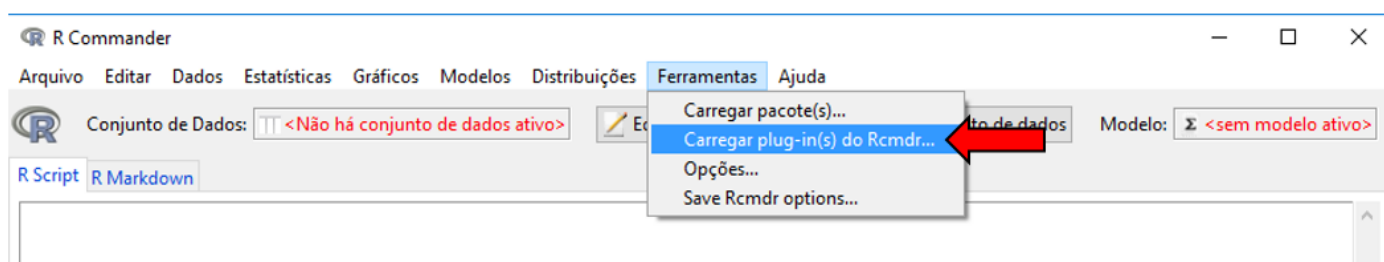


#-----
#Disciplina - Estatística
#Professores: Raquel de Vasconcellos C. de Oliveira
Julio Lima
Colaboração na elaboração: Fabiano Marcos de Lima e Luan Nóe da Silva
#-----

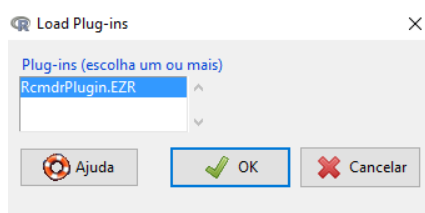
R Commander com RcmdrPuglin.EZR

Após carregar o ambiente do R Commander, vamos carregar o “RcmdrPlugin.EZR” no seguinte caminho:

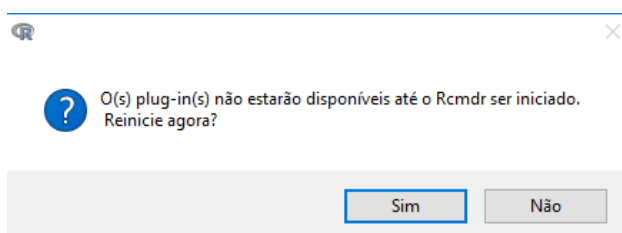
Ferramentas>Carregar plug-in(s) do Rcmdr..



Escolha o plugin desejado, no caso RcmdrPlugin.EZR e clique em OK



Aparecerá uma mensagem dizendo que o R será reiniciado clique em “Sim”



Então o R Commander irá reinicializar e abrir automaticamente.

Inferência Estatística

Intervalo de Confiança para a média

Exercício 1: Suponha que estejamos interessados no intervalo de confiança para a variável age (idade da mãe) no banco “birthwt” a um nível de confiança de 95%. Interprete.

Opção 1: Automaticamente pela biblioteca “epiDisplay”

Carregar a biblioteca:

```
library(epiDisplay)
```

Para obter um intervalo de confiança da média da variável idade (“age”) com nível de significância de 5%:

```
ci(banco$age,alpha=0.05)
```

Aparecerá o intervalo de confiança para a média de idade: [22.47779; 23.9984] anos.



```
Output
> library(epicalc)
> ci(banco$age,alpha=0.05)
      n  mean      sd      se lower95ci upper95ci
189 23.2381 5.298678 0.3854221 22.47779 23.9984
```

Opção 2: Manualmente no R

#Variância populacional desconhecida

```
n<-length(banco$age)
sdage<-sd(banco$age)
alpha<-0.01
tcritico<-qt(alpha/2,df=n-1,lower.tail=FALSE)
LI<-meanage-tcritico*(sdage/sqrt(n))
LSt<- meanage+tcritico*(sdage/sqrt(n))
cbind(LI,LSt)
```

#Variância populacional conhecida

```
n<-length(banco$age)
sdage<-sd(banco$age)
```

```
alpha<-0.01
zcritico<-qnorm(alpha/2,lower.tail=FALSE)
Llz<-meanage-zcritico*(sdage/sqrt(n))
LSz<- meanage+zcritico*(sdage/sqrt(n))
cbind(Llz,LSz)
```

Intervalo de Confiança para proporção

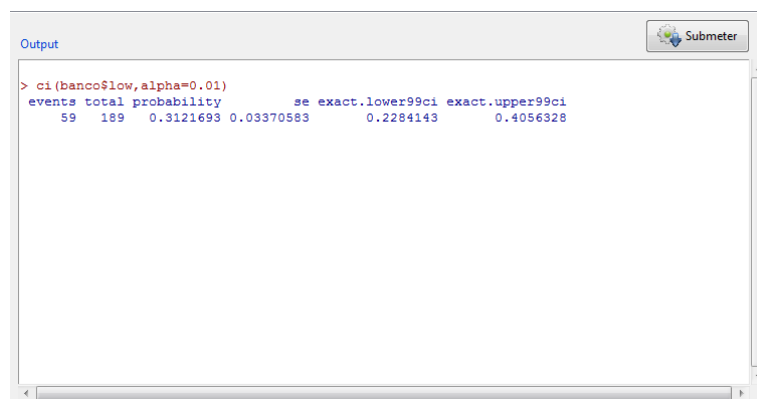
Exercício 2: Suponha que estejamos interessados no intervalo de confiança para a prevalência de baixo peso (variável “low”) no banco “birthwt” a um nível de confiança de 99%. Interprete.

Opção 1: Automaticamente pela biblioteca “epiDisplay”

```
ci(banco$low,alpha=0.01)
```

Atenção: A variável tem que ser dicotômica 0 e 1 para o comando funcionar! Não pode ser factor!

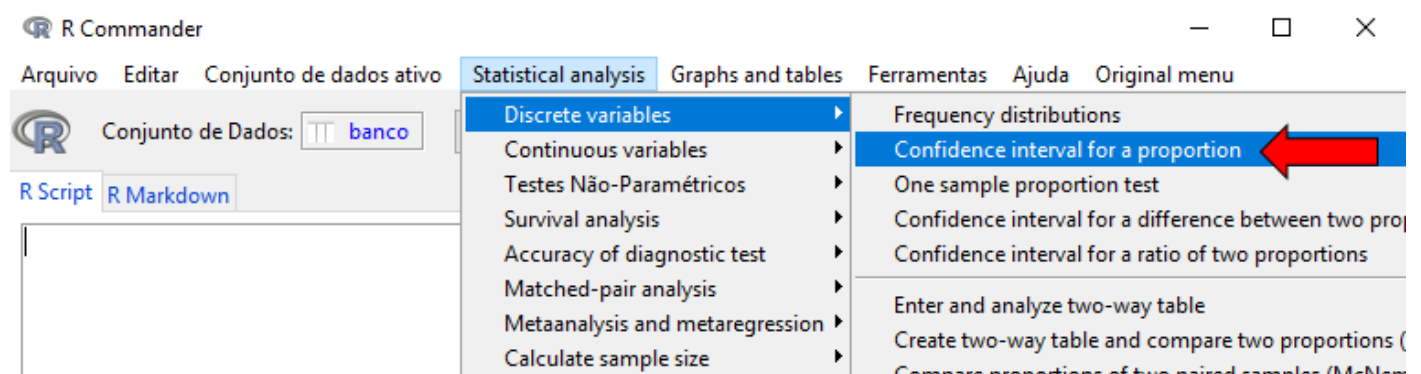
O intervalo de confiança para a proporção de baixo peso é [0.2284143;0.4056328]



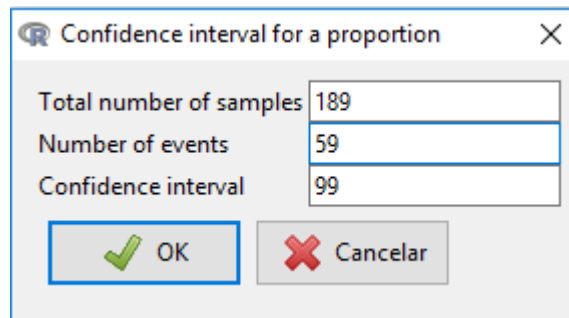
```
Output
> ci(banco$low,alpha=0.01)
events total probability      se exact.lower99ci exact.upper99ci
59    189    0.3121693 0.03370583      0.2284143      0.4056328
```

Opção 2: Pelo EZR

Statistical analysis > Discrete variables > Confidence interval for a proportion



Preencha as informações para o cálculo e dê ok.



Confidence interval for a proportion

Total number of samples: 189

Number of events: 59

Confidence interval: 99

OK Cancelar

Output

```
> #####Confidence interval for a proportion#####  
> prop.conf(59, 189, 99)  
[1] Probability : 0.312  
[1] 99% confidence interval : 0.228 - 0.406
```

Para desativar o menu do EZR terá que fechar o R e abrir novamente. Ou utilizar o menu antigo pela opção “Original menu” na janela de menu. O caminho dos comandos abaixo refere-se ao menu original, porém também existem opções pelo menu do EZR.

