

MAT02035 - Modelos para dados correlacionados

Apresentações

Rodrigo Citton Padilha dos Reis
`citton.padilha@ufrgs.br`

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DEPARTAMENTO DE ESTATÍSTICA

Porto Alegre, 2023

O professor

Olá!



Olá!

- ▶ Desde outubro de 2017 eu sou Professor do Departamento de Estatística e faço parte do Corpo Docente do Programa de Pós Graduação em Epidemiologia da Universidade Federal do Rio Grande do sul (UFRGS). Além disso, eu atuo como pesquisador no Estudo Longitudinal de Saúde do Adulto (ELSA-Brasil).
- ▶ Eu me formei Bacharel em Estatística pelo Departamento de Estatística da UFRGS em 2008, e Mestre (2010) e Doutor (2014) em Estatística pelo Programa de Pós Graduação em Estatística da Universidade Federal de Minas Gerais.
- ▶ A minha dissertação de mestrado, intitulada *Técnicas estatísticas para avaliação de novos marcadores de risco: aplicações envolvendo o Modelo de Cox*, foi orientada pelos Professores Enrico A. Colosimo e Maria do Carmo P. Nunes.

Olá!

- ▶ A minha tese de doutorado, intitulada *Análise hierárquica de múltiplos sistemas reparáveis*, foi orientada pelos Professores Enrico A. Colosimo e Gustavo L. Gilardoni.
- ▶ Os meus interesses de pesquisa são Inferência causal em epidemiologia, Análise de mediação, Técnicas de amostragem, Modelos de predição de risco e Análise de sobrevivência.
- ▶ Em estatística aplicada eu tenho interesse na epidemiologia do Diabetes Mellitus.

A disciplina

Objetivos

- ▶ Capacitar os alunos para identificar situações em que pode existir **correlação nos dados** observados.
- ▶ Apresentar aos alunos as técnicas de **Equações de Estimação Generalizadas** e **Modelos Mistos Lineares**.
- ▶ Capacitar os alunos no uso das técnicas para análise de dados correlacionados proporcionando **domínio teórico, aplicado e computacional** dos conceitos envolvidos.

Organização

- ▶ **Disciplina:** Modelos para dados correlacionados
- ▶ **Turma:** U
- ▶ **Modalidade:** Ensino presencial
- ▶ **Professor:** Rodrigo Citton Padilha dos Reis
 - ▶ e-mail: citton.padilha@ufrgs.br ou rodrigocpdosreis@gmail.com
 - ▶ Sala: B215 do Instituto de Matemática e Estatística

Aulas e material didático

- ▶ **Aulas** (teóricas e práticas)
 - ▶ Exposição e **discussão** dos conteúdos
 - ▶ Faremos leituras semanais de artigos e capítulos de livros
 - ▶ Exemplos
- ▶ **Notas de aula**
 - ▶ Slides
 - ▶ Arquivos de rotinas em R
- ▶ **Exercícios**
 - ▶ Listas de exercícios
 - ▶ Para casa
 - ▶ Questionários do Moodle
- ▶ **Canais de comunicação:**
 - ▶ Durante as aulas
 - ▶ Moodle: aulas, materiais, listas de exercícios
 - ▶ Sala de aula virtual: notas das avaliações
 - ▶ e-mail do professor

Aulas e material didático

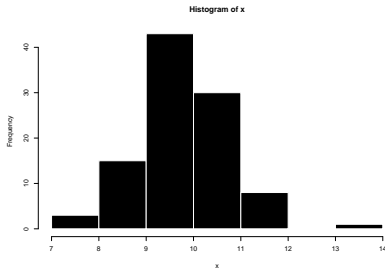
- ▶ **Aulas:** terças e quintas, das 8hs 30min às 10hs 10min, na Sala A114 do Instituto de Química e Matemática - Campus do Vale

Aulas e material didático



- ▶ Exemplos e exercícios com o apoio do computador:
 - ▶ R e RStudio

```
hist(x, col = 'black', border = 'white')
```



Conteúdo programático

Área 1

1. Introdução aos dados correlacionados
2. Dados longitudinais
3. Visão geral dos modelos lineares para dados longitudinais
4. Estimação e inferência estatística
5. Modelando a média
6. Modelando a covariância
7. Modelos lineares de efeitos mistos

Conteúdo programático (cont.)

