Ficha de Unidade Curricular A3ES
Técnicas de Estatística Multivariada
Licenciatura em Engenharia Informática e de Computadores
2024-25

1. Caracterização da Unidade Curricular

1.1 Designação
[4313] Técnicas de Estatística Multivariada / Multivariate Statistical Techniques

1.2 Sigla da área científica em que se insere
MAT

1.3 Duração
Unidade Curricular Semestral

1.4 Horas de trabalho
0h 00m

1.5 Horas de contacto
0h 00m

1.6 ECTS
6

1.7 Observações
Unidade Curricular Opcional

2. Docente responsável [1617] Paulo José Raimundo Ramos

3. Docentes e respetivas cargas letivas na unidade curricular [1617] Paulo José Raimundo Ramos | Horas Previstas: 60 horas

4. Objetivos de aprendizagem (conhecimentos, aptidões e competências a desenvolver pelos estudantes)

Após a aprovação na unidade curricular, o estudante deverá possuir a capacidade de:
1. Fornecer aos estudantes conceitos fundamentais de estatística multivariada, que permitam o tratamento de dados relacionados com a investigação científica de base experimental.
2. Desenvolver a capacidade de utilização do software estatístico R, que possibilite a aplicação dos métodos estatísticos abordados na unidade curricular.

4. Intended learning outcomes (knowledge, skills and competences to be developed by the students)

Upon approval, the student should be able to:
1. To provide students with fundamental concepts of statistics, which enables handling data related to scientific research on an experimental basis.
2. To develop the ability to use R statistical software, that enables the application of statistical methods presented in the curricular unit.

5. Conteúdos programáticos

1. Testes de hipóteses: Testes de aderência ou de qualidade de ajuste; Testes de homocedasticidade; ANOVA; Testes de comparação múltipla; Análise usando o R.
2. Análise descritiva de dados multivariados: Medidas descritivas; Análise usando o R.
3. Regressão linear múltipla: Pressupostos e modelo; Inferência estatística; Seleção de variáveis e construção do ?melhor? modelo; Análise usando o R.
4. Análise de componentes principais: Critérios para determinar o número de componentes principais a incluir; Rotação e interpretação das componentes principais; Análise usando o R.
5. Análise de clusters: Medidas de dissemelhança; Métodos hierárquicos e não hierárquicos; Avaliação da qualidade da solução encontrada; Análise usando o R.
6. Análise discriminante linear: Pressupostos, estimação de funções discriminantes lineares e interpretação dos resultados; Avaliação da significância das funções discriminantes lineares; Análise usando o R.

5. Syllabus

1. Tests of hypotheses: Goodness-of-fit tests; Tests for homogeneity of variance; ANOVA; Multiple comparison tests; Analysis using R.
2. Descriptive analysis of multivariate data: Descriptive measures; Analysis using R.
3. Multiple linear regression: Assumptions and model; Statistical inference; Variable selection and construction of the ?best? model ; Analysis using R.
4. Principal component analysis: Criteria to choose the number of principal components; Rotation and interpretation of principal components ; Analysis using R.
5. Cluster analysis: Dissimilarity measures; Hierarchical and non-hierarchical methods; Appreciation of the relative quality of clusters ; Analysis using R.
6. Linear discriminant analysis: Assumptions, estimation and interpretation of linear discriminant functions; Evaluation of the significance of linear discriminant functions ; Analysis using R.

6. Demonstração da coerência dos conteúdos programáticos com os objetivos de aprendizagem da unidade curricular

Os tópicos lecionados permitem ao aluno adquirir conceitos sobre análise descritiva de dados multivariados, sobre modelos de regressão linear multivariada e sobre métodos de análise de dados multivariados, usados no tratamento de dados relacionados com a investigação científica de base experimental. A utilização do software estatístico R, ao longo de todos os tópicos lecionados, permite ao aluno consolidar a aprendizagem dos conceitos lecionados e fornece uma importante ferramenta de análise de dados e cálculo. A conjugação destas duas abordagens permite cumprir os objetivos de aprendizagem 1. e 2. da unidade curricular.

