

Degree	Type	Year
2503852 Applied Statistics	OB	3

Errata

Contact: Llorenç Badiella Llorenç.Badiella@uab.cat

Contact

Name: Amanda Fernandez Fontelo

Email: amanda.fernandez@uab.cat

Teachers

Jordi Joan Tur Escandell

Teaching groups languages

You can view this information at the [end](#) of this document.

Prerequisites

Knowledge of data description and visualization methods, probability and statistical inference. A previous course in linear models is required.

Objectives and Contextualisation

This course aims to extend the use of linear combinations of a set of predictors to reduce the uncertainty of a response variable. In particular, we will work on the use of parametric models, beyond the normal law, for the response variable. Within this more general modelling framework, we will look more closely at how information can be incorporated, for example, information about the design of the experiment.

Learning Outcomes

1. CM09 (Competence) Assess the suitability of the models with the correct use and interpretation of indicators and graphs.

2. CM09 (Competence) Assess the suitability of the models with the correct use and interpretation of indicators and graphs.
3. CM10 (Competence) Modify the existing software if required by the statistic model, or create new software, if necessary.
4. KM12 (Knowledge) Provide the experimental hypotheses of modelling, considering the technical and ethical implications involved.
5. KM12 (Knowledge) Provide the experimental hypotheses of modelling, considering the technical and ethical implications involved.
6. KM13 (Knowledge) Detect interactions, co-linearity and importance between explanatory variables.
7. SM11 (Skill) Analyse the residuals of a statistical model.
8. SM12 (Skill) Interpret the results obtained to formulate conclusions about the experimental hypotheses.
9. SM13 (Skill) Compare the degree of adjustment between diverse statistical models.
10. SM14 (Skill) Use graphs to visualise the fit and suitability of the model.

Content

1. Generalized Linear Models:

- Introduction: Link function, exponential family, canonical function.
- Parameter estimation, hypothesis testing, model selection, and validation for:
 - Binary responses: Logistic regression.
 - Count responses: Poisson regression.
 - Regression for positive continuous responses: Gamma regression.

2. Statistical Regularization in Linear Models:

- Concepts of regularization and shrinkage. Ridge Regression: The ridge estimator and its properties, cross-validation methods to tune the smoothing parameter, etc.
- LASSO Regression: The estimation equation and its properties, interpretation, etc.

3. Generalized Additive Models:

- Smoothing: Concept of splines and other smoothing methods such as the nearest neighbour estimator, Nadaraya-Watson estimator, and loess.
- Methods for selecting the smoothing parameter: minimizing a selection criterion, cross-validation methods.
- Additive models.
- Generalized additive models.

If we have time, we will introduce linear models and generalized linear models with random effects (linear mixed models and generalized linear mixed models).

Activities and Methodology

Title	Hours	ECTS	Learning Outcomes
Type: Directed			
Computer Practices	50	2	CM09, CM10, KM12, KM13, SM11, SM12, SM13, SM14
Theory	50	2	CM09, CM10, KM12, KM13, SM11, SM12, SM13, SM14
Type: Supervised			

problems / exercises to solve	16	0.64	CM09, KM12, KM13, SM11, SM12, SM13, SM14
Type: Autonomous			
Preparation for the exam	10	0.4	CM09, KM13, SM11, SM12, SM13, SM14

The course material (theory notes, lists of problems and statements of practice) will be available at the virtual campus, progressively throughout the course.

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.

Assessment

Continuous Assessment Activities

Title	Weighting	Hours	ECTS	Learning Outcomes
Final test	40%	4	0.16	CM09, KM13, SM11, SM12, SM13, SM14
Partial exam	30%	4	0.16	CM09, KM13, SM11, SM12, SM13, SM14
Practices (deliveries or check)	30%	16	0.64	CM09, CM10, KM12, KM13, SM11, SM12, SM13, SM14

The subject will be assessed with assignments (exercise assignments, problem checks and/or practicals) and 2 exams. To obtain the weighted grade of continuous assessment you must have a minimum of 3/10 in each of the parts.

Students who have opted for the single assessment modality will have to complete an assessment that will consist of a theory exam, a problem test and the delivery of practical reports of the course. Assessment of submissions may require an assessment interview with the teacher. The student's grade will be the weighted average of the three previous activities, where the exam will account for 45% of the grade, the test 45% and the assignments 10%.

If the final grade does not reach 5/10, the student has another opportunity to pass the subject through the remedial exam that will be held on the date set by the degree coordinator. In this test you can recover 70% of the grade corresponding to the theory and the problems. The part of internships is not refundable.

Bibliography

Linear Mixed-Effects Models Using R A Step-by-Step Approach / by Andrzej Galecki, Tomasz Burzykowski
https://bibcercador.uab.cat/permalink/34CSUC_UAB/1eqfv2p/alma991010402935906709

Lee, Y., Nelder, J. and Pawitan, Y. (2006). Generalized Linear Models with Random Effects. Chapman & Hall. London.

John E. Freund, Irwin Miller, Marylees Miller. (2000) Estadística matemática con aplicaciones. Pearson Educación. (existeix castellà)

McCullagh, P. and Nelder, J. (1992). Generalized Linear Models. Chapman & Hall. London.

Daniel Peña; *Regresión y diseño de Experimentos*, Alianza Editorial (Manuales de Ciencias Sociales), 2002.

Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani; *An Introduction to Statistical Learning*, Springer texts in Statistics, 2013.

Christopher Hay-Jahans; *An R Companion to Linear Statistical Models*. Chapman and Hall, 2012.

John Fox and Sandord Weisberg; *An R Companion to Applied Regression*, 2nd edition, Sage Publications, 2011.

Software

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

Language list

Name	Group	Language	Semester	Turn
(PLAB) Practical laboratories	1	Catalan/Spanish	first semester	afternoon
(TE) Theory	1	Catalan/Spanish	first semester	afternoon