

Lego 2: Modelos de Regressão Linear e extensões

2º Semestre de 2021

Professor:

Rogério Jerônimo Barbosa
rogerio.barbosa@iesp.uerj.br

Monitora:

Luana Calzavara
luacalzvara@iesp.uerj.br

Professor Convidado para o Nivelamento:

Vinícius Pinheiro Israel (UniRio)
visrael@gmail.com

1. Apresentação e objetivos

Este curso tem por objetivo fornecer um panorama de diversos Modelos de Regressão e suas aplicações nas ciências sociais.

Na Parte 1, abordaremos o Modelo de Regressão Linear Múltipla. Revisaremos suas características básicas nas primeiras aulas, aprofundando a compreensão de seus pressupostos e propriedades matemáticas.

Na Parte 2, apresentamos extensões que permitem a aplicação para fenômenos não lineares e mais complexos, nos quais variáveis interagem entre si e modificam os efeitos umas das outras.

Na Parte 3, discutimos inferência causal e apresentamos modelos e estratégias que permitem aferir causalidade a partir de dados observacionais.

Na Parte 4, abordaremos modelos para dados longitudinais (i.e. nos quais observamos unidades analíticas ao longo do tempo).

2. Pré-requisitos

Este é um curso de nível intermediário no percurso de formação em métodos quantitativos. Assumimos que você:

- Já sabe estatística básica (medidas de tendência central, dispersão e associação; noções de probabilidade; noções de testes de hipótese e inferência estatística)
- Já estudou ou aplicou, ainda que rapidamente/superficialmente, modelos de regressão linear
- Já trabalha com softwares estatísticos que são operados por linha de comando (como R, Python ou Stata)

Se você cursou Lego 1, já tem tudo o que precisa ;-)

3. Avaliação prévia de alunos externos ou que não cursaram Lego 1

Se você não cursou Lego 1 ou está vindo de outra instituição, deverá fazer uma *pequena* lista de exercícios prévia, que tem a finalidade de prepará-lo para as aulas, fazendo-o revisar (e aprender) as matérias essenciais para o acompanhamento do curso.

Completá-la é obrigatório, mas sua pontuação será extra.

4. Software

Como em Lego 1, neste curso também utilizaremos o software R, que é livre, gratuito e aberto.

5. Logística do Curso

O curso é baseado em aulas expositivas, sessões de monitoria e realização de listas de exercício. As aulas serão ministradas online, por meio da plataforma Zoom.

6. Formas de Avaliação

- Leitura dos textos obrigatórios (10 pontos)
- Listas de Exercício Quinzenais (6 listas, 10 pontos cada)
- Trabalho Final (30 pontos)

7. Horários das aulas e monitorias

Lego 2 ocorre às terças-feiras, de 9hs às 12hs.

Haverá monitoria uma vez por semana, em horário a ser divulgado.

8. Nivelamento em Matemática para Lego 2

Uma semana antes do início das aulas, entre os dias 2 e 6 de agosto, de 9hs às 12hs, será ministrado um Nivelamento em Matemática para todos os alunos de Lego 2. A frequência a esse curso é obrigatória. Haverá duas listas de exercício, que renderão, cada uma, 5 pontos extras para a disciplina principal (10 pontos no total). Não há pré-requisitos.

O Nivelamento visa familiarizar o aluno com notações e ferramentas que serão imprescindíveis para o devido acompanhamento dos conteúdos. Lego 2 tem um caráter bastante prático e aplicado à pesquisa - no entanto, demonstrações e formas de argumentação exigem o domínio elementos de Matemática Superior.

