



Lego II - Modelo de Regressão Linear

Semestre: 2020/2

Thiago Moreira da Silva
thiagomoreira@iesp.uerj.br

Introdução

Este curso tem por objetivo fornecer um panorama do modelo de regressão linear e suas aplicações nas ciências sociais. Serão debatidos exemplos práticos, relacionados a temas atinentes à ciência política e à sociologia, bem como técnicas que autorizam o pesquisador a fazer previsões com bases amostrais. De início, os alunos terão contato com uma revisão de conhecimentos estatísticos básicos, como a probabilidade, o teste de hipóteses e os intervalos de confiança. Em seguida, fornecemos as bases teóricas para a compreensão da modelagem estatística e aprofundamos o conteúdo em torno de mecanismos mais sofisticados para a análise da regressão linear.

Pré-requisitos

Alunos que tiveram algum contato anterior com noções básicas de estatística e dominem alguma técnica de programação, especialmente em R, aproveitarão melhor o curso. No entanto, essas condições não são necessárias aos interessados. Isso posto, nas primeiras aulas faremos uma revisão dos preceitos estatísticos mais básicos.

Software

Neste curso usaremos o software R, a ferramenta mais popular de análise de dados para cientistas sociais atualmente. O *software* pode ser baixado gratuitamente em <https://cran.r-project.org/bin/windows/base/>. Recomendamos que os alunos façam o *download* da interface gráfica Rstudio, que facilita a operacionalização do software. O download pode ser feito em <https://www.rstudio.com/products/rstudio/download/>.

Avaliação

Com o objetivo de fixar os conhecimentos transmitidos, a avaliação será composta por listas de exercícios, enviadas aos alunos quinzenalmente.

Leituras

No curso, usaremos os seguintes livros:

Angrist, Joshua D. e Pischke, Jorn-Steffen, 2014. *Mastering Metrics: The Path from Cause to Effects*. The Princeton University Press.

Fox, John (1997). *Applied Regression Analysis, Linear Models, and Related Methods*. NY: Sage.

Fox, John (2015). *Applied regression analysis and generalized linear models*. Sage Publications.

Gelman, Andrew e Hill, Jennifer. 2007. *Data Analysis Using Regression and Multilevel/ Hierarchical Models*. Cambridge University Press.

Gujarati, Damodar e Porter, Dawn (2009). *Basic Econometrics*. The McGraw-Hill.

Monogan, J. E. (2015). *Political analysis using R*. Springer.

Morgan, Stephen e Winship, Christopher (2015). *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. The Cambridge University Press.

Wheelan, Charles (2013). Naked Statistics: Stripping the Dread from the Data. W.W. Norton.

Wooldridge, Jeffrey M (2015). Introductory Econometrics. New York: South-Western. 5th edition.

Programas das Aulas

**O cronograma pode ser modificado ao longo do curso.*

Aula 1: Introdução

1. Apresentação do curso.
2. Retomada de conceitos básicos.
3. Breve introdução ao R.

Aula 2: Regressão Linear Simples

1. Associação entre variáveis.
2. Expectância condicional.
3. Mínimos Quadrados Ordinários.

Leituras obrigatórias:

- . Wooldridge, cap. 1.
- . Agresti, cap. 12.

Leituras complementares:

- . Angrist e Pischke: capítulo 2.
- . Wheelan, cap. 11.

Aula 3: Regressão Multivariada

1. Adição de variáveis contínuas.
2. Variável omitida e multicolinearidade.

Leituras obrigatórias:

- . Wooldridge, cap. 3
- . Fox, cap. 7, pp. 128-141.

Leituras complementares:

- . Gujarati and Porter (Essentials of Econometrics), pp. 102-104; 113-116, pp. 116-131, 160-163.
- . Wheelan, cap. 12 e 13.

Aula 4: Pressupostos Básicos do Modelo Linear

1. Distribuição amostral do estimador MQO,
2. Variância amostral do estimador MQO,
3. Propriedades do MQO para amostras grandes,
4. Inferência para MQO,
5. Ajuste do modelo.

Leituras obrigatórias:

- . Wooldridge, capítulo 2.
- . Gujarati, caps. 3 e 4.

Leituras complementares:

- . Agresti, cap. 12.

Aula 5 - Estimação de dados deficientes (I)

1. Não-linearidade,
2. Transformações,
3. Heteroscedasticidade,
4. *Outlier* e casos influentes.

Leituras obrigatórias:

- . Fox (2015), caps.: 12 e 13.
- . Wooldridge, cap. 3.

Aula 6 - Estimação de dados deficientes (II)

1. Diagnósticos e Soluções.

Leituras obrigatórias:

- . Fox (2015), caps.: 11.
- . Wooldridge, cap. 3.

Aula 7 - Interpretação de Variáveis Categóricas

1. Adição de variáveis binárias,
2. Adição de variáveis multinomiais.

Leituras obrigatórias:

- . Fox (2015), caps.: 7, pp. 128-141.
- . Wooldridge, cap. 3.

Aula 8 - Interações

1. Interações entre variáveis,
2. Formas funcionais não-lineares.

Leituras obrigatórias:

- . Wooldridge, cap. 6.2, 8 e 9.5.
- . Fox (2015), cap. 7, pp. 140-150.

Aula 9 - Visualização dos Efeitos Preditos

1. Variáveis contínuas,
2. Variáveis binárias,
3. Interações.

Aula 10 - Variáveis Instrumentais

1. Utilização de variáveis instrumentais,
2. Endogeneidade.

Leituras obrigatórias:

- . Angrist and Pischke, Cap. 4.
- . Gujarati, cap. 9.
- . Gujarati and Porter, cap. 13, seção 13.5; Cap. 20.

Leitura complementar:

Acemoglu, Johnson, and Robinson. 2001. The Colonial Origins of Comparative Development: An Empirical Investigation. The American Economic Review 91(5): 1369-1401. URL: <http://goo.gl/57CUYJ>.

Aula 11 - Modelos Generalizados

1. Regressão logística,
2. Interpretação dos coeficientes,
3. Diagnósticos e comparações.

Leitura obrigatória:

Wooldridge, cap. 17.

Fox (2015), cap. 14