

Modelos de Regressão e Previsão

Lista 1

Prof. Carlos Trucíos
carlos.trucios@facc.ufrj.br
ctruciosm.github.io

Questão 1

Sejam $\hat{\beta}_0$ e $\hat{\beta}_1$ os estimadores MQO da regressão y sobre x . Mostre que os estimadores MQO da regressão cy sobre x são $c\hat{\beta}_0$ e $c\hat{\beta}_1$, respectivamente.

Questão 2

Utilizando o software R, rode o seguinte código:

```
library(MASS)
beta0 = 0.2
beta1 = 0.5
mSigma = matrix(c(1,0,0,1),2)
vMu = c(0,0)
dados_simulados = mvrnorm(2000, mu = vMu, Sigma = mSigma)
x = dados_simulados[,1]
u = dados_simulados[,2]
y = beta0 + beta1*x + u
lm(y~x)
```

- Os valores de $\hat{\beta}$ são próximos dos valores de β ?
- E se `mSigma = matrix(c(1,0.3,0.3,1),2)`
- E se `mSigma = matrix(c(1,-0.5,-0.5,1),2)`
- E se `mSigma = matrix(c(1,0.7,0.7,1),2)`
- Segundo os resultados obtidos, a que conclusões poderia chegar?

Questão 3

Mostre que $\sum_{i=1}^n \hat{u}_i^2/n$ é um estimador viesado para σ^2

Questão 4

Seja a regressão linear simples através da origem, *i.e.* $y = \beta_1 x + u$. Derive o estimador MQO $\hat{\beta}_1$

Questão 5 (C2 do livro texto, pag 65)

O conjunto de dados **CEOSAL2.txt** contém informações sobre CEOs de empresas dos US. A variável *salary* é a compensação anual (em milhares de USD) e a variável *ceoten* é o número prévio (em anos) como CEO da empresa.

- Encontre o salário médio e a permanência média dos CEOs
- Quantos CEOs estão no seu primeiro ano na empresa? (*ceoten* = 0)
- Calcule a regressão $\log(\text{salary}) = \beta_0 + \beta_1 \text{ceoten} + u$
- Qual o aumento percentual previsto no salário dos CEOs se tem um ano a mais como CEO na empresa?

Questão 6

O conjunto de dados **WAGE1.txt** contém informações pessoas na força de trabalho em 1996. A variável *wage* é o salário médio por hora (em USD) e a variável *expert* é o número de anos de experiencia.

- Faça uma análise exploratória de dados de ambas as variáveis
- Construa um gráfico de dispersão e plote a reta de regressão
- Calcule a regressão $wage = \beta_0 + \beta_1 expert + u$
- Interprete os resultados
- Calcule a regressão $\log(wage) = \beta_0 + \beta_1 expert + u$
- Interprete os resultados

Questão 7

O conjunto de dados **catholic** contém informações de pontuações de testes de estudantes dos U.S. que cursaram a oitava série em um determinado ano. As variáveis *mate12* e *read12* são notas padronizadas de matemática e leitura respectivamente.

- Quanto estudantes existem na amostra? Encontre as médias e desvio padrão de cada variável.
- Construa um boxplot de ambas variáveis. O que poderia dizer ao olhar o gráfico?
- Calcule a regressão *mate12* sobre *read12* ($mate12 = \beta_0 + \beta_1 read12 + u$). Como é a qualidade do ajuste do modelo?
- Interprete os resultados
- Faça um graficos de *mate12* vs. $\widehat{mate12}$

Nota: o conjunto de dados **catholic** esta no pacote *wooldridge* do **R**

```
install.packages("wooldridge")
library(wooldridge)
data(catholic)
str(catholic)
```