UNIVERSITÀ
DEGLI STUDI
DI TORINO

010082

COURSES BROCHURE

M.Sc. in Stochastics and Data Science

Indice

Analysis (Course A)

Analysis (Course A)

Academic year:	2017/2018
Course ID:	MAT0032
Teachers:	Prof. Elena Cordero Prof. Joerg Seiler
Teacher contacts:	0116702803, *elena.cordero@unito.it*
Year:	1st year
Type:	D.M. 270 TAF B - Distinctive
Credits/recognition:	9
Course SSD (disciplinary sector):	MAT/05 - analisi matematica
Delivery:	Formal authority
Language:	English
Attendance:	Mandatory
Type of examination:	Written and oral

PREREQUISITES
Calculus and mathematical Analysis in one and several real variables. Ordinary Differential Equations.

COURSE OBJECTIVES

The course introduces the participants to the theory of infinite-dimensional vector spaces and of linear operators between them, with a special focus on the concepts of normed vector spaces, completeness, compactness, and the different topologies which characterize the infinite dimensional vector spaces. Applications concern various spaces of functions and operators between them (in particular, integral and differential operators). The course presents basic tools of modern mathematical analysis which are of fundamental importance in many branches of pure and applied mathematics, in particular in probability theory, statistics, numerical analysis, partial differential equations and dynamical systems.

COURSE AIMS

The student will acquire knowledge of many basic tools which are of common use in the analysis of infintite dimensional vector spaces. In particular he will learn the theory of Banach and Hilbert spaces and their dual spaces, of linear, bounded, and compact operators, and he will know the theory of distributions (generalized functions). He will be able to apply this knowledge to solve simple problems and exercises related to the theory (in particular, to solve simple integral or differential equations) and he will be able to rigorously prove main results of the theory.

COURSE DELIVERY

Standard lectures in classroom

LEARNING ASSESSMENT METHODS

The assessment consists in a written test followed by an oral examination, after completion of the course.

The written test consists in open questions and exercises on the topics treated in class and has a duration of 180 minutes. The mark will be expressed in thirtieth; the single points (30 in total) will be distributed to the questions and exercises on the basis of their importance and length; the final score will be given by summing up the partial scores of each question and exercise.

The oral examination is scheduled after the written test and can be given only after having passed the written test with a mark of 18 or better. The oral examination consists of questions on the written test and on the topics treated in class and listed in the examination programme (which is available to the participants on the web-site of the course).

Both written test and oral examination will result in a final mark expressed in thirtieth; the minimal mark allowed for successful assessment is 18. Otherwise, the student's performance is considered insufficient and the student has to repeat the examination (both written test and oral examination).

Both written test and oral examination must be achieved in the same examination period.

SYLLABUS

- Banach spaces.
- Linear operators.
- Hilbert spaces, projections, orthonormal basis.
- Generalized Fourier series.
- Dual spaces: linear functionals, weak convergence.
- Compactness in finite dimensional spaces.
- Compact operators and applications to integral equations.
- Fundamentals of spectral theories
- Distributions (generalized functions)
- Fourier transform
- Laplace transform

SUGGESTED TEXTBOOKS AND READINGS

- Bryan P. Rynne and Martin A. Youngson, Linear Functional Analysis, Second Edition, Springer, 2008.
- Dudley, R. M., Real Analysis and Probability, Cambridge University Press.
- Royden, H.L. Real Analysis. MacMillan.

Additional Lecture Notes will be made available to the students.

Course webpage: http://www.master-sds.unito.it/do/corsi.pl/Show?_id=mohh

---

Analysis (Course B)

Analysis (Basics)

Academic year:	2017/2018
Course ID:	MAT0033
Teacher:	Prof. Bertrand Lods
Teacher contacts:	*bertrand.lods@unito.it*
Year:	1st year
Type:	D.M. 270 TAF B - Distinctive
Credits/recognition:	9
Course SSD (disciplinary sector):	MAT/05 - analisi matematica
Delivery:	Formal authority
Language:	English
Attendance:	Mandatory
Type of examination:	Written

PREREQUISITES
A good knowledge of basic calculus and real analysis.

COURSE OBJECTIVES

This course is a 9-credit course aimed at introducing and developing many of the mathematical tools necessary in many fields of Probablity, statistics and applied mathematics. It introduces in particular the theory of infinite-dimensional vector spaces with a special focus on the concepts of normed vector spaces, completeness, compactness, and other characteristic properties of infinite dimensional vector spaces. Concrete applications to spaces of functions will be provided. Application of this theory to the study of ordinary differential equation and Fourier analysis will be also given.

COURSE AIMS

At the end of the course, the student is expected to be capable of:

- - using the basic tools and results to pose, formalize and solve a complex mathematical problem of applied interest.
- - being able to think about possible and useful generalizations of the various results studied during the lectures.
- - being able to communicate such findings using appropriate and clear mathematical notation and language

COURSE DELIVERY

The course is articulated in 72 hours of formal in-class lecture time, and in at least as many hours of at-home work solving practical exercises.

LEARNING ASSESSMENT METHODS

The course grade is determined solely on the basis of a written examination. The examination (2 hours and 45 minutes) test the student's ability to do the following:

- Present briefly the main ideas, concepts and results developed in the course, also explaining intuitively the meaning and scope of the definitions and the arguments behind the validity of the result. Students will be required to know the definitions, the statements of the theorems, the idea behing the proofs and their applications.

- Use effectively the concepts and the result to answer questions pertaining to functional analysis.

The above is accomplished by asking the student to answer 5-6 questions. Each of the questions has an essay part, and some of the questions also have a more practical ("exercise ") part.

SUPPORT ACTIVITIES

The course includes exercises classes; extra exercises are suggested as homework.

SYLLABUS
The course is divided in 4 parts:

  - Reminders of multivariate calculus

- Functions of several variables, differential calculus;
- Linear algebra: matrices, determinants, diagonalization
- Series of numbers and series of functions;
- Integral calculus for functions of several variables.

  - Abstract vector spaces

- Topological metric spaces, compactness;
- Metric spaces: properties, continuity of functions;
- Banach spaces: fundamental properties, examples;
- Hilbert spaces; fundamental properties, projection theorem.
- L_p spaces.
- The space of continuous functions, Ascoli-Arzela Theorem.

  - Ordinary Differential Equations

- Cauchy Lipschitz theory;
- Classical methods of integration of ODE's (linear equations, Laplace transform, separation of variables).

  - Fourier Analysis.

- Fourier series, Fourier transform;
- Applications

SUGGESTED TEXTBOOKS AND READINGS

KHURI, A.I. Advanced calculus with applications in statistics. Wiley Series in Probability and Statistics.

- ROYDEN, H.L. Real Analysis. MacMillan.

- RYNNE, B. P. and Martin A. YOUNGSON, M.A., Linear Functional Analysis, Second Edition, Springer 2008

- DUDLEY, R. M., Real Analysis and Probability, Cambridge University Press.

- Additional Lecture Notes will be made available to the students.

Course webpage: http://www.master-sds.unito.it/do/corsi.pl/Show?_id=3h10

Bayesian nonparametric statistics (not offered in 2017/2018)

Bayesian nonparametric statistics (not offered in 2017/2018)

Academic year:	2016/2017
Course ID:	MAT0042
Teacher:	Prof. Antonio Canale Prof. Matteo Ruggiero Ramses H. Mena
Teacher contacts:	011 670 5724, *antonio.canale@unito.it*
Year:	2nd year
Type:	D.M. 270 TAF C - Related or integrative
Credits/recognition:	6
Course SSD (disciplinary sector):	SECS-S/01 - statistica
Delivery:	Formal authority
Language:	English
Attendance:	Optional
Type of examination:	Oral

PREREQUISITES
STOCHASTIC MODELLING FOR STATISTICAL APPLICATIONS

COURSE OBJECTIVES

The course aims at providing a modern overview of Bayesian nonparametric statistical methods.

COURSE AIMS

Students will learn how to model statistical problems with Bayesian nonparametric tools, study theoretical properties of the involved objects and devise appropriate computational algorithms for their implementation.

COURSE DELIVERY

The course consists mainly of class lectures, with some additional computer lab sessions using R.

LEARNING ASSESSMENT METHODS

Oral examination and optional paper presentation or discussion of an essay elaborated by the student.

SYLLABUS

This course covers the fundamentals of Bayesian nonparametric inference and focuses on the key probabilistic concepts and stochastic modelling tools at the basis of the most recent advances in the field:
• foundations of Bayesian nonparametric inference: exchangeability and de Finetti's representation theorem
• the Dirichlet process
• models beyond the Dirichlet process
• mixture models for density estimation and clustering
• random partitions
• dependent priors for partially exchangeable data
• elements of Bayesian asymptotics

16 hours of the course will be taught by Visiting Professor Ramses Mena on Random partitions and dependent processes.