Opção 3: Manualmente no R

```
p<-mean(banco$low)  
n<-length(banco$low)  
q<-1-p  
alpha<-0.01  
zcritico<-qnorm(alpha/2,lower.tail=FALSE)  
LIp<-p-zcritico*(sqrt(p*q/n))  
LSp<- p+zcritico*(sqrt(p*q/n))  
cbind(LIp,LSp)
```

Teste de Hipóteses

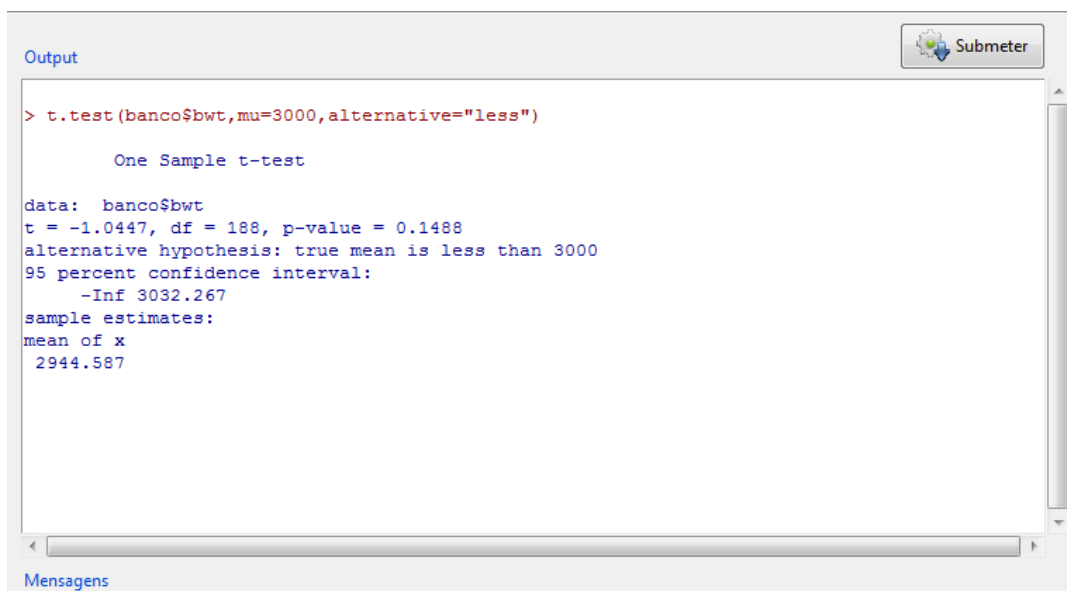
O teste de hipóteses pretende realizar inferência sobre certos parâmetros populacionais a partir dos dados amostrais assumindo hipóteses pré-definidas pelo pesquisador do estudo. Os testes de hipóteses podem ser realizados para médias, proporções, variâncias e outros diversos parâmetros. Os testes de hipóteses dividem-se em paramétricos (assumem normalidade dos dados) e não-paramétricos. Os testes paramétricos possuem uma maior precisão com mesmo tamanho amostral, além da maior facilidade no seu cálculo.

No caso paramétrico, os cálculos são diferentes se as variâncias populacionais são conhecidas ou desconhecidas. O caso mais comum é da variância populacional desconhecida e, portanto, utiliza-se a distribuição t de student nos cálculos. Abaixo, seguem os testes de hipóteses para variâncias desconhecidas.

Exercício 3: Suponha que estejamos interessados em verificar se o peso médio ao nascer dos bebês no banco “birthwt” é menor que 3000g. Que conclusão chega a um nível de significância de 5%?

Execute o seguinte comando na aba “R Script”:

```
t.test(banco$bwt,mu=3000,alternative="less")
```



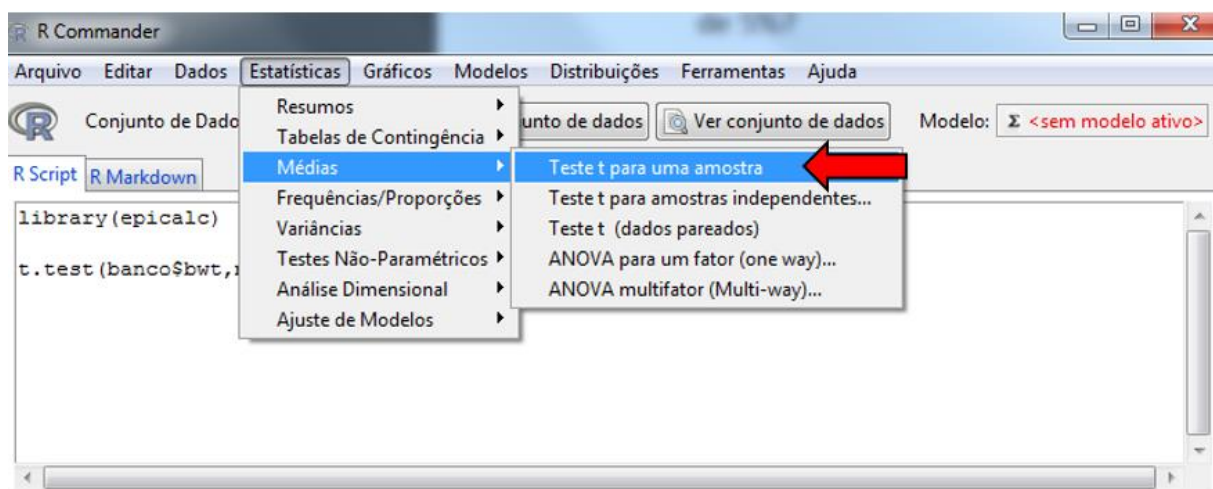
```
Output
> t.test(banco$bwt,mu=3000,alternative="less")

One Sample t-test

data:  banco$bwt
t = -1.0447, df = 188, p-value = 0.1488
alternative hypothesis: true mean is less than 3000
95 percent confidence interval:
 -Inf 3032.267
sample estimates:
mean of x
 2944.587
```

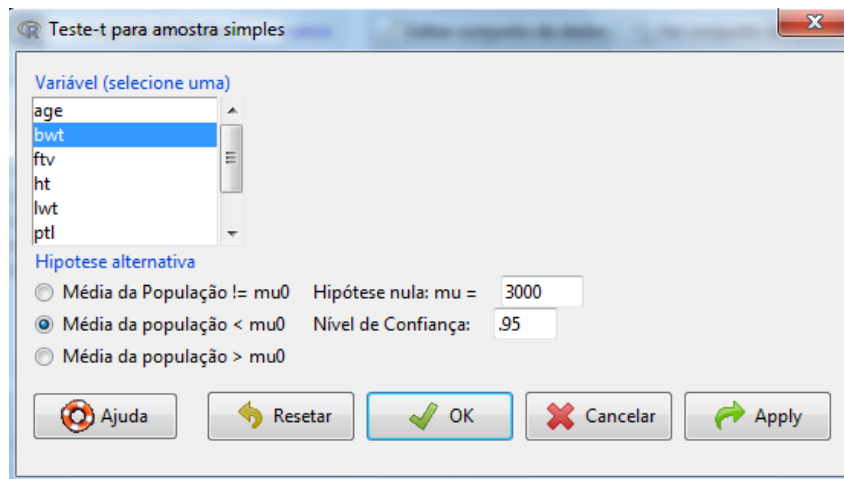
No R Commander a mesma função pode ser executada da seguinte maneira:

“Estatísticas > Médias > Teste t para uma amostra”



Aparecerá as seguintes opções para serem preenchidas com os dados informados no problema a ser realizado o teste estatístico.

Dados $\mu=3000$ e $\text{confiança}=.95$



O resultado será exatamente o mesmo do comando gerado no R Script

```

Output
Submiter

> with(banco, (t.test(bwt, alternative='less', mu=3000, conf.level=.95)))

One Sample t-test

data:  bwt
t = -1.0447, df = 188, p-value = 0.1488
alternative hypothesis: true mean is less than 3000
95 percent confidence interval:
 -Inf 3032.267
sample estimates:
mean of x
 2944.587

```

Difícilmente trabalhamos com testes de hipóteses em um único grupo, como no exemplo acima, já que estamos interessados em verificar diferenças de parâmetros (exemplo: média de idade) em dois grupos (exemplo: se tem ou não baixo peso). Primeiramente, para testar se existe diferenças de parâmetros populacionais entre dois ou mais grupos, deve-se estabelecer se os grupos (amostras) são pareados ou independentes.

TESTES DE HIPÓTESES PARAMÉTRICO PARA MÉDIAS– 2 amostras independentes– Variâncias populacionais desconhecidas

Exercício 4: Suponha que estejamos interessados em verificar se a idade média das mães dos recém-nascidos é diferente entre as crianças com baixo peso ou peso normal. Que conclusão chega a um nível de significância de 5%? Dica: Lembre-se que os cálculos são diferentes quando as variâncias são desconhecidas, mas supostas iguais ou diferentes.

