

**MI513 - Modelos Lineares Generalizados**  
**Segundo Semestre de 2024**  
**Programa**

**Professor:** Caio L. N. Azevedo  
**Sala do professor:** 210, IMECC  
**e-mail:** [cnaber@ime.unicamp.br](mailto:cnaber@ime.unicamp.br)  
**Página do curso:** [link](#)

**Sistemática**

**1. Aulas**

- (a) Em princípio, serão presenciais.
- (b) Dias, horários e local: Segundas e Quartas (Sala 322 IMECC), das 10h00 às 12h00.

**2. Tira-dúvidas**

- (a) Professor:
  - i. Durante os dias e horários das aulas (Segundas e Quartas das 10h00 às 12h00), poderão ser dirimidas dúvidas, consoante o andamento das aulas.
  - ii. Também poderá ser solicitado atendimento (das 16h00 as 17h00, às Terças), conforme descrito abaixo, os quais ocorrerão de forma remota (em breve enviarei o link que será usado durante todo o semestre).
  - iii. Solicitação de atendimento: no dia anterior, enviar um e-mail (até as 18h00 do dia anterior, nos moldes apresentados neste Programa, veja abaixo). Por exemplo, se houver interesse em agendar atendimento para o dia 27/08, deve-se enviar e-mail no dia 26/08, até as 18h00.
- (b) Obs: Não haverá monitor.

**3. Comunicação**

- (a) Aluno-Professor:
  - i. Via e-mail do professor (informado acima, não enviar e-mail via moodle).
  - ii. Durante as aulas.
  - iii. Deverá ser enviado e-mail ao Professor, somente: em caso de correções necessárias nos slides/programas em R/site do curso, para informar problemas com links de acesso/arquivos, para solicitação de Prova Substitutiva e Exame (veja Seção “Critérios de Avaliação”), bem como para tirar dúvidas sobre o conteúdo da

disciplina e regras (veja item 2, ou seja, procure tirar suas dúvidas durante as aulas/atendimento) e solicitar atendimento. Utilize seu e-mail acadêmico. O título do e-mail deve conter o assunto seguido de (MI513) e o corpo do e-mail deve conter (ao final) o nome completo e RA do aluno. E-mails fora desse padrão não serão respondidos.

(b) Professor-aluno:

- i. Durante as aulas e atendimento; moodle (para disponibilização de informações não apresentadas na página do curso e para enviar e-mail para todos os alunos, quando necessário); página do curso (slides, programas em R, referências, informações adicionais etc).

(c) Recomenda-se acessar a página do curso/moodle, diariamente.

## **Ementa**

Resumo: Família exponencial e terminologia de modelos lineares generalizados (MLG). Princípio de modelos lineares. Modelos para variáveis binárias; modelos multinomiais; modelos log-lineares. Teoria assintótica. Funções de estimação e quasi-verossimilhança. Modelos para respostas dependentes. GEE-estimação de equações generalizadas para modelos marginais. Verossimilhança condicional e outras verossimilhanças. Análise de tabelas 2x2. Modelagem conjunta de médias e dispersões. Componentes de dispersão. Modelos não Lineares. Diagnósticos. As suposições de normalidade e homocedasticidade dos dados são reconhecidamente inválidas em diversos exemplos na literatura e situações reais. Ademais, estruturas (modelos) lineares, muitas vezes, não são apropriadas para modelar a média da variável resposta em função de covariáveis de interesse. Sendo assim, relaxar tais suposições é de extrema importância para que uma análise mais realista seja conduzida. A classe de modelos de regressão lineares generalizados (MLG), e outras classes dela derivadas, constituem alternativas de análise que contemplam tal relaxamento. Portanto, o estudo da classe de MLG proverá aos alunos ferramentas adicionais em termos de análise de dados. Sendo assim, nesta disciplina, pretendemos apresentar e discutir os principais modelos da classe dos lineares generalizados bem como mecanismos de inferência, estimação, testes de hipótese e validação de modelos sob a ótica frequentista. Estudaremos modelos para dados positivos assimétricos, para dados binários, para dados binomiais com sobredispersão, para dados de contagem, para dados de contagem com sobredispersão, para dados de contagem com excesso de zeros e para alguns tipos de dados dependentes. Consideraremos o pacote R para concretização da análise de dados.