SUGGESTED TEXTBOOKS AND READINGS

GHOSAL and VAN DER VAART (2016). Theory of nonparametric Bayesian inference. Cambridge University Press.
HJORT, HOLMES, MUELLER and WALKER (eds.) (2010). Bayesian Nonparametrics. Cambridge University Press.
GHOSH, RAMAMOORTHI. (2003). Bayesian Nonparametrics. Springer.

NOTE

This course will be delivered at the ESOMAS Department.

Course webpage: http://www.master-sds.unito.it/do/corsi.pl/Show?_id=meln

---

Bayesian statistics

Bayesian statistics

Academic year:	2017/2018
Course ID:	MAT0070
Teacher:	Prof. Matteo Ruggiero Prof. Raffaele Argiento
Teacher contacts:	011 670 5758, *matteo.ruggiero@unito.it*
Year:	2nd year
Type:	D.M. 270 TAF C - Related or integrative
Credits/recognition:	6
Course SSD (disciplinary sector):	SECS-S/01 - statistica
Delivery:	Formal authority
Language:	English
Attendance:	Optional
Type of examination:	Oral

PREREQUISITES
STOCHASTIC MODELLING FOR STATISTICAL APPLICATIONS

COURSE OBJECTIVES

The course aims at providing a modern overview of Bayesian statistical methods, covering the fundamentals of both the parametric and the nonparametric approach. The course will focus on the key probabilistic concepts, stochastic modelling tools and most widely used computational strategies at the basis of the most recent advances in the field.

A short module of the course will be taught by Visiting Professor Jim Griffin on a topic TBA (see International visiting professors)

COURSE AIMS

Students will learn how to model statistical problems with Bayesian parametric and nonparametric tools, study the theoretical properties of the involved objects and devise appropriate computational algorithms for their implementation.

COURSE DELIVERY

The course consists of roughlyt 80% of class lectures, and 20% of computer lab sessions.

LEARNING ASSESSMENT METHODS

Oral examination on the material covered in class, plus optional paper presentation or discussion of an essay elaborated by the student.

SYLLABUS

- Motivation and foundations of Bayesian inference: exchangeability and de Finetti's representation theorems
- Conjugacy, posteriors and parametric families of conjugate models
- Markov chain Monte Carlo methods for parametric inference
- The Bayesian nonparametric approach
- The Dirichlet process: definition, properties and constructions

- Hierarchical priors derived from the Dirichlet process
- Models beyond the Dirichlet process
- Markov chain Monte Carlo methods for nonparametric inference.

If time allows, the course will also cover a brief introduction of the following topics:

- Elements of Bayesian asymptotics
- Dependent priors for partially exchangeable data

SUGGESTED TEXTBOOKS AND READINGS

Lecture notes will be made available. Additional suggested reading are:

HOFF, P.D. (2009). A First Course in Bayesian Statistical Methods. Springer.
GHOSAL, S. and VAN DER VAART, A. (2017). Theory of nonparametric Bayesian inference. Cambridge University Press.
HJORT, N., HOLMES, C., MUELLER, P. and WALKER, S.G. (eds.) (2010). Bayesian Nonparametrics. Cambridge University Press.
GHOSH, J.K. and RAMAMOORTHI, R.V. (2003). Bayesian Nonparametrics. Springer.

NOTE

This course will be delivered at the ESOMAS Department.

Course webpage: http://www.master-sds.unito.it/do/corsi.pl/Show?_id=22bc

Complex networks

Complex networks

Academic year:	2017/2018
Course ID:	MAT0049
Teacher:	Prof. Giancarlo Francesco Ruffo
Teacher contacts:	0116706771, *ruffo@di.unito.it*
Year:	2nd year
Type:	D.M. 270 TAF C - Related or integrative
Credits/recognition:	6
Course SSD (disciplinary sector):	INF/01 - informatica
Delivery:	Formal authority
Language:	English
Attendance:	Optional
Type of examination:	Mixed

PREREQUISITES
A strong working knowledge of probability and linear algebra (at the level of a bachelor degree in a scientific discipline) will certainly be helpful, as is some mathematical maturity. The ability to write code is important, because programming skills are required for the coursework project.

COURSE OBJECTIVES

This module introduces the fundamental concepts, principles and methods in the interdisciplinary field of network science, with a particular focus on analysis techniques, modeling, and applications for the World Wide Web and online social media. Topics covered include graphic structures of networks, mathematical models of networks, common networks topologies, structure of large scale graphs, community structures, epidemic spreading, PageRank and other centrality measures, dynamic processes in networks, graphs visualization.

COURSE AIMS

On successful completion of this module the students will be able to:

- Define and calculate basic network graphic metrics.
- Describe structural features of socio-technical networks.
- Relate graphic properties to network functions and evolution.
- Relate local properties to global emerging patterns.
- Explore new angles to understand network collective behaviours.
- Design and conduct analysis on large network datasets.
- Visualize networks to highlight structural and global features.
- Use network analysis tools, such as igraph library (R and Python), and GePhi.

COURSE DELIVERY

A Moodle webpage is created for the course. All course materials, such as lecture notes and online resources will be shared. By using the Moodle, students will also be able to discuss ideas and questions with the lecturer and other students.

LEARNING ASSESSMENT METHODS

Oral examination (60%).

Coursework I (20%): essay writing (2000-3000 words).

Coursework II (20%): individual project on network data analysis (programming is usually required).

To pass the module students must achieve a pass mark of 60% when all elements are combined.

SYLLABUS

Network science

• Introduction to complex networks

• Graph Theory and network metrics

• Random networks

• Small-world networks

• Scale-free networks

• Evolving networks

• Degree correlations

• Communities

• Spreading phenomena

• Learning and games on networks

Case studies and applications

• Internet core structure – evolution and modelling

• Structure of the Web – PageRank and document networks

• Online social media networks - Twitter, Facebook, Amazon, …

• Network visualizations

• Similarity networks and recommendation systems

• "Rich gets richer" phenomenon

• Link, neighbourhood and community

• Cascades and epidemics

• Network structure balance

• Sentimental, temporal and spatial analysis of social media networks

SUGGESTED TEXTBOOKS AND READINGS

A.-L. Barabási, Network Science, Cambridge University Press, 2015 (online version: http://barabasi.com/networksciencebook/)

D. Easley and J. Kleinberg. Networks, Crowds, and Markets: Reasoning About a Highly Connected World, Cambridge University Press, 2010.

M. E. J. Newman. Networks: An Introduction, Oxford University Press, 2010.

NOTE

This course is borrowed from Complex Networks, delivered at the Computer Science Department.

Borrowed from: See notes below.

Course webpage: http://www.master-sds.unito.it/do/corsi.pl/Show?_id=iw8h

---

Computational methods for statistics

Computational methods for statistics

Academic year:	2017/2018
Course ID:	MAT0069
Teacher:	Prof. Raffaele Argiento
Teacher contacts:	011 670 6095, *raffaele.argiento@unito.it*
Year:	2nd year
Type:	D.M. 270 TAF C - Related or integrative
Credits/recognition:	6
Course SSD (disciplinary sector):	SECS-S/01 - statistica
Delivery:	Class Lecture
Language:	English
Attendance:	Optional
Type of examination:	Written

COURSE OBJECTIVES

This course aims at introducing the students with computational statistics methods. The program includes some computationally intensive methods in statistics, such as Monte Carlo methods, bootstrap, and permutation tests. An important part of the course will be devoted to practicals: all the methods discussed during the course will be will be implemented in the R software.

COURSE AIMS

After this course the students will be familiar with pseudo-random number generators and with the statistical software R. They will know how to sample an independent and identically distributed sequence or (pseudo) random number with a given distribution, and will be able to implement a Monte Carlo integration algorithm in R. Moreover, students will learn some of the most common statistical methods based on sampling strategies (e.g., Bootstrap, Jackknife, Bayesian estimation).

COURSE DELIVERY

Half of the lectures will be devoted to the theoretical aspects of simulation and Monte Carlo Integration and the remaining half to their practical implementation in the R software considering both the related numerical and computational issues. Exercises will be assigned during lectures and lab sessions.

LEARNING ASSESSMENT METHODS

The exam consists of two parts: the first part is a written exam on theory; the second part is a practical session with R.

