GUSTAVO PRIMOLAN DE CARA

PROBABILIDADE E ESTATÍSTICA

Questão 1:

Suponha que foi realizado um teste a cego objetivando-se avaliar a qualidade de um produto que será lançado no mercado, para isto, selecionamos 100 participantes e os dividimos de forma aleatória em dois grupos, primeiro grupo recebeu o produto padrão que esta disponível no mercado, por outro lado o segundo grupo recebeu o novo produto. Podemos afirmar com um nível 95% de significância que há diferença entre os produtos.

Primeiro grupo:

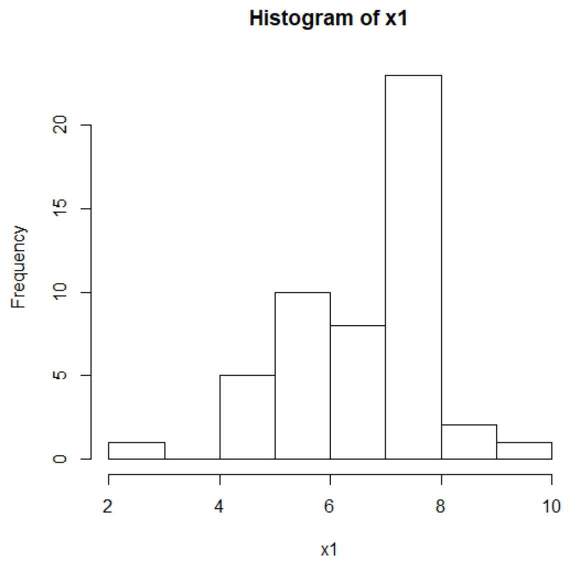
Avaliação do primeiro grupo:

(6.9,6.6,7.1,6.7,7.2,4.7,7.0,7.9,2.7,7.1,4.4,7.4,6.5,4.9,7.2,6.0,7.1,8.0,5.1,
5.9,9.4,7.6,5.4,7.8,5.1,8.0,7.1,8.1,7.8,5.3,8.2,7.7,7.6,8.0,5.9,6.5,5.1,6.3,7.6,7.1,7.8,7.6,4.
6,7.8,4.1,7.2,5.8,5.7,7.7,6.3)

```
summary(x1)
 Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
2.700   5.825   7.100   6.652   7.675   9.400
```

Média do grupo 1: 6.652

Mediana: 7.10



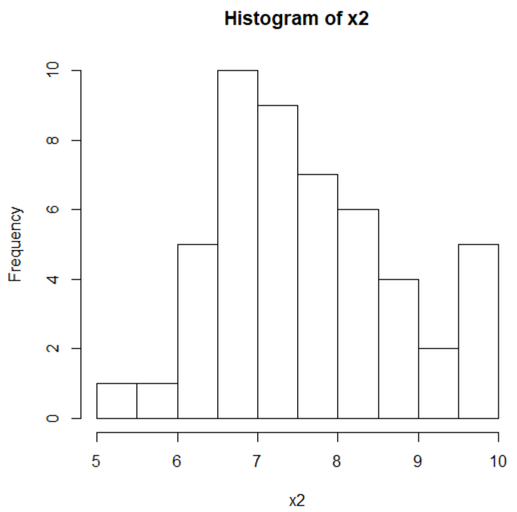
Possível notar, segundo o histograma, que a avaliação está com maior frequência é a 7.

Grupo 2:

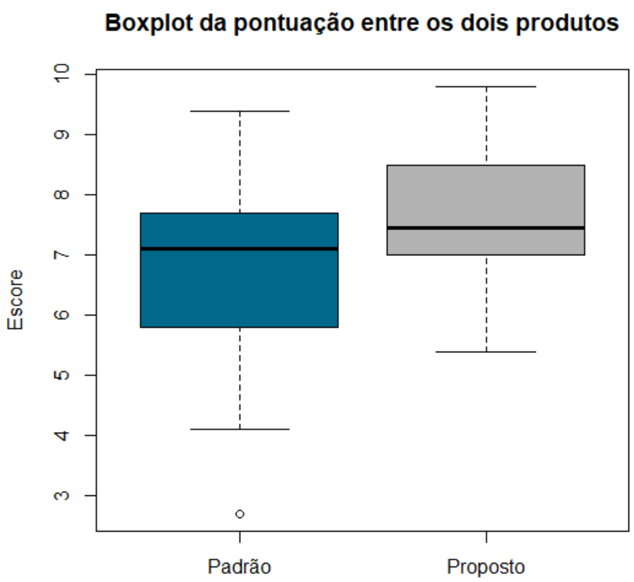
(8.8,7.0,9.2,9.8,6.9,7.1,8.9,7.0,9.4,9.6,6.1,6.1,7.4,7.6,8.5,9.8,7.9,8.0,7.4,8.1,6.7,6.3,8.6,7.6,5.7,9.6,7.1,7.1,7.2,6.5,8.5,5.4,8.5,7.0,7.0,6.8,7.3,8.3,7.5,7.0,7.6,9.6,6.6,8.0,6.4,8.2,7.4,7.6,8.8,7.0)

```
 Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 5.40    7.00    7.45    7.67    8.50    9.80
```

Possível notar que o segundo grupo possui a média de 7.67 e mediana de 7.45.



O grupo dois possui uma destruição mais proporcional em relação ao primeiro.



É possível notar que pelo gráfico boxplot mostrando os dois grupos, que o primeiro grupo possui um outlier, e o segundo grupo possui notas mais altas que o primeiro.

```
> shapiro.test(x1)

        Shapiro-Wilk normality test

data:  x1
W = 0.94335, p-value = 0.01832

>
> ##Teste de Normalidade para X2##
> shapiro.test(x2)

        Shapiro-Wilk normality test

data:  x2
W = 0.96671, p-value = 0.1696
```

Como o x1 não é normal, não é possível aplicar o teste.

Com o teste de Wilcoxon, é possível observar que, através do p-value < 0.05, não há evidências que os dois grupos são iguais, possibilitando a conclusão de que o grupo que recebeu os novos produtos, ficou mais satisfeito que o segundo grupo.

```
> wilcox.test(x1,x2)

        Wilcoxon rank sum test with continuity correction

data:  x1 and x2
W = 781, p-value = 0.001226
alternative hypothesis: true location shift is not equal to 0
```

Se os dados são normais:

Foram gerados dois grupos com distribuição normais.

```
> y1<-rnorm(100,5,1)

> y2<-rnorm(100,7,1)
```

Após isso, foi executado o teste de Shapirol-Wilk para observarmos a normalidade dos grupos.

```
> shapiro.test(y1)

        Shapiro-Wilk normality test

data:  y1
W = 0.98505, p-value = 0.3202
```

```
> shapiro.test(y2)

        Shapiro-Wilk normality test

data:  y2
W = 0.98996, p-value = 0.6612
```

Após concluirmos que os dois grupos possuem uma distribuição normal, executamos o teste de variância. Como o p-value > 0.05 significa que eles possuem variâncias iguais, aplicamos o teste t, passando o parâmetro var.equal = TRUE

```
> var.test(y1,y2)

        F test to compare two variances

data:  y1 and y2
F = 0.99083, num df = 99, denom df = 99, p-value = 0.9635
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.6666698 1.4726009
sample estimates:
ratio of variances
         0.9908272
```

Teste t com o parâmetro indicando a variância verdadeira.

```
> t.test(y1,y2, var.equal = TRUE)

        Two Sample t-test

data:  y1 and y2
t = -14.645, df = 198, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -2.387371 -1.820733
sample estimates:
mean of x mean of y
 5.021192  7.125244
```

Suponha que uma empresa esteja interessada em melhorar seu nível de satifação com seus clientes. Alguns procedimentos foram definidos para serem implementados objetivando-se melhorar tal experiência para isto, um grupo de 40 candidatos foram selacionados e apresentaram uma avaliação de satisfação antes e após a alteração dos procedimentos da empresa. Aplique um teste de hipotese para confirmar ou rejeitar tal hipótese

