UNIVERSITÀ CATTOLICA del Sacro Cuore

Statistical modelling

PROF. GUIDO CONSONNI; PROF. RAFFAELE ARGIENTO

Course aims and intended learning outcomes

The course concentrates on Bayesian Statistics, a powerful methodology that over the years has gained wide recognition because it is sound, flexible, produces clear answers, and makes use of a variety of information. Key features are: models and computation, using the free and open-source language and environment R, and the JAGS/STAN software.

The following learning abilities are provided and expected to be achieved by participants at the end of the course:

Knowledge and understanding
Knowledge of the foundations and methods of Bayesian inference, also in relation to frequentist inference. Knowledge of computational methods, in particular Markov Chain Monte Carlo, that now represent an essential ingredient of the Bayesian methodology.

Applying knowledge and understanding
Ability to carry out Bayesian inference in standard univariate as well as regression settings with conjugate priors, and in hierarchical modelling setups. This entails: devising a model formulation and implementing a computational algorithm, typically through a package. Knowledge should be adequate to understand models and output of an empirical report.

Making judgements
Evaluate a statistical model, and the resulting findings.

Communication skills
Ability to describe with an appropriate statistical language a model used in the analysis and to communicate the results of empirical findings using appropriate summaries.

Lifelong learning skills
Skills for data-analysis need methodological and computational understanding in order to learn proficiently future developments, which will be crucial for effective data-driven decision-making. Coding skills to build upon in future activities.

Course content

The course is divided in two modules

Module 1

- Introduction to the R-programming language.
- Bayes' rule: Prior, posterior and predictive distributions.
- Bayesian inference for binomial, Poisson, normal model using conjugate priors.
- Fisher information, non informative priors and Jeffreys priors.
- Bayesian hypothesis testing and model selection.

Module 2

- Monte Carlo Integration for Bayesian analysis.
- Introduction to computational methods for non conjugate Bayesian models and Markov Chain Monte Carlo methods.

UNIVERSITÀ CATTOLICA del Sacro Cuore

- Basics of the Metropolis-Hastings and the Gibbs sampler algorithms, and criteria for assessing convergence.
- Bayesian hierarchical modelling.
- Bayesian analysis of the linear and generalized linear regression model.
- Overview of software environments for Bayesian statistical modelling (JAGS/STAN).

Reading list

P.D. HOFF, *A first course in bayesian statistical methods*, Springer, 2009.
I. NTZOUFRAS, *Bayesian modeling using WinBUGS,* John Wiley & Sons, 2011.

Language and environment:
R - http://www.r-project.org/
JAGS - http://mcmc-jags.sourceforge.net/
Stan - https://mc-stan.org/

Class notes, coding and further material will be posted on the University platform Blackboard.

Teaching method

A blend of lectures, coding, data analysis and discussion in class (60 hours). Exercise and lab sessions (20 hours).

Assessment method and criteria

Written examination on part I including questions on methods and exercises (50% weight)

Applied project on part II making use of suitable statistical software (50% weight)

Students enrolling in this course are expected to know data analysis, probability and frequentist inference, at the level of Statistics courses usually taught in a bachelor degree in Economics; see for instance the topics covered in 'Statistica (analisi dei dati e probabilità) and 'Statistica applicata' (or 'Statistics' and 'Applied Statistics') at this University. These topics will be presented in a preliminary course in Statistics which will be held in the week before the start of the lectures. Students whose knowledge of statistical inference is weak are suggested to attend this course.

Notes and prerequisites

Prerequisites

Students enrolling in this course are expected to know data analysis, probability and frequentist inference at the level of Statistics courses usually taught in a bachelor degree in Economics. (More specific Notes at the bottom).