SYLLABUS

- Introduction to the R statistical software.
- Pseudo-random number generator. Linear congruential generators.
- Methods for Generating Random Variables: the inverse transform method, the acceptance-rejection method, the transformation methods.
- Monte Carlo integration methods.
- Variance Reduction, the importance sampling (sampling importance resampling) and the stratified sampling.
- Monte Carlo methods in Inference in a Bayesian and frequentist framework.

- Bootstrap and Jackknife.
- Permutation Tests for Equal Distributions.

SUGGESTED TEXTBOOKS AND READINGS

- Rizzo, M.L. (2015) "Statistical Computing with R (Second Edtion)" -- Chapman & Hall/CRC The R Series.
- Ross. S.M. (2006) "Simulation 4th edition" -- Academic Press.
- Jones, O., Maillardet, R. and Robinson A. (2009). "Introduction to scientific programming and simulation usig R" -- Chapman and Hall/CRC;

NOTE

Class schedule available here.

Borrowed from: NUMERICAL AND STATISTICAL METHODS FOR FINANCE (ECO0152)
Quantitative Finance and Insurance

Course webpage: http://www.master-sds.unito.it/do/corsi.pl/Show?_id=2c4q

---

Database and algorithms

Database and algorithms

Academic year:	2017/2018
Course ID:	MAT0040
Teacher:	Prof. Rosa Meo Dott. Rossano Schifanella
Teacher contacts:	+39 011 670 68 17, *meo@di.unito.it*
Year:	1st year
Type:	D.M. 270 TAF C - Related or integrative
Credits/recognition:	12
Course SSD (disciplinary sector):	INF/01 - informatica
Delivery:	Class Lecture
Language:	English
Attendance:	Mandatory
Type of examination:	Mixed

PREREQUISITES
Knowledge on programming.

PROPEDEUTIC FOR
Complex networks, Introduction to Data Mining

COURSE OBJECTIVES

The objectives are formalized for each of the two parts of the course.

Databases

This course will teach the fundamentals of relational theory, SQL language and its relationships with relational algebra, design of data in relational databases by means of the conceptual and logical design of databases. The course will introduce the students to the basic notions of NoSQL databases, important for the new generation of databases and the management of big data. In the laboratory the students will be able to work with a practical database management system.

Algorithms

In this course students will learn several fundamental principles of algorithm design and how to implement some fundamental data structures (e.g., graphs, arrays, trees, hash tables). This course aims at providing a solid methodological background for the analysis of algorithms in terms of their correctness, complexity (in time and in space), and tractability.

COURSE AIMS

Databases

After the course students will be able to design data for relational databases, formulate a query in SQL or relational algebra, interact with a real database management system and will have the basic notions of NoSQL databases.

Algorithms

After the course students will be able to approach a problem through the design, analysis and implementation of appropriate algorithms and data structures.

COURSE DELIVERY

This course consists in two parts: the former is on Databases and the latter is on Algorithms.

Databases

The course will consist of 32 hours of frontal lessons and 16 hours of practical assignments at the computer or at assigned exercises. Personal training on the assigned exercises on both the theory and practice modules is fundamental to successfully pass the final exam.

Algorithms

The course will consist of 32 hours of frontal lessons and 16 hours of practical assignments at the computer. Personal training on the assigned exercises on both the theory and practice modules is fundamental to successfully pass the final exam.

LEARNING ASSESSMENT METHODS

Databases and Algorithms

The final exam will consist of a written test and a following oral discussion. In the written test the candidate will be asked to solve some data design problems, write in SQL a data retrieval request, present and discuss the practical assignments implemented during the course. Usually the examination on the two parts of the course (Databases and Algorithms) are held in the same day (one in the morning and the other one in the afternoon), but can be overcome separately (but in the same year).
At the end, both the tests on the two parts of the course (Databases and Algorithms) must be satisfactory to allow the student to overcome the overall examination.

During the following, oral exam (planned some days after the written part) the student will be asked to discuss the presented solution in the written part.

Students are required to pass the written test before to be admitted the oral part.

SUPPORT ACTIVITIES

The laboratory will consists in assignments that will be solved by means of practical activities at the computer that will support the theorical notions learnt during the course.

SYLLABUS

Databases

Databases; database management systems; data models; database languages; the relational model and its languages; integrity constraints; relational algebra; SQL;

Database design methodologies and models; the database design process; the Entity-Relationship model; conceptual design; requirement collection and analysis; general data modelling criteria; design strategies; qualities of a conceptual schema; a general methodology for database design; CASE tools for database design; logical design; translation towards the relational model;

Normalization theory for database design; redundancies and anomalies; functional dependencies; Boyce- Codd normal form; qualities of decompositions; third normal form; normalization and the design process

Algorithms

Problems and algorithms: solvability, correctness, complexity. Termination and non-termination. Unsolvable problems: undecidability of the halting problem. Correctness of algorithms. Mathematical induction principles: simple induction, complete induction, structural induction.

Analysis of algorithms. Complexity (in time and in space) of algorithms. Complexity of problems. Tractable vs. intractable problems and algorithms.

Data structures, Abstract Data Types (ADT), structure invariants. Sequences, arrays, linked lists, stacks, queues, trees, dictionaries, hash tables. Graph representation and primitives.

Sorting and selection.

SUGGESTED TEXTBOOKS AND READINGS

Databases

Database Systems - Concepts, Languages and Architectures
Paolo Atzeni, Stefano Ceri, Stefano Paraboschi and Riccardo Torlone
McGraw-Hill
(http://dbbook.dia.uniroma3.it)

Algorithms

Suggested book:

Introduction to Algorithms. T Cormen, C Leiserson, R Rivest, C Stein

Course webpage: http://www.master-sds.unito.it/do/corsi.pl/Show?_id=99p1

Decision and Uncertainty

Decision and Uncertainty

Academic year:	2017/2018
Course ID:	MAT0071
Teacher:	
Teacher contacts:	
Year:	2nd year
Type:	D.M. 270 TAF C - Related or integrative
Credits/recognition:	6
Course SSD (disciplinary sector):	SECS-S/06 - metodi matematici dell'economia e delle scienze att. e finanz.
Delivery:	Formal authority
Language:	English
Attendance:	Optional
Type of examination:	Written

Borrowed from: DECISION AND UNCERTAINTY (SEM0067)
Corso di studio in Quantitative Finance and Insurance

Course webpage: http://www.master-sds.unito.it/do/corsi.pl/Show?_id=nxxs

Decision theory (deactivated)

Decision theory (deactivated)

Academic year:	2016/2017
Course ID:	MAT0046
Teacher:	Prof. Paolo Ghirardato
Teacher contacts:	011 6705220, *paolo.ghirardato@unito.it*
Year:	2nd year
Type:	D.M. 270 TAF C - Related or integrative
Credits/recognition:	6
Course SSD (disciplinary sector):	SECS-S/06 - metodi matematici dell'economia e delle scienze att. e finanz.
Delivery:	Formal authority
Language:	English
Attendance:	Optional
Type of examination:	Written

NOTE

This course will be delivered at the ESOMAS Department.

In the a.y. 2017/18 the course's name and code will change to MAT0071 Decisions and Uncertainty.

Borrowed from: DECISION AND UNCERTAINTY (SEM0067)
Corso di studio in Quantitative Finance and Insurance

Course webpage: http://www.master-sds.unito.it/do/corsi.pl/Show?_id=y9cs

Econometrics

Econometrics

Academic year:	2017/2018
Course ID:	MAT0045
Teacher:	Prof. Alessandro Sembenelli
Teacher contacts:	011 6706059, *alessandro.sembenelli@unito.it*
Year:	2nd year
Type:	D.M. 270 TAF C - Related or integrative
Credits/recognition:	6
Course SSD (disciplinary sector):	SECS-P/05 - econometria
Delivery:	Formal authority
Language:	English
Attendance:	Optional
Type of examination:	Written

COURSE OBJECTIVES

The main purpose of this course is to give a general and comprehensive overview of the different econometric methodogies and approaches, focusing on what is relevant for doing and understanding empirical work on large data-bases. The number of econometric techniques that can be used is numerous and their validity often depends crucially upon the validity of the underlying assumption. This course attempts to guide students through this array of estimation and testing procedures by also offering several computer-lab sessions where students will face real world empirical cases.