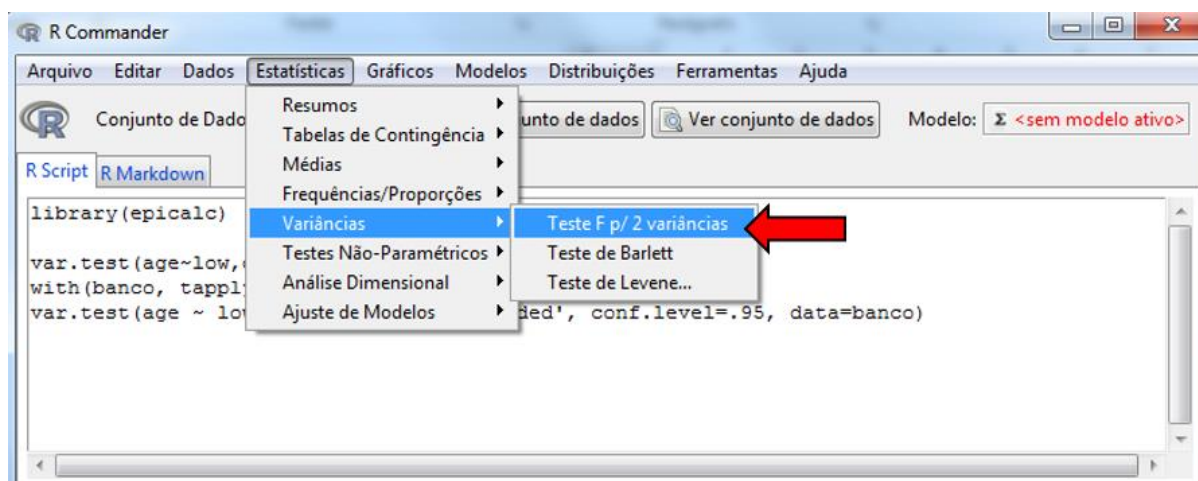
Primeiramente verificar se as variâncias são supostas iguais ou diferentes.

Opção 1: Na aba “R Script”

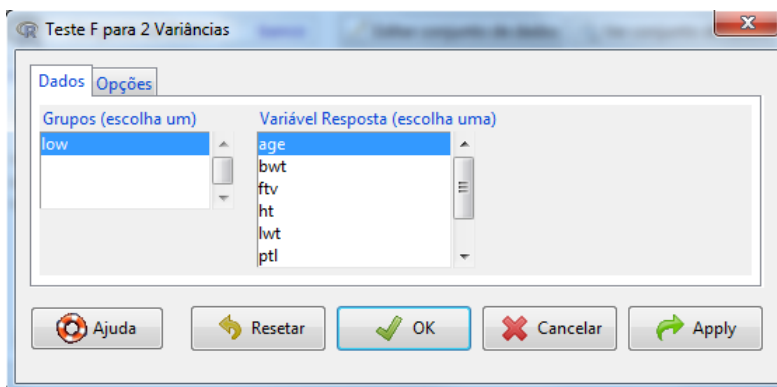
```
var.test(age~low,data=banco)
```

Opção 2: No point-and-click do RCommander

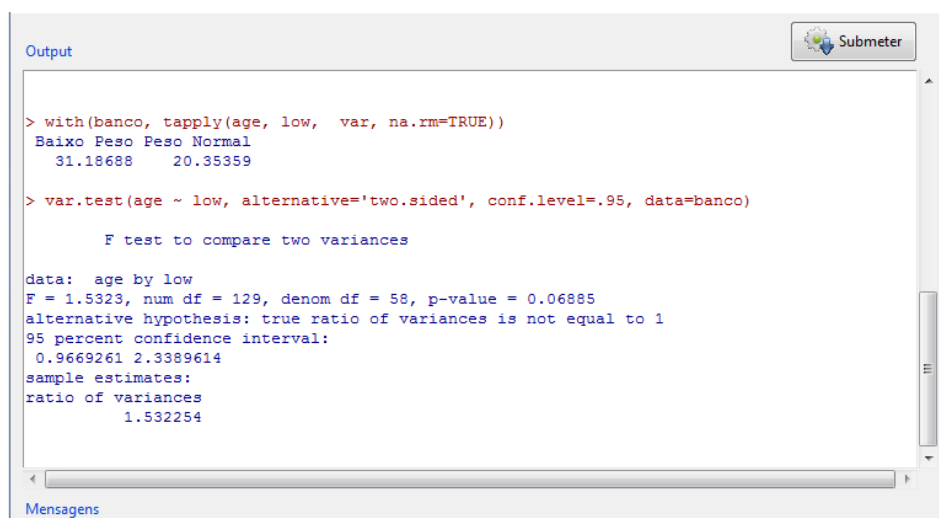
“Estatísticas > Variância > Teste F p/ 2 variâncias”



Selecionando as variáveis “low” e “age” para verificar se há diferença entre as variáveis.



Clique em Ok e aparecerá o seguinte resultado:



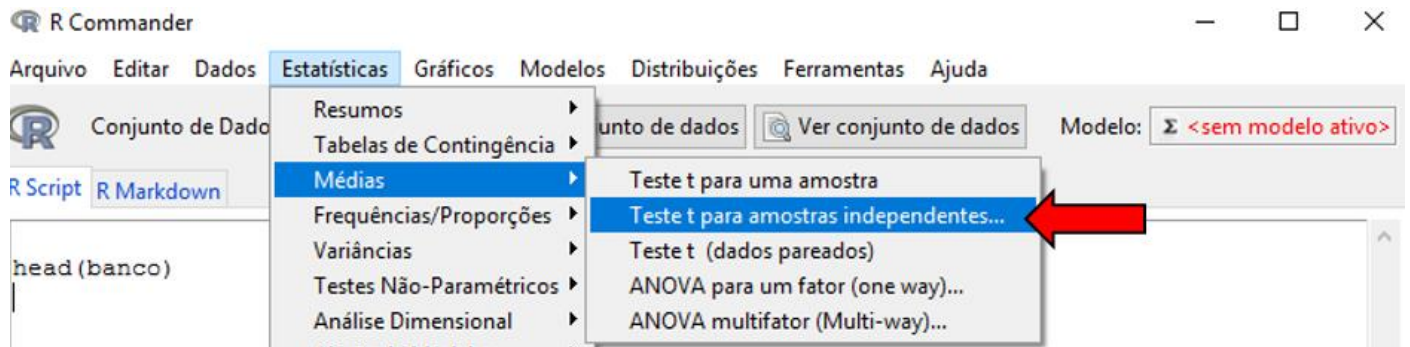
Para testar as médias supondo variâncias iguais:

Opção 1: Na aba “R Script”

```
t.test(age~low,data=banco,var.equal=TRUE,alternative="two.sided",paired=FALSE,conf.level=0.95)
```

Opção 2: No point-and-click do RCommander

“Estatísticas > Médias > Teste t para amostras independentes”



```
> t.test(age~low, alternative='two.sided', conf.level=.95, var.equal=FALSE, data=banco)

Welch Two Sample t-test

data:  age by low
t = 1.7737, df = 136.941, p-value = 0.07834
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.1558349  2.8687423
sample estimates:
mean in group >2.5 mean in group <2.5
      23.66154      22.30508
```


TESTES DE HIPÓTESES PARAMÉTRICO PARA MÉDIAS– 2 amostras pareadas

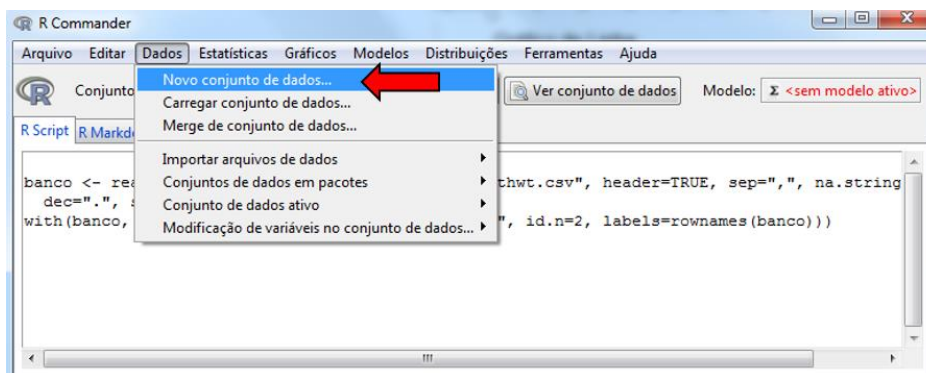
Ao utilizar dados pareados, consideramos a diferença entre os dois grupos e todo o teste de hipóteses é conduzido como se fosse para uma única amostra.

Exercício 5: Suponha que esteja interessado em examinar os efeitos da transição da circulação fetal para a pós-natal entre bebês prematuros. Para cada um dos 14 recém-nascidos saudáveis, a taxa respiratória é medida em dois diferentes momentos – uma vez quando o bebê tem menos de 15 dias e outra quando tem mais que 25 dias.

