

PLANO DE ENSINO (2021/1)

Disciplina: Análise Multivariada no R (BOT00170)

Carga horária: 45 horas

Créditos: 3

Modalidade de ensino: Ensino Remoto Emergencial

Professor: Daniel Dutra Saraiva

SÚMULA

Fundamentos básicos da análise multivariada de dados; tipos de dados e variáveis; tipos de técnicas multivariadas e suas características; medidas de distância multivariada; orientações básicas para a aplicação da análise multivariada na pesquisa científica; uso de diferentes técnicas de análise no ambiente R.

OBJETIVOS

Motivar, instrumentar e capacitar os alunos para trabalharem com ferramentas estatísticas multivariadas no ambiente R. Espera-se que ao final da disciplina, eles conheçam as principais ‘suposições estatísticas’ dos modelos específicos, sejam capazes de identificar e manusear o modelo apropriado para a questão de pesquisa, e sejam capazes de interpretar o resultado da análise usando o modelo proposto.

CONTEÚDO PROGRAMÁTICO

1. Introdução

- 1.1. O que é análise multivariada?
- 1.2. Conceitos básicos
- 1.3. Tipos de dados e variáveis em análise multivariada
- 1.4. Erro de medida e medida multivariada
- 1.5. Significância estatística *versus* poder estatístico
- 1.6. Apresentação do ambiente R

2. Classificação dos métodos de análise multivariada

- 2.1. Métodos dependentes
- 2.2. Métodos interdependentes
- 2.3. Critérios para a escolha do método adequado

3. Diretrizes básicas para a análise multivariada

- 3.1. Significância prática *versus* significância estatística
- 3.2. Relação entre tamanho da amostra e significância estatística
- 3.3. Validação de resultados

4. Medidas de distância multivariada no R

- 4.1. Medidas de similaridade
- 4.2. Medidas de dissimilaridade (diversidade beta)
- 4.3. Critérios para a escolha da medida adequada

5. Técnicas de ordenação de dados multivariados no R

- 5.1. Análise de Componentes Principais
- 5.2. Análise de Coordenadas Principais
- 5.3. Escalonamento Multidimensional Não-Métrico

- 5.4. Análise de Redundância
- 5.5. Análise de Redundância baseada em distância
- 5.6. Análise de Correspondência Canônica

6. Técnicas de agrupamento de dados multivariados no R

- 6.1. Método de ligação simples
- 6.2. Método de ligação completa
- 6.3. Método de ligação média (UPGMA)
- 6.4. Método de Ward

7. Técnicas para comparação de grupos multivariados no R

- 7.1. Análise de Variância Multivariada Permutacional
- 7.2. Análise de Covariância Multivariada Permutacional
- 7.3. Teste de homogeneidade de dispersões multivariadas

8. Particionamento de variação de dados multivariados no R

- 8.1. Particionamento de variação usando uma matriz explanatória
- 8.2. Particionamento de variação usando duas ou mais matrizes explanatórias

MÉTODO DE ENSINO

Aulas em tempo real via plataforma **Google Meet**. A disciplina terá um viés prático, baseado na análise exploratória de dados biológicos, no ambiente de programação e computação estatística **R** (<https://www.r-project.org>). As aulas práticas serão realizadas no **RStudio Desktop** (<https://www.rstudio.com>), um software gratuito que fornece a integração entre o **R** e um **editor de texto** avançado para a edição de comandos. Os alunos terão acesso a uma apostila elaborada para a disciplina, contendo explicações sucintas sobre as análises apresentadas e um breve roteiro com os comandos necessários para execução dessas análises (e gráficos complementares) no R, usando exemplos e exercícios complementares. As aulas teóricas expositivas serão gravadas e disponibilizadas aos alunos.

AVALIAÇÃO

Os alunos serão avaliados através de exercícios de aprendizagem (30%) durante a aula e apresentação de um seminário (70%) ao final da disciplina. O seminário consistirá em exposição oral do aluno sobre os resultados de uma análise estatística usando dados preferencialmente relacionados à sua dissertação ou tese. Os conceitos finais serão aplicados com base nos seguintes intervalos de notas: conceito A (de 9,0 a 10,0), conceito B (de 7,5 a 8,9), conceito C (de 6,0 a 7,4) e conceito D (< 6,0).

BIBLIOGRAFIA BÁSICA

- EVERITT, B.; HOTHORN, T. 2008. An introduction to applied multivariate analysis with R. Springer, New York. 273p.
- HAIR, J. F. HAIR JR.; ANDERSON, R. E.; TATHAM, R. L.; BLACK, W. C. 2009. Análise multivariada de dados. 6ª ed. Bookman, Porto Alegre. 593 p.
- McCUNE, B.; GRACE, J. B. 2002. Analysis of ecological communities. 2ª ed. MJM Software Design, Oregon. 300 p.

BIBLIOGRAFIA COMPLEMENTAR

- BORCARD, D.; GILLET, F.; LEGENDRE, P. 2011. Numerical Ecology with R. Springer, New York. 306p.
- GOTELLI, N. J.; ELLISON, A. M. 2011. Princípios de estatística em ecologia. Artmed, Porto Alegre. 528 p.

- JONGMAN, R. H. G.; TER BRAAK, C. J. F.; VAN TONGEREN, O. F. R (Eds.). 1995. Data analysis in community and landscape ecology. Cambridge University Press, United Kingdom. 299 p.
- LEGENDRE, P.; LEGENDRE, L. 2012. Numerical Ecology. Vol. 24, 3^a ed, Elsevier. 853p.