

# ECONOMETRIA I - LISTA 4

## MODELO DE REGRESSÃO LINEAR MÚLTIPLA: INFERÊNCIA

Mateus Cardoso

28/06/2021

**1)** (Baseado em Wooldridge, Cap. 4) Quais são as hipóteses do modelo linear clássico (MRLC)?

**2)** (Wooldridge, Cap. 4) Quais dos seguintes itens podem fazer com que as estatísticas  $t$  de MQO não sejam válidas (isto é, que não tenham distribuições  $t$  sob  $H_0$ )?

- (i) Heteroscedasticidade.
- (ii) Um coeficiente de correlação de 0,95 entre duas variáveis independentes que estão no modelo.
- (iii) Omitir uma variável explicativa importante.

**3)** As taxas de aluguel são influenciadas pela população de estudantes em uma cidade onde há universidades? Seja *rent* o aluguel médio mensal pago pela unidade alugada em uma determinada cidade, onde há universidades. Seja *pop* o total da população da cidade, *avginc*, a renda média da cidade e *pctstu*, a população de estudantes como um percentual da população total. Um modelo para testar uma relação é

$$\log(\text{rent}) = \beta_0 + \beta_1 \log(\text{pop}) + \beta_2 \log(\text{avginc}) + \beta_3 \text{pctstu} + u.$$

- (i) Formule a hipótese nula de que o tamanho da população estudantil relativo à população das cidades não tem efeito *ceteris paribus* sobre os aluguéis mensais. Formule a alternativa de que há um efeito.
- (ii) Quais sinais você espera para  $\beta_1$  e  $\beta_2$ ?
- (iii) A equação estimada, usando dados de 1990 de 64 cidades com universidades do arquivo RENTAL, é

$$\widehat{\log(\text{rent})} = 0,043 + 0,066 \log(\text{pop}) + 0,507 \log(\text{avginc}) + 0,0056 \text{pctstu}$$

$$\begin{array}{cccc} (0,844) & (0,039) & (0,081) & (0,0017) \\ n = 64, & R^2 = 0,458. \end{array}$$

que está errado com a seguinte afirmação: “Um aumento de 10% na população está associado a um aumento de cerca de 6,6% no aluguel”?

(iv) Teste a hipótese formulada no item (i) no nível de 1%.

4) (Wooldridge, Cap. 4) Considere a equação estimada do Exemplo 4.3 que poderia também ser usada para estudar os efeitos de faltar às aulas sobre a GPA em curso superior:

$$\begin{array}{cccc} \widehat{colGPA} = 1,39 + 0,412hsGPA + 0,15ACT - 0,083skipped \\ (0,33) & (0,094) & (0,011) & (0,026) \\ n = 141, & R^2 = 0,234. \end{array}$$

- (i) Usando a aproximação normal padronizada, encontre o intervalo de confiança de 95% para  $\beta_{hsGPA}$ .
- (ii) Você pode rejeitar a hipótese  $H_0 : \beta_{hsGPA} = 0,4$  contra a hipótese alternativa bilateral no nível de 5%?
- (iii) Você pode rejeitar a hipótese  $H_0 : \beta_{hsGPA} = 1$  contra a hipótese alternativa bilateral no nível de 5%?

## Exercícios no R

1) (Wooldridge, Cap. 4) O modelo a seguir pode ser usado para estudar se gastos de campanha afetam os resultados das eleições:

$$voteA = \beta_0 + \beta_1 \log(expendA) + \beta_2 \log(expendB) + \beta_3 prtystroA + u,$$

em que  $voteA$  é a porcentagem de votos recebidos pelo Candidato A,  $expendA$  e  $expendB$  são os gastos de campanha dos Candidatos A e B, e  $prtystroA$  é uma medida da força do partido do Candidato A (uma porcentagem dos votos presidenciais mais recentes que foram para o partido A).

- (i) Qual é a interpretação de  $\beta_1$ ?
- (ii) Em termos de parâmetros, defina a hipótese nula de que um aumento de 1% nos gastos de A seja compensado por um aumento de 1% nos gastos de B.

- (iii) Estime o modelo dado usando as informações do arquivo VOTE1 e registre os resultados na forma usual. Os gastos do Candidato A afetam os resultados? E os gastos do Candidato B? Você poderia usar esses resultados para testar a hipótese do item (ii)?
- (iv) Estime um modelo que dê diretamente a estatística t para testar a hipótese do item (ii). O que você conclui? (Use uma alternativa bilateral.)

**2)** (Wooldridge, Cap. 4) O conjunto de dados do arquivo 401KSUBS contém informações sobre os ativos financeiros líquidos (*nettfa*), idade dos respondentes da pesquisa (*age*), renda familiar anual (*inc*), tamanho da família (*fsize*) e participação em determinados planos de pensão para pessoas dos Estados Unidos. As variáveis de patrimônio e renda são registradas em milhares de dólares. Para essa questão, use somente os dados de residências com apenas uma pessoa (assim, *fsize* = 1).

- (i) Quantas residências com apenas uma pessoa existem no conjunto de dados?
- (ii) Use os MQO para estimar o modelo

$$nettfa = \beta_0 + \beta_1 inc + \beta_2 age + u,$$

e reporte os resultados usando o formato usual. Certifique-se de usar somente residências com apenas uma pessoa na amostra. Interprete os coeficientes de inclinação. Houve alguma surpresa em relação às estimativas de inclinação?

- (iii) O intercepto da regressão do item (ii) teve um significado interessante? Explique.
- (iv) Encontre o p-valor do teste  $H_0 : \beta_2 = 1$  contra  $H_1 : \beta_2 < 1$ . Você rejeita  $H_0$  ao nível de significância de 1%?
- (v) Se você fizer uma regressão simples de *nettfa* sobre *inc*, o coeficiente estimado de renda seria muito diferente do estimado do item (ii)? Por que sim ou por que não?