

Exemplos de Econometria Aplicada em R

Tainan Boff

31 de agosto de 2016

Abstract

Ao iniciar o estudo de Econometria, os alunos se deparam com uma grande diversidade de softwares estatísticos disponíveis para uso e, muitas vezes, têm dúvidas sobre qual programa deveriam aprender. Para tomar esta decisão, em geral, avaliam se o software é gratuito (ou disponibilizado pela universidade) ou pago, o *trade-off* existente entre facilidade de uso vs. flexibilidade e o quanto um ou outro programa tem seu uso mais difundido entre pesquisadores e profissionais da área de interesse do aluno. Entre os pacotes mais conhecidos, podemos citar Gretl, Minitab, EvIEWS, Stata, Matlab, SAS, SPSS, R, etc. Entre as vantagens em utilizar o R estão o fato de ele ser um software livre, com código aberto, sendo possível encontrar funções prontas para um grande número de análises ou desenvolver suas próprias funções. Entre as desvantagens, o R utiliza uma interface de linha de comando que pode ser desconfortável para usuários leigos em linguagem de programação e sua curva de aprendizado é lenta no começo. Este documento fará uma breve apresentação de como o R pode ser usado para replicar os exemplos vistos nas aulas da disciplina de Econometria Aplicada. Ele está voltando sobretudo para aqueles alunos que não estão familiarizados com este software.

Download do R e do R Studio

<http://cran-r.c3sl.ufpr.br>

<https://www.rstudio.com/products/rstudio/download3/>

Exemplo: Regressão simples

Defina o diretório de trabalho do R, de modo que ele saiba onde salvar / procurar os seus dados. Atenção às barras: o R aceitará “/” ou “\”.

```
setwd("/Users/tainanboff/Documents/Graduação/Econometria aplicada/excelfiles")
```

Neste exemplo, utilizaremos a base de dados WAGE1 disponibilizada pelo autor Jeffrey M. Wooldridge. Para fazer o download de dados da internet, utilize a seguinte função:

```
download.file('http://fmwww.bc.edu/ec-p/data/wooldridge/wage1.dta', 'wage1.dta', mode='wb')
```

Como argumentos da função “download.file”, inserimos, entre aspas, a url onde os dados foram disponibilizados, o nome que este arquivo receberá em nosso computador e o modo como o arquivo será gravado (formato binário). Para mais detalhes sobre os argumentos de uma função, utilize a aba “help” do R Studio.

Agora que salvamos a base de dados no computador, o próximo passo consiste em “ler” o arquivo no R. Como o nosso arquivo possui a extensão .dta, que é um formato do Stata (outro software estatístico), precisaremos baixar um pacote específico no R, o qual possui uma função que permite ler arquivos neste formato.

Para instalar um pacote, utilizamos a função install.packages. Observe que escolhemos o servidor da UFPR, pois é o mais próximo de Porto Alegre, reduzindo a carga de rede.

```
install.packages("foreign", repos = "http://cran-r.c3sl.ufpr.br/")
```

Para carregar este pacote, podemos utilizar duas funções diferentes:

```
require(foreign) # ou
library(foreign)
```

Agora, estamos prontos para ler o arquivo. Como não queremos apenas visualizar a base de dados, mas sim utilizá-la para a análise de regressão, vamos criar um objeto chamado `wage1` e atribuir os dados a este objeto. Observe que ele irá aparecer no canto superior direito do R Studio (ambiente).

```
wage1 <- read.dta("wage1.dta")
```

Para que possamos chamar cada uma das variáveis que compõem essa base de dados pelo seu nome (ex.: `wage`, `married`, etc.), utilizamos a seguinte função:

```
attach(wage1)
```

O nome das variáveis está em inglês. Se quisermos alterá-los, podemos usar a seguinte função:

```
colnames(ceosal1) = c("nome da coluna 1", "nome da coluna 2", etc. )
```

Mas nesse exemplo, vamos utilizar os nomes originais.

A base `WAGE1` contém dados de corte transversal de 526 trabalhadores no ano de 1976. As variáveis incluem salários em dólares por hora (*wage*), anos de educação (*educ*), anos de experiência potencial da força de trabalho (*exper*), o tempo de permanência no mesmo emprego (*tenure*), uma variável dummy para sexo feminino (*female*), uma para estado civil igual a casado (*married*), etc.

