



UNIVERSIDADE FEDERAL DE UBERLÂNDIA



FICHA DE COMPONENTE CURRICULAR

CÓDIGO:	COMPONENTE CURRICULAR: ESTATÍSTICA MULTIVARIADA E MINERAÇÃO DE DADOS	
UNIDADE ACADÊMICA OFERTANTE: Faculdade de Matemática		SIGLA: FAMAT
CH TOTAL TEÓRICA: 60 horas	CH TOTAL PRÁTICA: 30 horas	CH TOTAL: 90 horas

1. OBJETIVOS

Objetivo Geral: Lidar com observações com mais do que uma variável e a interdependência inerente entre as variáveis. Descrever extensões de ideias e técnicas vistas a nível multivariado. Lidar com problemas que não têm similar compatível no caso univariado.

Objetivos específicos: Capacitar o aluno na análise conjunta de dados multivariados, possibilitando que o mesmo estabeleça conclusões a partir das amostras estudadas. O mesmo será capaz de analisar qualitativamente por meio de técnicas multivariadas a análise de variáveis categóricas ou do tipo de medida nominal, ordinal, intervalar ou proporcional. Utilização das técnicas apropriadas e interpretação dos resultados de forma ética, assim como a verificação dos requisitos para o uso destas.

2. EMENTA

Introdução a estatística multivariada. Distribuição normal multivariada. Inferência sobre o vetor de média. Análise de componentes principais. Análise fatorial. Análise de correlação canônica, análise de agrupamentos. Análise discriminante.

3. PROGRAMA

1. Introdução a estatística multivariada.
 - 1.1. Aplicações da análise multivariada
 - 1.2. Organização dos dados
 - 1.3. Vetores aleatórios, vetores de médias e matrizes de covariâncias e correlação
 - 1.4. Vetores de médias amostrais e matrizes covariâncias e correlações amostrais
 - 1.5. Variância generalizada e variância total
 - 1.6. Teste de esfericidade de Bartlett
2. Distribuição normal multivariada
 - 2.1. Função densidade
 - 2.2. Propriedades da distribuição
 - 2.3. Estimação dos parâmetros da distribuição normal multivariada
 - 2.4. Testes de aderência para a normal multivariada
3. Inferência sobre o vetor de média
 - 3.1. Teste de hipóteses para o vetor de médias,
 - 3.2. Região de confiança para a média,
 - 3.3. Comparação entre vetores médios de duas populações, região de confiança para a diferença entre dois vetores médios,
 - 3.4. Comparação entre vetores médios de várias populações (MANOVA): modelos a um critério e dois critérios de classificação.
4. Análise de componentes principais
 - 4.1. Componentes principais populacionais
 - 4.2. Componentes principais via matriz de covariância amostral
 - 4.3. Componentes principais via matriz de correlação amostral
 - 4.4. Critérios para determinar o número de componentes
 - 4.5. Representação gráfica
 - 4.6. Inferência em componentes principais para grandes amostras
 - 4.7. Aplicação em Mineração de Dados
5. Análise fatorial exploratória
 - 5.1. Modelo de fatores ortogonais
 - 5.2. Estimação dos fatores pelos métodos de componentes principais
 - 5.3. Estimação dos fatores pelos métodos da máxima verossimilhança
 - 5.4. Critérios para determinar o número de fatores
 - 5.5. Rotação de fatores: rotações ortogonais e oblíquas
 - 5.6. Estimação dos escores dos fatores: método de mínimos quadrados e método de regressão
 - 5.7. Qualidade do ajuste modelo de fatores ortogonais
 - 5.8. Aplicação em Mineração de Dados
6. Análise de agrupamentos
 - 6.1. Medidas de similaridade: variáveis quantitativas e qualitativas
 - 6.2. Métodos hierárquicos: método de ligação simples, ligação completa, ligação média, do centroide, e de Ward

- 6.3. Critérios para determinar o número de agrupamentos
- 6.4. Métodos não hierárquicos: método das k-médias (k-means) e método fuzzy
- 6.5. Validação dos agrupamentos
- 6.6. Agrupamento de variáveis
- 6.7. Aplicação em Mineração de Dados

7. Análise discriminante

- 7.1. Discriminação e classificação em duas populações;
- 7.2. Avaliação da funções discriminantes;
- 7.3. Discriminação e classificação em várias populações;
- 7.4. Função discriminante de Fischer;
- 7.5. Análise discriminante não paramétrica;
- 7.6. Aplicação em Mineração de Dados.

8. Mineração de dados

- 8.1 Mineração de dados e análise estatística
- 8.2 Importância do banco de dados
- 8.3 O Processo KDD
- 8.4 Modelos supervisionados de classificação
- 8.5 Modelos pontuais e predição em mineração de dados

4. BIBLIOGRAFIA BÁSICA

FÁVERO, L.P. *et al.* **Análise de dados**: modelagem multivariada para tomada de decisões. Rio de Janeiro: Elsevier, 2009.

HAIR, J. F *et al.* **Análise multivariada de dados**. 6. ed. Porto Alegre: Bookman. 2009.

HAN J. **Data mining: concepts and techniques**. 3rd ed. Burlington: Elsevier; c2011

JOHNSON, R. A. WICHERN, D. W. **Applied multivariate statistical analysis**. 4. ed. New Jersey. Prentice Itall Inc. 1998.

WITTEN, IH. **Data mining: practical machine learning tools and techniques with Java implementations**. San Francisco: Morgan Kaufman; c2000.

5. BIBLIOGRAFIA COMPLEMENTAR

ANDERSON, T. W. **An introduction to multivariate statistics**. New York: John Wiley, 1984.

DILLON, W. R. GOLDSTEIN. **Multivariate analysis: methods and application**. New York. John Wiley, 1984.

JACKSON, J. E. **A users guide to principal components**. New York: John Wiley, 1991.

JOBSON, J. D. **Applied multivariate data analysis**. vols I. e II, New York: Springer Verlag, 1992.

JOURNAL OF MULTIVARIATE ANALYSIS. Amsterdam: Elsevier. 1980-. ISSN 0047-259X. KACHIGAN, Acesso via Portal Capes.

MINGOTI, S. A., **Análise de dados através de métodos de estatística multivariada**. Belo Horizonte: UFMG. 2007.

RENCHER, A. C. **Methods of multivariate analysis**. New York: John Wiley, 1995.

6. APROVAÇÃO



Documento assinado eletronicamente por **Vinicius Vieira Favaro, Diretor(a)**, em 26/05/2022, às 23:29, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **José Fausto de Moraes, Coordenador(a)**, em 28/05/2022, às 15:33, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **2401910** e o código CRC **645033DB**.