

Regressão (MI 406)

Segundo semestre de 2014

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, para o envio do artigo relativo ao seminário e para 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: Segundas e Quartas, Sala 225 (IMECC), 16h00-18h00

Atendimento :

- Quartas-feiras, das 13h às 14h (na supracitada sala). Não haverá atendimento fora desse horário.
- O(a) aluno(a) deverá enviar um e-mail (para o supracitado endereço) com pelo menos 72 horas de antecedência solicitando o agendamento de atendimento para um ou mais dias específicos. Por exemplo, se ele(a) quiser atendimento no dia 24/09, deverá enviar um e-mail até às 13h00 do dia 21/09. 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 agendamentos de atendimento.

Página na internet do curso:

http://www.ime.unicamp.br/~cnaber/Material_Reg_POS_2S_2014.html

Monitoria: Não há

1. Programa

Resumo:

Apresentar os principais aspectos dos modelos de regressão lineares (com ênfase no uso da distribuição normal para modelar a variável resposta): propriedades dos estimadores, intervalos de confiança, testes de hipóteses e avaliação da qualidade do ajuste de modelos. Serão discutidos e apresentados os principais resultados em termos do comportamento dos estimadores, intervalos de confiança e estatísticas do teste. Discutiremos as definições e propriedades das distribuições: normal multivariada, T de Student, F de Snedcor e Qui-quadrado, estas três últimas nas suas versões centrais e não centrais. Modelos que consideram transformações nas variáveis (resposta e explicativas) serão abordados, assim como alguns modelos não gaussianos. Estudar-se-á o ajuste dos modelos por mínimos quadrados (ordinários, ponderados e generalizados) e máxima verossimilhança. Metodologias de avaliação e seleção de modelos (covariáveis) serão discutidas. As questões relativas à inferência (estimação, intervalo de confiança e testes de hipótese) serão apresentadas sob o enfoque frequentista. Apresentar-se-á uma introdução aos modelos com erros nas variáveis. Serão apresentadas/discutidas análise de dados (aplicações) com o apoio computacional do pacote R.

1. Motivação

1.1 Exemplos motivadores e discussão sobre algumas idéias de modelagem.

2. Revisão de Álgebra de Matrizes

2.1 Definições e propriedades básicas.

3. Distribuição normal multivariada

3.1 Apresentação.

3.2 Propriedades: distribuições marginais, condicionais e função característica.

4. Introdução aos modelos de regressão lineares

4.1 Modelo de regressão linear simples (MRLS).

4.2 Estimação dos parâmetros por mínimos quadrados ordinários e propriedades dos estimadores (MRLS).

4.3 Modelo de regressão linear múltiplo (MRLM).

4.4 Estimação dos parâmetros por mínimos quadrados ordinários e propriedades dos estimadores (MRLM).

4.5 Forma matricial do MRLM.

4.6 Estimação por mínimos quadrados ordinários.

4.7 Estimação por mínimos quadrados generalizados.

4.8 Aplicações.

5. Distribuições de probabilidade não centrais

5.1 Distribuição t de Student não central.

5.2 Distribuição Qui-quadrado não central.

5.3 Distribuição F de Snedcor não central.

6. Intervalos de Confiança e Testes de Hipótese

6.1 Distribuição de formas quadráticas Gaussianas.

6.2 Intervalos de confiança e testes de hipótese (MRLS).

6.3 Intervalos de confiança e testes de hipótese: Tabela ANOVA e Testes do tipo

$$H_0: C\beta = M \text{ vs } H_1: C\beta \neq M \text{ (MRLM).}$$

6.4 Aplicações.

7. Mecanismo de validação/comparação de modelos

7.1 Análise residual.

7.2. Comparação se seleção de modelos.

7.3 Aplicações.

8. Transformações de variáveis e métodos de estimação alternativos

8.1 Transformações nas variáveis explicativas.

8.2. Transformações na variável resposta.

8.3 Alguns modelos não gaussianos.

8.4 Estimação por máxima verossimilhança.

8.5 Aplicações.

9. Introdução aos modelos de regressão lineares normais com erros nas variáveis

9.1 Definição.

9.2 Estimação.

9.3 Intervalos de Confiança e Testes de hipótese.

9.4 Aplicações.

2. Bibliografia

• Básica

- Draper, N. R. and Smith, H. (1998). **Applied regression analysis, third edition. New York, NY: John Wiley & Sons, disponível na biblioteca.**
- Fox, J. (2008). **Applied regression analysis and generalized linear models, second edition. Los Angeles, CA, Sage, disponível na biblioteca.**
- Rawlings, J. O, Pantula, J. O. and Dickey, D. A. (1998). **Applied regression analysis: a research tool, second edition, New York, NY: Springer, disponível na biblioteca.**

• Complementar

- Magnus, J. R. and Neudecker, H. (1998). **Matrix differential calculus with applications in Statistics and econometrics, first edition, Chichester, John Wiley.**
- 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 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).
 - 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 ou em dupla. Elas podem ser manuscritas (salvo quando se tratar de questões relativas à análise de dados) ou digitadas. 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.
- Média Global (MG):

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

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

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

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%