Indivíduo	Taxa respiratória (respiração/min)	
	Momento 1	Momento 2
1	62	46
2	35	42
3	38	40
4	80	42
5	48	36
6	48	46
7	68	45
8	26	40
9	48	42
10	27	40
11	43	46
12	67	31
13	52	44
14	88	48

Avalie a hipótese nula de que a diferença mediana nas taxas respiratórias para os dois momentos é igual a 0. Qual a sua conclusão ao nível de 10%?

Carregue o novo banco de dados “questao1.csv” ou crie um novo conjunto de dados



Para criar um novo conjunto de **Dados** -> digite os valores:

Realizando o teste de hipóteses da diferença das médias:

Opção 1: Na aba “R Script”

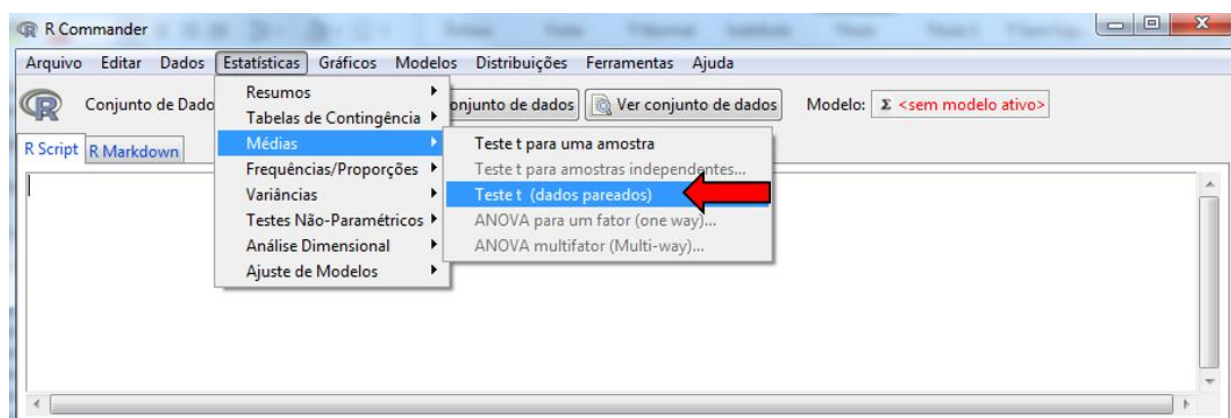
`t.test(questao1$momento1,questao1$momento2,paired=TRUE,conf.level=0.90)`

ou

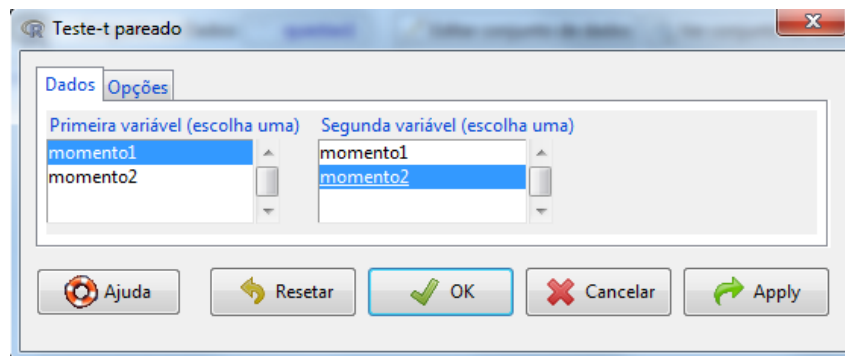
`t.test(questao1$momento2-questao1$momento1,conf.level=0.90)`

Opção 2: No point-and-click do RCommander

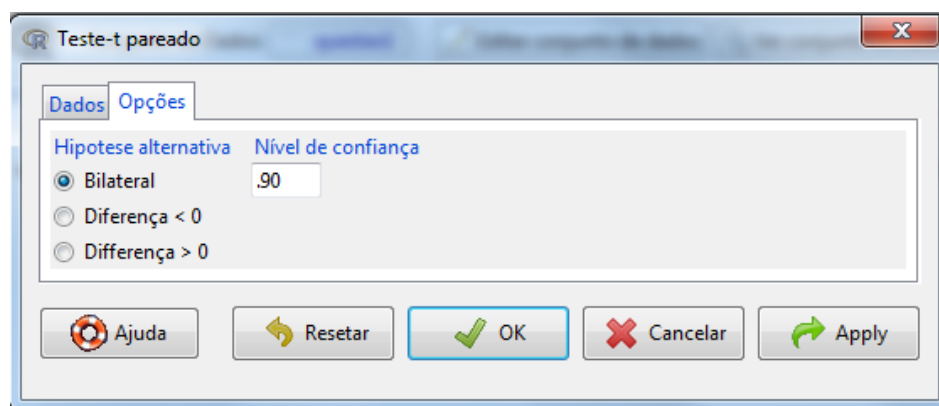
Estatísticas > Média > Teste t (dados pareados)



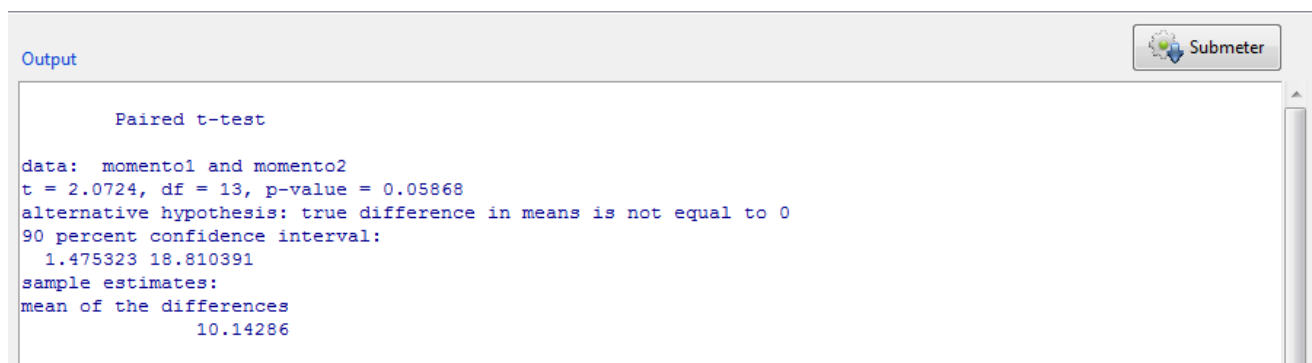
Selecione as variáveis que deseja fazer o teste



Na aba de opções selecione os parâmetros desejados para o teste



Resultado:



Testes de normalidade

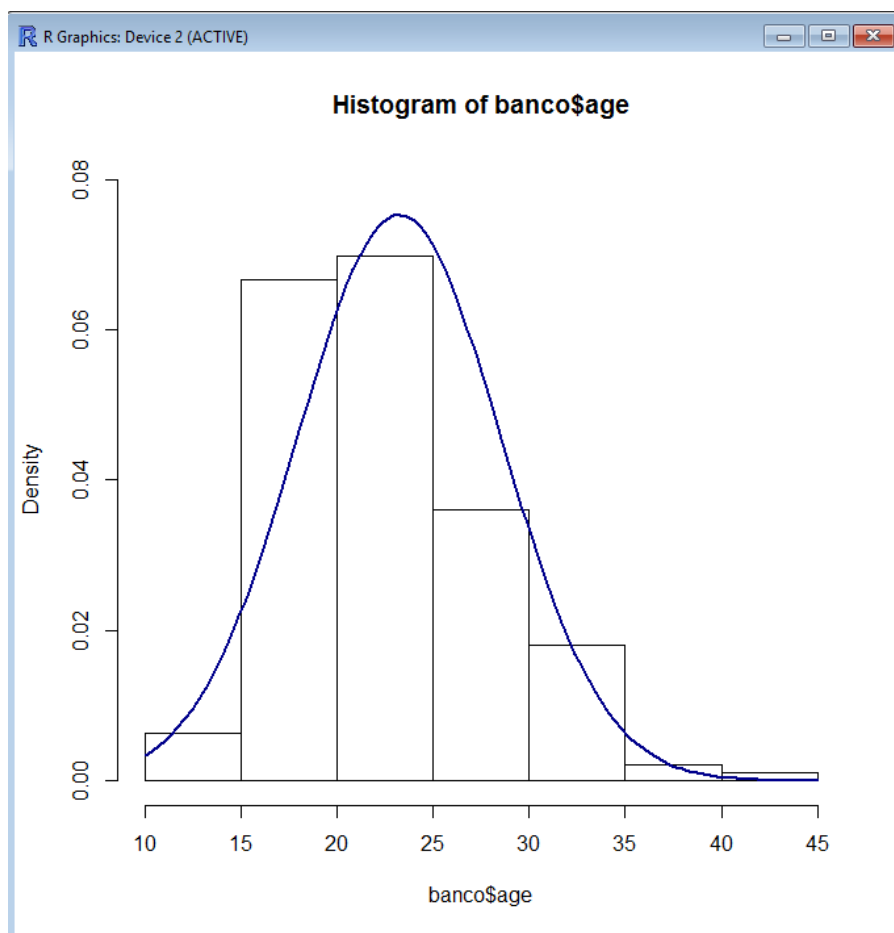
Em alguns casos, o pressuposto de normalidade dos dados é refutado para as variáveis contínuas, o que impede a utilização de testes paramétricos, como o teste T. Existem diversos testes de normalidade de dados, porém os mais conhecidos são os testes de Kolmogorov-Smirnov (com e sem correção de Lilliefors) e Shapiro-Wilk. O teste de Kolmogorov-Smirnov supõe a comparação da distribuição observada acumulada com a distribuição esperada dispomos da média ou variância populacional, o que torna o teste não-aplicável. O teste de Lilliefors utiliza o mesmo acumulada ao supor distribuição normal. A principal desvantagem desse método é que em muitos casos a média e variância populacional são desconhecidas. O teste de Shapiro-Wilk supõe diversas ponderações para as diferenças entre os dados observados e é o teste mais eficiente, mesmo em tamanhos amostrais reduzidos. Muitas vezes são realizados os dois, na “esperança” que um deles não rejeite a hipótese de normalidade dos dados.

