

MI 634 – Análise de Dados Longitudinais

Segundo semestre de 2015

Professor: Caio L. N. Azevedo

Sala: 210 - IMECC

e-mail: cnaber@ime.unicamp.br

- Preferencialmente, procurar o Professor na supracitada sala, dentro do horário de atendimento. Não serão dirimidas dúvidas via e-mail.
- O e-mail deverá ser utilizado somente para: solicitação de agendamento de atendimento, justificativa de ausência em um atendimento agendado, justificativa de ausência em aula, o envio do artigo relativo ao seminário e o envio de eventuais correções constantes no site do curso, incluindo os materiais disponibilizados. Dúvidas serão sanadas somente durante o atendimento e durante as aulas.

Aulas: Terças e Quintas, Sala 225 (IMECC), 10h00 – 12h00

Atendimento :

- Quintas-feiras, das 13h às 14h (na sala 210 - IMECC). Não haverá atendimento fora desse horário.
- O(a) aluno(a) deverá enviar um e-mail (para o supracitado endereço) com 24 horas de antecedência, solicitando o agendamento de atendimento para o dia posterior. Por exemplo, se ele(a) quiser atendimento no dia 03/09, deverá enviar um e-mail no dia 02/09 até as 13h00. O(a) aluno(a) que não comparecer à um atendimento agendado, e não justificar devidamente (por e-mail) o motivo de sua ausência, não poderá mais solicitar agendamento de atendimento.

Página do curso: http://www.ime.unicamp.br/~cnaber/Material_ADL_POS_2S_2015.html

Monitoria: Não há

1. Programa

Resumo:

A suposição de independência (ou não correlação) entre observações é reconhecidamente inválida em diversos exemplos na literatura e situações reais. Ademais, a desconsideração de estruturas desse tipo na análise, pode levar ao comprometimento da qualidade dos resultados inferenciais. Sendo assim, considerar tais estruturas é de extrema importância para que uma análise mais realista seja realizada. Existem diversas metodologias estatísticas que permitem levar em consideração estruturas de dependência (correlação) úteis, de modo apropriado. Neste curso, estudaremos algumas das principais classes de modelos lineares gaussianos e não gaussianos para a análise de dados longitudinais. Discutiremos a modelagem das estruturas de dependência e correlação. Apresentaremos, também, alguns exemplos de modelos não lineares. Apresentaremos e discutiremos os principais aspectos inferenciais: estimação, testes de hipótese e verificação da qualidade de ajuste dos modelos. Trataremos, essencialmente, da inferência sob a ótica frequentista. Consideraremos o pacote R (WinBUGS/OpenBUGS) para concretização da análise de dados. O estudo das metodologias a serem abordadas proverá aos alunos uma maior gama de ferramentas na análise de dados.

1. Introdução aos dados longitudinais: exemplos

2. Análise exploratória de dados longitudinais

3. Modelos lineares multivariados para análise de dados longitudinais

- 3.1. Introdução.
- 3.2. Estimação.
- 3.3. Testes de hipótese.
- 3.4. Verificação da qualidade de ajuste do modelo.
- 3.5. Modelagem da estrutura de dependência/covariância.

4. Modelos lineares mistos para análise de dados longitudinais

- 4.1. Introdução.
- 4.2. Estimação .
- 4.3. Testes de hipótese.
- 4.4. Verificação da qualidade de ajuste do modelo.
- 4.5. Modelagem da estrutura de dependência/covariância.
- 4.6. Modelos marginais.

5. Modelos lineares generalizados mistos para análise de dados longitudinais

- 5.1. Introdução.
- 5.2. Estimação.
- 5.3. Testes de hipótese.
- 5.4. Verificação da qualidade de ajuste do modelo.
- 5.5. Modelagem da estrutura de dependência/covariância.

6. Outros tópicos

- 6.1. Inferência Bayesiana.
- 6.2. Modelos não lineares.
- 6.3. Equações de estimação generalizadas (modelos marginais).

