

Bachelor's Degree Final Project

Code: 103166
ECTS Credits: 12

| Degree | Type | Year | Semester |
|----------------------------|------|------|----------|
| 2503852 Applied Statistics | OB | 4 | 0 |

The proposed teaching and assessment methodology that appear in the guide may be subject to changes as a result of the restrictions to face-to-face class attendance imposed by the health authorities.

Contact

Name: Lluís Antoni Quer Sardanyons
Email: LluísAntoni.Quer@uab.cat

Use of Languages

Principal working language: catalan (cat)
Some groups entirely in English: No
Some groups entirely in Catalan: Yes
Some groups entirely in Spanish: No

Teachers

Alejandra Cabaña Nigro

Prerequisites

The rules of permanence establish a minimum of 160 ECTS of the degree passed to be able to enroll in the Final Project.

Objectives and Contextualisation

See the Catalan version.

Competences

- Analyse data using statistical methods and techniques, working with data of different types.
- Critically and rigorously assess one's own work as well as that of others.
- Design a statistical or operational research study to solve a real problem.
- Formulate statistical hypotheses and develop strategies to confirm or refute them.
- Identify the usefulness of statistics in different areas of knowledge and apply it correctly in order to obtain relevant conclusions.
- Make efficient use of the literature and digital resources to obtain information.
- Select and apply the most suitable procedures for statistical modelling and analysis of complex data.
- Select statistical models or techniques for application in studies and real-world problems, and know the tools for validating them.
- Select the sources and techniques for acquiring and managing data for statistical processing purposes.
- Students must be capable of collecting and interpreting relevant data (usually within their area of study) in order to make statements that reflect social, scientific or ethical relevant issues.
- Students must be capable of communicating information, ideas, problems and solutions to both specialised and non-specialised audiences.
- Students must develop the necessary learning skills to undertake further training with a high degree of autonomy.

- Use quality criteria to critically assess the work done.

Learning Outcomes

1. Apply techniques of descriptive statistics, inferential statistics and/or data mining to perform the relevant analyses.
2. Choose and apply the most suitable procedures for statistical modelling and analysis of complex data.
3. Choose suitable sources of information for developing the bachelor's degree final project.
4. Choose the most suitable statistical techniques for analysing the data obtained.
5. Critically assess the work done on the basis of quality criteria.
6. Design suitable statistical studies to solve the problems formulated in the bachelor's degree final project.
7. Determine which statistical procedures are the most suitable for performing the task at hand.
8. Formulate statistical hypotheses and develop strategies to confirm or refute them.
9. Justify the choice of certain techniques rather than others.
10. Make effective use of references and electronic resources to obtain information.
11. Reappraise one's own ideas and those of others through rigorous, critical reflection.
12. Students must be capable of collecting and interpreting relevant data (usually within their area of study) in order to make statements that reflect social, scientific or ethical relevant issues.
13. Students must be capable of communicating information, ideas, problems and solutions to both specialised and non-specialised audiences.
14. Students must develop the necessary learning skills to undertake further training with a high degree of autonomy.

Content

The final degree works (TFG) may be rather theoretical (some topic of statistics that is not worked on any of the subjects of the degree) or of a more practical nature (to study in depth a problem and / or specific data). In the first case it will have to contain examples of practical application of the results studied. In the second case, it must contain an adequate theoretical foundation of the results that are used.

The student and the tutor will determine the content of the TFG when this subject begins. The work can be chosen from those proposed by the teachers of the degree or can be proposed by the same student within a line of interest offered by the professors of the Department of Mathematics or Sociology. In both cases you must have the approval of the degree coordinator.

The extension of the TFG can be variable but it is recommended between fifteen and thirty pages. The work can be presented in Catalan, Spanish or English. The first page will include a title, author and tutor, place and dates where the work is carried out. It will then follow a summary that will be in the same language of the text and with its English language version. Non-original content must have been clearly referenced in the bibliography that will appear at the end of the text.