Vamos iniciar reproduzindo o exemplo 2.4 de Wooldridge (2013), no qual para estudar a relação que existe entre os anos de educação e o salário por hora, podemos estimar o seguinte modelo de regressão linear:

$$wage = \beta_0 + \beta_1 educ + u$$

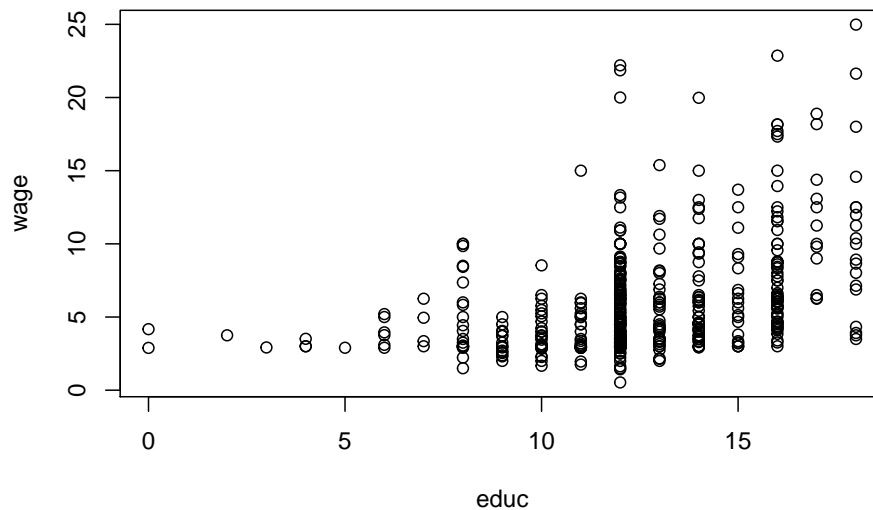
Podemos iniciar o nosso estudo analisando as estatísticas descritivas das variáveis (mínimo, primeiro quartil, mediana, média, terceiro quartil, máximo) e o gráfico com a nuvem de pontos:

```
summary(wage1)
```

```
##      wage      educ      exper      tenure
##  Min.   : 0.530   Min.   : 0.00   Min.    : 1.00   Min.    : 0.000
##  1st Qu.: 3.330   1st Qu.:12.00   1st Qu.: 5.00   1st Qu.: 0.000
##  Median : 4.650   Median :12.00   Median :13.50   Median : 2.000
##  Mean   : 5.896   Mean    :12.56   Mean    :17.02   Mean    : 5.105
##  3rd Qu.: 6.880   3rd Qu.:14.00   3rd Qu.:26.00   3rd Qu.: 7.000
##  Max.    :24.980   Max.    :18.00   Max.    :51.00   Max.    :44.000
##  nonwhite  female  married  numdep
##  Min.    :0.0000   Min.    :0.0000   Min.    :0.0000   Min.    :0.000
##  1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.000
##  Median :0.0000   Median :0.0000   Median :1.0000   Median :1.000
##  Mean    :0.1027   Mean    :0.4791   Mean    :0.6084   Mean    :1.044
##  3rd Qu.:0.0000   3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:2.000
##  Max.    :1.0000   Max.    :1.0000   Max.    :1.0000   Max.    :6.000
##  smsa      northcen  south      west
##  Min.    :0.0000   Min.    :0.0000   Min.    :0.0000   Min.    :0.0000
##  1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.0000
```

```
## Median :1.0000 Median :0.000 Median :0.0000 Median :0.0000
## Mean :0.7224 Mean :0.251 Mean :0.3555 Mean :0.1692
## 3rd Qu.:1.0000 3rd Qu.:0.750 3rd Qu.:1.0000 3rd Qu.:0.0000
## Max. :1.0000 Max. :1.000 Max. :1.0000 Max. :1.0000
## construc ndurman trcommpu trade
## Min. :0.00000 Min. :0.0000 Min. :0.00000 Min. :0.0000
## 1st Qu.:0.00000 1st Qu.:0.0000 1st Qu.:0.00000 1st Qu.:0.0000
## Median :0.00000 Median :0.0000 Median :0.00000 Median :0.0000
## Mean :0.04563 Mean :0.1141 Mean :0.04373 Mean :0.2871
## 3rd Qu.:0.00000 3rd Qu.:0.0000 3rd Qu.:0.00000 3rd Qu.:1.0000
## Max. :1.00000 Max. :1.0000 Max. :1.00000 Max. :1.0000
## services profserv profocc clerocc
## Min. :0.0000 Min. :0.0000 Min. :0.0000 Min. :0.0000
## 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.0000
## Median :0.0000 Median :0.0000 Median :0.0000 Median :0.0000
## Mean :0.1008 Mean :0.2586 Mean :0.3669 Mean :0.1673
## 3rd Qu.:0.0000 3rd Qu.:1.0000 3rd Qu.:1.0000 3rd Qu.:0.0000
## Max. :1.0000 Max. :1.0000 Max. :1.0000 Max. :1.0000
## servocc lwage expersq tenursq
## Min. :0.0000 Min. : -0.6349 Min. : 1.0 Min. : 0.00
## 1st Qu.:0.0000 1st Qu.: 1.2030 1st Qu.: 25.0 1st Qu.: 0.00
## Median :0.0000 Median : 1.5369 Median : 182.5 Median : 4.00
## Mean :0.1407 Mean : 1.6233 Mean : 473.4 Mean : 78.15
## 3rd Qu.:0.0000 3rd Qu.: 1.9286 3rd Qu.: 676.0 3rd Qu.: 49.00
## Max. :1.0000 Max. : 3.2181 Max. : 2601.0 Max. : 1936.00
```