## Conteúdo

### 1. Introdução

- (a) Motivação.
- (b) Notação.
- (c) Revisão sobre modelos de regressão normais lineares homocedásticos.
- (d) Família exponencial.

### 2. Modelos Lineares Generalizados

- (a) Definição.
- (b) Função desvio.
- (c) Estimação dos parâmetros.
- (d) Teste de hipóteses.
- (e) Técnicas de diagnóstico.
- (f) Seleção de modelos.
- (g) Teoria Assintótica/Reamostragem.
- (h) Aplicações.

### 3. Modelos para dados positivos

- (a) Modelos com resposta gama.
- (b) Modelos com resposta inversa Gaussiana.
- (c) Estimação, teste de hipóteses, técnicas de diagnóstico e seleção de modelos.
- (d) Aplicações.

### 4. Modelos para dados binários

- (a) Modelos para dados agrupados (Bernoulli).
- (b) Modelos para dados não agrupados (binomial).
- (c) Regressão logística linear.
- (d) Outras funções de ligação.
- (e) Modelos para dados com sobredispersão (beta-binomial)
- (f) Estimação, teste de hipóteses, técnicas de diagnóstico e seleção de modelos.
- (g) Aplicações.

5. Modelos para dados de contagem
  - (a) Modelos com resposta Poisson.
  - (b) Modelos com resposta binomial negativa (dados com sobredispersão).
  - (c) Estimaco, teste de hipteses, tcnicas de diagnstico e seleo de modelos.
  - (d) Aplicaes.
  - (e) Modelos para dados com excesso de zeros.
  - (f) Estimaco, teste de hipteses, tcnicas de diagnstico e seleo de modelos.
  - (g) Aplicaes.
6. Funes de estimaco e quasi-verossimilhana.
7. Modelos com disperso varivel.
8. Modelos no lineares.
9. Modelos para dados dependentes e Equaes de Estimaco Generalizadas.
10. Verossimilhana condicional e outras verossimilhanas.

### Bibliografia (em ordem alfabética)

**As referências que estão marcados com (\*) são as mais relevantes (minimamente) para um aprendizado básico (fundamental) do conteúdo**

1. Agresti, A. (2015) Foundations of Linear and Generalized Linear Models, first edition, Wiley. (não encontrei disponível em nenhuma biblioteca da Unicamp)
2. Azevedo, C. L. N. (2021) Notas de aula de regressão linear e referências nele contidas (\*, como pré-requisito) ([link](#)).
3. Azevedo, C. L. N. (2024) Notas de aula disponíveis no site do curso (e outras referências a serem disponibilizadas/divulgadas no próprio site) (\*).
4. Atkinson, A. C. (1985) Plots, Transformations, and Regressions, Oxford Science Publications: An Introduction to Graphical Methods of Diagnostic Regression Analysis, first edition, Oxford. (não encontrei disponível em nenhuma biblioteca da Unicamp)
5. Cameron, A. C., Trivedi, P. K. (2013) Regression Analysis of Count Data, Cambridge. (disponível na Biblioteca do IMECC)
6. Cordeiro, G. M., Neto, E. A. L. (2006) Modelos Paramétricos, 248p. ([link](#)).
7. Cordeiro, G. M., Demétrio, C. G. B., Moral, R. A. (2024) Modelos Lineares Generalizados e Aplicações, Blucher.
8. Dobson, A. J., Barnett, A. G. (2018) An Introduction to Generalized Linear Models, fourth edition, CRC Press.(disponível na Biblioteca do IMECC, third edition)
9. Dunn, P. K., Smyth, G. K. (2018) Generalized Linear Models With Examples in R, first edition, Springer (\*). (não encontrei disponível em nenhuma biblioteca da Unicamp)
10. Faraway, J. J. (2014). Linear Models with R, second Edition, Chapman and Hall. (não encontrei disponível em nenhuma biblioteca da Unicamp)
11. Faraway, J. J. (2016). Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models, second edition, Chapman and Hall. (disponível na Biblioteca do IMECC, first edition)
12. Hardin, J. W., Hilbe, J. M. (2018) Generalized Linear Models and extensions, fourth edition, Stata press.(disponível na Biblioteca do IMECC, second edition)
13. Hosmer, D. W., Lemeshow, S., Sturdivant, R. X. (2013) Applied Logistic Regression, third edition, Wiley Series. (disponível na Biblioteca do IMECC, third edition)