Exercício 6: Verifique a normalidade da variável idade da mãe (age) no banco de dados dos recém-nascidos (birthwt). A que conclusão chega a partir dos gráficos e dos testes de normalidade?

#Realizando gráficos para explorar a normalidade dos dados

Opção 1: Na aba “R Script”

```
hist(banco$age,freq=FALSE,ylim=c(0,0.08))  
curve(dnorm(x,mean=mean(banco$age),sd=sd(banco$age)), add=TRUE, col="darkblue", lwd=2)
```



Opção 2: No point-and-click do RCommander vá no menu de Gráficos

#Testes de normalidade para a variável age

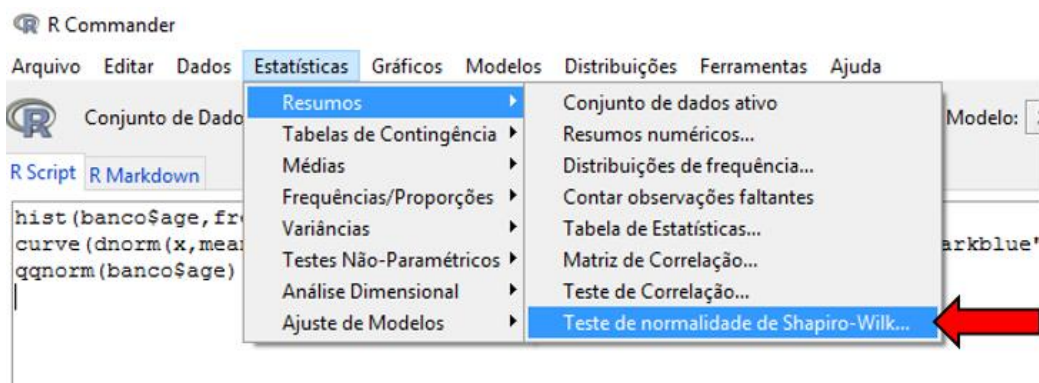
Opção 1: Na aba “R Script”

```
library(nortest)
ks.test(banco$age,"pnorm",mean(banco$age),sd(banco$age)) #KS sem correção
lillie.test(banco$age) #Kolmogorov-Smirnov com correção de Lilliefors
shapiro.test(banco$age) #Shapiro-Wilk
```

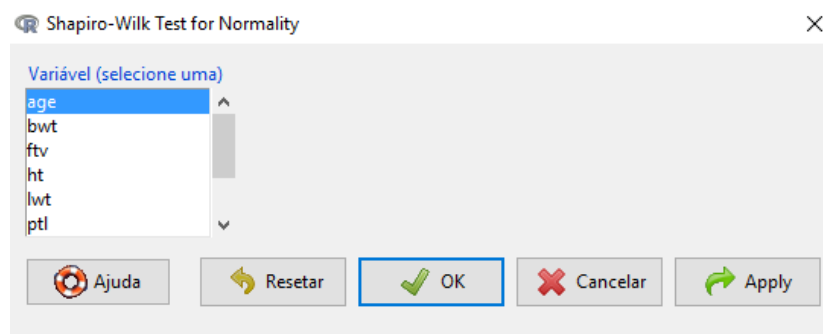
Pergunta: O que aconteceu? Como interpreto?

Opção 2: No point-and-click do RCommander

Estatística > Teste de normalidade Shapiro-Wilk



Selecione a variável “age”



```
Output
Shapiro-Wilk normality test
data:  age
W = 0.9598, p-value = 3.189e-05
```

Pergunta: O que aconteceu? Como interpreto?

Se tiver mais de duas amostras independentes, pode utilizar o teste de análise de variância (ANOVA) para comparar as médias.

Teste Não-Paramétricos

No caso de refutação da normalidade dos dados ($p\text{-valor} < \text{nível de significância}$), não podemos utilizar os testes paramétricos que supõem normalidade dos dados. A principal desvantagem de utilizar testes não-paramétricos reside na dificuldade de alcançar significância estatística para o mesmo tamanho amostral que utilizaria num teste paramétrico. A principal vantagem é não precisar supor qualquer distribuição para os dados.

Exercício 7: Pelo que vimos nos testes de normalidade, a variável idade da mãe (age) parece não seguir a normalidade dos dados no banco de dados “birthwt”. Entretanto, isso não foi verificado separadamente para cada grupo de crianças (baixo peso e peso normal). Verificar separadamente a normalidade da variável age, para confirmar a aplicabilidade dos testes paramétricos ao nível de 5%.

#Verificando a normalidade dos dados dentro de cada categoria (Baixo Peso e Normal)

Opção 1: Na aba “R Script”

```
library(nortest)
by(banco$age,banco$low,lillie.test)
by(banco$age,banco$low,shapiro.test)
```

```
Output
D = 0.0884, p-value = 0.301

> by(banco$age, banco$low, shapiro.test)
banco$low: >2.5

      Shapiro-Wilk normality test

data:  dd[x, ]
W = 0.9497, p-value = 0.000108

-----
banco$low: <2.5

      Shapiro-Wilk normality test

data:  dd[x, ]
W = 0.9818, p-value = 0.521
```

#Confirmada a rejeição da normalidade dos dados no grupo de recém-nascidos com peso normal ->
#Utilização de testes não-paramétricos

Opção 2: No point-and-click do RCommander não existe essa opção

Exercício 8: Suponha que estejamos interessados em verificar se a idade média das mães dos recém-nascidos é diferente entre as crianças com baixo peso ou peso normal. Qual a conclusão ao nível de 10%?

TESTES DE HIPÓTESES NÃO-PARAMÉTRICO - 2 amostras independentes (Mann-Whitney)

Opção 1: Na aba “R Script”

```
wilcox.test(age~low,data=banco,paired=FALSE,alternative="two.sided")
by(banco$age,banco$low,quantile)
```

```
R Script R Markdown

wilcox.test(age~low,data=banco,paired=FALSE,alternative="two.sided")
by(banco$age,banco$low,quantile)

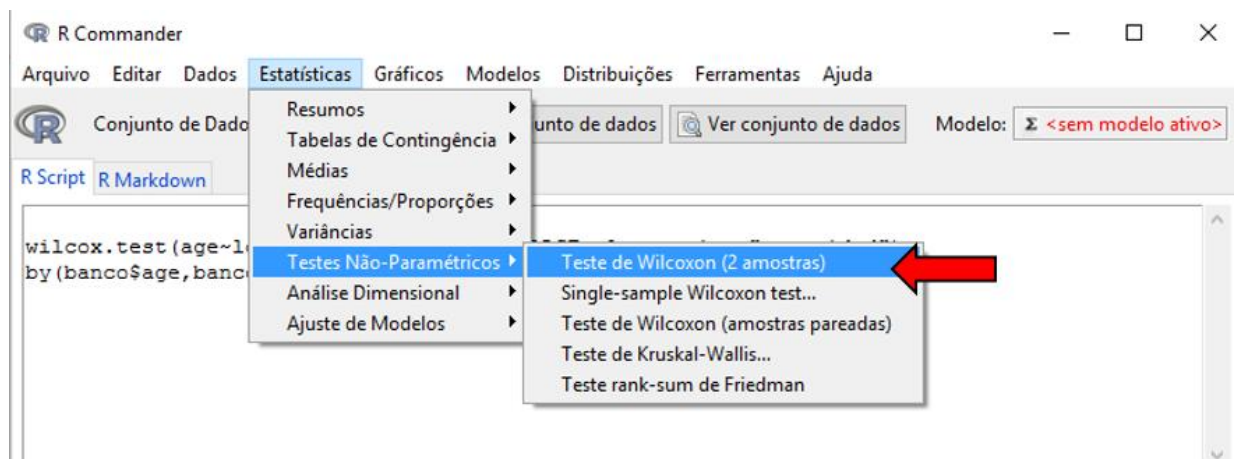
Output
Wilcoxon rank sum test with continuity correction

data: age by low
W = 4238, p-value = 0.2471
alternative hypothesis: true location shift is not equal to 0

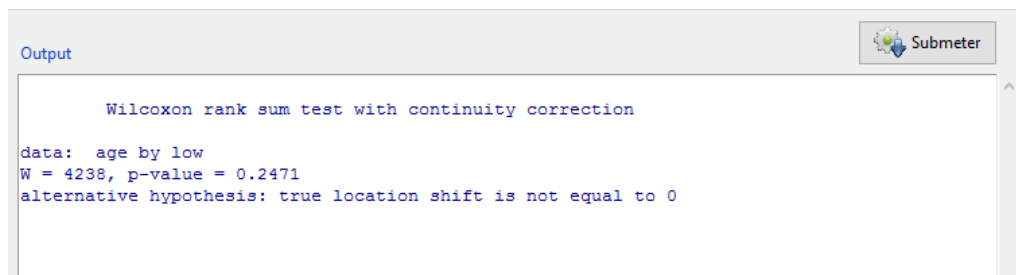
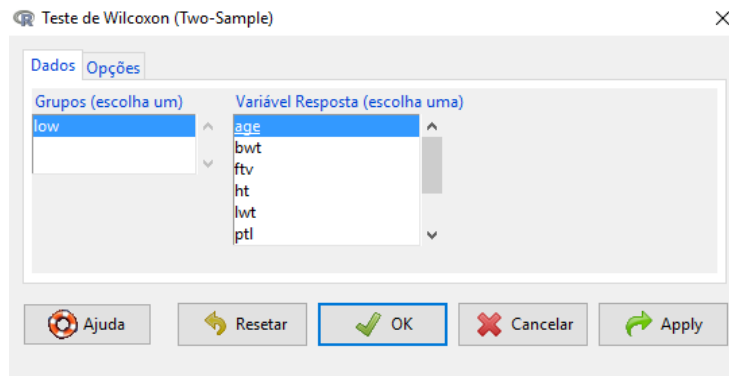
> by(banco$age,banco$low,quantile)
banco$low: 0
 0%  25%  50%  75% 100%
 14  19  23  28  45
-----
banco$low: 1
 0%  25%  50%  75% 100%
14.0 19.5 22.0 25.0 34.0
```