2. Bibliografia (em ordem alfabética)

- Demidenko, E. (2004). ***Mixed models: theory and applications, Hoboken, NJ: Wiley-Interscience.***
- Diggle, P.; Heagerty, P.; Liang, K-Y.; Zeger, S. (2013). ***Analysis of Longitudinal Data, second edition, Oxford University Press, Oxford.***
- Galecki, A.; Burzykowski, T. (2013). ***Linear Mixed-Effects Models Using R: A Step-by-Step Approach, Springer-Verlag, New York, NY.***
- Hedeker, D.; Gibbons, R. D. (2006). ***Longitudinal Data Analysis, Wiley & Sons, New York, NY.***
- Magnus, J. R. and Neudecker, H. (1998). ***Matrix differential calculus with applications in Statistics and econometrics, first edition, Chichester, John Wiley.***
- Pinheiro, J.; Bates, D. (2009). ***Mixed-Effects Models in S and S-PLUS, Springer-Verlag, New York, NY.***
- Rocha, F. M. M. (2004), ***Seleção de estruturas de covariância para dados com medidas repetidas, Dissertação de Mestrado, IME-USP.***
- Rocha, F. M. M. (2009), ***Efeito da estrutura de covariância na análise de dados longitudinais, Tese de Doutorado, IME-USP.***
- Singer, J. M.; Andrade, D. F (2000). ***Analysis of Longitudinal Data. In: C R Rao; P K Sen. (Org.). Handbook of Statistics: Bio-Environmental and Public Health Statistics. Amsterdam: North Holland, v. 17, p. 115-160.***
- Weiss, R. E. (2010). ***Modeling Longitudinal Data. Springer-Verlag, New York, NY.***
- Materiais extras a serem, eventualmente, disponibilizados no site.
- Referências adicionais a serem divulgadas, eventualmente, ao longo do curso, no site.

3. Critérios de avaliação

- Metodologia de avaliação: duas provas, listas de exercícios e um seminário.
- Seminários: apresentação de um artigo que verse sobre modelos e/ou metodologias que não foram ou não serão apresentadas em sala de aula, mas que estejam relacionados(as) ao conteúdo da disciplina.
 - Os seminários serão individuais.
 - Cada aluno(a) deve enviar um artigo como sugestão para que seja apresentado por ele/ela. É vedado a apresentação de artigos em que um dos autores seja aluno(a) desta disciplina e/ou o professor dela. O artigo pode estar apenas “aceito” (sem ter sido publicado ainda), contanto que ele esteja disponível no site da respectiva revista científica. O Professor julgará a adequabilidade do artigo sugerido e, caso ele não seja aceito, o aluno(a) deverá enviar outro. Naturalmente, o mesmo artigo não pode ser apresentado por mais de um aluno(a). Será disponibilizado no site do curso uma lista com os artigos a serem apresentados, com o(a) respectivo(a) apresentador(a), a medida que os artigos forem sendo aceitos para a apresentação.
 - Regras adicionais, relativas aos seminários serão divulgadas posteriormente.
- Listas de exercícios serão disponibilizadas periodicamente. Delas serão selecionados exercícios para serem entregues, valendo nota. As listas poderão ser resolvidas e entregues individualmente, em dupla ou em trio. Elas podem ser manuscritas (salvo quando se tratar de questões relativas à análise de dados) ou digitadas (editor de textos à escolha). As listas devem ser entregues de forma impressa (não poderão ser entregues por e-mail). O formato é livre, mas as resoluções das questões têm de ser legíveis e organizadas e as listas devem ser entregues com capa indicando seu número bem como os nomes e os RA's dos integrantes das equipes. O(s) pacote(s) computacional(ais) a ser(em) usado(s) nas análises é(são) de livre escolha.
- Média Global (MG):

- Se o aluno(a) comparecer ao todos os seminários e apresentar o seu

$$MG = 0,55*MP + 0,25*ML + 0,20*NS$$

$$MP = 0,40*NP_i + 0,60*NP_i, NP_i: \text{ é a nota da } i\text{-ésima prova } (i=1,2)$$

ML: é a média aritmética simples das notas das listas.

NS : nota do seminário.

- Se o aluno(a) não comparecer a pelo menos um dos seminários (sem a devida justificativa) e/ou não apresentar o seu

$$MG = 0,6*(0,55*MP + 0,25*ML + 0,20*NS).$$

- Conceito Final

| Média Global | Conceito |
|---------------|----------|
| 8,5 --- 10,0 | A |
| 7,0 --- 8,4 | B |
| 5,0 --- 6,9 | C |
| 0,0 --- 4,9 | D |

- Frequência mínima para aprovação é de 75%