UCLouvain

lstat2185

2025

Numerical Methods for Statistics: Optimization, Simulations and the Bootstrap

The version you're consulting is not final. This course description may change. The final version will be published on 1st June.

5.00 credits	30.0 h + 15.0 h	Q1

Language :	English > French-friendly
Place of the course	Louvain-la-Neuve
Prerequisites	Concepts and tools equivalent to those taught in teaching units LSTAT2020 Logiciels et programmation statistique de base LSTAT2120 Linear models LSTAT2100 Modèles linéaires généralisés et données discrètes
Main themes	The course focuses on numerical methods and on computer intensive and iterative techniques that allow for parameter estimation and valid inferential procedures for statistical models. The course introduces concepts such as constrained and unconstrained optimization, popular numerical algorithms such as Newton, quasi-Newton methods and the EM algorithm (among others), setting up simulation studies to evaluate the performance of multiple competitors, as well as inferential techniques based on the bootstrap principle.
Learning outcomes	
Evaluation methods	January session: • During the semester the student must submit 3 compulsory assignments (short, 2-3 pages maximum per assignment), counting for 1.5 point of the final grade (each assignment = 0.5 points). The assignments are to be solved individually or in groups of 2. A mark will be assigned per group. Assignments arriving after the deadline are not considered. • A project (written in French/English in min 6 and max 12 pages in the template on Moodle, appendices not included) which will illustrate the methods of the course for 5 points. This (written) project will be submitted before the exam session and discussed with the teacher during the exam session. The evaluation of the project is done on the basis of the written report and on the basis of the answers in an oral discussion (without slides) on the results and methodology used for the report, during the exam session. The project is to be solved individually or in groups of 2. A score will be awarded per group. Projects arriving after the deadline are not considered. • An oral exam (~45min), in which the teacher will assess the mastery of the material covered in class (13.5 points), the quality of the project and the homework. The final grade for the LSTAT2185 course in January is given by the points obtained for the assignments + the points obtained for the project + the points obtained for the mastery of the subject. Attention: To validate the course, the student needs a final mark of 10 or more. August session: • A project (written in French/English in min 6 and max 12 pages in the template on Moodle, appendices not included) which will illustrate the methods of the course for 5 points. This (written) project will be submitted before the exam session and discussed with the teacher during the exam session. The evaluation of the project is done on the basis of the written report and on the basis of the answers in an oral discussion (without slides) on the results and methodology used for the report, during the exam session. The project is to be solved individually or in groups of 2. A score will be awarded per group. Projects arriving after the deadline are not considered. • An oral exam (~45min), in which the teacher will assess knowledge about the material covered in class (15 points) and the quality of the project. The final grade for the LSTAT2185 course in August is given by the points obtained for the project + the points obtained for the mastery of the subject. Attention: To validate the course, the student needs a final mark of 10 or more. Attention: Any usage of artificial intelligence software for producing part of text, code, figures or equations that are included in the final project or homework is strictly forbidden. All projects and homework will be analyzied with specialized software.

Teaching methods	The class consists of lectures (30h) and exercise sessions (15h). Attendance at lectures and practical exercise sessions is highly encourgared, almost MANDATORY!

Content	The course outline is as follows:
	• Part I:
	  • Basics of one-dimensional function optimization. Special case: the likelihood function.
	  • Global vs local optima; numerical convergence and approximation errors.
	  • Challenges of optimizing multi-dimensional functions. Special cases: linear and generalized linear models and computing the multivariate normal density (link with LSTAT2120, 2100, 2110, 2040).
	  • Newton's method, Fisher scoring and IRLS.
	  • Non-linear functions and numerical differentiation.
	  • Case study: ?optim(), ?nlm(), ?deriv() and friends.
	  • The EM algorithm. Special cases: missing data, normal mixture models and linear mixed models (link with LSTAT2210).
	• Part II:
	  • Setting up controlled simulation studies: competitor selection, performance metrics and reproducibility.
	  • Sampling from distributions and DGPs.
	  • Case study I: Sample mean (mean, median, trimmed mean from normal and skewed distributions); German tank problem (estimators from slide LSTAT 2040) - mean, bias, MSE, RE.
	  • Case study II: t-test (size, power, coverage and length for CIs) and Binomial test (Tables from slides LSTAT 2040).
	  • Case study III: GLM variable selection (TPR, FPR, FDR).
	  • Parallel computing in R: doParallel, foreach, mclapply and friends to illustrate Case study I-III.
	• Part III:
	  • Bootstrap and resampling methods.
	  • Bias and variance approximation based on resampling.
	  • Bootstrap confidence intervals and hypothesis testing.
	  • Other techniques: Permutation tests and Jacknife

Inline resources	Slides and notes will be distributed during the semester. Moodle website of the class :LSTAT2185 - Numerical Methods for Statistics: Optimization, Simulations and the Bootstrap https://moodle.uclouvain.be/course/view.php?id=5785
Bibliography	Givens, G.H. and Hoeting. J.A. (2013). Computational Statistics (2nd ed). Wiley. Rizzo, M.L. (2007). Statistical Computing with R (2nd ed). Chapman & Hall /CRC. Gentle, J.E. (2009). Computational Statistics. Springer. Lange, K. (2010). Numerical Analysis for Statisticians (2nd ed). Springer. Peng, R.D. (2020+). Advanced Statistical Computing. Available at https://bookdown.org/rdpeng/advstatcomp/ Chernick, M.R. (2008). Bootstrap methods : a guide for practitioners and researchers, Wiley Series in Probability and Statistics. Davison, A.C. et Hinkley, D.V. (1997). Bootstrap Methods and their Applications, Cambridge University Press. Efron, B. et Tibshirani, R.J. (1993). An Introduction to the Bootstrap, Chapman and Hall. Hall, P. (1992). The Bootstrap and Edgeworth Expansion, Springer. Mammen, E. (1992). When does bootstrap work ? Springer.
Other infos	Software: R/Python
Faculty or entity in charge	LSBA

Programmes containing this learning unit (UE)

Program title	Acronym	Credits	Prerequisite	Learning outcomes
Master [120] in Data Science : Statistic	DATS2M	5		
Master [120] in Statistics: Biostatistics	BSTA2M	5		
Master [120] in Statistics: General	STAT2M	5		
Certificat d'université : Statistique et science des données (15/30 crédits)	STAT2FC	5		