Methodology

See the Catalan version.

Annotation: Within the schedule set by the centre or degree programme, 15 minutes of one class will be reserved for students to evaluate their lecturers and their courses or modules through questionnaires.

Activities

| Title | Hours | ECTS | Learning Outcomes |
|----------------|-------|------|-------------------|
| Type: Directed | | | |

| | | |
|-----------------------|-----|------|
| Bibliographic inquiry | 15 | 0.6 |
| Type: Autonomous | | |
| autonomous learning | 59 | 2.36 |
| work completion | 225 | 9 |

Assessment

The assessment will consist of two parts: evaluation of the memory (80%) and evaluation of the presentation (20%). In both cases, a PDF document must be submitted to the Virtual Campus.

Memory evaluation. A court consisting of the tutor and two external evaluators awards up to 8 points to the memory presented by the student. The tutor can award up to 4 points and the two external evaluators can award up to 2 points each.

Evaluation of the presentation: The presentation can be evaluated by the tutor (with a maximum of 1.5 points) or, in case the memory obtains a rating greater than 7 (over 8), with a public presentation before a single court the day scheduled for this purpose. Those that obtain 7 or more points on the memory can obtain up to 8.5 if they decide that their tutor evaluates them. Alternatively, the qualification obtained can be increased up to 2.5 points if they are submitted to the court's assessment.

Assessment Activities

| Title | Weighting | Hours | ECTS | Learning Outcomes |
|--------------|-----------|-------|------|---|
| Contents | 80% | 0.5 | 0.02 | 7, 11, 1, 5, 6, 4, 8, 9, 14, 13, 12, 2, 3, 10 |
| Debate | 10% | 0.25 | 0.01 | 11, 5, 9, 10 |
| Presentation | 10% | 0.25 | 0.01 | 11, 5, 9, 10 |

Bibliography

Recommended bibliography:

GENERAL

Snedecor, G. W. and Cochran, W. G. (1989) Statistical Methods - The Iowa State University Press

Steel, R. and Torrie, J. H. (1976) Introduction to Statistics -McGraw-Hill

Steel, R. and Torrie, J. H. (1985) Bioestadística: principios y procedimientos - McGraw-Hill

ANOVA

Box, P., Hunter and W., Hunter, J (1988) Estadística para investigadores. Introducción al diseño de experimentos, análisis de datos y construcción de modelos - Barcelona:Editorial Reverté.

Cochran, W.G. and Cox, G.M. (1957) Experimental Designs - second. ed, New York: John Wiley & Sons, Inc.

Fisher, R.A. (1925) Statistical Methods for Research Workers - Edinburgh: Oliver & Boyd.

Snedecor, G.W. and Cochran, W.G. (1980) Statistical Methods - seventh ed, IA: Iowa State University Press.

MODELS FOR CATEGORICAL DATA

Agresti, A. (1990) Categorical Data Analysis - New York: John Wiley & Sons, Inc.

Andersen, P.K. and Borgan, O. (2000) Statistical Models BAseD on Counting Processes - Springer-Verlag New

York, Inc.

Cox, D.R. and Snell, E.J. (1989) The Analysis of Binary Data - second ed, London: Chapman and Hall.

MODELS OF REGRESSION

Amemiya, T. (1985) Advanced Econometrics - Oxford Basil Blackwell

Draper, N. and Smith, H. (1981) Applied Regression Analysis - second ed, New York: John Wiley & Sons, Inc.

Jobson, J.D. (1991) Applied Multivariate Data Analysis (Volume I: Regression and Experimental Design) - Springer-Verlag New York, Inc.

Rao, C.R. (1973) Linear Statistical Inference and Its Applications - second ed, New York: John Wiley & Sons, Inc.

LOGISTIC REGRESSION

Agresti, A. (1984) Analysis of Ordinal Categorical Data - New York: John Wiley & Sons, Inc.

Christensen, R. (1990) Log-Linear Models - Springer-Verlag New York, Inc.

