teste T para Cidades apresenta como resultados:

0.1059217 0.2079245

sample estimates:

mean of x

0.1569231

o Teste T para Part apresenta:

$t = 5.4658$, $df = 9$, $p\text{-value} = 0.0003974$

alternative hypothesis: true mean is not equal to 0

95 percent confidence interval:

0.08850457 0.21349543

sample estimates:

mean of x

0.151