COURSE AIMS

Knowledge and understanding: this course will provide students with a deep and up-to-date knowledge of modern econometric theories and related estimation and testing techniques.
- Applying knowledge and understanding: students will learn how to apply econometrics techniques to actual economic problems. To this aim students will be introduced to a professional econometric software which will be used for the computations presented in this course.
- Making judgements: the students will learn how to assess the validity of the assumptions of a wide range of econometric models with the purpose of realizing potential drawbacks or dangers in their application to relevant empirical economic questions.
- Communication skills: students will learn how to effectively organize ideas both in written and oral form, possibly with the help of presentation of scientific papers during the course.
- Learning skills: this course will enable students to understand the recent developments in econometrics and will be a suitable basis for further research work in the area.

COURSE DELIVERY

The course consists of 46 lecture hours. Strong interaction between teachers and students is warmly encouraged. Part of the course will be given at the Computer Lab.

LEARNING ASSESSMENT METHODS

75 m. (max.) written exam with closed books at the end of the course

SYLLABUS

- The Classical Linear Regression Model and Its Violations (chap. 2-3-4)

- Endogeneity, Instrumental Variables and GMM (chap. 5)
- MaximumLikelihood Estimation and Specification Tests (chap. 6)
- Models with Limited Dependent Variables (chap. 7)

SUGGESTED TEXTBOOKS AND READINGS

The course is mostly based on Verbeek's A Guide to Modern Econometrics (4th edition, 2012). For most topics lecture notes with further references will be also circulated .

NOTE

This course will be delivered at the ESOMAS Department.

Borrowed from: ECONOMETRICS (SEM0083)
Corso di studio in Economics

Course webpage: http://www.master-sds.unito.it/do/corsi.pl/Show?_id=aawf

---

Game theory

Game theory

Academic year:	2017/2018
Course ID:	MAT0047
Teacher:	
Teacher contacts:	
Year:	2nd year
Type:	D.M. 270 TAF C - Related or integrative
Credits/recognition:	6
Course SSD (disciplinary sector):	SECS-P/01 - economia politica
Delivery:	Class Lecture
Language:	English
Attendance:	Optional
Type of examination:	Written

NOTE

This course will be delivered at the ESOMAS Department.

Borrowed from: GAME THEORY (SEM0062)
Corso di studio in Economics

Course webpage: http://www.master-sds.unito.it/do/corsi.pl/Show?_id=ms52

Information theory

Information theory

Academic year:	2017/2018
Course ID:	MAT0052
Teacher:	Prof. Matteo Sereno
Teacher contacts:	0116706718, *matteo.sereno@unito.it*
Year:	2nd year
Type:	D.M. 270 TAF C - Related or integrative
Credits/recognition:	6
Course SSD (disciplinary sector):	INF/01 - informatica
Delivery:	Formal authority
Language:	English
Attendance:	Optional
Type of examination:	Written and oral

COURSE OBJECTIVES

The course represents an introduction to classical results of Shannon information theory.

COURSE AIMS

At the end of the course the student will have the capacity to apply information theory tools and approaches to both theoretical and practical problems related to information management, coding, representation, protection and information metrics.

COURSE DELIVERY

The course will be based on theretical lessons followed by in class exercises and computer based experiments. Personal training on assigned exercises is important for the success in this class.

LEARNING ASSESSMENT METHODS

The assesment comprises a written test followed by an oral examination.

SYLLABUS

The course is structured in two parts.

The first part of the course is devoted to the classical information theory. In particular, the addressed topics are: definition of information and source types, the concept of entropy, source coding, Shannon's first theorem (source coding), uniquely decodable codes, optimality of Huffman coding, models of noisy channels, definition of the channel capacity according to Shannon's theorem (channel coding).

The second part of the course is devoted to the study of source coding and channel coding algorithms used in many applications, communication systems and networks. The selected topics include arithmetic coding, the Lempel-Ziv-Welch algorithms and state of the art standards for image and video compression. As far as channel coding is regarded the course will introduce linear block codes, cyclic codes, convolutional codes and fountain codes.

SUGGESTED TEXTBOOKS AND READINGS

R. W. Yeung, "Information Theory and Network Coding, ISBN: 978-0-387-79233-0

Thomas M. Cover, Joy A. Thomas, "Elements of Information Theory, 2nd Edition", ISBN: 978-0-471-24195-9

NOTE

This course will be delivered at the Computer Science Department.

Course webpage: http://www.master-sds.unito.it/do/corsi.pl/Show?_id=tsx1

Introduction to data mining

Introduction to data mining

Academic year:	2017/2018
Course ID:	MAT0051
Teacher:	Roberto Esposito Prof. Rosa Meo
Teacher contacts:	0116706714, *roberto.esposito@unito.it*
Year:	2nd year
Type:	D.M. 270 TAF C - Related or integrative
Credits/recognition:	6
Course SSD (disciplinary sector):	INF/01 - informatica
Delivery:	Class Lecture
Language:	English
Attendance:	Optional
Type of examination:	Written

PREREQUISITES
Databases and Algorithms, Programming

COURSE OBJECTIVES

The objectives of the course will be introduce students to the field of Data Mining and Machine Learning, that merge together competencies of statistics and computer science.

The course will teach the differences between tasks and models and will introduce the students to some of the popular models in Machine Learning such as binary classification and related tasks, transformation of a binary classification model into a multiple class model, concept learning by means of logical formulas, tree models and their purposes, rule models, subgroup discovery, linear models (least squares, regression), perceptron, Support Vector Machines, Kernel methods.

The course will introduce the algorithms for the training of the models.

The laboratory part of the course will introduce the students to a practical open software suite that includes the algorithms of learning of the models seen during the course (and much more).

COURSE AIMS

The results of the learning outcomes will be mastering some the main concepts in Data Mining and Machine Learning and using them in the context of a practical open software suite for data analysis and machine learning.

COURSE DELIVERY

The course lessons will be both theorical and practical.

LEARNING ASSESSMENT METHODS

The final exam will be oral in which the students will be asked to show that they master the theorical lessons (knowledge of the models and of their purposes) and use of the practical software suite (Weka) for data analysis in some use cases.

SUPPORT ACTIVITIES

Machine learning experiments in Laboratory with a software suite for Data Mining.

The laboratory will be a practical support to the learning of the theorical lessons by means of practical data analysis assignments on public data-sets.

SYLLABUS

Tasks and models; Binary classification and related tasks; Beyond binary classification (transformation of a binary classification model into a multiple class model; Concept learning by means of logical formulas; Version Space; learning hypothesis by means of Horn clauses; Tree models (decision trees, regression trees, features trees, ranking trees); rule models (list of rules and sets of rules); subgroup discovery; linear models (least squares, regression); perceptron; Support Vector Machines; Kernel methods;

SUGGESTED TEXTBOOKS AND READINGS

Peter Flach, Machine Learning - The Art and Science of Algorithms that Make Sense of Data, Cambridge University Press, 2012.

NOTE

This course is borrowed from Machine Learning and Intelligent Data Analysis and will be delivered at the Computer Science Department.

Course webpage: http://www.master-sds.unito.it/do/corsi.pl/Show?_id=f8ey

---

Multivariate statistical analysis

Multivariate statistical analysis

Academic year:	2017/2018
Course ID:	MAT0041
Teacher:	Prof. Pierpaolo De Blasi
Teacher contacts:	*pierpaolo.deblasi@unito.it*
Year:	1st year
Type:	D.M. 270 TAF C - Related or integrative
Credits/recognition:	6
Course SSD (disciplinary sector):	SECS-S/01 - statistica
Delivery:	Formal authority
Language:	English
Attendance:	Mandatory
Type of examination:	Written

PREREQUISITES
Probability Theory

PROPEDEUTIC FOR
Statistical Machine Learning

COURSE OBJECTIVES

The course aims at introducing multivariate analysis in statistical modeling. All the methods will be implemented on real datasets in the R language.

COURSE AIMS

The student will learn the basic techniques for analyzing multi-dimensional data (including visualization), study multivariate distributions and their properties, discuss various methods for dimension reduction.

COURSE DELIVERY

The course is composed of 48 hours of class lectures. Examples and exercises will be dealt with at class through the R language.

LEARNING ASSESSMENT METHODS

Problem Sets:
There will be 2 problem sets assigned throughout the course. They will be posted in due time on
https://sites.google.com/a/carloalberto.org/pdeblasi/teaching
together with an indication of the deadline.
Problem sets must be submitted and there are no late submissions. They are an essential part of the course, providing students with a guide on how well they are grasping the material on a "real time" basis. They request the solution of exercises, solution which might require the use of a statistical software. Students are encouraged to work in groups on the problem sets. However, students should understand the material on their own, and hand in their own problem sets.