Hosmer, D.W. Jr and Lemeshow, S. (1989) Applied Logistic Regression - John Wiley & Sons, Inc.

GENERALIZED LINEAR MODELS

McCullagh, P. and Nelder, J.A. (1989) Generalized Linear Models - London: Chapman and Hall.

Rao, C.R. (1973) Linear Statistical Inference and Its Applications - New York: John Wiley & Sons, Inc.

MULTIVARIANT ANALYSIS

Escofier, B. and Pagès, J. (1988) Análisis factoriales simples y múltiples: objetivos, métodos e interpretación - Servicio editorial de la Universidad del País Vasco.

Greenacre, M.J. (1984) Theory and Applications of Correspondence Analysis - London: Academic Press.

Lebart, L., Morineau, A. and Warwick, K.M. (1984) Multivariate Descriptive Statistical Analysis:

Correspondence Analysis and Related Techniques for Large Matrices - New York: John Wiley & Sons, Inc.

DISCRIMINANT ANALYSIS

Hand, D.J. (1981) Discrimination and Classification - New York: John Wiley & Sons, Inc.

Lachenbruch, P.A. (1975) Discriminant Analysis - New York: Hafner.

ANALYSIS OF CONGLOMERATES

Duran, B.S. and Odell, P.L. (1974) Cluster Analysis - New York: Springer-Verlag.

Everitt, B.S. (1980) Cluster Analysis - second ed, London: Heineman Educational Books Ltd.

Hartigan, J.A. (1975) Clustering Algorithms - New York: John Wiley & Sons, Inc.

SURVIVABILITY ANALYSIS

Collet, D. (1994) Modelling survival data in medical research - Chapman & Hall.

Cox, D.R. and Oakes, D. (1984) Analysis of Survival Data - London: Chapman and Hall.

Kalbfleisch, J.D. and Prentice, R.L. (1980) The Statistical Analysis of Failure Time Data - New York: John Wiley & Sons, Inc.

Klein, J. and Moeschberger, M. (1997) Survival Analysis: Techniques for censored and truncated data - New York: Springer

Lawless, J.E. (1982) Statistical Models and Methods for Lifetime Data - New York: John Wiley & Sons, Inc.

STATISTICAL SAMPLE

Kish, L. (1965) Survey Sampling - New York: John Wiley & Sons, Inc.

Wolter, K. M. (1985) Introduction to Variance Estimation - New York: Springer-Verlag Inc.

NON PARAMETRIC ANALYSIS

Conover, W.J. (1980) Practical Nonparametric Statistics - second ed, New York: John Wiley & Sons, Inc.

Hollander, M. and Wolfe, D.A. (1973) Nonparametric Statistical Methods - New York: John Wiley & Sons, Inc.

STRUCTURAL EQUATIONS

Bollen, K.A. (1989) Structural Equations with Latent Variables - New York: John Wiley & Sons, Inc.
Wiley, D.E. (1973) The Identification Problem for Structural Equation Models with Unmeasured Variables in Goldberger A.S. and Duncan, O.D. eds. Tructural Equation Models in the Social Sciences - New York: Academic Press.

Time Series Analysis

Fuller, W.A. (1976) Introduction to Statistical Time Series - New York: John Wiley & Sons, Inc.

MIXED MODELS

Littell, R.C., Milliken, G.A., Stroup, W.W., and Wolfinger, R.D. (1996) SAS System for Mixed Models - Cary, NC: SAS Institute Inc.
Verbeke, G. and Molenberghs, G. (1997) Linear Mixed Models in Practice: A SAS-Oriented Approach - New York: Springer.

BOOTSTRAP

Good, P. (2000) Permutation Tests: A practical guide to resampling methods for testing hypotheses - Springer Verlag New York, Inc.

DATAMINING

Hastie, T. and Tibshirani, R. (2001) The Elements of Statistical Learning: data mining, inference and prediction - Springer- Verlag New York, Inc

Software

The software required for the TFG.