Course offered for the STAT PhD program starting from a.y. 2019/2020

| | |
|---|---|
| TITLE | **Basics of applied statistics and probability: applications with R** |
| Lecturer | Barani Simone, Chiantore Mariachiara, Spallarossa Daniele |
| Duration and Credits | 6 CFU (2 + 2 + 2) |
| Course description | The course aims at providing the student with basic knowledge of standard tools for statistical analysis of data. In particular, most of course focuses on the use of a number of tools included in the R package. Practical exercises will be preceded by a set of theoretical lessons aimed at presenting basic concepts of probability and statistics. These are fundamental to a correct understanding of the R tools. |
| Course organization | The course is divided into three moduli:<br>1) The first modulus deals with basics of probability and statistics from a theoretical perspective. The following topics will be discussed: discrete vs continuous random variables; discrete probability distributions (e.g., binomial distribution, Poisson distribution); continuous probability distributions (e.g., normal and lognormal distributions, Gumbel distributions); definition of major statistical descriptors (mean, mode, and quantiles); uncertainty vs variability; introduction to statistical hypothesis tests (e.g., chi-square test, K-S test); introduction to regression analysis.<br>2) The second modulus is the first practical modulus based on the use of R and R studio. It will present conventional tools for data processing, visualization, and statistical analysis. The first part of the modulus is an introduction to R, including the use of the R console, workspace use, how to navigate the help page, reading of input data. The second part provides basics of programming in the R language, including definition of variables, data vectors and arrays, mathematical and conditional operators, data visualization.<br>3) The third modulus will address fundamental uni- and multi-variate analyses, such as one way and multi way ANOVA, regression, model selection in case of multiple regression, cluster analysis and ordination (PCA and MDS) and multivariate hypothesis testing (PERMANOVA). All the theoretical lessons will be complemented with practical sessions using R packages. |
| Teaching period | September or January-February |