Exam:
There will be a final exam, check out for dates on
http://www.master-sds.unito.it

The final examination consists of a written test, either a short or a long test according to the problem sets. Specifically,

(1) First 2 exam dates: the course grade is determined by the problem sets and the final exam. The final exam consists of a short written test (1h) on the part of the program not covered by problem sets followed by an oral examination. The final grade will be a combination of the problem sets grades (70%), and the final exam grade (45%). For students who have failed to submit the solutions of the problem sets, case (2) below applies.

(2) From the 3rd exam date on: the final exam consists of a long written test (3h) on the whole program and the final grade will be determined solely by it (100%).

SYLLABUS

- Introduction
    - summary statistics for multivariate data
    - multivariate data visualization
    - multivariate Gaussian distributions
- Principal Component Analysis (PCA):
    - geometric and algebraic basics of PCA
    - calculation and choice of components
    - plotting PCs, interpretation
- Factor Analysis (FA):
    - model definition and assumptions
    - estimation of loadings and communalities
    - choice of the number of factors
    - factor rotation
- Canonical Correlation Analysis:
    - computation and interpretation
    - relationship with multiple regression
- Discriminant Analysis and Classification:
    - classification rules
    - linear and quadratic discrimination
    - error rates
- Cluster Analysis:
    - measure of similarity
    - hierarchical clustering
    - K-means clustering
    - model based clustering

SUGGESTED TEXTBOOKS AND READINGS

The bibliography, to be confirmed at the beginning of the course, is:

- R.A. Johnson and D.W. Wichern (2007). Applied Multivariate Statistical Analysis. Prentice-Hall, 6th Ed.

Suggested readings:
- Hastie, Tibshirani, Friedman (2009). The Elements of Statistical Learning, 2nd ed., Springer
- Afifi A., May S., Clark V.A. (2012). Practical Multivariate Analysis, 5th ed., Chapman & Hall/CRC
- Everitt B. (2005). An R and S-PLUS Companion to Multivariate Analysis. Springer
- Rencher A. C., Christensen W. F. (2012). Methods of multivariate analysis, 3rd ed., Wiley
- Rencher A.C. (1992). Interpretation of canonical discriminant functions, canonical variates and principal components. The American Statistician 46, 217-225.

Course webpage: http://www.master-sds.unito.it/do/corsi.pl/Show?_id=u23n

Partial differential equations

Partial differential equations

Academic year:	2017/2018
Course ID:	TCPDE16-17
Teacher:	Dott. Sandro Coriasco (Tutor)
Teacher contacts:	0116702803, *sandro.coriasco@unito.it*
Year:	2nd year
Type:	D.M. 270 TAF F - Other activities
Credits/recognition:	0
Course SSD (disciplinary sector):	MAT/05 - analisi matematica
Delivery:	Class Lecture
Language:	English
Attendance:	Optional
Type of examination:	Practice test

COURSE OBJECTIVES

To give a general overview of various methods to solve (free, boundary value, initial value) problems associated with PDEs commonly met in modelling and applications.

COURSE AIMS

It is expeced that the students are able, at the end of the lectures, to solve basic exercises of the types examined during the lectures.

COURSE DELIVERY

Standard lectures and tutorial activities in classroom.

LEARNING ASSESSMENT METHODS

Solution of exercises (Practice test, optional).

SUPPORT ACTIVITIES

.

SYLLABUS

The method of characteristics. First order linear PDEs in two variables, First order linear PDEs in n variables, Semilinear first order PDEs in n variables.

Second order linear, semilinear and quasilinear PDEs. Elliptic, parabolic and hyperbolic PDEs.

Methods based on (generalized) Fourier series for boundary value problems and initial value problems associated with linear PDEs.

The Fourier-Laplace method for free problems, boundary value problems and initial value problems associated with linear PDEs. Laplacian, heat and wave operators.

SUGGESTED TEXTBOOKS AND READINGS

- L.C. Evans, Partial Differential Equations. AMS (2010)

- S. Salsa, Equazioni a derivate parziali. Springer (2010)

NOTE

Lectures will take place in classroom Vercelli (13-14/9) and Novara (19-21/9). Please, notice that the first lecture on tue 13/9 will take place from 11:00 to 13:00.

The final assessment test will consist in the solution of some exercises of the types examined during the lectures. It has a self-evaluation character, it is fully optional and provides no credits.

Course webpage: http://www.master-sds.unito.it/do/corsi.pl/Show?_id=kefo

---

Probability theory

Probability theory

Academic year:	2017/2018
Course ID:	MAT0034
Teacher:	Prof. Laura Sacerdote Prof. Federico Polito
Teacher contacts:	+39 011 6702919, *laura.sacerdote@unito.it*
Year:	1st year
Type:	D.M. 270 TAF B - Distinctive
Credits/recognition:	9
Course SSD (disciplinary sector):	MAT/06 - probabilita' e statistica matematica
Delivery:	Formal authority
Language:	English
Attendance:	Mandatory
Type of examination:	Written and oral

PREREQUISITES
An undergrauate level class in Probability and good knowledge of real analysis. Good abilities in elementary probabilistic problem solving are also necessary for the success in this class.

PROPEDEUTIC FOR
Stochastic Processes, Statistics for Stochastic Processes and EDS-Stochastic Dfferential Equations use concepts and tools introduced in this course.

COURSE OBJECTIVES

Topics taught in this class are essential tools required to a statistician and a probabilist. They are fundamental for any modern mathematician. Students re-think to subjects of their undergraduate studies with a different level of abstraction.This new approach allows them to control some advanced methods of probability theory, useful for applications as wel as for research.

COURSE AIMS

Students attain a detailed knowledge of the foundations of the theory of probability and related topics in measure theory. They attain good ability in probabilistic problem solving becoming able to deal both with theoretical and applied problems related with conditional expectation, convergence features, characteristic functions and martingales.
They become able to prove new results related with the studied theory, furthermore they become used to learn using different textbooks.

COURSE DELIVERY

There will be 72 hours of lessons, including 16 hours of in class exercises. Personal training on assigned exercises is important for the success in this class.

LEARNING ASSESSMENT METHODS

The final exam includes both a written and an oral tests. The two tests are scheduled on different dates. The written test is valid until the following oral exam. The written test requires the solution of two exercises and the proof of a theorem (selected from those discussed during classes). It is mandatory to pass this test to be admitted to the oral test. The use of textbooks and personal notes during the written test is not allowed. The oral examination includes a discussion on the written test as well as two question, taken at random by the student. The list of the possible questions for the oral examination will be provided in advance.

SUPPORT ACTIVITIES

The course include exercises classes; extra exercises are suggested as homework.

SYLLABUS

Overview of elementary probability. Construction of probability measures on R and random variables. Integrals over probability measures. Independent random variables. Distributions on Rn. Sums of random variables. 0-1 Laws. Convergence of sequences of random variables. Weak convergence and characteristic functions. Laws of large numbers and central limit theorem. Conditional expectations. Discrete time martingales, optional stopping , Doob decomposition and martingale inequalities. Convergence properties of discrete time martingales. Introduction to continuous time martingales.

SUGGESTED TEXTBOOKS AND READINGS

Textbooks:

- Çınlar, E., "Probability and Stochastics", Springer, 2011.
- Durrett, R. "Probability: Theory and Examples", Cambridge, 2010.

Further suggested books:

- Williams, D., "Probability with Martingales", Cambridge University Press, 2001;
- Shiryaev, A.N., "Probability", Springer, 1996;
- Billingsey, P., "Probability and Measure", Wiley-Interscience, 1995;
- Feller, W. "Introduction to Probability Theory and Applications", 2 Volumes, Wiley, 2008;
- Varadhan, S.R.S., "Probability Theory", AMS, 2001;
- Jacod, J., Protter, P., "Probability Essentials", Springer, 2004.

NOTE

Note that the exams' rules have changed since last academic year. These set of rules will be applied from the session of January 2018.

The old set of rules (academic year 2016/2017 - see on the corresponding tab on the top of the page) will be valid until the session of September 2017 included.

Attention: students must register on the exams web-page to be admitted to the written/oral exams (in time). Not registered students will not be admitted being impossible to register their final mark.

Course webpage: http://www.master-sds.unito.it/do/corsi.pl/Show?_id=ej1a

---

Programming for data science

Programming for data science

Academic year:	2017/2018
Course ID:	MAT0036
Teacher:	Dott. Marco Beccuti
Teacher contacts:	+39 011 6706780, *marco.beccuti@unito.it*
Year:	1st year
Type:	D.M. 270 TAF F - Other activities
Credits/recognition:	3
Course SSD (disciplinary sector):	INF/01 - informatica
Delivery:	Formal authority
Language:	English
Attendance:	Mandatory
Type of examination:	Mixed

PREREQUISITES
ELEMENTS OF STATISTICS Basic knowledge in Calculus as provided by the first year Mathematics course. ELEMENTS OF COMPUTER SCIENCE No specific computer science knowledge is required.