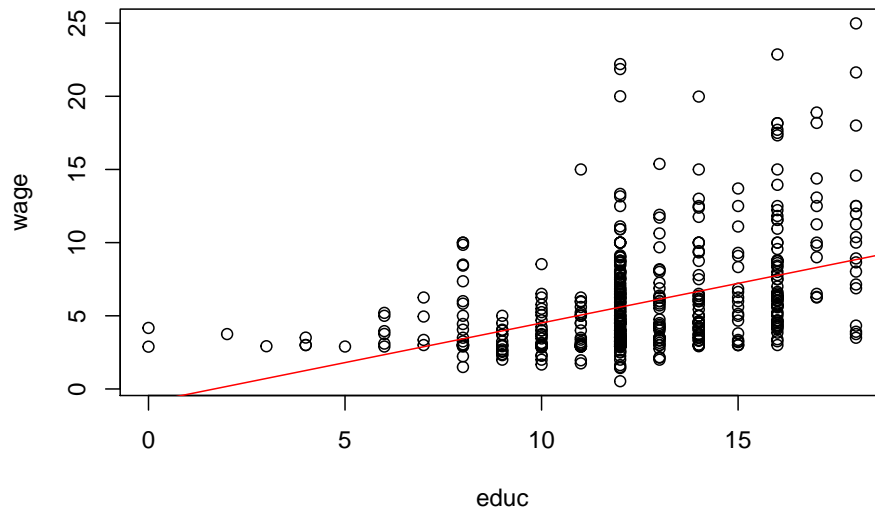
```
plot(educ, wage)
```



Para estimar um modelo de regressão no R, utilizamos a função `lm()`. Podemos estimar o modelo com ou sem intercepto e adicionar a reta de regressão à nuvem de pontos:

```
regressao1 <- lm(wage ~ educ)
regressao2 <- lm(wage ~ educ - 1) # ou
regressao2 <- lm(wage ~ 0 + educ)

plot(educ, wage)
abline(regressao1, col="red")
```



Vejamos um resumo dos resultados obtidos nos modelo com intercepto e sem intercepto:

```
summary(regressao1)
```

```
##
## Call:
## lm(formula = wage ~ educ)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5.3396 -2.1501 -0.9674  1.1921 16.6085
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.90485    0.68497  -1.321   0.187
## educ         0.54136    0.05325  10.167 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.378 on 524 degrees of freedom
## Multiple R-squared:  0.1648, Adjusted R-squared:  0.1632
## F-statistic: 103.4 on 1 and 524 DF, p-value: < 2.2e-16
```

```
summary(regressao2)
```

```
##
## Call:
## lm(formula = wage ~ 0 + educ)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5.142 -2.246 -1.066  1.154 16.528
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## educ  0.47266    0.01146   41.25 <2e-16 ***
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.381 on 525 degrees of freedom
## Multiple R-squared:  0.7642, Adjusted R-squared:  0.7637
## F-statistic: 1701 on 1 and 525 DF,  p-value: < 2.2e-16
```

```
anova(regressao1)
```

```
## Analysis of Variance Table
##
## Response: wage
##           Df Sum Sq Mean Sq F value    Pr(>F)
## educ         1 1179.7  1179.73   103.36 < 2.2e-16 ***
## Residuals  524  5980.7    11.41
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova(regressao2)
```

```
## Analysis of Variance Table
##
## Response: wage
##           Df Sum Sq Mean Sq F value    Pr(>F)
## educ         1 19445.7 19445.7  1701.3 < 2.2e-16 ***
## Residuals  525   6000.6    11.4
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Podemos obter um intervalo de confiança de 95% para o valor dos parâmetros e um intervalo de previsão para os valores ajustados:

```
confint(regressao1)
```

```
##           2.5 %    97.5 %
## (Intercept) -2.2504719 0.4407687
## educ         0.4367534 0.6459651
```

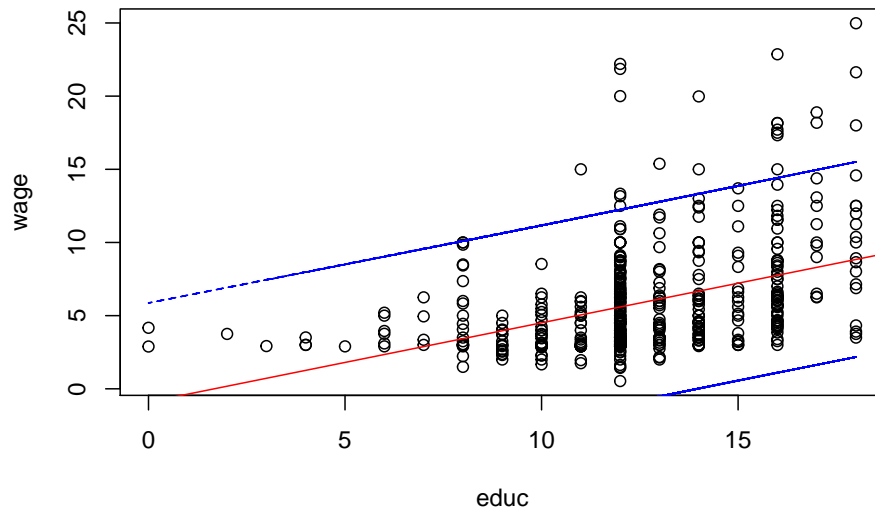
```
IP <- predict(regressao1, interval="predict")
```

```
## Warning in predict.lm(regressao1, interval = "predict"): predictions on current data refer to _future_
```

```
head(IP)
```

```
##           fit          lwr          upr
## 1 5.050100 -1.595068 11.69527
## 2 5.591459 -1.051958 12.23488
## 3 5.050100 -1.595068 11.69527
## 4 3.426022 -3.234258 10.08630
## 5 5.591459 -1.051958 12.23488
## 6 7.756896  1.104016 14.40978
```

```
plot(educ, wage)
abline(regressao1, col="red")
matlines(educ, IP[, c("lwr", "upr")], col = "blue")
```



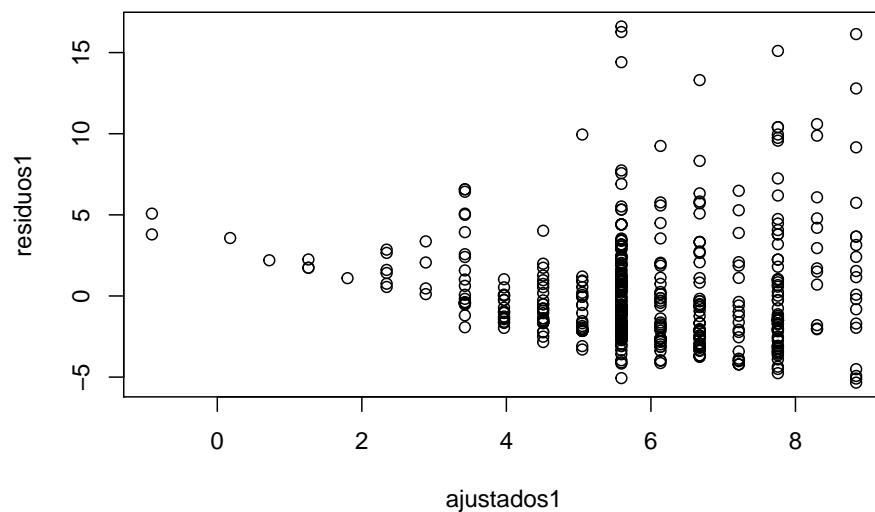
Vamos fazer um diagnóstico do modelo através de análise gráfica. Para investigar se a função de regressão é linear e se os resíduos apresentam variância constante, podemos fazer um gráfico dos resíduos amostrais contra o regressor ou contra os valores ajustados. Um bom modelo gera um gráfico em que os resíduos não apresentam um padrão e estão em torno de zero.

Em primeiro lugar, vamos salvar os resíduos do modelo e os valores ajustados:

```
residuos1 <- regressao1$residuals
ajustados1 <- regressao1$fitted.values
```

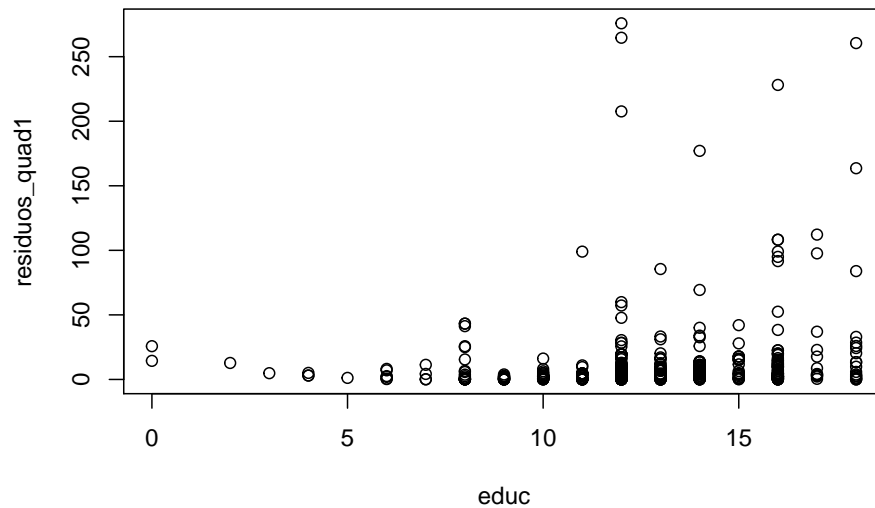
E, então, imprimir os gráficos:

```
plot(ajustados1, residuos1)
```



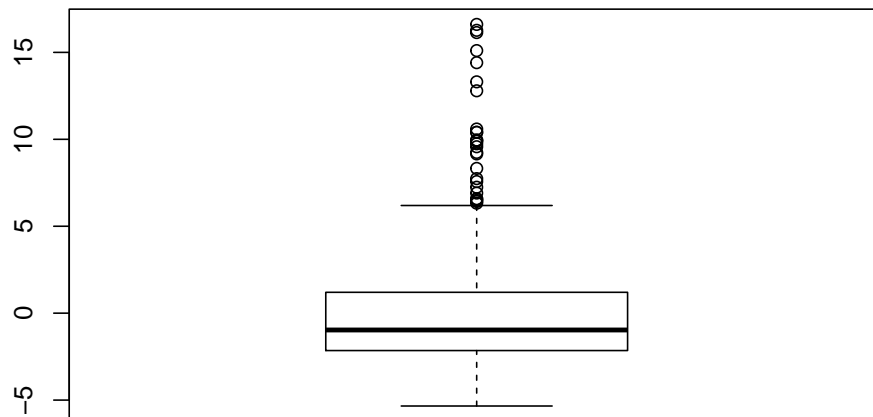
Ainda, para verificar se a sequência de resíduos apresenta variância constante, podemos imprimir um gráfico dos resíduos amostrais ao quadrado (ou em valor absoluto) contra o regressor:

```
residuos_quad1 = (residuos1)^2  
plot(educ, residuos_quad1)
```



A presença de *outliers* pode ser avaliada através de um Box-Plot dos resíduos:

```
boxplot(residuos1)
```



Uma forma simples de obtermos informações sobre possíveis *outliers* / pontos de alavancagem / observações influentes é utilizarmos a função `<influence.measures>`, que inclui os seguintes resultados:

- `dfbetas`: visa medir a influência de uma observação nas estimativas dos parâmetros;
- `dffits` e distância de Cook: medem o quanto o valor ajustado de y_i é afetado ao excluir a observação i do ajuste;
- `covariance ratios`: medem o efeito da exclusão de uma observação na variância das estimativas dos parâmetros;
- elementos diagonais da matriz H ou matriz chapéu (h_{ii}): indicam o quanto y_i afeta o y estimado.

Observações influentes com respeito a qualquer uma destas medidas são marcadas com um asterisco.

Como vimos em aula, se x_i , em uma regressão simples, é tal que $h_{ii} > \frac{4}{n}$, então x_i é um ponto de alavancagem, que pode ser bom ou mau.