Opção 2: No point-and-click do RCommander

“Estatística > Testes Não-Paramétricos > Teste de Wilcoxon (2 amostras)”



Escolha as variáveis “low” e “age” e clique em OK



O teste de Mann-Whitney é uma particularidade do teste de Wilcoxon em amostras independentes, por isso o software chama de Wilcoxon rank sum test, porém o nome que utilizamos é Mann-Whitney.

Se tiver mais de duas amostras independentes, pode utilizar o teste de análise de variância de Kurskal-Wallis para comparar as médias, no caso de amostras independentes, e teste de Friedman, no caso pareado.

TESTES DE HIPÓTESES NÃO-PARAMÉTRICO – 2 amostras relacionadas (Wilcoxon)

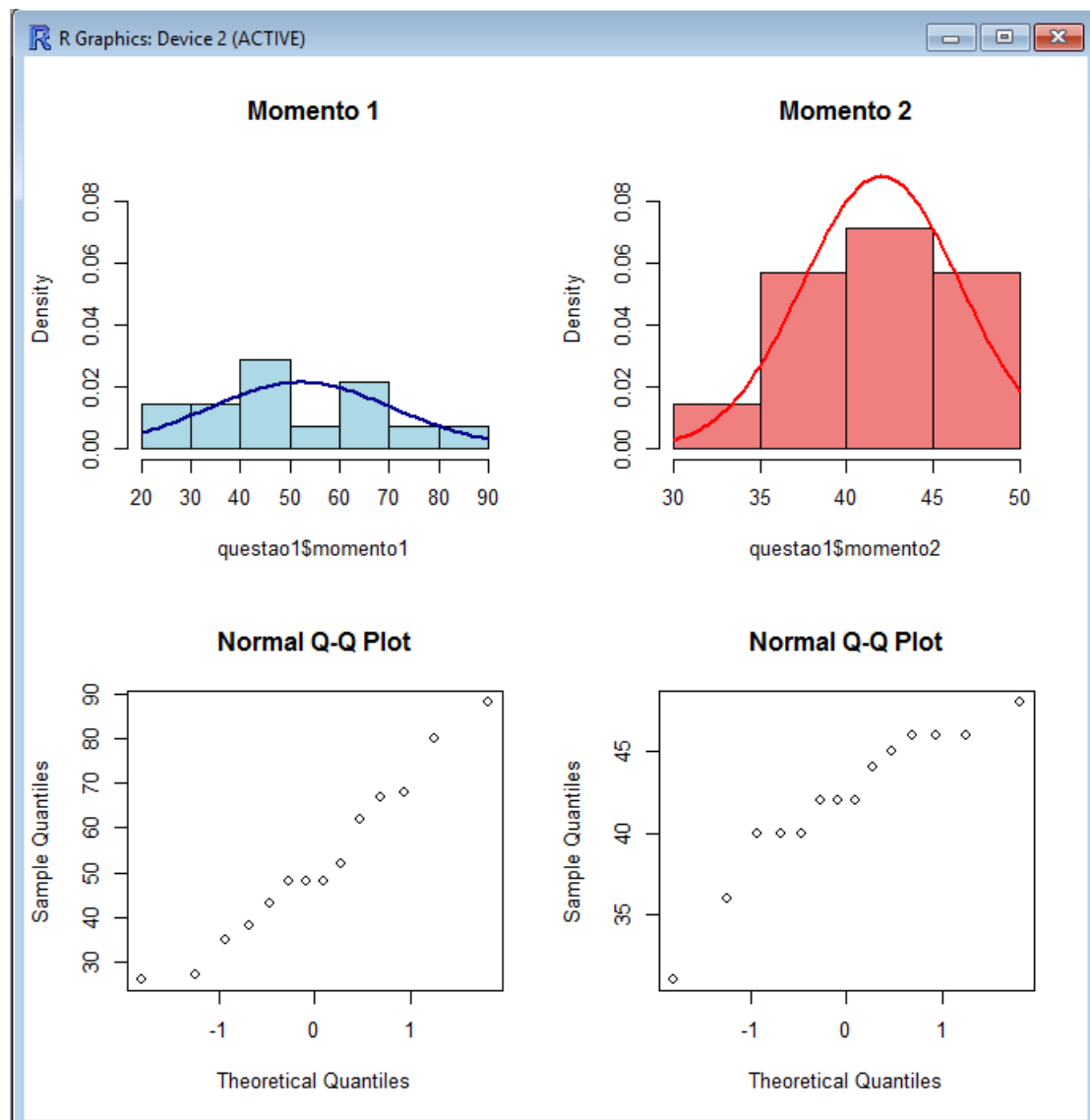
Exercício 9: Suponha que esteja interessado em examinar os efeitos da transição da circulação fetal para a pós-natal entre bebês prematuros (dados “questao1.csv”). Para cada um dos 14 recém-nascidos saudáveis, a taxa respiratória é medida em dois diferentes momentos – uma vez quando o bebê tem menos de 15 dias e outra quando tem mais que 25 dias.

Opção 1: Na aba “R Script”

```
par(mfrow=c(2,2))
hist(questao1$momento1,freq=FALSE,ylim=c(0,0.09),col="lightblue",main="Momento 1")
curve(dnorm(x,mean=mean(questao1$momento1),sd=sd(questao1$momento1)), add=TRUE,
col="darkblue", lwd=2)
hist(questao1$momento2,freq=FALSE,ylim=c(0,0.09),col="lightcoral",main="Momento 2")
curve(dnorm(x,mean=mean(questao1$momento2),sd=sd(questao1$momento2)), add=TRUE, col="red",
lwd=2)
qqnorm(questao1$momento1)
qqnorm(questao1$momento2)
library(nortest)
```

```
lillie.test(questao1$momento1)
lillie.test(questao1$momento2)
shapiro.test(questao1$momento1)
shapiro.test(questao1$momento2)
```