Conteúdo do Curso de Nivelamento em Matemática para Lego 2:

Dias 2 a 6 de agosto, 9hs às 12hs

Aula 1: Fundamentos de Álgebra Linear (Parte 1) - 2 de agosto

- Vetores e Matrizes
- Sistemas de Equações Lineares e solução por eliminação
- Produto escalar e ortogonalidade

Aula 2: Fundamentos de Álgebra Linear (Parte 2) - 3 de agosto

- Matrizes Especiais: triangulares, diagonais, Identidade
- Determinantes
- Matrizes Inversas

Aula 3: Fundamentos de Álgebra Linear (Parte 3) - 4 de agosto

- Matrizes Inversas e a Solução de Sistemas de Equação
- Regressão linear múltipla: uma visão matricial

Aula 4: Fundamentos de Cálculo (Parte 1) - 5 de agosto

- Funções
- Uma visão geral sobre o Cálculo: Limites, Derivadas e Integrais

Aula 5: Fundamentos de Cálculo (Parte 2) - 6 de agosto

- Regras de Derivação

Parte 1: Regressão Linear, o básico

Aula 1: Regressão Linear Múltipla - 10 de agosto

- Revisão:
 - Regressão como média (esperança) condicional
 - Leitura dos coeficientes
 - Coeficiente de Determinação (R^2)
 - Inferência Estatística

Aula 2: Estimadores da Regressão Linear, duas abordagens: o Método dos Mínimos Quadrados Ordinários e o Método dos Momentos - 17 de agosto

- Regressão como “melhor” ajuste aos dados
- Regressão como expressão de um processo gerador de dados
 - Resíduos como variáveis omitidas
 - O pressuposto da exogeneidade e a leitura causal da regressão
- Pressupostos do modelo de regressão

Aula 3: Problemas à vista? Multicolinearidade e Heterocedasticidade - 24 de agosto

- Multicolinearidade *vs.* Colinearidade Perfeita
- Heterocedasticidade, WLS e Erros de Huber-White

Aula 4: Qualidade do Ajuste e Diagnósticos - 31 de agosto

- Análise de Resíduos:
 - Análise Gráfica
 - Outliers e Leverage
 - Medidas de Influência

[7 de setembro – Não haverá Aula]

Aula 5: Regressões parciais (Frish-Waugh) e o significado de “controlar” por variáveis em Regressão Múltipla - 14 de setembro

Parte 2: Regressão Linear, extensões

Aula 6: Interações e não linearidade (Parte 1) - 21 de setembro

- Regressões lin-lin, log-lin, lin-log e log-log
- Variáveis dependentes limitadas e o Modelo Linear de Probabilidade

Aula 7: Interações e não linearidade (Parte 2) - 28 de setembro

- Polinômios e interpretações gráficas
- Efeitos marginais

Aula 8: Microsimulação como estratégia analítica - 5 de outubro

- Utilidade das microsimulações para testes de teoria e políticas públicas
- Cenários simulados são contrafactuais?

[12 de outubro- Não haverá Aula]

Parte 3: Regressão e Causalidade

Aula 9: Variáveis Instrumentais, uma introdução - 19 de outubro

- Regressão Múltipla e o pressuposto da seleção em observáveis

Aula 10: Múltiplas variáveis instrumentais e regressão em dois estágios (2SLS) - 26 de outubro

- Modelos subidentificados, justamente identificados e sobre identificados

Aula 11: Diagramas Causais Acíclicos e estratégias para seleção de variáveis - 9 de novembro

- Modelos teóricos e suas representações
- Causalidade por condicionamento *vs.* variáveis instrumentais
- Colliders e Confounders

Aula 12: Regressão Descontínua - 16 de novembro

Parte 4: Regressões para dados longitudinais (em painel)

Aula 13: Estruturas de dados longitudinais (*long* e *wide*) e modelos de regressão - 23 de novembro

- Modelo pooled
- Modelo de Efeitos Fixos
- Modelo de Primeiras Diferenças
- Inclusão de Tendências Temporais
- Variáveis com *lag/forward*

Aula 14: Afinal, efeitos fixos ou aleatórios? - 30 de novembro

- Vantagens e desvantagens
- Endogeneidade dos Efeitos Aleatórios
- Relação entre Modelos Multinível e modelos de efeitos aleatórios (*mixed effects*)

Aula 15: Causalidade em Modelos de Painel - 7 de dezembro

- O Estimador de Diferenças-em-Diferenças em Painel
- Estratégias “criativas” de identificação