

Modelos de Regressão e Previsão

Lista 2

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

- Caso se indique o contrário, considere sempre um nível de significância de $\alpha = 0.05$
- Os comandos *summary()*, *confint()*, *anova()* e *predict()* lhe ajudarão a responder as perguntas
- Verifique o pacote *wooldridge* está instalado e carregado

```
install.packages("wooldridge")  
library(wooldridge)
```

Questão 1

No seguinte modelo:

```
modelo = lm(log(bwght) ~ npvis + I(npvis^2), data = bwght2)
```

- As variáveis são estatisticamente significativas?
- Interprete os betas
- Calcule intervalos de confiança 90% para os betas

Questão 2

No seguinte modelo:

```
modelo = lm(log(wage)~educ+exper+tenure+nonwhite+female+married+numdep, data = wage1)
```

- Qual percentagem da variabilidade de $\log(wage)$ é explicada pelo modelo?
- Quais variáveis são estatisticamente significativas?
- Teste a hipótese: $H_0 : \beta_{nonwhite} = 0, \beta_{numdep} = 0$ vs. $H_1 : \beta_{nonwhite} \neq 0, \beta_{numdep} \neq 0$. Podemos rejeitar H_0 ?
- No modelo de regressão original, inclua o termo de interação $nonwhite*female*married$. Alguma interação é estatisticamente significativa?

Questão 3

Sejam os seguintes modelos:

```
modelo1 = lm(log(wage)~educ+exper+tenure+nonwhite+female+married+numdep, data = wage1)  
modelo2 = lm(log(wage)~educ+exper+tenure+female+married + smsa + construc +  
             ndurman + trade + services + profserv + profocc + servocc, data = wage1)  
modelo3 = lm(log(wage)~educ+exper+ I(exper^2) + tenure+female+married +  
             smsa + trade + services + profocc + servocc + female*married, data = wage1)
```

- Se seu interesse é explicar a variabilidade de $\log(wage)$, qual modelo escolheria? (porquê?)
- Se seu interesse é explicar a variabilidade de $wage$, qual modelo escolheria? (porquê?)

Questão 4

No seguinte modelo:

```
modelo = lm(bwght~. -moth -foth, data = bwght2)
```

- Qual percentagem da variabilidade de *bwght* é explicada pelo modelo?
- Observa alguma coisa errada na modelagem?
- Reestime o modelo excluindo também a variável *lbwght*
- Qual percentagem da variabilidade de *bwght* é explicada pelo modelo?
- Quais variáveis são estatisticamente significativas?
- Explique essa grande diferença na percentagem de variabilidade explicada pelos modelos.

Questão 5

No seguinte modelo:

```
modelo = lm(lavgsal ~ bs, data = benefits)
```

- Seja $H_0 : \beta_{bs} = 0$ vs $H_1 : \beta_{bs} \neq 0$, podemos rejeitar H_0 ?
- Teste a hipótese $H_0 : \beta_{bs} = -1$ vs $H_1 : \beta_{bs} > -1$, podemos rejeitar H_0 ?
- Estime o modelo

$$\text{lavgsal} = \beta_0 + \beta_1 \text{bs} + \beta_2 \text{lenroll} + \beta_3 \text{lstaff} + \beta_4 \text{lunch} + u$$

- Quais variáveis são estatisticamente significativas?
- Interprete os betas
- Calcule intervalos de confiança 95% para os betas
- adicione a variável *lunch*² e verifique se a qualidade do ajuste do modelo melhorou ou não.

Para exercícios adicionais, veja as Seções **Exercícios em computador** dos Capítulos 4, 6 e 7 do livro texto.