#Alguns gráficos são possíveis de serem realizados no menu do RCommander!!!



```
Output

Lilliefors (Kolmogorov-Smirnov) normality test

data:  questao1$momento1
D = 0.1588, p-value = 0.4407

> lillie.test(questao1$momento2)

Lilliefors (Kolmogorov-Smirnov) normality test

data:  questao1$momento2
D = 0.1863, p-value = 0.2081

> shapiro.test(questao1$momento1)

Shapiro-Wilk normality test

data:  questao1$momento1
W = 0.9558, p-value = 0.6536

Mensagens

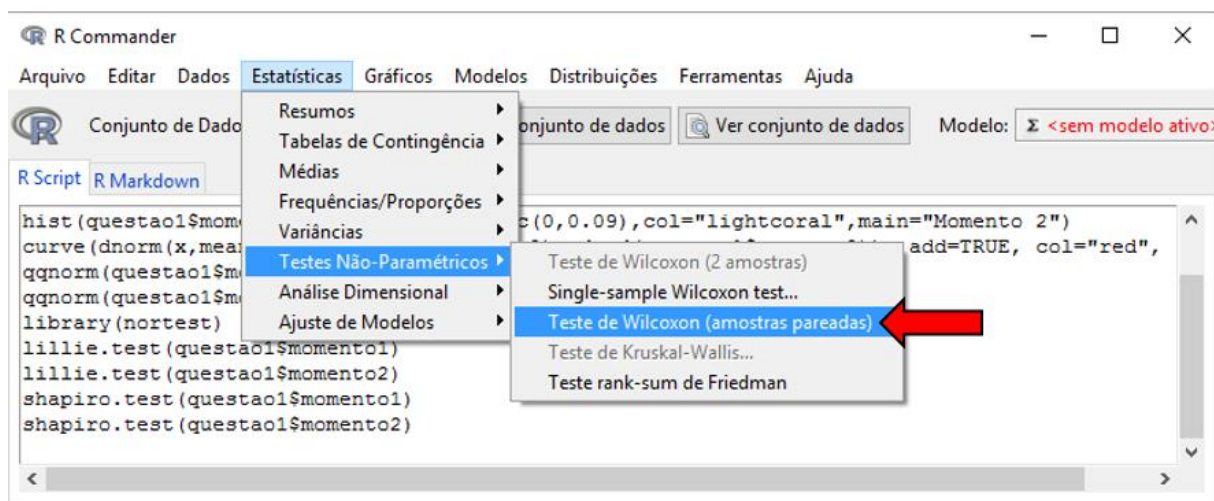
y is empty or has only NAs
[21] NOTA: Os dados questao1 tem 14 linhas e 2 colunas.
```

#Apesar dos testes não-rejeitarem a hipótese de normalidade, os gráficos apresentam grande assimetria -> testes não-paramétricos

`wilcox.test(questao1$momento1,questao1$momento2,paired=TRUE)`

Opção 2: No point-and-click do RCommander

“Estatística > Teste Não-Paramétricos > Teste de Wilcoxon (amostras pareadas)”



Escolha as variáveis na ordem que deseja realizar os testes. No nosso exemplo as variáveis “momento1” e “momento2”, clique em OK

Test de Wilcoxon Pareado

Dados Opções

Primeira variável (escolha uma) Segunda variável (escolha uma)

momento1 momento1

momento2 momento2

Ajuda Resetar OK Cancelar Apply

```

Output
Submiter

> with(questao1, median(momento1 - momento2, na.rm=TRUE)) # median difference
[1] 7

> with(questao1, wilcox.test(momento1, momento2, alternative='two.sided', paired=TRUE))

    Wilcoxon signed rank test with continuity correction

data:  momento1 and momento2
V = 78.5, p-value = 0.1093
alternative hypothesis: true location shift is not equal to 0
  
```

Muitas vezes o R o erro mostra erro devido a empates nos dados amostrais, sendo indicada correção do p-valor encontrado no teste estatístico de Wilcoxon.

#Correção do teste de Wilcoxon quando ocorrem empates

Opção 1: Na aba “R Script”

```

library(exactRankTests)
wilcox.exact(questao1$momento1,questao1$momento2,paired=TRUE, alternative = "two.sided",exact = TRUE)
  
```

```

Output
Submiter

> library(exactRankTests)

> wilcox.exact(questao1$momento1,questao1$momento2,paired=TRUE, alternative = "two.sided",exact = TRUE)

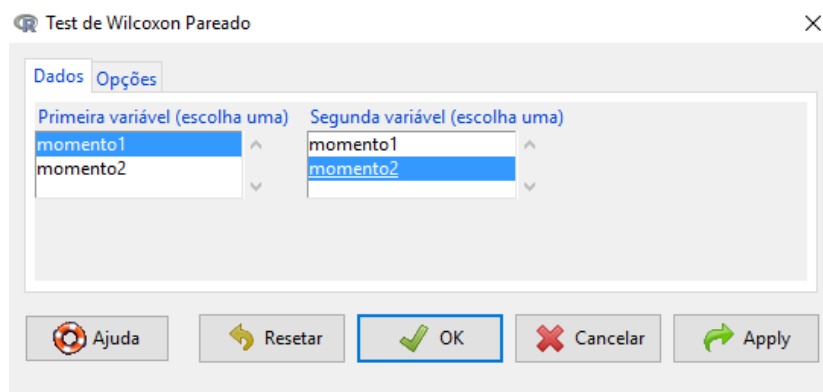
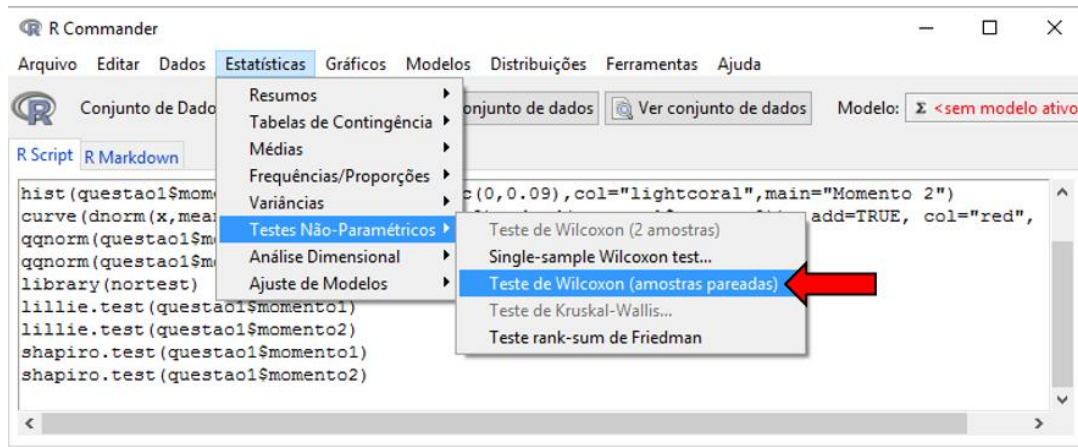
    Exact Wilcoxon signed rank test

data:  questao1$momento1 and questao1$momento2
V = 78.5, p-value = 0.1075
alternative hypothesis: true mu is not equal to 0
  
```

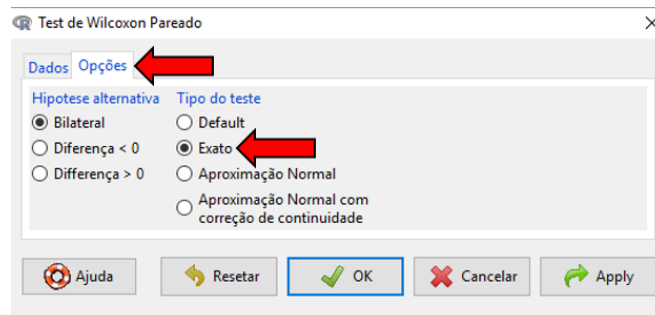
Mensagens

Opção 2: No point-and-click do RCommander

“Estatística > Teste Não-Paramétricos > Teste de Wilcoxon (amostras pareadas)”



Na aba “Opções” selecione a opção “Tipo de teste” - “Exato” e clique em OK



Output



```
> with(questaol, median(momento1 - momento2, na.rm=TRUE)) # median difference
[1] 7

> with(questaol, wilcox.test(momento1, momento2, alternative='two.sided', exact=TRUE,
+   paired=TRUE))

      Wilcoxon signed rank test with continuity correction

data:  momento1 and momento2
V = 78.5, p-value = 0.1093
alternative hypothesis: true location shift is not equal to 0
```

Tabelas de Contingência

As tabelas de contingência são utilizadas para relacionar dados de variáveis qualitativas. Os testes de hipóteses envolvem a comparação das contagens observadas nas células com as contagens esperadas se a hipótese de independência fosse verdadeira ($P(A \cap B) = P(A) \times P(B)$). Os testes de hipóteses de tabelas de contingência são chamados de testes de associação, sendo o teste de qui-quadrado o mais utilizado.