14. Lindsey, J. K. (1997) *Applying Generalized Linear Models*, first edition, Springer. (disponível na Biblioteca do IMECC, first edition)
15. McCullagh, P., Nelder, J. A. (1989) *Generalized Linear Models*, second Edition, Chapman and Hall (\*). (disponível na Biblioteca do IMECC, second edition)
16. McCulloch, C. E., Searle, S. R., Neuhaus, J. M. (2008) *Generalized, Linear, and Mixed Models*, second edition, Wiley. (disponível na Biblioteca do IMECC, second edition)
17. Paula, G. A. (2024) *Modelos de Regressão: Com Apoio Computacional, Versão Preliminar*, 463 p., IME-USP ([link](#)) (\*).
18. Robak, P., Legler, J. (2024) *Beyond Multiple Linear Regression: Applied Generalized Linear Models And Multilevel Models in R*, Chapman and Hall. (não encontrei disponível em nenhuma biblioteca da Unicamp)
19. Stroup, W. W., Ptukhina, M., Garai, J. (2024) *Generalized Linear Mixed Models, Modern Concepts, Methods and Applications*, first edition, CRC. (disponível na Biblioteca do IMECC, second edition)

Materiais e referências adicionais serão, eventualmente, disponibilizados no site, inclusive sobre temas específicos.

Para acessar os livros em formato digital (alguns deles estão disponíveis nesse formato) de fora da Unicamp, você poderá precisar do VPN. Veja instruções de instalação [aqui](#). Veja também: Tutorial da Plataforma Minha Biblioteca - versão para alunos: [link](#)

### **Critérios de avaliação**

1. Metodologia de avaliação: duas provas, um trabalho, um seminário e um exame (este último, se necessário).
  - (a) 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 tem de ser resolvido e entregue individualmente. Ele poderá ser manuscrito (salvo quando se tratar de questões relativas à análise de dados) ou digitado (editor de textos de livre escolha). O trabalho deverá ser entregue pessoalmente, até a data limite de entrega (na sala de aula ou na sala do professor). Maiores instruções sobre o modo de entrega serão divulgadas posteriormente. O trabalho deverá ser entregue com capa contendo o nome completo o RA do aluno. O(s) pacote(s) computacional(ais) a ser(em) usado(s) nas análises é(são) de livre

escolha. A data de entrega 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).

- (b) Seminários: apresentação de um artigo que verse sobre modelos e/ou metodologias que não foram e/ou não serão apresentadas em sala de aula, mas que esteja(m) relacionados(as) ao conteúdo da disciplina.
  - i. Os seminários devem ser desenvolvidos e apresentados, oral e individualmente, acompanhados de slides apropriados..
  - ii. Cada aluno deve enviar (por e-mail) um artigo como sugestão para sua apresentação. É vedada a apresentação de artigos em que um dos autores seja aluno desta disciplina. O artigo pode estar apenas “aceito” (sem ter sido publicado ainda), contanto que ele esteja disponível no site da respectiva revista científica (não são aceitos trabalhos do [Arxiv](#)). O Professor julgará a adequabilidade do artigo sugerido e, caso ele não seja aceito, o aluno deverá enviar outro. Naturalmente, o mesmo artigo não pode ser apresentado por mais de um aluno. Será disponibilizado no site do curso uma lista com os artigos a serem apresentados, com o respectivo apresentador, à 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 06/11/2024. Ou seja, de modo a haver tempo para que o professor avalie e aceite ou não a sugestão.
  - iii. Regras adicionais, relativas aos seminários serão divulgadas posteriormente.
- (c) Provas (e o Exame): Em princípio, serão presenciais, durante o horário das aulas.
- (d) Conteúdo das Provas (Exame): em princípio, todo o conteúdo apresentado até a última aula antes da prova poderá ser cobrado.
- (e) O professor não é, direta ou indiretamente, responsável por questões relativas ao funcionamento dos sistemas computacionais, de e-mail, de impressão etc. Favor dirigir-se aos respectivos responsáveis.

