

5.00 credits

30.0 h + 15.0 h

Q1

Teacher(s)	Pircalabelu Eugen ;
Language :	English
Place of the course	Louvain-la-Neuve
Prerequisites	<p>Concepts and tools equivalent to those taught in teaching units</p> <p>LSTAT2020 Logiciels et programmation statistique de base</p> <p>LSTAT2120 Linear models</p> <p>LSTAT2100 Modèles linéaires généralisés et données discrètes</p>
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	<p>At the end of this learning unit, the student is able to :</p> <p>to understand the basic algorithms presented throughout the class and will be able to implement and apply these concepts to perform estimation/inference for classical models in statistics.</p> <p>to set up a simulation study to investigate the performance of multiple competitors and decide on the appropriate performance metrics, depending on the objectives of the study.</p>
Evaluation methods	<p>Session de janvier :</p> <ol style="list-style-type: none"> 1. Pendant le quadrimestre l'étudiant doit rendre 2 devoirs obligatoires (courts, à 2 pages maximum par devoir), comptant pour 2 points de la note finale (chaque devoir = 1 points). Les devoirs sont à résoudre individuellement ou en groupe de 2. Une note sera attribuée par groupe. Les devoirs qui arrivent après la date limite ne sont pas prises en compte. 2. Un projet (rédigé en français/anglais en min 6 et max 12 pages dans le template sur Moodle, annexes non incluses) qui illustrera les méthodes du cours pour 5 points. Ce projet (écrit) sera remis avant la session d'examen et discuté avec le professeur lors de la session d'examen. L'évaluation du projet se fait sur base du rapport écrit et sur base des réponses dans une discussion orale (sans diapositives) sur les résultats et méthodologie utilisée pour le rapport, pendant la session d'examens. Le projet est à résoudre individuellement ou en groupe de 2. Une note sera attribuée par groupe. Les projets qui arrivent après la date limite ne sont pas prises en compte. 3. Un examen oral (~45min), auquel le professeur évaluera la maîtrise de la matière vue au cours (13 points), la qualité du projet et des devoirs. <p>La note finale pour le cours LSTAT2185 en janvier est donné par les points obtenues pour les devoirs + les points obtenues pour le projet + les points obtenues pour la maîtrise de la matière.</p> <p>Attention : Pour valider le cours, l'étudiant a besoin d'une note finale de 10 ou plus. L'étudiant ne peut pas valider le cours si la note au projet est < 3/5 (sans arrondi) ou si la note à l'examen est < 6/13 (sans arrondi).</p> <p>Session d'août :</p> <ol style="list-style-type: none"> 1. Un projet (rédigé en français/anglais en min 6 et max 12 pages dans le template sur Moodle, annexes non incluses) qui illustrera les méthodes du cours pour 5 points. Ce projet (écrit) sera remis avant la session d'examen et discuté avec le professeur lors de la session d'examen. L'évaluation du projet se fait sur base du rapport écrit et sur base des réponses dans une discussion orale (sans diapositives) sur les résultats et méthodologie utilisée pour le rapport, pendant la session d'examens. Le projet est à résoudre individuellement ou en groupe de 2. Une note sera attribuée par groupe. Les projets qui arrivent après la date limite ne sont pas prises en compte. 2. Un examen oral (~45min), auquel le professeur évaluera la maîtrise de la matière vue au cours (15 points) et la qualité du projet. <p>La note finale pour le cours LSTAT2185 en août est donnée par les points obtenues pour le projet + les points obtenues pour la maîtrise de la matière.</p> <p>Attention : Pour valider le cours, l'étudiant a besoin d'une note finale de 10 ou plus. L'étudiant ne peut pas valider le cours si la note au projet est < 3/5 (sans arrondi) ou si la note à l'examen écrit est < 6/15 (sans arrondi).</p>

Teaching methods	The class consists of lectures (30h) and exercise sessions (15h)
Content	<p>The course outline is as follows:</p> <ul style="list-style-type: none"> • Part I: <ul style="list-style-type: none"> • 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: <code>?optim()</code>, <code>?nlm()</code>, <code>?deriv()</code> and friends. • The EM algorithm. Special cases: missing data, normal mixture models and linear mixed models (link with LSTAT2210). • Part II: <ul style="list-style-type: none"> • 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: <code>doParallel</code>, <code>foreach</code>, <code>mclapply</code> and friends to illustrate Case study I-III. • Part III: <ul style="list-style-type: none"> • Bootstrap and resampling methods. • Bias and variance approximation based on resampling. • Bootstrap confidence intervals and hypothesis testing. • Other techniques: Permutation tests and Jackknife
Inline resources	<p>Slides and notes on Moodle. Moodle website of the class :LSTAT2185 - Numerical Methods for Statistics: Optimization, Simulations and the Bootstrap https://moodle.uclouvain.be/course/view.php?id=5785</p>
Bibliography	<p>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.</p>
Other infos	Software: R
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		