```
influence_measures <- influence.measures(regressao1)
summary(influence_measures)
```

```
## Potentially influential observations of
##   lm(formula = wage ~ educ) :
##
##      dfb.1_ dfb.educ dffit    cov.r    cook.d hat
## 10  -0.17   0.21    0.24_*  0.98_*  0.03  0.01
## 15   0.09  -0.04    0.22_*  0.92_*  0.02  0.00
## 16  -0.12   0.16    0.20_*  0.98_*  0.02  0.00
## 38   0.10  -0.09    0.10    1.02_*  0.00  0.02_*
## 59  -0.29   0.33    0.37_*  0.96_*  0.07  0.01
## 66  -0.05   0.09    0.20_*  0.95_*  0.02  0.00
## 80   0.04  -0.04   -0.05    1.01_*  0.00  0.01
## 98   0.04  -0.02    0.10    0.99_*  0.00  0.00
## 105 -0.03   0.06    0.12    0.98_*  0.01  0.00
## 107  0.04  -0.02    0.10    0.99_*  0.01  0.00
## 112 -0.37   0.42    0.47_*  0.93_*  0.11  0.01
## 131  0.00   0.00    0.00    1.01_*  0.00  0.01
## 139  0.19  -0.18    0.19_*  1.03_*  0.02  0.03_*
## 160  0.04  -0.05   -0.06    1.01_*  0.00  0.01
## 165  0.00   0.00   -0.01    1.01_*  0.00  0.01
## 170  0.01   0.02    0.12    0.98_*  0.01  0.00
## 186  0.09  -0.04    0.22_*  0.92_*  0.02  0.00
## 229 -0.20   0.25    0.32_*  0.93_*  0.05  0.00
## 245 -0.14   0.17    0.22_*  0.97_*  0.02  0.00
## 252  0.09  -0.09    0.10    1.01_*  0.00  0.01_*
## 260 -0.20   0.24    0.26_*  0.98_*  0.03  0.01
## 278 -0.19   0.22    0.26_*  0.97_*  0.03  0.01
## 298  0.05  -0.05    0.05    1.02_*  0.00  0.01_*
## 305  0.02  -0.02    0.02    1.02_*  0.00  0.01_*
## 306  0.07  -0.07    0.07    1.02_*  0.00  0.02_*
## 309  0.00   0.00    0.00    1.01_*  0.00  0.01
## 311  0.02  -0.02   -0.02    1.01_*  0.00  0.01
## 326 -0.14   0.17    0.22_*  0.97_*  0.02  0.00
## 330  0.01  -0.01    0.01    1.01_*  0.00  0.01
## 343  0.10  -0.07    0.15    0.97_*  0.01  0.00
## 379  0.32  -0.31    0.32_*  1.04_*  0.05  0.04_*
## 396  0.05  -0.04    0.05    1.02_*  0.00  0.01_*
## 397  0.03  -0.02    0.03    1.02_*  0.00  0.01_*
## 431 -0.03   0.04    0.04    1.01_*  0.00  0.01
## 433 -0.03   0.03    0.03    1.01_*  0.00  0.01
## 440  0.08  -0.04    0.19_*  0.94_*  0.02  0.00
## 465  0.10  -0.10    0.10    1.03_*  0.01  0.02_*
## 470  0.07  -0.07    0.07    1.02_*  0.00  0.02_*
## 476 -0.13   0.16    0.20_*  0.98_*  0.02  0.00
## 488  0.09  -0.08    0.09    1.01_*  0.00  0.01_*
## 503  0.24  -0.23    0.24_*  1.04_*  0.03  0.04_*
## 504  0.04  -0.04    0.04    1.02_*  0.00  0.02_*
```



```
## 505 -0.13  0.16      0.21_*  0.98_*  0.02  0.00
## 514  0.06 -0.05      0.06    1.01_*  0.00  0.01
```

Para detectar um mau ponto de alavancagem, usamos os resíduos padronizados. Os resíduos padronizados nos dizem quantos desvios-padrão estimados um dado ponto se encontra distante da reta estimada. Como regra de bolso, dizemos que, em uma amostra grande, um mau ponto de alavancagem tem resíduo padronizado fora do intervalo $(-4, 4)$.

