UNIVERSIDADE DE LISBOA
FACULDADE DE LETRAS
DEPARTAMENTO DE LINGUÍSTICA GERAL E ROMÂNICA

Tratamento Estatístico de Dados Linguísticos

S1, 2015-2016

Sónia Frota e Joseph Butler (4h/semana)

sonia.frota@mail.telepac.pt
http://labfon.letras.ulisboa.pt/personal/sfrota/index.htm
jbutler@letras.ulisboa.pt
http://labfon.letras.ulisboa.pt/joe_butler.htm

Objectivos de aprendizagem (conhecimentos, aptidões e competências a desenvolver pelos estudantes):

Pretende-se que o aluno (i) compreenda a importância do tratamento estatístico de dados linguísticos (ii) adquira conhecimentos básicos de estatística descritiva e inferencial, habitualmente utilizados na investigação em linguística experimental; (iii) saiba utilizar estatística paramétrica e não paramétrica; e (iv) saiba escolher o modelo estatístico ajustado a diferentes tipos de dados. Os alunos deverão ser capazes de desenvolver um trabalho de aplicação dos conhecimentos a novos dados quantitativos.

Learning outcomes of the curricular unit:

Students are expected to (i) understand the relevance of statistical analysis of linguistic data; acquire basic knowledge in descriptive and inferential statistics, which is usually employed in research in experimental linguistics; (iii) know how to use parametric and non-parametric statistics; and (iv) be able to choose the most adequate statistical model to each type of data. Students should be able to write a paper where the knowledge, procedures and tools explored in the course are applied to new quantitative data.

Conteúdos programáticos:

Relevância do tratamento estatístico dos dados linguísticos, ilustrada por estudos exemplificativos em áreas diversas como a fonética, sintaxe, psicolinguística, aquisição da linguagem, sociolinguística ou variação linguística. Elementos básicos de estatística descritiva e inferencial. Tipos de distribuição, medidas de tendência central, medidas de dispersão, desvio padrão. Testes estatísticos, teste de hipóteses, correlações. Estatística paramétrica e não-paramétrica. Poder estatístico. Tamanho das amostras. Diferentes modelos estatísticos para diferentes tipos de dados. Aplicações exemplificativas em diversas áreas da linguística. Programas informáticos de análise estatística, incluindo SPSS e R.

Syllabus:

Demonstration of the relevance of statistical treatment of linguistic data, illustrated by studies in various areas, such as phonetics, syntax, psycholinguistics, language acquisition, sociolinguistics or language variation. Fundamentals of descriptive and inferential statistics. Types of distributions, measures of central tendency, measures of dispersion, standard deviation. Statistical tests, hypothesis testing, correlations. Parametric and non-parametric statistics. Statistical power. Sample size. Different statistical models for different types of data. Illustrative applications in various areas of linguistics. Statistical analysis software packages, including SPSS and R.

Metodologias de ensino (avaliação incluída):

Ensino teórico-prático dos conceitos e procedimentos fundamentais em estatística, sempre acompanhado de casos exemplificativos em diferentes domínios da linguística experimental, bem como de exercícios sistemáticos permitindo o treino desses conceitos e procedimentos. Exploração de estudos exemplificativos disponíveis na literatura em linguística. Será enfatizada a aplicação de conhecimentos a novos conjuntos de dados e a utilização, pelo aluno, de pacotes de software de análise estatística.

A avaliação da unidade curricular assenta no trabalho realizado em aula (20%), na elaboração de um trabalho de aplicação dos conhecimentos a novos dados quantitativos (30%) e num teste individual de avaliação de conhecimentos (50%).

Teaching methodologies (including evaluation):

Theoretical and practical teaching of basic foundations and fundamental procedures in statistics, illustrated with copious examples of studies from various areas of experimental linguistics, and systematically accompanied by practical exercises. Application of the concepts and procedures to new data, and use of statistical analysis software packages.

Students are evaluated on the basis of the quality and amount of work done/reflected in class (20%), a paper where the knowledge and procedures worked in the course are applied to new quantitative data (30%), and a written test (50%).

Bibliografia de consulta

Field, Andy (2013) *Discovering Statistics using IBM SPSS*. 4th ed. Los Angeles: Sage

Gómez, Pascual Cantos (2013) *Statistical Methods in Language and Linguistic* Research. London: Equinox.

Johnson, Keith (2008) *Quantitative Methods In Linguistics*. New York: Wiley-Blackwell.

Linebach, Jared A., Brian P. Tesch & Lea M. Kovacsics (2014) *Nonparametric Statistics for Applied Research*. New York: Springer.

Rietveld, Tony & Roeland van Hout (2005) *Statistics in language research: Analysis of variance*. New York: Mouton de Gruyter.

http://www.statisticshell.com/html/dsus.html

http://eu.wiley.com/WileyCDA/WileyTitle/productCd-1405144246.html