COURSE OBJECTIVES

Aim of the course is to introduce methods, techniques and related computer science instruments for the analysis of experimental data.

It provides the basic knowledge to use computer science applications as Spreadsheet (e.g. Excel, Calc,...) and programming languages for statistical computing and graphics (e.g. R programming language)

COURSE AIMS

KNOWLEDGE AND UNDERSTANDING – Completing the course students will be able to:

1) use suitable descriptive and inferential statistics techniques to describe and understand the phenomena being studied;

2) manage suitable computer science instruments such as worksheet or dedicated software programs for statistical data analysis.

APPLYING KNOWLEDGE AND UNDERSTANDING – Students will perform the statistical analyses required by the problem under study by selecting the most computationally and graphically suitable computer science support.

MAKING JUDGEMENTS – Students will decide which statistical techniques to use according to the available data sets to describe and understand the phenomena under consideration.

COMMUNICATION – The student will be able to justify the choices for the analysis to be performed and to give a synthetic description of the techniques employed and of the results obtained.

COURSE DELIVERY

ELEMENTS OF COMPUTER SCIENCE

The course consists of 10 hours of lectures, and 14 hours of laboratories . Laboratories include exclusively practical activities.

The slides presented during lectures are available to students as online materials.

Attendance to lessons is not mandatory, but highly recommended due to the necessity of learning and employing specific computer science instruments.

LEARNING ASSESSMENT METHODS

The exam consists of a written test and requires a practice exercise on R programming languages

WRITTEN EXAMINATION:

- ten multiple choice questions on course topics (4 options, with the possibility of 0-4 correct options);

- a practice exercise on R programming languages

The maximum possible score is 30 cum laude.

SYLLABUS

ELEMENTS OF COMPUTER SCIENCE

- Introduction to R programming language;
- Basic R functionalities:
  - Data structures: vector, matrix, array, list, data.frame \ldots;
  - Apply operators;
  - Input/output operator;
  - Package and library.
- Programming with R:
  - Function;
  - Flow control: if,for, while, break ... statements;
  - Debugging in R.
- Probability distributions:
  - Densities;
  - Cumulatives,
  - Quantiles;
  - Randon numbers.
- Statistical graphics:
  - Graphical devices;
  - High level plot;
  - Low level plots.
- Statistical functions

SUGGESTED TEXTBOOKS AND READINGS

- P. Dalgaard, Introductory Statistics with R, Springer 2008

- The R Manuals: An Introduction to R (http://cran.r-project.org/doc/manuals/r-releas /Rintro.pdf)

The teaching material used for lessons and a series of practical exercises are available on the web site of the course.

Course webpage: http://www.master-sds.unito.it/do/corsi.pl/Show?_id=h783

Simulation

Simulation

Academic year:	2017/2018
Course ID:	MAT0050
Teacher:	Prof. Gianfranco Balbo
Teacher contacts:	(+39) 011 670 n6740, *gianfranco.balbo@unito.it*
Year:	2nd year
Type:	D.M. 270 TAF C - Related or integrative
Credits/recognition:	6
Course SSD (disciplinary sector):	INF/01 - informatica
Delivery:	Class Lecture
Language:	English
Attendance:	Optional
Type of examination:	Mixed

PREREQUISITES
The basis of Probability Theory and Elements of Statistics are assumed to be known by the students. The knowledge of a general purpose programming language is required in order to implement the simulators required as part of the homework exercises and of the final project.

COURSE OBJECTIVES

Simulation is one of the most common techniques used for the evaluation of the performance and if the reliability of Discrete Event Dynamic Systems (DEDS) often modelled with Stochastic Processes. Discrete Event Simulation consists on the execution of a program which results in the production of a realization of a stochastic process driven by Monte-Carlo methods. Learning how to construct a simulator is the main objective of this course, together with the development of the techniques needed for the statistical analysis of the simulation output. To deeply understand the difficulty of writing an efficient simulator equipped with the output analysis components, students will be required to write a few simple simulators "from scratch" without using available tools and libraries.

COURSE AIMS

At the end of the course the students will be able to perform the simulation of non-trivial Discrete Event Sysrtems. The exercises and the final project will provide the students with the capability of writing the simulators using a general purpose programming language of their choice. Having developed the simulators "fromn scratch" will allow the students to understand the potentials and the limits of the Discrete Event Simultauion technique, thus providing them with the capability of using professional simulators with competence

COURSE DELIVERY

The course will be based on theretical lessons as well as on the soltion of class exercises. Computer implementations will be required as homework assignments. Personal training on assigned exercises is important for the success in this class.

LEARNING ASSESSMENT METHODS

The final examination will consist in the discussion of a project developed individually by the students used as the basis for asking questions on the theoretical aspects of the exercise. Students will not be required to be able to reproduce the derivations used to obtain the results discussed during the course, but will have to know the definitions and the applications of the theory.

The final grade will be out of thirty.

SUPPORT ACTIVITIES

Exercises will be assigned as homework. The course will include sessions devoted to the discussion of the solutions of selected homework, as well as to the solution of additional exercises.

SYLLABUS

Introduction

- Discrete Event Dynamic Systems nodelling and performance indices
- Formalisms for System Modelling

Operational Analysis

- Introduction, measurable entities and operational variables
- Flow analysis in queuing networks
- Balance equations
- Queuing networks with product form solution
- Computational algorithms for product form solution

Simulation

- Introduction to Discrete Event Simulation
- Construction of a simple simulator
- Random number generators
- Generating instances of random variables
- Data structures and basic architecture of a simulator
- Statistical analysis of simulation output
- Validation

SUGGESTED TEXTBOOKS AND READINGS

-Lawrence M. Leemis and Stephen K. Park, "Discrete-Event Simulation: a first course", Pearson Prentice Hall, 2006.

-George S. Fishman, "Principles of Discrete Event Digital Simulation", John Wiley & Sons, 1978.

-Hisashi Kobayashi, "Modeling and Analysis: An Introduction to System Performance Evaluation Methodology", Addison-Wesley, 1978.

-Kishor S. Trivedi, "Probability and Statistics with Reliability, Queueing and Computer Science Applications",

Prentice Hall, 1982.

-Giuseppe Iazeolla, "Introduzione alla Simulazione Discreta", Boringhieri.

-Additional Lecture Notes will be made available to the students.

NOTE

This course is borrowed from Simulation and Modelling and will be delivered at the Computer Science Department.

Course webpage: http://www.master-sds.unito.it/do/corsi.pl/Show?_id=3kh8

Simulation models for economics

Simulation models for economics

Academic year:	2017/2018
Course ID:	MAT0048
Teacher:	Prof. Pietro Garibaldi
Teacher contacts:	0116706079, *pietro.garibaldi@unito.it*
Year:	2nd year
Type:	D.M. 270 TAF C - Related or integrative
Credits/recognition:	6
Course SSD (disciplinary sector):	SECS-P/01 - economia politica
Delivery:	Formal authority
Language:	English
Attendance:	Optional
Type of examination:	Mixed

NOTE

This course will be delivered at the ESOMAS Department.

Borrowed from: NUMERICAL METHODS IN ECONOMICS (SEM0080)
Corso di studio in Economics

Course webpage: http://www.master-sds.unito.it/do/corsi.pl/Show?_id=wpkt

Statistical inference

Statistical inference

Academic year:	2017/2018
Course ID:	MAT0035
Teacher:	Prof. Stefano Favaro
Teacher contacts:	+39 011 6705724, *stefano.favaro@unito.it*
Year:	1st year
Type:	D.M. 270 TAF C - Related or integrative
Credits/recognition:	9
Course SSD (disciplinary sector):	SECS-S/01 - statistica
Delivery:	Class Lecture
Language:	English
Attendance:	Mandatory
Type of examination:	Written

PREREQUISITES
Mathematical, probabilistic and statistical tools acquired in the three-year undergraduate program. A detailed list of the required backgroud will be provided during the first lecture.

COURSE OBJECTIVES

Ability to apply statistical concepts and statistical techniques with respect to the point estimation, hyphotesis testing and confidence sets.

COURSE AIMS

Knowledge and understanding
Advances knowledge of statistical modeling via point estimation, hypothesis testing and confidence intervals.

Applying knowledge and understanding
Ability to convert various problems of practical interest into statistical models and make inference on it.

Making judgements
Students will be able to discern the different aspects of statistical modeling.

