

Tema central: Introdução à análise de regressão para pesquisas sociais

Ministrante: Thiago Cordeiro Almeida

E-mail de contato: thiagocordalmeida@gmail.com

Datas de realização: 15 a 18 de setembro

Horário das aulas: 19 às 22 horas

Objetivo(s) do curso:

Este curso tem por objetivo introduzir pesquisadoras/es na utilização de modelos de análise de regressão à investigação de fenômenos sociais. Serão abordadas técnicas para análise de dados contínuos (regressão linear) e categóricos (regressão logística), com a exposição de suas propriedades, procedimentos de implementação usando o software R e a realização de exercícios práticos a partir de bases de dados disponíveis gratuitamente online. Ao fim do curso, os/as participantes estarão familiarizados com os procedimentos necessários para a implementação de modelos de regressão, de forma que possam adaptar para cada problema de pesquisa que trabalhem. Para participar do curso é necessário ter noções básicas dos softwares Excel, R e ter instalado o RStudio em seu computador. É recomendado ter noções de estatística básica (frequências, tabelas cruzadas, médias, medianas, etc.).

Requisitos:

Noções básicas de Excel, R e ter instalado em sua máquina o RStudio.

Estrutura do curso (temas e subtemas previstos e materiais complementares):

DIA 1 – Fundamentos para a modelagem de dados quantitativos: conceitos, definições e pressupostos

1. Aula expositiva e prática

Introdução ao curso e apresentações

- Estrutura do curso
- Apresentação

Algumas definições acerca de dados quantitativos

- Natureza dos dados
- Tipos de dados quantitativos

Conceitos estatísticos

- População e amostra
- Variáveis
- Tipos de medidas
- Tipos de análises

Relação entre fenômenos sociais: associação ou efeito?

- Quando e como podemos dizer que um fenômeno social interage com outro?
- Correlação
- Causalidade
- **Exercício:** Caso empírico para aplicação das noções vistas anteriormente.

Modelagem de fenômenos sociais

- Intuição da modelagem de fenômenos sociais
- Motivação para se modelar fenômenos de diferentes naturezas
- Qual(is) o(s) limite(s) de uma modelagem?
- Fundamentos de uma modelagem

2. Materiais de referência

(Textos marcados com “” são aqueles sugeridos como maior importância)*

*Morettin & Bussab. Estatística básica, 9ª edição. **Capítulos 2, 3 e 4.**

*Chein (2019). **Apresentação e Capítulo 1.**

Leituras complementares

Texto sobre regressão linear aplicada ao *machine learning*:
<https://medium.com/@lauradatalks/regress%C3%A3o-linear-6a7f247c3e29>.

Livro “Como mentir com estatística”, *principalmente capítulo 1*, de Darrell Huff, 2016.

3. Parte prática

Materiais para a parte prática vão ser adicionados a esta pasta.

4. Vídeo(s)

- Vídeo de como instalar o R e o Rstudio.
- Vídeos sobre “Introdução ao R”, também enviados por e-mail.
 - Curso R para iniciantes, do Didática Tech: material até a aula 15 (são vídeos curtos) – [clique aqui para ir para o vídeo](#).
 - Aula de Introdução ao R, disponibilizada no youtube pelo pesquisador Bruno Montezano (trechos entre as marcações 06:12-47:37) – [clique aqui para ir para o vídeo](#).

DIA 2 – Modelagem de dados contínuos – regressão linear simples e multivariada

1. Aula expositiva e prática

Modelagem de fenômenos sociais contínuos

- Fluxo de análise de dados
- Conceitos e definições econométricas importantes
- Exemplo empírico

Modelo de Regressão linear simples

- Intuição dos modelos
- Formalização na população
- Função de distribuição associada
- Propriedades e pressupostos
- Avaliação dos modelos
- Interpretação dos resultados
- Exemplo empírico

Regressão linear multivariada

- Intuição
- O que muda entre a regressão linear simples e multivariada?
- Interpretação em um cenário de regressão multivariada
- **Exercício: exemplo empírico**

2. Materiais de referência

(Textos marcados com “” são aqueles sugeridos como maior importância)*

Hoffman, R. 2016. **Capítulo 1 (1.1 e 1.2).**

Morettin & Bussab. Estatística básica, 9ª edição. **Capítulos 16 (16.1 e 16.2).**

*Chein (2019). **Capítulo 2.**

Leituras complementares

Livro “R para Ciência de Dados”, Capítulo “Visão Geral”:
<https://pt.r4ds.hadley.nz/whole-game.html>.