Área 2

8. Modelos de efeitos fixos versus efeitos aleatórios
9. Diagnóstico e análise de resíduos
10. Revisão de modelos lineares generalizados (MLG)
11. Modelos marginais: introdução e visão geral
12. Modelos marginais: Equações de Estimação Generalizadas (GEE)
13. Dados ausentes: conceitos básicos
14. Modelos multiníveis
15. Tópicos complementares

Avaliação

- ▶ Serão realizadas quatro atividades de avaliação (pelo menos uma em de cada área):
 - ▶ duas provas (P_1 e P_2) presenciais e individuais;
 - ▶ dois trabalhos em grupo (T_1 e T_2).
- ▶ Cada atividade de avaliação vale 10 pontos
- ▶ Será realizado uma prova presencial e individual como atividade de recuperação (PR)
 - ▶ Para os alunos que não atingirem o conceito mínimo
 - ▶ **Esta prova abrange todo o conteúdo da disciplina**

Avaliação

$$MF = \frac{P_1 + P_2 + T_1 + T_2}{4}$$

- ▶ **A:** $9 \leq MF \leq 10$
- ▶ **B:** $7,5 \leq MF < 9$
- ▶ **C:** $6 \leq MF < 7,5$
- ▶ **D:** $MF < 6$
- ▶ **FF:** se o aluno tiver frequência inferior a 75% da carga horária prevista no plano da disciplina

Avaliação

- ▶ Se $MF < 6$ e frequência mínima de 75% o aluno poderá realizar a prova de recuperação e neste caso

$$MF' = MF \times 0,4 + PR \times 0,6$$

- ▶ C: $MF' \geq 6$
- ▶ D: $MF' < 6$

Referências bibliográficas



Principal

Fitzmaurice, G., Laird, N., Ware, J. **Applied Longitudinal Analysis**. John Wiley & Sons, 2011, 2ed.

Complementares

Singer, J. M., Nobre, J. S., Rocha, F. M. M. **Análise de dados longitudinais**. USP, 2018.

Diggle, P., Heagerty, P., Liang, K. Y., Zeger, S. **Analysis of Longitudinal Data**. Oxford University Press, 2002, 2ed.

Referências bibliográficas

<https://content.sph.harvard.edu/fitzmaur/ala2e/>

Applied Longitudinal Analysis, 2nd Edition

by [Garrett Fitzmaurice](#), [Nan Laird](#) & [James Ware](#)

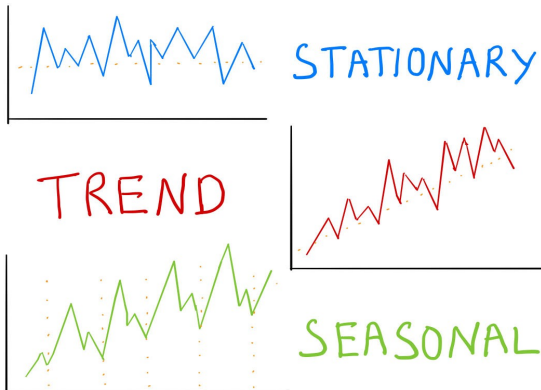
[Preface](#)
[Table of Contents](#)
[Author Information](#)
[Datasets](#)
[Sample SAS Programs](#)
[Sample Stata Programs](#)
[Sample R Programs](#)
[Lecture Slides](#)
[Book Reviews](#)
[Errata](#)
[Order Online](#)