TESTES QUI-QUADRADO E EXATO DE FISHER

Opção 1: Na aba “R Script”

```
table(banco$smoke,banco$low)  
prop.table(banco$smoke,banco$low)
```

#Tabelas de contingência pelo epiDisplay

```
library(epiDisplay)  
tabpct(banco$smoke,banco$low,graph=FALSE)
```

#Qui-quadrado de Pearson

```
chisq.test(table(banco$smoke,banco$low),correct=FALSE)
```

#Qui-quadrado de Pearson com correção de Yates

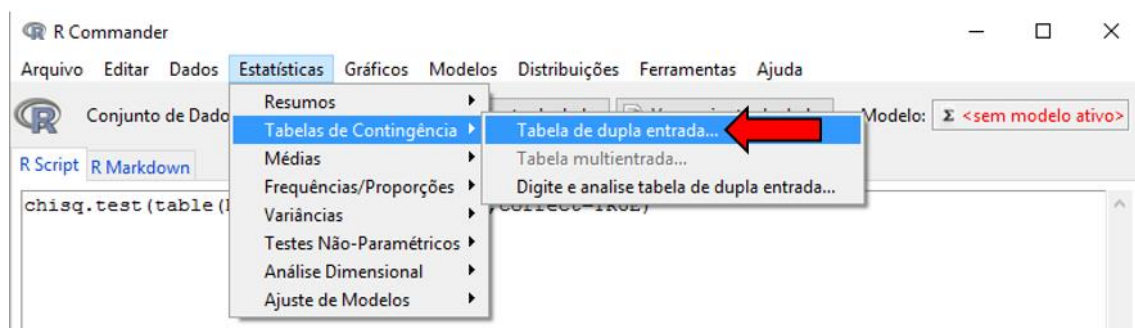
```
chisq.test(table(banco$smoke,banco$low),correct=TRUE)
```

#Teste exato de Fisher

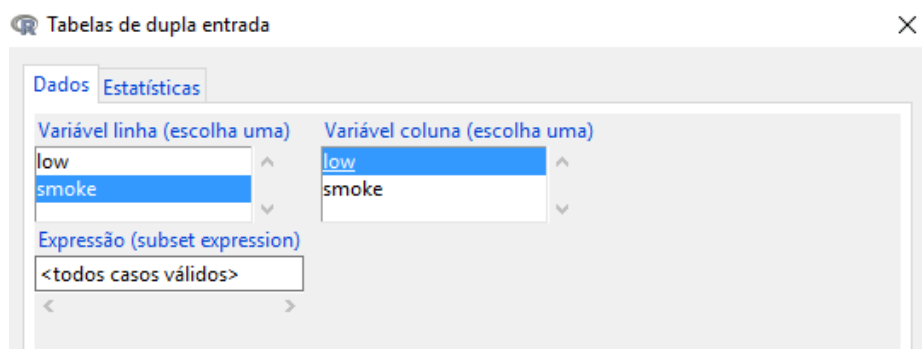
```
fisher.test(table(banco$smoke,banco$low))
```

Opção 2: No point-and-click do RCommander

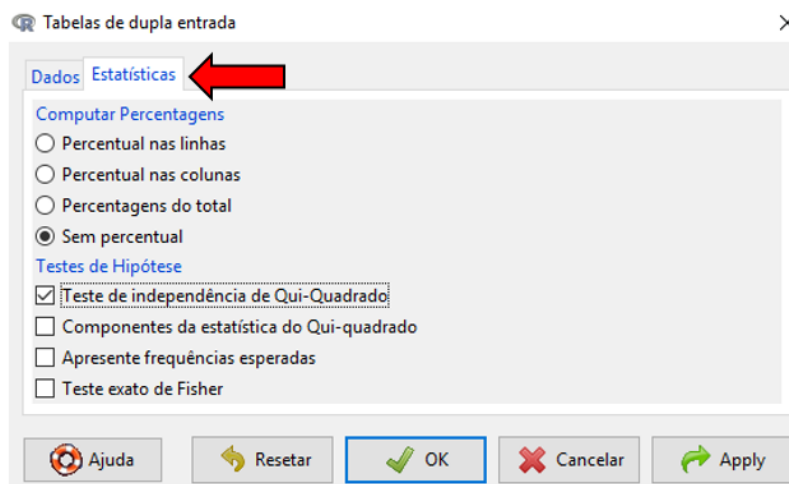
“Estatística > Tabela de Contingência > Tabela de dupla entrada...”



Selecione as variáveis “smoke” e “low”



Na aba “Estatísticas” selecione as opções desejadas para o teste.



Kappa Simples e Ponderado

O Kappa é uma medida de concordância utilizada para avaliar confiabilidade intra ou inter-observadores. Ele dá um grau de concordância que pode ser classificado em graus variando de pobre a perfeita (Landis & Koch, 1977).

Nesse caso, o teste-qui-quadrado não é aconselhável. Por pior que seja a concordância sempre existirá uma tendência de agrupamento das contagens na diagonal principal, o que faz com que o qui-quadrado sempre seja significativo.

O Kappa é denominado simples quando trabalhamos com tabelas quadradas com duas linhas e duas colunas e ponderado quando trabalhamos com tabelas maiores. O kappa ponderado dará maiores pesos para discordâncias mais distantes e menores pesos para discordâncias mais próximas, já que assume o uso de variáveis ordinais (exemplo: Concordo totalmente, Concordo, Nem concordo nem discordo, discordo, discordo totalmente). A ponderação pode ser linear ou quadrática, de acordo com o critério do pesquisador.

Exercício 11: Suponha que queiramos saber a concordância (kappa simples) na avaliação de dois observadores quanto a cura numa determinada doença X.

Observador 1	Observador 2	
	Curado	Não curado
Curado	30	5
Não curado	10	50

Testes de Kappa Simples

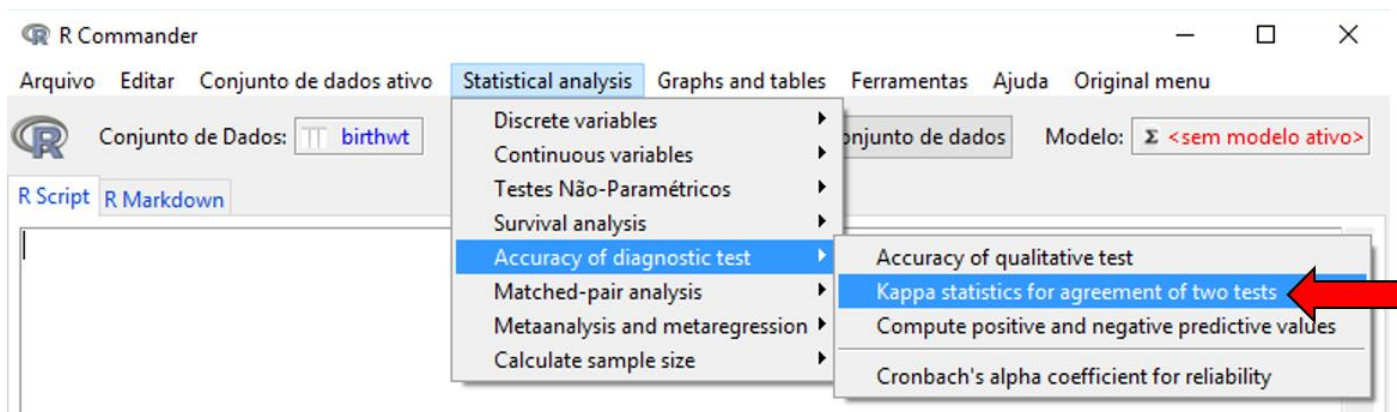
Opção 1: Na aba “R Script”, no pacote psych

```
tabela <- matrix(c(30,10,5,50), nrow = 2, dimnames = list(Obs1 = c("Curado", "Não Curado"), Obs2 =  
c("Curado", "Não Curado")))  
library(psych)  
  
cohen.kappa(tabela)
```

O comando pode ser alterado para calcular kappa ponderado

Opção 2: No point-and-click do RCommander

Statistical analysis > Accuracy diagnostic test > Kappa statistics agreement of two test



Aparecerá a seguinte tabela a ser preenchida com os dados para o teste

The dialog box 'Kappa statistics for agreement of two tests' is shown. It contains a table for inputting data:

Number	Test2 (+)	Test2 (-)
Test1 (+)		
Test1 (-)		

Below the table are buttons for 'Ajuda', 'OK', and 'Cancelar'.

Preencha a tabela com os dados (exemplo abaixo) e clique em ok.

The dialog box is shown with the following data entered:

Number	Test2 (+)	Test2 (-)
Test1 (+)	98	0
Test1 (-)	1	1

The 'OK' button is highlighted with a blue border.

```
Output
Submiter

> .Table
      Test2 (+) Test2 (-)
Test1 (+)      98      0
Test1 (-)       1      1

> res <- epi.kappa(.Table, conf.level = 0.95)

> remove(.Table)

> colnames(res$kappa) <- gettext(domain="R-RcmdrPlugin.EZR", colnames(res$kappa))

> res[1]
$kappa
      est      lower      upper
1 0.6621622 0.003331236 1.320993

> remove(res)

Mensagens
[10] NOTA: Aviso em data(birthwt) : data set 'birthwt' not found
[11] NOTA: Os dados birthwt tem 189 linhas e 10 colunas.
```

Opção 3: Na aba “R Script”, no pacote epiDisplay

A função para que realiza o teste de kappa é a seguinte:

kap(x, ...) onde **x** é a matriz com os valores utilizado no teste.

Antes de calcular o coeficiente no R script, temos que criar uma matriz.
Para o nosso exemplo a matriz é construída da seguinte forma:

x<- matrix(c(98, 0, 1, 1), byrow = TRUE, 2, 2)

Após criar a matriz podemos realizar o teste:

library(epiDisplay)

kap(x)

```

> x<- matrix(c(98, 0, 1, 1), byrow = TRUE, 2, 2)
>
> kap(x)

Table for calculation of kappa
  A B
A 98 0
B  1 1

Observed agreement = 99 %
Expected agreement = 97.04 %
Kappa = 0.662
Standard error = 0.094 , Z = 7.035 , P value = < 0.001

```

Exercício 12: Explore o menu do EZR para realizar acurácia de teste diagnóstico e cálculo amostral.