Communication skills
Students will properly use statistical and probabilistic language arising from the classical statistics.

Learning skills
The skills acquired will give students the opportunity of improving and deepening their knowledge of the different aspects of statistical modeling.

COURSE DELIVERY

Main lectures are devoted to the theorerical aspects of statistical inference. Exercises will be assigned during these lectures. Lecture devoted to exercises are included in the course.

LEARNING ASSESSMENT METHODS

The exam consists of two parts: the first part is a formal discussion of one of the main topics of statistical infence; the second part consists of two exercises, typically with more than two questions.

SYLLABUS

Properties of random samples: random samples and their distributions; functions of random samples; Hoeffding's and Bernstein's inequality; Efron-Stein inequality; generating random samples; the likelihood function and the formal likelihood principle; exponential families of distributions.

Estimators and principle of data reduction: sufficient statistics; minimal sufficient statistics; Fisher factorization and Lehmann-Scheffé theorem; finite-sample properties of estimators; Cramer-Rao lower bound and Rao-Blackwell theorem; large-sample properties of estimators.

Point estimation: moment-based estimators; maximum likelihood estimators; the expectation-maximization algorithm; finite-sample properties of maximum likelihood estimators; large-sample properties of maximum likelihood estimators; Cramer theorem.

Hypothesis testing: probabilistic structure of hypothesis tests; Neyman-Pearson lemma; likelihood ratio test; the Karlin-Rubin test; asymptotics for likelihood ration test; other large-sample hypothesis tests; hypothesis testing under the Gaussian model; oneway analysis of variance

Regression models; simple and multiple linear regression; least squares estimators and maximum likelihood estimators; Gauss-Markov theorem; hypothesis testing for regression models; generalized linear regression; the logistic regression model; the poisson regression models.

SUGGESTED TEXTBOOKS AND READINGS

Bickel, P.J. and Doksum, K.A. (2015). Mathematical Statistics: basic ideas and selected topics. Chapman and Hall/CRC

Casella, G. and Berger, R.L. (2008). Statistical inference. Duxbury Press

Course webpage: http://www.master-sds.unito.it/do/corsi.pl/Show?_id=0b44

Statistical machine learning

Statistical machine learning

Academic year:	2017/2018
Course ID:	MAT0043
Teacher:	Prof. Matteo Ruggiero (Lecturer)
Teacher contacts:	011 670 5758, *matteo.ruggiero@unito.it*
Year:	2nd year
Type:	D.M. 270 TAF C - Related or integrative
Credits/recognition:	6
Course SSD (disciplinary sector):	SECS-S/01 - statistica
Delivery:	Formal authority
Language:	English
Attendance:	Optional
Type of examination:	Mixed

COURSE OBJECTIVES

The course introduces methods and models to extract important patterns and trends from big amount of data, and presents basic concepts of machine learning and data mining from a statistical perspective. All the methods will be introduced from a theoretical point of view and implemented on real datasets in the R language.

COURSE AIMS

Knowledge and understanding

- Advances knowledge of parametric and nonparametric models for prediction and classification

Applying knowledge and understanding

- Ability to convert various problems and data into statistical models to perform several type of prediction/classification.

Making judgements

- Students will be able to discern the different aspects of statistical learning in modern settings.

Communication skills

- Students will properly use statistical language to comunicate the results of their findings.

Learning skills

- The skills acquired will give students the opportunity of improving and deepening their knowledge of statistical modeling.

COURSE DELIVERY

Half of the lectures are devoted to the theorerical aspects of statistical machine learning and the remaining half to their practical implemetation in the R software considering both the related numerical and computational issues. Exercises will be assigned during lectures and lab sessions.

LEARNING ASSESSMENT METHODS

The exam consists of three parts: the first part is a written exam on theory; the second part is a practical session with R; the last part is an oral discussion.

SYLLABUS

Introduction

- Context and motivations;
- Trade-off between goodness-of-fit and model complexity (i.e. variance and bias);
- Model selection techiniques (AIC, BIC, cross validation);
- Training and test set;

Regression

- Variable selection and shrinkage
- Elements of nonparametric regression
- Structured nonparametric regression

Classification:

- Logistic and multilogit regression;
- Elements of nonparametric classification
- Ensable techniques (bagging, boosting, random forest);

Miscellanea:

- Tools for data visualization;
- Computational tools (parallel computing, recursive estimations);

SUGGESTED TEXTBOOKS AND READINGS

- AZZALINI, SCARPA. Data analysis and data mining . Oxford University Press
- HASTIE, TIBSHIRANI AND FRIEDMAN. The elements of statistical learning: data mining, inference and prediction. Springer-Verlag.

NOTE

This course will be delivered at the ESOMAS Department.

Course webpage: http://www.master-sds.unito.it/do/corsi.pl/Show?_id=mb7n

---

Statistics for stochastic processes

Statistics for stochastic processes

Academic year:	2017/2018
Course ID:	MAT0038
Teacher:	Prof. Elvira Di Nardo
Teacher contacts:	0116702862, *elvira.dinardo@unito.it*
Year:	1st year
Type:	D.M. 270 TAF B - Distinctive
Credits/recognition:	6
Course SSD (disciplinary sector):	MAT/06 - probabilita' e statistica matematica
Delivery:	Formal authority
Language:	English
Attendance:	Optional
Type of examination:	Mixed

PREREQUISITES
Good knowledge of probability theory and the basics of stochastic processes. In more details you will need - laws of large numbers and central limit theorems - measure theory - conditional expectations - L^p spaces with respect to a probability measure - Hilbert spaces (some introductory material on this topic is present in the text books)

COURSE OBJECTIVES

The goal of lectures is to introduce statistical inference for time series taking into account both the theoretical/mathematical aspects and their practical application to data analysis.

Time series are considered, aiming to characterize properties, asymptotic behavior, estimations and forecasting, spectral analysis as well as decomposition in trend and seasonal components. Such concepts are applied to the analysis of simulated data or existing databases in order to infer and validate a model supporting the data.

COURSE AIMS

At the end of the course, students will have understood how to model time series with focus on forecasting and estimation of the moments, of the spectrum and of the parameters of time series models.

Moreover they will know which are the main steps of the analysis of a dataset, and which tools are available to this aim:
- descriptive statistics, moment and spectrum estimation
- formulation of models, parameter estimation, model selection, model verification
- forecasting

COURSE DELIVERY

We will mainly deliver frontal lectures, but a computer lab is also included. During the lectures we will alternate a formal presentation of some topics, including proofs and technical details, with a more informal part where we will introduce some concepts that are useful for the analysis of data sets. In the computer lab we will use R to simulate and analyse datasets from ARMA processes or existing databases. We refer to some particular packages useful to deal with simulations, decompositions and forecasting.

LEARNING ASSESSMENT METHODS

Who wants to be examined on the syllabus of

a.a.<2015/16: send an e-mail to Elvira Di Nardo, one week before the practical session, to organize the methods

a.a.=2015/16: a practical session on the analysis of a dataset in the computer lab is followed by writing a short essay on one of the arguments introduced by Prof. Sirovich. The final evaluation with a regular oral examination will be after the correction of this essay and the analysis in the computer lab a couple of days later.

a.a.=2016/17: a practical session on the analysis of a dataset in the computer lab is followed by writing a short essay on one of the arguments introduced by Prof. Rinott. The final evaluation with a regular oral examination will be after the correction of this essay and the analysis in the computer lab a couple of days later.

SUPPORT ACTIVITIES

Computer lab.

SYLLABUS

Time series: ARMA processes, covariance and spectrum. Estimation and elimination of trend, seasonal components and periodicities. Linear filtering, causality and smoothing. Recursive methods for computing the best linear predictors: Durbin-Levinson algorithm, innovations algorithms. Spectral representation of simple processes. Herglotz Theorem. Spectral density, the relation to characteristic functions and their inversion in probability. Computing the spectral density for ARMA models. Applying the spectral density to obtain causal invertible models. Stochastic integrals: definition, existence, examples, properties, relation to spectral distributions. Spectral representation of stationary processes by stochastic integrals and applications to prediction in ARMA. Estimation of the mean, the covariance, the partial autocorrelation. Estimation of the parameters and model selection. Diagnostic tools. Asymptotic theory: m-dependent variables and the associated CLT. Computer lab: simulation and statistical analysis of time series with R.

SUGGESTED TEXTBOOKS AND READINGS

Lectures will not adhere to the material of any single text, but the students can find material on the topics we teach on different books. References for each topic will be made available during the course.

