# Statistical inference

## Prof. Guido Consonni

***COURSE AIMS AND LEARNNG OUTCOMES***

The course represents a thorough study of principles and methods of statistical inference. It presents both the frequentist approach based on the likelihood alone, as well as the Bayesian approach which combines the likelihood with a prior distribution. The teaching method will involve theoretical concepts, including proofs of important results, exercises to advance students’ knowledge, use of computing based on the free and open-source language and environment R. The ultimate aim is to provide students with: i) concepts and methods to understand statistical analyses, and ii) a scientific language and framework to learn new statistical or data science methods and techniques.

At the end of the course the student is expected to:

* have acquired a good knowledge of the principal aspects of frequentist inference, as well as of the basic features of Bayesian inference. Topics of special relevance are: likelihood, sufficiency, statistical information, prior and posterior distribution, model selection. – *Knowledge and understanding*;
* understand and interpret the assumptions behind models for empirical analyses, to evaluate the statistical properties of the methods, and to understand their implications from a statistical perspective. – *Applying knowledge and understanding;*
* appreciate uncertainty of statistical inference based on probabilistic sampling. To evaluate courses of actions to address an applied problem. – *Making judgements*;
* be able to describe with an appropriate statistical language the assumptions behind a model, and to communicate the results of empirical findings. – *Communication skills*;
* acquire a conceptual framework to advance his/her studies in the methods and techniques of Statistics and Data Science. – *Lifelong learning skills*.

***COURSE CONTENT***

– Introduction to the R-language

* Likelihood and log-likelihood; maximum likelihood (ML) estimation; score function; sufficiency
* Frequentist inference: unbiasedness, consistency, standard errors and confidence intervals; significance tests and p-values
* Frequentist properties of the likelihood: expected Fisher information; score statistic
* The asymptotic distribution of the ML estimator: Cramér-Rao lower bound, consistency of the ML estimator,
* Likelihood ratio statistic
* Likelihood inference in multiparameter models: score vector and Fisher information matrix

– Bayesian inference: Bayes’ theorem, prior and posterior distribution for basic models (Bernoulli/Binomial, Normal), Bayesian point estimate, credible interval.

– Decision theory: loss and risk functions; minimax and Bayes estimators

* Model selection: Akaike’s Information Criterion (AIC); Cross validation; Bayesian Information Criterion (BIC).

***READING LIST***

L. Held-D. Sabanés-Bové. *Likelihood and Bayesian Inference.* Second edition. Springer, 2020.

Class notes, coding and further material will be posted on Blackboard

***TEACHING METHOD***

A blend of lectures and coding (60 hours), exercise sessions (20 hours) and lab-sessions on R.

***ASSESSMENT METHOD***

Written examination with open-ended questions on methods, exercises and the R-language. No mid-term exam is envisaged.

***NOTES AND 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; 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 offered in the week before lectures begin. Students whose knowledge of statistical inference is weak are suggested to attend this course.

*Office hours*

Dipartimento di Scienze Statistiche Edificio Lanzone 18; by appointment through email