

MI 634 – Análise de Dados Longitudinais

Segundo semestre de 2018

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 **ensino aberto (moodle)** será usado somente para enviar e-mail (avisos) aos alunos. Caso o(a) aluno(a) queira se comunicar via e-mail, faça-o através do supracitado endereço (não enviar e-mails, posts etc, através do ensino aberto (moodle)).
- 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 e o envio de eventuais correções relativas ao site do curso, incluindo os materiais disponibilizados e, eventualmente, para comunicações específicas, como aquela destacada abaixo sobre a prova substitutiva. Dúvidas serão sanadas somente durante o atendimento e durante as aulas.

Aulas: Segundas e Quartas, Sala 224 (IMECC), 10h00 – 12h00

Atendimento :

- Segundas-feiras, das 13h às 14h, sala 210 IMECC.
- **O(a) aluno(a) deverá enviar um e-mail (para o supracitado endereço) em um determinado dia, solicitando o agendamento de atendimento para o dia posterior. Por exemplo, se ele(a) quiser atendimento para o dia 13/08, deverá enviar um e-mail no dia 12/08 até as 18h00. O(a) aluno(a) que não comparecer a 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_2018.html

Monitoria: Não háverá

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 estruturais 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 (eventualmente WinBUGS/OpenBUGS, para uma breve introdução sobre modelagem bayesiana) 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.

Objetivos: Prover ao aluno conhecimento, em nível de Doutorado, acerca da análise de dados longitudinais, desde suas estruturas mais básicas, passando pela a análise descritiva, análise inferencial e escolha de estruturas de dependência apropriadas. Assim, espera-se que os alunos tornem-se mais capacitados a analisar dados com estruturas de dependência.

- 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)

- Davis, C. S. (2003). **Statistical Methods for the Analysis of Repeated Measurements**, Springer, New York, NY.
- Davidian, M., Giltinan, D. M. (2017). **Nonlinear Models for Repeated Measurement Data**, Chapman & Hall/CRC, Boca raton, FL.
- 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.
- Hoffman, L. (2014). **Longitudinal Analysis: Modeling Within-Person Fluctuation and Change (Multivariate Applications Series)**, Routledge, New York, NY.
- Long, J. D. (2011). **Longitudinal Data Analysis for the behavioal sciences using R**, Sage publications, Los Angeles, CA.
- Magnus, J. R. and Neudecker, H. (1998). **Matrix differential calculus with applications in Statistics and econometrics**, first edition, Chichester, John Wiley.
- Molenberghs, G., Verbeke, G. (2005). **Models for Discrete Longitudinal Data**, Springer-Verlag, New York, NY.
- 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 Holand, v. 17, p. 115-160.
- Verbeke, G., Molenberghs, G. (2013). **Linear Mixed Models for Longitudinal Data**, Springer-Verlag, New York, NY.
- 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, um trabalho, um seminário e uma prova substitutiva (esta última, se necessário).
- 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. 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.
 - As sugestões de artigos devem ser enviadas, ao supracitado e-mail, em tempo hábil para que os artigos sejam escolhidos, até o dia 30/09/2018. Ou seja, que haja tempo para que o professor avalie e aceite ou não as sugestões.
 - Regras adicionais, relativas aos seminários serão divulgadas posteriormente.
- Listas de exercícios serão disponibilizadas periodicamente. Recomenda-se resolvê-las na íntegra. Delas serão selecionados exercícios para serem entregues, valendo nota, em forma de trabalho. O trabalho poderá ser resolvido e entregue individualmente, em dupla ou em trio. Ele poderá ser manuscrito (salvo quando se tratar de questões relativas à análise de dados) ou digitado (editor de textos à escolha). O trabalho deve ser entregue 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 o trabalho deve ser entregue com capa indicando 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. O trabalho deverá ser entregue impresso (pessoalmente) na secretaria de graduação, em data a ser anunciada. Essa data será definida com no máximo 10 dias de antecedência. Portanto, é de suma importância que os alunos resolvam as listas com antecedência (além claro, do óbvio ganho no aprendizado que terão).

- Média Global (MG):
 - Se o aluno(a) comparecer a todos os seminários, apresentar o seu e entregar o trabalho completo

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

$$MP = 0,5*N_{Pi} + 0,5*N_{Pi}, N_{Pi}: \text{é a nota da } i\text{-ésima prova (} i=1,2)$$

NT: nota do trabalho.

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 e/ou não entregar o trabalho completo

$$MG = 0,5*(0,55*MP + 0,25*NT + 0,20*NS).$$

- 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

- A prova substitutiva (PS) será feita mediante solicitação, via e-mail, (até o dia 29/11/2018, as 18h00), o qual deve conter: nome completo, RA, o pedido, nome da disciplina e a(s) justificativa(s) da(s) ausência(s) na(s) prova(s), quando pertinente. Uma vez solicitada, caso o aluno não realize a PS, sua nota, nela, será igual a zero.
- Caso o aluno tenha feito as duas provas regulares (PR), a nota da PS substituirá a menor delas, se $MG < 7,0$. Ou seja, neste caso, a MG poderá diminuir. Caso $MG \geq 7,0$, a nota da PS só substituirá a menor das notas das PR, caso esta seja menor do que aquela. Caso contrário, a MG permanecerá inalterada. Se o aluno não tiver feito pelo menos uma delas (PR), a nota da PS será usada para substituir a(s) nota(s) faltante(s).

- Provas não realizadas equivalem a ter, nessas provas, nota zero.
- Recomenda-se a leitura: <https://www.dac.unicamp.br/portal/vida-academica/graduacao/avaliacao-e-frequencia/abono-de-faltas-e-exercicios-domiciliares>.
- Frequência mínima para aprovação é de 75%