6. Evidence of the syllabus coherence with the curricular unit´s intended learning outcomes	The topics taught allow the student to acquire concepts about descriptive analysis of multivariate data, multivariate linear regression models and multivariate data analysis methods, used in the treatment of data related to experimental-based scientific research. The use of the statistical software R, throughout all the topics taught, allows the student to consolidate the learning of the concepts taught and provides an important tool for data analysis and calculation. The combination of these two approaches allows learning outcomes 1. and 2. to be fulfilled.
7. Metodologias de ensino (avaliação incluída)	Aulas teórico-práticas, com apresentação de teoria, seguida pela resolução de exercícios e utilização do software estatístico R. A avaliação de conhecimentos é efetuada por avaliação distribuída com exame final. A avaliação distribuída ao longo do período letivo, compreende a realização de dois testes escritos de avaliação distribuída (TE) e um trabalho de grupo teórico-prático considerado pedagogicamente fundamental (TP), com apresentação e discussão obrigatória. Os estudantes ficam dispensados do exame final (EF), caso tenham obtido avaliação positiva na avaliação distribuída. Para obter aprovação, a classificação mínima em qualquer um dos TE é 8.00 valores, com média simples mínima de 9.50 valores; a classificação mínima do TP é 8.00 valores; a classificação mínima do EF é 9.50 valores. A classificação final ponderada mínima (CF) é 9.50 valores e é obtida por uma das fórmulas: CF=0.7TE+0.3TP ou CF=0.7EF+0.3TP. A avaliação de conhecimentos não contempla a realização de exames parciais.
7. Teaching methodologies (including assessment)	Theoretical-practical classes, with presentation of theory, followed by exercise resolution and using the statistical software R. Knowledge assessment is carried out through distributed assessment with a final exam. Distributed assessment throughout the academic period comprises two written distributed assessment tests (WT) and theoretical-practical group work considered pedagogically fundamental (PW), with mandatory presentation and discussion. Students are exempted from the final exam (FE) if they have obtained a positive evaluation in the distributed assessment. To obtain approval, the minimum classification in any of the WT is 8.00 values, with minimum simple average of 9.50 values; the minimum PW classification is 8.00 values; the minimum FE classification is 9.50 values. The minimum weighted final classification (FC) is 9.50 values and is obtained by one of the formulas: FC=0.7WT+0.3PW or FC=0.7FE+0.3PW. Knowledge assessment does not include partial exams.

8. Demonstração da coerência das metodologias de ensino com os objetivos de aprendizagem da unidade curricular

A metodologia adotada permite dotar os estudantes com competências ao nível do tratamento de dados multivariados, bem como da utilização do software estatístico R, que poderão ser usados na análise de dados multivariados, em futuros trabalhos de investigação.
Os métodos de avaliação permitem averiguar se o aluno adquiriu conhecimentos suficientes para atingir os objetivos propostos na unidade curricular.

8. Evidence of the teaching methodologies coherence with the curricular unit´s intended learning outcomes

The adopted methodology provide students with skills for the treatment of multivariate data, as well as the use of the statistical software R, which can be used in the analysis of multivariate data in future research work.
The evaluation methods allow to find if the student has acquired sufficient knowledge to achieve the intended learning outcomes of the curricular unit.

9. Bibliografia de consulta/existência obrigatória

1. Everitt, B., & Dunn, G. (2001). *Applied Multivariate Data Analysis* (2nd ed.). Wiley.
2. Harrell, F. (2015). *Regression Modeling Strategies with Applications to Linear Models, Logistic and ordinal Regression, and Survival Analysis* (2nd ed.). Springer.
3. Hothorn, T., & Everitt, B. (2014). *A Handbook of Statistical Analyses using R* (3rd ed.). CRC Press .
4. Hair, J., Babin, B., Anderson, R., & Black, W. (2018). *Multivariate Data Analysis* (8th ed.). Cengage.
5. Johnson, R., & Wichern, D. (2023). *Applied Multivariate Statistical Analysis* (6th ed.). Pearson .
6. Maroco, J. (2021). *Análise Estatística com o SPSS Statistics* (8ª ed.). Report Number .
7. de Sá, J. (2007). *Applied Statistics Using SPSS, STATISTICA, MATLAB and R* (2nd ed.). Springer.
8. Reis, E. (2006). *Estatística Multivariada Aplicada* (2ª ed.). Edições Sílabo .
9. Wang, C., Zhao, L., Sun, W., Xue, J., & Xie, Y. (2018). Identifying redundant monitoring stations in an air quality monitoring network. *Atmospheric Environment* , *190* , 256-268.

10. Data de aprovação em CTC «INFORMAÇÃO NÃO DISPONIVEL»

11. Data de aprovação em CP «INFORMAÇÃO NÃO DISPONIVEL»