Popper, K. A lógica da pesquisa científica.

3. Parte prática

Materiais para a parte prática vão ser adicionados a esta pasta.

4. Vídeo(s)

Se necessário, vídeo de alguma parte que não tenhamos coberto em aula.

DIA 3 – Modelagem de dados contínuos – regressão linear multivariada com variáveis categóricas binárias/nominais/ordinais

1. Aula e prática

Não se aplica.

2. Materiais de referência

Interpretação dos modelos de regressão linear múltipla

- Exercício iniciado em aula (para fazer em casa)
 - Referências
 - Scripts e slides trabalhados em aula

Variáveis independentes qualitativas

- Variável binária (*dummy*) em modelos de regressão linear
 - Referências
 - i. Chein (2019). **Capítulo 3 (3.9)**.
 - ii. Hoffman, R. (2016). Capítulo 5 (5.1 e 5.2).

3. Parte prática

Materiais para a parte prática vão ser adicionados a esta pasta.

DIA 4 – Modelagem de variáveis dependentes binárias – regressão logística

1. Aula expositiva e prática

Modelagem de fenômenos sociais categóricos

- Fluxo de análise de dados
- Conceitos e definições
- Exemplo prático

Regressão logística

- Família de modelos de regressão linear generalizados
- Modelo de regressão logística - *logit*
 - Intuição
 - Formalização na população
 - Propriedades e pressupostos
 - Interpretação
 - Avaliação do modelo
- Exemplo da aula – aplicação na prática!

Regressão logística multivariada – exemplo empírico

- Intuição, semelhanças e diferenças com os modelos anteriores
- Formalização do modelo populacional
- **Exemplo:**
- Interpretação em um cenário de regressão multivariada

Próximas etapas

- Perspectivas futuras nos estudos de regressões lineares
- Perspectivas futuras nos estudos de regressões não lineares (MLG)

2. Materiais de referência

Livro “Probabilidade e Estatística”, UFRGS. Capítulo 8 (8.4):
https://www.ufrgs.br/probabilidade-estatistica/livro/livro_completo/ch8-reg-mult-log.html.

Leituras complementares

Livro “R para Ciência de Dados”, Capítulo “Visão Geral”:
<https://pt.r4ds.hadley.nz/whole-game.html>.

Popper, K. A lógica da pesquisa científica.

3. Parte prática

Materiais para a parte prática vão ser adicionados a esta pasta.

4. Vídeo(s)

Se necessário, vídeo de alguma parte que não tenhamos coberto em aula.

DIA 5 – Aplicação empírica

1. Aula expositiva e prática

Não se aplica.

2. Materiais de referência

Todos os materiais utilizados ao longo das partes anteriores do curso.

3. Parte prática

Agora é hora de botarmos a mão na massa em um exercício prático de como o trabalho com modelagem de dados utilizando regressão linear é feito em pesquisa social.

Todas as orientações estão presentes no arquivo “**dia 5 > Exercício final**”.

4. Vídeo(s)

Todos os vídeos utilizados ao longo das partes anteriores do curso.

Referenciais bibliográficos

Chein, Flavia. 2019. **Introdução aos modelos de regressão linear**: um passo inicial para compreensão da econometria como uma ferramenta de avaliação de políticas públicas. Brasília: ENAP, 2019, 76 p.

Hoffmann, Rodolfo. 2016. **Análise de regressão: uma introdução à econometria**. 5. ed. Piracicaba: O Autor, 2016, 393 p.

Matos, Manuel António. **Manual Operacional para Regressão Linear**. Porto: FEUP, 1995, 26p.

Morettin, Pedro Alberto; Bussab, Wilton O. 2017. Estatística básica. 9ª ed. São Paulo: Saraiva, 2017, 453 p.

Curso-R. 2021. **Ciência de Dados em R**. Uma adaptação do livro em inglês “R for Data Sciences”, do Garrett Golemund e Hadley Wickham.