```
rstandard <- rstandard(regressao1)
summary(rstandard)
```

```
##      Min.   1st Qu.   Median     Mean   3rd Qu.     Max.
## -1.588000 -0.637200 -0.287100  0.000376  0.353300  4.921000
```

```
badleverage <- which(rstandard >= 4)
badleverage
```

```
##  15 112 186 229 440
##  15 112 186 229 440
```

Podemos explorar estes pontos comparando seus valores com as estatísticas descritivas da amostra:

```
wage1[c(15, 112, 186, 229, 440),c(1, 2)]
```

```
##      wage educ
## 15  22.20   12
## 112 24.98   18
## 186 21.86   12
## 229 22.86   16
## 440 20.00   12
```

```
summary(wage)
```

```
##      Min. 1st Qu.  Median     Mean 3rd Qu.     Max.
##  0.530   3.330   4.650   5.896   6.880   24.980
```

```
summary(educ)
```

```
##      Min. 1st Qu.  Median     Mean 3rd Qu.     Max.
##   0.00   12.00   12.00   12.56   14.00   18.00
```

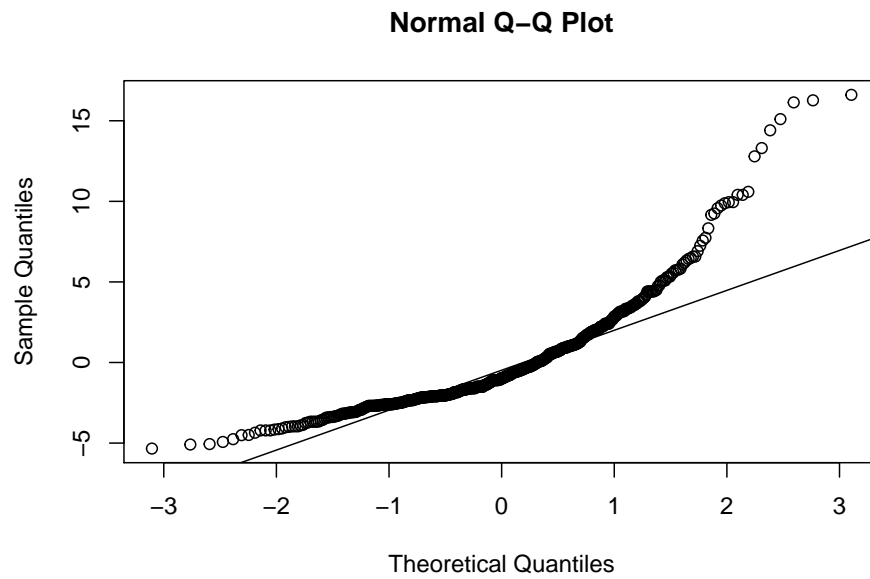
E se quisermos excluir os maus pontos de alavancagem?

```
wage2 <- wage1[-c(15, 112, 186, 229, 440), ]
# attach(wage2)
```

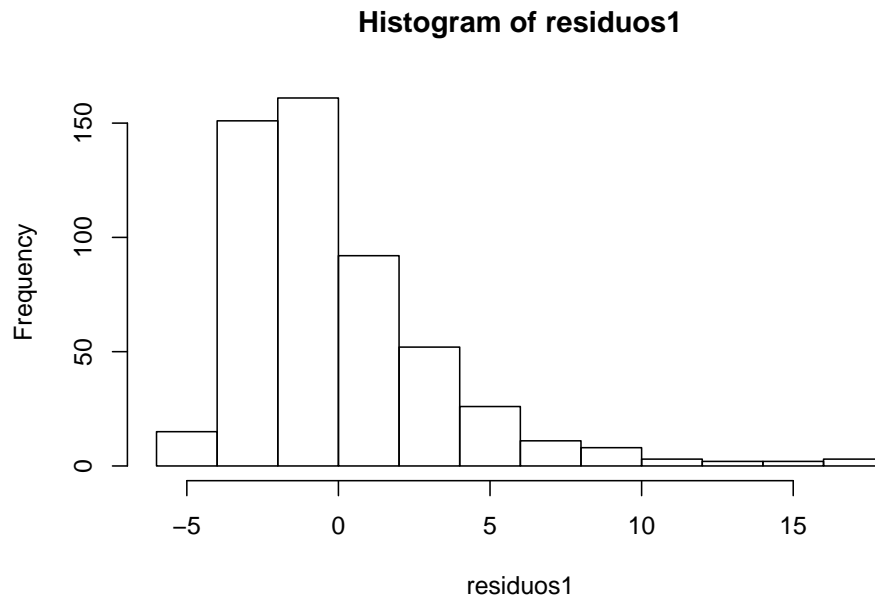
Wage2 é uma subamostra de wage1, da qual foram retiradas 5 linhas e nenhuma coluna.

Para testar se os resíduos são (aproximadamente) normalmente distribuídos, podemos usar o QQ-Plot, ou Gráfico de Distribuição Normal, e o histograma.

