

ME 714 – Análise de Dados Discretos

Primeiro Semestre de 2014

Professor: Caio L. N. Azevedo

Sala do professor: 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 será usado somente para enviar e-mail aos alunos. Caso o(a) aluno(a) queira se comunicar via e-mail, faça-o através do supracitado endereço (não enviar e-mails através do ensino aberto).

Atendimento : Quartas-feiras, das 13h às 14h (na supracitada sala).

Aulas: Segundas e Quartas, das 16h00 às 18h00, sala CB02

Página na internet do curso: http://www.ime.unicamp.br/~cnaber/Material_ADD_1S_2014.htm

Monitoria: Não haverá

1. Programa

Resumo:

Apresentar as principais idéias relativas à análise de dados discretos, incluindo alguns tópicos de análise de dados categorizados (categóricos). Revisão sobre variáveis aleatórias discretas e funções geradoras (geratrizes). Testes de hipótese e medidas de associação relativas à tabelas de contingência. Modelos de regressão linear com variáveis categorizadas na matriz de planejamento. Modelos de regressão para tabelas de contingência. Modelos de regressão para respostas binárias e politômicas (nominais e ordinais). Modelos lineares generalizados. Modelo de regressão de Poisson. Métodos de estimação por máxima verossimilhança e mínimos quadrados. Validação e seleção de modelos.

O suporte computacional será desenvolvido/apresentado na plataforma R (eventualmente o SAS pode ser utilizado).

1. Revisão sobre variáveis aleatórias discretas

1.1 Definição e funções geradoras (ou geratrizes) de momentos e de probabilidade.

1.2 Algumas distribuições discretas: Binomial, Multinomial, Poisson, Geométrica, Hipergeométrica, Binomial Negativa.

2. Tabelas de Contingências

2.1 Características e análises de interesse.

2.2 Testes de homogeneidade e independência

2.3 Teste exato de Fisher e teste de McNemar

2.4 Testes de Mantel Haenszel e Teste de Pearson.

2.5 Medidas de associação

2.6 Aplicações

3. Revisão sobre modelos de regressão (normais) lineares

3.1 Definição e propriedades

3.2 Regressão linear com variáveis categorizadas (categóricas) na matriz de planejamento

3.3 Equivalência entre as parametrizações da matriz de planejamento com variáveis categorizadas

3.4 Aplicações

4. Modelos de regressão para variáveis categorizadas (categóricas)

4.1 Modelos de regressão para tabelas de contingências.

4.2 Modelos de regressão para variáveis binárias.

4.3 Modelos de regressão para variáveis politômicas.

5. Modelos de regressão para outros tipos de variáveis discretas

5.1 Modelos lineares generalizados

5.2 Modelos de regressão de Poisson

5.3 Validação e seleção de modelos

5.4 Aplicações

2. Bibliografia Básica

- **Principal**

- Agresti, A. (2007). *An introduction to categorical data analysis, second edition*. New Jersey, Wiley-Interscience. Disponível na biblioteca.
- Paula, G. A. (2013). *Modelos de regressão com apoio computacional, versão preliminar*. Disponível em: http://www.ime.usp.br/~giapaula/texto_2013.pdf.
- Agresti, A. (1990). *Categorical data analysis*. New York, John Wiley. Disponível na biblioteca.
- Hosmer, D. W. and Lemeshow, S. (2000). *Applied logistic regression, second edition*, New York, John Wiley. Disponível na biblioteca.
- Finney, D. J. (1971). *Probit analysis*, Cambridge, University Press. Disponível na biblioteca.

- **Complementar**

- McCullagh, P. and Nelder, J. A. (1989). *Generalized linear models, second edition*, Chapman & Hall/CRC. Disponível na biblioteca.
- Paulino, C. D. e Singer, J. M. (2006). *Análise de Dados Categorizados*, Edgard Blücher. Disponível na biblioteca.
- Materiais extras a serem disponibilizados no site.

3. Critérios de avaliação

- Metodologia de avaliação: duas provas e um trabalho.
- Periodicamente serão disponibilizadas listas de exercícios. Recomenda-se, fortemente, resolvê-las na íntegra.
- O trabalho consistirá em análises de conjuntos de dados, através de metodologias vistas em sala e/ou a serem pesquisadas pelos alunos, conforme sugestão do professor. Os conjuntos de dados poderão estar presentes (eventualmente de modo antecipado), também, nas listas de exercícios. O trabalho deverá ser entregue na forma de relatório o qual deverá incluir uma descrição do conjunto de dados e análises descritivas (além das análises inferenciais). Futuramente, todos os detalhes acerca do trabalho (data de entrega, formação dos grupos, conjuntos de dados etc) serão divulgados na página do curso e discutidos em sala.
- Média Global (MG):
 - Se o aluno(a) entregar o trabalho completo: $MG = 0,75*MP + 0,25*NT + 0,5$.
 - Se o aluno(a) não entregar o trabalho ou entregar o trabalho incompleto: $MG = 0,6*MP$,

em que, $MP = 0,40*NP1 + 0,60*NP2$; NPi : nota da i-ésima prova, $i=1,2$ e NT : nota do trabalho.
 - Se $MG \geq 7,0$ o(a) aluno(a) estará aprovado(a), caso contrário, terá de fazer EXAME. OBS: o(a) aluno(a) aprovado(a) poderá fazer o EXAME, para melhorar sua nota. Entretanto, ele(a) deverá comunicar sua decisão até uma semana (sete dias) antes, pessoalmente, na sala do Professor.
- Média Final (MF):
 - Se $MG \geq 7,0$; $MF = \text{máximo}(MG, ME)$, caso contrário $MF = ME$, em que,

 $ME = 0,5*MG + 0,5*NE$; NE : nota do exame.
 - Se $ME \geq 5,0$, o(a) aluno(a) estará aprovado(a), caso contrário, estará reprovado(a).
- A frequência mínima para aprovação é de 75%