## 2. Média Global (MG):

- (a) Se o aluno entregar o trabalho completo e apresentar seu respectivo seminário:

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

em que  $MP = 0,5 \times NP_1 + 0,5 \times NP_2$ ,  $NP_i$ : é a nota da  $i$ -ésima prova ( $i=1, 2$ ), NT: nota do trabalho, NS: nota do seminário.

- (b) Se o aluno não entregar o trabalho completo e/ou não apresentar seu seminário e/ou não assistir a pelo menos um dos outros seminários

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

(c) Conceito Final

- i. Valores: “A” se  $MG \in [8,5; 10,0]$ , “B” se  $MG \in [7,0; 8,4]$ , “C” se  $MG \in [5,0; 6,9]$ , “D” se  $MG \in [0,0; 4,9]$ .
- ii. Se  $MG \geq 5,0$ , estará aprovado, se  $MG < 2,5$ , estará automaticamente reprovado e se  $2,5 \leq MG < 5,0$ , terá de fazer EXAME.
- iii. OBS: o aluno aprovado poderá fazer o EXAME, para melhorar sua nota. Entretanto, ele deverá comunicar sua decisão até uma semana (sete dias corridos) antes, por e-mail.
- iv. Média Final (MF) (em caso de Exame):  $MF = 0,5 \times MG + 0,5 \times NE$ ; NE: nota do exame. Se  $MF \geq 5,0$ , o aluno estará aprovado, caso contrário, estará reprovado. Caso o aluno tenha sido aprovado com a MG e tenha feito exame, sua média final será  $MF = \max(MG, MF)$

(d) Sobre frequência mínima e abono de faltas:

- i. A frequência mínima para aprovação é de 75% ([link](#)).
- ii. Somente nos casos listados [aqui](#), desde que devidamente comprovado, o aluno terá direito à abono de falta. Neste caso, o aluno deverá enviar um e-mail ao professor, apresentando a devida justificativa e os respectivos documentos comprobatórios, solicitando o(s) abono(s), mencionado o(s) dia(s) específicos de ausência. Se tudo estiver correto, o(s) abono(s) é (serão) concedido(s).
- iii. Para situações não contempladas no item ii) (acima) pode-se enviar um e-mail ao professor, apresentando a devida justificativa e os respectivos documentos comprobatórios, solicitando o(s) abono(s), mencionado o(s) dia(s) específicos de ausência. Neste caso, mesmo se tudo estiver correto, o pedido será avaliado.
- iv. A quantidade de faltas só poderá ser vista mediante solicitação de atendimento (veja item 2)-(a)-iii)) (Tira-dúvidas). Recomenda-se que cada aluno procure monitorar suas respectivas faltas.

(e) Provas substitutivas (somente em caso de ausência em alguma(s) prova(s), não é para substituir nota de provas realizadas):

- i. Em caso de ausência por algum(ns) do(s) motivo(s) listados [aqui](#), o aluno deve enviar um e-mail ao professor, solicitando a realização da prova substitutiva (indicando a que prova o pedido se refere), bem como apresentando a devida justificativa e os respectivos documentos comprobatórios. Nesse caso, será feita uma prova específica substituindo a que não fora realizada. Por exemplo, se o aluno não tiver feito a Prova II, ser-lhe-á aplicada uma prova equivalente (conteúdo até aquela prova).

ii. Para situações não contempladas no item i) (acima) excepcionalmente, o Exame poderá servir como avaliação substitutiva para pelo menos uma das Provas (I e II). Ou seja, a nota do Exame reporia a(s) nota(s) da(s) prova(s) não realizada(s). Nesse caso o aluno deve enviar um e-mail ao professor, solicitando a realização do Exame para o fim em questão (indicando a(s) que prova(s) o pedido se refere), bem como apresentando a devida justificativa e os respectivos documentos comprobatórios para que o pedido seja avaliado (pode ou não ser deferido). Neste caso a nota do Exame não poderá servir para aumentar a MF (como descrito na Seção 2) Média Global (MG).

3. Casos omissos serão decididos pelo Professor da Disciplina.

#### **Suporte médico/psicológico:**

1. CECOM: Centro de Saúde da Comunidade [link](#).
2. SAPPE: Serviço de Assistência Psicológica e Psiquiátrica ao Estudante [link](#).