Grupo 1:

```
> w1<-c(5.7,4.4,4.6,5.3,6.8,5.2,6.1,6.6,6.9,6.2,6.4,6.0,3.6,5.3,4.0,4.8,5.4,3.8$
+ 6.2,3.9,5.9,5.5,3.7,6.6,6.1,7.7,4.6,3.9,5.8,6.7,5.8,6.4,5.6,7.4,7.6,5.8,6.1,
+ 5.5,6.0)
```

Grupo 2:

```
> w2<-c(7.0,5.8,6.6,6.7,8.6,6.2,7.7,7.7,7.5,7.8,8.0,7.0,5.1,5.9,5.7,6.8,5.6,5.3,5.7,
+ 8.0,5.4,8.0,6.6,5.6,7.9,7.0,9.0,5.7,5.2,6.7,7.7,7.4,8.4,6.4,9.1,9.0,7.2,7.5,
+ 7.3,8.0)
```

Executado o Shapiro teste, para verificarmos a normalidade dos grupos

```
> shapiro.test(w1)

        Shapiro-Wilk normality test

data:  w1
W = 0.95873, p-value = 0.1515
```

Normalidade do grupo 2.

```
> shapiro.test(w2)

        Shapiro-Wilk normality test

data:  w2
W = 0.95943, p-value = 0.1601
```

Como os dois eram normais, foi executado o teste de variância dos grupos

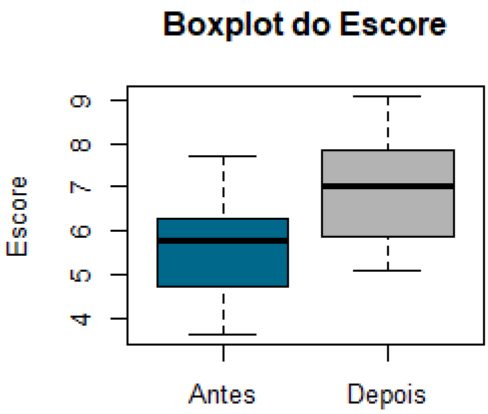
```
> var.test(w1,w2)

        F test to compare two variances

data:  w1 and w2
F = 0.94794, num df = 39, denom df = 39, p-value = 0.8683
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.5013642 1.7922861
sample estimates:
ratio of variances
         0.9479389
```

Após identificarmos o teste de variância comprovar que os dados possuem variâncias iguais. É executado o t.test

```
> t.test(w1,w2,paired = TRUE, var.equal = TRUE)

        Paired t-test

data:  w1 and w2
t = -20.169, df = 39, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -1.537653 -1.257347
sample estimates:
mean of the differences
                -1.3975
```

Bloxplot dos dois grupos.

```
> boxplot(w1,w2, ylab="Escore", range = 1.5,col=c("deepskyblue4","gray70"),
+ main="Boxplot do Escore",names=c("Antes", "Depois"), horizontal=FALSE)
```



Intervalos de Confiança

Função para calcular o intervalo de confiança

```
> ICmedia<-function(x,alpha,desvio){
+ n<-length(x)
+ z<-qnorm(alpha/2, lower.tail = FALSE)
+ if(missing(desvio)){
+ desvio<-sd(x)
+ z<-qt(1-alpha/2, n-1)
+ }
+ LI<-mean(x)-z*desvio/sqrt(n)
+ LS<-mean(x)+z*desvio/sqrt(n)
+ return(c(LI,LS))
+ }
```

Função para calcular o intervalo de confiança para proporção

```
> ICprop<-function(y,n,alpha){
+ p<-y/n
+ z<-qnorm(alpha/2, lower.tail = FALSE)
+ LI<-max(p-z*sqrt((p*(1-p))/n),0)
+ LS<-min(p+z*sqrt((p*(1-p))/n),1)
+ return(c(LI,LS))
+ }
```