```
qqnorm(residuos1)
qqline(residuos1)
```



```
hist(residuos1)
```



Em nosso exemplo, cada ano adicional de educação acarreta um acréscimo de 54 cents ao salário/hora, independentemente do total de anos de educação do indivíduo. Provavelmente, o problema fica melhor caracterizado em termos de uma variação percentual do salário/hora. Um modelo que retorna um efeito percentual constante é:

```
regressao3 <- lm(lwage ~ educ)
summary(regressao3)
```

```
##
## Call:
## lm(formula = lwage ~ educ)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.21158 -0.36393 -0.07263  0.29712  1.52339
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.583773    0.097336   5.998 3.74e-09 ***
## educ         0.082744    0.007567  10.935 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4801 on 524 degrees of freedom
## Multiple R-squared:  0.1858, Adjusted R-squared:  0.1843
## F-statistic: 119.6 on 1 and 524 DF,  p-value: < 2.2e-16
```

Podemos incluir mais variáveis explicativas e estimar um modelo de regressão múltipla. Novamente, iremos salvar os resíduos para realizar diagnósticos do modelo.

```
regressao4 <- lm(lwage ~ educ + exper + tenure)
summary(regressao4)
```

```
##
## Call:
## lm(formula = lwage ~ educ + exper + tenure)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.05802 -0.29645 -0.03264  0.28788  1.42809
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.284360    0.104190   2.729  0.00656 **
## educ         0.092029    0.007330  12.555 < 2e-16 ***
## exper        0.004121    0.001723   2.391  0.01714 *
## tenure       0.022067    0.003094   7.133 3.29e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4409 on 522 degrees of freedom
## Multiple R-squared:  0.316, Adjusted R-squared:  0.3121
## F-statistic: 80.39 on 3 and 522 DF,  p-value: < 2.2e-16
```

```
anova(regressao4)
```

```
## Analysis of Variance Table
##
## Response: lwage
##           Df Sum Sq Mean Sq F value    Pr(>F)
```

```
## educ      1  27.561 27.5606 141.802 < 2.2e-16 ***
## exper     1   9.424  9.4244  48.490 1.002e-11 ***
## tenure    1   9.889  9.8891  50.881 3.294e-12 ***
## Residuals 522 101.456  0.1944
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
residuos4 <- regressao4$residuals
```

Existem diversos testes para heterocedasticidade. Entre os quais, o teste de Goldfeld-Quandt e o teste de Breusch-Pagan. Para realizar estes testes, precisaremos instalar/carregar alguns pacotes no R:

```
require(lmtest)
gqtest(regressao4, point = 0.5, fraction = 0)
```

```
##
## Goldfeld-Quandt test
##
## data: regressao4
## GQ = 0.69448, df1 = 259, df2 = 259, p-value = 0.9983
```

```
bptest(regressao4)
```

```
##
## studentized Breusch-Pagan test
##
## data: regressao4
## BP = 10.761, df = 3, p-value = 0.01309
```

Os argumento “point” da função <gqtest> pode ser interpretado como o percentual de observações do começo e do final da amostra que serão usados para a comparação das variâncias. “Fraction” refere-se ao percentual de observações centrais a serem omitidas.

Obs.: O teste de Goldfeld-Quandt tem hipótese nula de homocedasticidade. O teste de Breusch-Pagan tem hipótese nula de heteroscedasticidade.

Para investigar a existência de multicolinearidade, podemos visualizar a matriz de correlação das variáveis explicativas:

```
x = wage1[, c(2,3,4)]
head(x)
```

```
##   educ exper tenure
## 1   11     2      0
## 2   12    22      2
## 3   11     2      0
## 4    8    44     28
## 5   12     7      2
## 6   16     9      8
```

```
cor(x)
```

```
##          educ      exper      tenure
## educ      1.00000000 -0.2995418 -0.05617257
## exper -0.29954184  1.00000000  0.49929145
## tenure -0.05617257  0.4992914  1.00000000
```

Aqui encerramos este exemplo, que procurou cobrir todos os testes apresentados em aula até o momento. Nosso objetivo não foi explorar a qualidade do ajuste do modelo aos dados, mas sim apresentar algumas funções que serão úteis para realizar análise de regressão usando o software R.

References

Wooldridge, Jeffery M. 2013. *Introductory Econometrics : A Modern Approach*. 5th ed. South-Western Cengage Learning.