

## ME 731- Métodos em Análise Multivariada

Segundo Semestre de 2015

**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).
- 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 constantes no site do curso, incluindo os materiais disponibilizados. Dúvidas serão sanadas somente durante o atendimento e durante as aulas.

**Atendimento :** Terças, das 13h às 14h (na supracitada sala).

- 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 25/08, deverá enviar um e-mail no dia 24/08 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.

**Aulas:** Segundas (sala CB 10) e Quartas (sala CB 09), das 14h00 às 16h00.

**Página do curso:** [http://www.ime.unicamp.br/~cnaber/Material\\_AM\\_2S\\_2015.htm](http://www.ime.unicamp.br/~cnaber/Material_AM_2S_2015.htm)

**Monitoria:** Não haverá

## **1. Programa**

Resumo:

Apresentar e discutir as principais ideias relativas à análise multivariada de dados, através de algumas das principais metodologias. As metodologias apresentadas são: Distribuição normal multivariada; Análise de variância multivariada; Análise de componentes principais; Análise fatorial; Análise de correspondência; Análise de correlações canônicas; Análise discriminante. Muitas dessas técnicas baseiam-se na suposição de normalidade multivariada dos dados. No entanto serão discutidas, muito brevemente, algumas extensões dessas metodologias a fim de relaxar tal suposição. O suporte computacional será desenvolvido/apresentado na plataforma R. Naturalmente, o conteúdo referente aos pré-requisitos para esta disciplina, será necessário para o seu desenvolvimento.

### **1. Distribuição normal multivariada**

- 1.1. Caso bivariado.
- 1.2. Caso multivariado, densidade e função de distribuição acumulada.
- 1.3. Função geradora de momentos e função característica.
- 1.4. Distribuições marginais e condicionais.
- 1.5. Estimação e testes de hipótese
- 1.6. Matriz de dados e verificação da normalidade multivariada dos dados.

### **2. Análise de variância multivariada.**

- 2.1. Modelo linear geral.
- 2.2. Estimação.
- 2.3. Testes de Hipótese.
- 2.4. Verificação da qualidade de ajuste do modelo.

### **3. Análise de componentes principais.**

3.1. Introdução.

3.2. Obtenção das componentes principais.

3.3. Determinação do número das componentes e sua posterior utilização.

### **4. Análise fatorial.**

4.1. Modelo de análise fatorial.

4.2. Estimação das cargas fatoriais.

4.3. Determinação do número de fatores.

4.4. Predição dos fatores.

4.5. Verificação da qualidade de ajuste do modelo.

4.6. Comparação com a metodologia de componentes principais.

### **5. Análise de correspondência.**

5.1. Definição e paralelo com a metodologia de análise fatorial.

5.2. Obtenção dos fatores e determinação da quantidade de fatores a serem utilizados.

5.3. Gráficos de análise.

### **6. Análise de correlações canônicas.**

6.1. Definição e interpretação.

6.2. Equações canônicas.

6.3. Interpretações das equações canônicas.

## 7. Análise discriminante.

7.1. Definição.

7.2. Métodos de análise discriminante.

7.3. Verificação da qualidade de classificação.

## 2. Bibliografia (em ordem alfabética)

- Brian, E. (2011). **An introduction to applied multivariate analysis with R**, New York, NY : Springer. (disponível na biblioteca do IMECC)
- Johson, R. A. and Wichern, D. W. (2002). **Applied Multivariate Statistical Analysis, 5ª edição**, Upper Saddle River, NJ : Prentice-Hall. (disponível na biblioteca do IMECC)
- Mardia, K. V., Kent, J. T. and Bibby, J. M. (1992). **Multivariate Analysis**, London : Academic Press. (disponível na biblioteca do IMECC).
- Rencher, A. C. and Christensen, W. F. (2012). **Methods of Multivariate Analysis, 3ª edição**, John Wiley & Sons, New Jersey, NJ. (disponível na biblioteca do IMECC).
- Notas de aula e material adicional a serem disponibilizados na página do curso.

## 3. Critérios de avaliação

- Metodologia de avaliação: duas provas e dois trabalhos.
- Periodicamente serão disponibilizadas listas de exercícios. Recomenda-se, fortemente, resolvê-las na íntegra.
- Cada 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. Cada trabalho deverá ser feito em equipe, necessariamente, de 2 a 4 alunos. O(s) pacote(s) computacional(ais) a ser(em) usado(s) nas análises é(são) de livre escolha.

Futuramente mais detalhes acerca dos trabalhos (data de entrega, formato, número de páginas, conjuntos de dados etc) serão divulgados na página do curso e discutidos em sala.

- Média Global (MG):
  - Se o aluno(a) entregar os dois trabalhos (completos):  $MG = 0,7*MP + 0,3*MT$ , em que  $MP = 0,4*NP1 + 0,6*NP2$ ,  $NP_i$  é a nota da  $i$ -ésima prova,  $i=1,2$  e  $MT = 0,5*NT1 + 0,5*NT2$ ,  $NT_i$  é a nota do  $i$ -ésimo trabalho,  $i=1,2$ . Caso contrário, ou seja, se ele(a) não entregar pelo menos um dos trabalhos completo,  $MG = 0,6*(0,7*MP + 0,3*MT)$ .
  - 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%