UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

Statistical Models and Bayesian Inference

2627-2-F8205B011

Learning objectives

The course is structured into two modules belonging to the fields of statistical sciences, computer science, and social sciences. Its aim is to provide students with theoretical knowledge and practical skills for advanced statistical analysis, with particular attention to the reproducibility and replicability of analyses, as well as to the effective communication of results.

The first module introduces the main advanced methods for statistical inference, including bootstrap methods, mixture models, and forecasting methods. Teaching activities involve the use of the R software and the RMarkdown environment for the analysis of real and simulated data, enabling students to develop autonomy in statistical reasoning, problem-solving skills, and competencies in the production of reproducible reports.

The second module explores the Bayesian approach to statistical inference, integrating it with the classical methods presented in the first module. Students acquire the theoretical foundations of Bayesian models, Markov chain Monte Carlo computational methods, and the use of software tools such as R and SAS for model estimation, evaluation, and comparison. This module also places particular emphasis on the reproducibility of analyses and on the production of integrated documents containing code, results, and their interpretation.

For a detailed description of the contents, please refer to the syllabus published on the webpages of the two respective course modules.

Overall, the course enables students to acquire solid theoretical foundations and applied skills for addressing advanced statistical analysis problems in various fields, including biostatistics, epidemiology, genetics, public health, and the social sciences. By the end of the course, thanks to its practice-oriented approach and its focus on reproducibility and scientific communication, students will be able to further explore the topics covered independently and apply the knowledge acquired in research and professional contexts.

Expected learning outcomes — Dublin Descriptors

DdD 1 – Knowledge and understanding

By the end of the course, students will be able to:
• describe the theoretical foundations of the main advanced statistical models, including mixture models and Bayesian models;
• understand and explain the assumptions underlying the models, the estimation procedures, including maximum likelihood with the expectation-maximization algorithm and Bayesian inference, and the main criteria for model evaluation;
• recognize the role of reproducibility and replicability in scientific research through the integrated use of R, RStudio, RMarkdown, and SAS.

DdD 2 – Applying knowledge and understanding

By the end of the course, students will be able to:
• develop simulations from univariate and multivariate distributions;
• analyze real data from different applied fields using R, RMarkdown, and SAS;
• apply advanced statistical techniques, including bootstrap methods, mixture models, Bayesian inference, diagnostics, model selection and validation, as well as forecasting methods;
• produce reproducible reports integrating code, results, interpretations, and critical discussion of the analyses.

DdD 3 – Making judgements

By the end of the course, students will be able to:
• critically assess the validity of model assumptions and the appropriateness of statistical methods with respect to research questions;
• compare alternative models using appropriate diagnostic criteria and measures of predictive accuracy;
• critically interpret the results obtained, discussing their strengths, limitations, and implications in different applied contexts.

DdD 4 – Communication skills

By the end of the course, students will be able to:
• effectively communicate the results of statistical analyses through structured, documented, and reproducible reports;
• present analyses, interpretations, and conclusions in written form using clear and rigorous scientific language, providing justification for the methodological choices adopted.

DdD 5 – Learning skills

By the end of the course, students will be able to:
• independently continue the study of advanced statistical methods using the bibliography, teaching materials, and code examples provided;
• update and broaden their methodological and computational skills by applying data science tools and statistical inference methods in interdisciplinary and research contexts.

Contents

For a detailed description of the contents, please refer to the syllabus published on the webpages of the two respective course modules.

Detailed program

For a detailed description of the contents, please refer to the syllabus published on the webpages of the two respective course modules.

Prerequisites

For a detailed description of the contents, please refer to the syllabus published on the webpages of the two respective course modules.

Teaching methods

For a detailed description of the contents, please refer to the syllabus published on the webpages of the two respective course modules.

Assessment methods

For a detailed description of the contents, please refer to the syllabus published on the webpages of the two respective course modules.

Textbooks and Reading Materials

Albert, J. (2009). Bayesian computation with R (2nd ed.). Springer.

Albert, J., & Hu, J. (2019). Probability and Bayesian modeling. Chapman and Hall/CRC.

Bartolucci, F., Farcomeni, A., & Pennoni, F. (2013). Latent Markov models for longitudinal data. Chapman and Hall/CRC.

Bishop, Y. M. M., Fienberg, S. E., & Holland, P. W. (2007). Discrete multivariate analysis: Theory and practice. Springer.

Blitzstein, J. K., & Hwang, J. (2014). Introduction to probability. Chapman and Hall/CRC.

Dey, D. K., Ghosh, S. K., & Mallick, B. K. (Eds.). (2000). Generalized linear models: A Bayesian perspective. CRC Press.

Gentle, J. E., Härdle, W. K., & Mori, Y. (Eds.). (2004). Handbook of computational statistics: Concepts and methods. Springer.

Lange, K. (2010). Numerical analysis for statisticians (2nd ed.). Springer.

Migon, H. S., Gamerman, D., & Louzada, F. (2014). Statistical inference: An integrated approach (2nd ed.). Chapman and Hall/CRC.

Pennoni, F. (2026). Dispensa di Inferenza Bayesiana: Teoria e applicazioni con R e SAS. Dipartimento di Statistica e Metodi Quantitativi, Università degli Studi di Milano-Bicocca.

Pennoni, F. (2026). Dispensa di Modelli Statistici II: Teoria e applicazioni con R. Dipartimento di Statistica e Metodi Quantitativi, Università degli Studi di Milano-Bicocca.

R Core Team. (2026). R: A language and environment for statistical computing. R Foundation for Statistical Computing.

Robert, C. P., & Casella, G. (2004). Monte Carlo statistical methods (2nd ed.). Springer.
SAS Institute Inc. (2012). SAS/STAT 12.1 user's guide: The MCMC procedure. SAS Institute Inc.

Semester

For a detailed description of the contents, please refer to the syllabus published on the webpages of the two respective course modules.

Teaching language

For a detailed description of the contents, please refer to the syllabus published on the webpages of the two respective course modules.

Sustainable Development Goals

PARTNERSHIPS FOR THE GOALS