Publisher: [John Wiley & Sons](#), August 2011

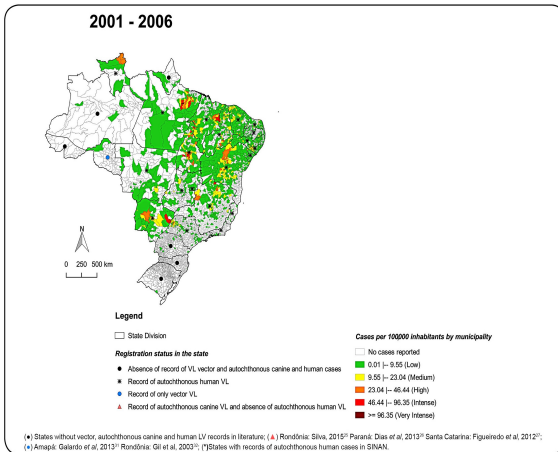
Os dados correlacionados

Séries temporais



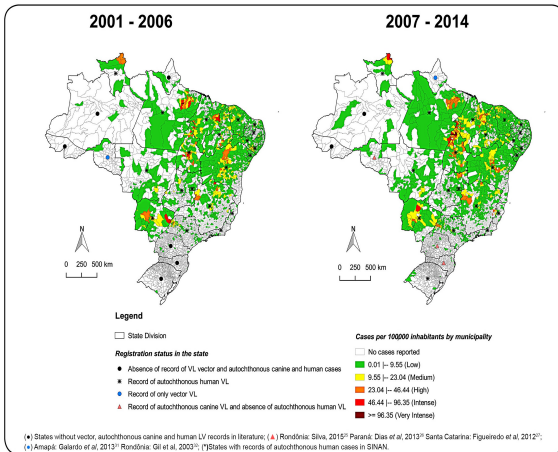
- ▶ Não discutiremos!
 - ▶ Disciplina MAT02263 - ANÁLISE DE SÉRIES TEMPORAIS A

Dados espaciais



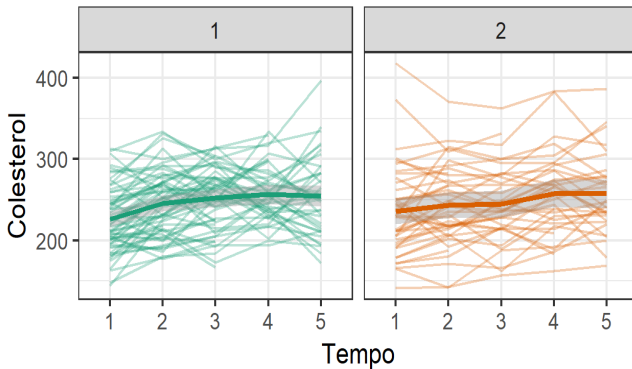
- ▶ Não discutiremos!
 - ▶ Disciplina MAT02040 - ESTATÍSTICA ESPACIAL

Dados espaço-temporais



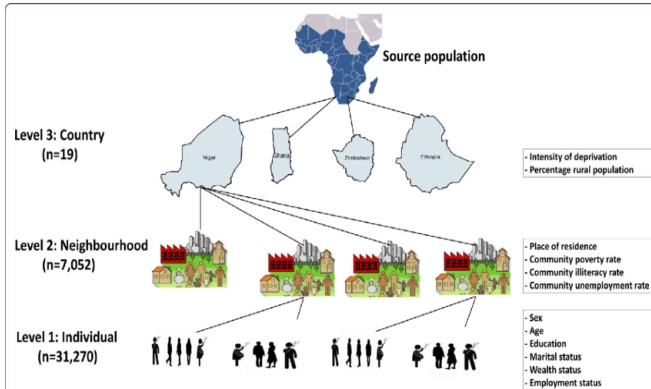
- ▶ Não discutiremos!
- ▶ Disciplina MAT02040 - ESTATÍSTICA ESPACIAL

Dados longitudinais



► Discutiremos!

Dados agrupados



► Discutiremos!

O que veremos nesta disciplina?

Daremos atenção a estruturas de **dados longitudinais** e **dados agrupados (multiníveis ou hierárquicos)** para:

- ▶ Descrever os dados;
 - ▶ Gráficos e resumos numéricos;
- ▶ Modelar os dados;
 - ▶ Aplicar modelos estatísticos para representar de maneira compacta a estrutura dos dados;
 - ▶ Fazer estimação e inferência (construção de intervalos de confiança e testes de hipóteses) com respeito às quantidades de interesse (coeficientes de regressão, componentes de variância, etc.);
 - ▶ Avaliar o modelo que melhor descreve os dados através da comparação de modelos;

Próxima aula

- ▶ Introdução aos dados longitudinais e agrupados.

Por hoje é só! Sejam todos bem-vind@s!