Considere a situação em que o interesse é obter o intervalo de confiança ao nível de significância de 5 para a proporção de alunos da instituição que utilizam produtos da Apple. Tendo como resultado da amostragem que 9 dos 25 alunos responderam afirmativamente. obtenha o intervalo de confiança para proporção

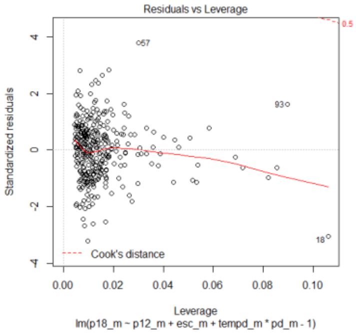
```
> ICprop(9,25,0.05)
[1] 0.1718435 0.5481565
```

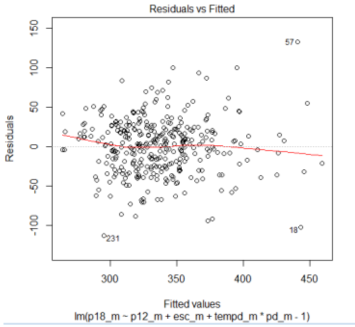
```
> cor(p18_m,p12_m)
[1] 0.6234603
```

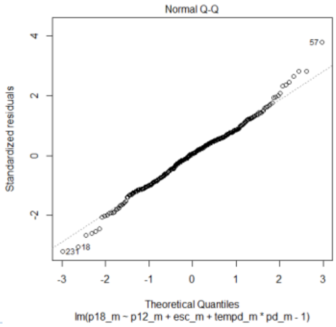
```
> cor.test(p18_m,p12_m)

        Pearson's product-moment correlation

data:  p18_m and p12_m
t = 14.66, df = 338, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.5538236 0.6844358
sample estimates:
      cor
0.6234603
```

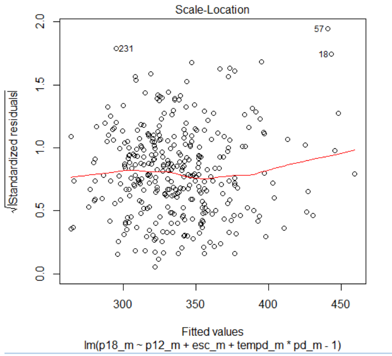
Regressão Linear

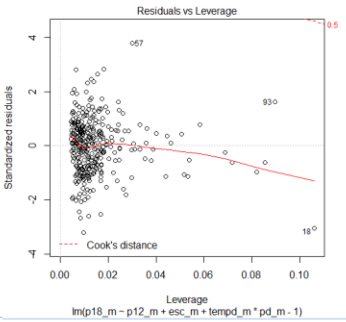
```
> ajuste_m2<-lm(p18_m ~  p12_m + esc_m +tempd_m * pd_m - 1)
> summary (ajuste_m2)

Call:
lm(formula = p18_m ~ p12_m + esc_m + tempd_m * pd_m - 1)

Residuals:
    Min      1Q  Median      3Q     Max 
-113.05  -24.31    2.28   21.02  132.36 

Coefficients:
               Estimate Std. Error t value Pr(>|t|)    
p12_m         5.678e-01  5.639e-02  10.069  < 2e-16 ***
esc_m        -1.252e+01  3.232e+00  -3.873 0.000129 ***
tempd_m       7.841e-01  8.732e-02   8.980  < 2e-16 ***
pd_m          1.686e+00  1.439e-01  11.716  < 2e-16 ***
tempd_m:pd_m -5.275e-03  4.431e-04 -11.904  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 35.54 on 335 degrees of freedom
Multiple R-squared:  0.9892,    Adjusted R-squared:  0.9891 
F-statistic:  6158 on 5 and 335 DF,  p-value: < 2.2e-16
```












```
> shapiro.test(ajuste_m2$residuals)

        Shapiro-Wilk normality test

data:  ajuste_m2$residuals
W = 0.99176, p-value = 0.0557
```

```
> AIC(ajuste_m2)
[1] 3399.876
```