- Brockwell and Davis, Introduction to Time Series and Forecasting, Second Edition. Springer texts in statistics. 2002

- Brockwell and Davis, Time Series, theory and methods, Springer (collana SSS), New York, 1991

- Priestley, Spectral Analysis and Time Series, Academic Press, Vol I, 1981

- Shumway and Stoffer, Time series Analysis and Its Applications, Springer, 2011.

For the Lab, refer to www.stat.pitt.edu/stoffer/tsa4/

Course webpage: http://www.master-sds.unito.it/do/corsi.pl/Show?_id=i6y2

Stochastic differential equations

Stochastic differential equations

Academic year:	2017/2018
Course ID:	MAT0044
Teacher:	Prof. Enrico Priola
Teacher contacts:	0116702883, *enrico.priola@unito.it*
Year:	2nd year
Type:	D.M. 270 TAF B - Distinctive
Credits/recognition:	6
Course SSD (disciplinary sector):	MAT/05 - analisi matematica
Delivery:	Class Lecture
Language:	English
Attendance:	Mandatory
Type of examination:	Oral

COURSE OBJECTIVES

The course aims to put the student in a position to understand the mathematical formulation of various models of applied sciences and financial mathematics which involve stochastic differential equations. The course uses probabilistic concepts and tools that are developed in the course ``Probability Theory'' and elements of Functional Analysis (see ``Analysis''); these concepts are briefly mentioned in the first lectures. The proofs of the main results of the course are carried out completely. They show important links between Analysis and Probability. To improve the skills of reading and study the teacher proposes the reading of some scientific articles. Together with the course ``Stochastic Processes'' it suggests an approach to the research in stochastic environments. The course also provides basic concepts on parabolic equations of Kolmogorov type.

COURSE AIMS

Knowledge of the Ito stochastic integral and the related stochastic differential equations. Knowledge of the relations between stochastic differential equations and Kolmogorov equations. Ability to apply stochastic differential equations to solve problems in applied sciences.

COURSE DELIVERY

Lessons in the classroom.

LEARNING ASSESSMENT METHODS

Oral examination. Questions on the program (theorems, remarks and examples). Concerning the proofs we require to know in details 3 important proofs. Such required proofs are given in the folder ``Teaching material'' below. This folder also contains more information on the examination.

SYLLABUS

- Reminder of basic notions on measure theory and probability theory. Multidimensional Gaussian distributions.

- Brownian motion (its construction by means of Haar functions; regularity properties of trajectories); the Wiener measure.

- The Doob L^p estimates for martingales with continuous paths.

- The Ito stochastic integral (basic properties; comparison between the stochastic integral and the Riemann-

Stieltjes integral)

- The Ito formula and its applications

- Stochastic differential equations (existence and uniqueness theorems)

- Markov property of solutions of stochastic differential equations; connections between stochastic differential equations and parabolic Kolmogorov equations

- Possible applications of stochastic differential equations to Mathematical Finance and Population Dynamics

SUGGESTED TEXTBOOKS AND READINGS

- Lectures notes

- I. Karatzas and S. E. Shreve, Brownian Motion and Stochastic Calculus, Springer-Verlag, New York, Second Edition, 1991.

- Arnold, L., Stochastic Differential Equations, Theory and Applications, New York. John Wiley & Sons. 1974

- P. Baldi: Equazioni differenziali stocastiche e applicazioni, Pitagora Ed., Bologna, 2000.

NOTE

This course will be delivered at the ESOMAS Department.

Course webpage: http://www.master-sds.unito.it/do/corsi.pl/Show?_id=0waw

---

Stochastic modelling for statistical applications

Stochastic modelling for statistical applications

Academic year:	2017/2018
Course ID:	MAT0039
Teacher:	Prof. Matteo Ruggiero
Teacher contacts:	011 670 5758, *matteo.ruggiero@unito.it*
Year:	1st year
Type:	D.M. 270 TAF B - Distinctive
Credits/recognition:	6
Course SSD (disciplinary sector):	MAT/06 - probabilita' e statistica matematica
Delivery:	Class Lectures
Language:	English
Attendance:	Optional
Type of examination:	Oral

PREREQUISITES
PROBABILITY THEORY (MAT0034)

PROPEDEUTIC FOR
BAYESIAN STATISTICS (MAT0070)

COURSE OBJECTIVES

The course introduces to the theory of Markov chains, in discrete and continuous time, and Lévy processes. These are nowadays considered essential probabilistic instruments which should be part of a modern statistician's toolbox. As an illustrative application, some time will be devoted to introduce the basics of Markov chain Monte Carlo methods, with a few examples of the most widely used strategies.

COURSE AIMS

The student will possess a quite detailed knowledge of Markov chain theory in discrete and continuous time, knowing how to formulate a model relative to the required task or application and how to analyse its properties, and will have acquired sufficient familiarity with Levy processes and Markov chain Monte Carlo methods to be able to autonomously comprehend a scientific paper on those topics.

COURSE DELIVERY

The course is composed of 48 hours of class lectures.

LEARNING ASSESSMENT METHODS

The final assessment consists in an oral examination on the material covered during the course.

The possibility of presenting a scientific paper whose content is coherent with the course's syllabus will be discussed at the beginning of the course.

SYLLABUS

- Introduction: stochastic processes; finite dimensional distributions; existence theorem; classes of stochastics processes based on path properties.

- Markov chains: transition matrices, Chapman-Kolmogorov equations, strong Markov property, classification of states, invariant measures, reversibility, convergence to equilibrium.

- Elements of Markov chain Monte Carlo methods: Monte Carlo principle; Markov chain Monte Carlo principle; Metropolis-Hastings algorithm; Gibbs sampler; slice sampler.

- Continuous time Markov chains: transition functions and Chapman-Kolmogorov equations; transition rates and infinitesimal generators; backward and forward Kolmogorov equations; embedded chains and holding times; uniformisation; stationarity; reversibility; scaling limits and diffusion approximations.

- Levy processes: definition; infinite divisibility; Levy-Khintchine formula; Levy-Ito decomposition; Poisson random measures.

The material introduced will be throughly discussed and illustrated with numerous examples.

An 8 hours module, included in the courseload, will be taught by visiting professor Dario Spanò on "Introduction to stochastic modelling in Population Genetics".

SUGGESTED TEXTBOOKS AND READINGS

Main references:

- NORRIS, J.R. Markov chains. Cambridge Series in Statistical and Probabilistic Mathematics.
- BREMAUD, P. Markov Chains. Springer.

Further suggested readings:

- BILLINGSLEY, P. Probability and measure. Wiley.
- GRIMMETT, G.R. and STIRZAKER, D.R. Probability and random processes. Oxford University Press.
- KARLIN and TAYLOR. A first Course in Stochastic Processes. Academic Press.
- KARLIN and TAYLOR. A second Course in Stochastic Processes. Academic Press.

Course webpage: http://www.master-sds.unito.it/do/corsi.pl/Show?_id=lrta

---

Stochastic processes

Stochastic processes

Academic year:	2017/2018
Course ID:	MAT0037
Teacher:	Prof. Laura Sacerdote Prof. Cristina Zucca
Teacher contacts:	+39 011 6702919, *laura.sacerdote@unito.it*
Year:	1st year
Type:	D.M. 270 TAF B - Distinctive
Credits/recognition:	6
Course SSD (disciplinary sector):	MAT/06 - probabilita' e statistica matematica
Delivery:	Formal authority
Language:	English
Attendance:	Mandatory
Type of examination:	Written and oral

PREREQUISITES
Good knowledge of Probability and Analysis

COURSE OBJECTIVES

The course is aimed at giving the students the skills to use diffusion processes to represent different realities of practical interest. The student should use the different techniques for carrying out the analysis of the models. The student will demonstrate both the ability of self-study of advanced topics, connected to the content of the course, and the ability to collaborate. Students should also use the software Mathematica to perform some assigned simulations.

COURSE AIMS

At the end of the course, students will know several important methods to study stochastic models of applied interest. They will know some of the important classes of stochastic processes and will be able to study their main functional and features.

COURSE DELIVERY

Lessons (48 hours, 6 CFU) are given in lecture rooms.

LEARNING ASSESSMENT METHODS

During the course homeworks are assigned. Solution of these exercises is part of the final exam. Teamwork is allowed for this part of the work. Exam is oral. Students that do not make homeworks will solve exercises immediately before the oral exam.

The evaluation of homeworks is valid only for the Summer exam session. From September session students are required to solve exercises immediately before the oral exam.

SYLLABUS

Brownian Motion: Markov property, existence of the Brownian motion; maximum and first passage time distribution; arcosine law; iterated logarithm law; Reflected Brownian motion; Heat equation and Brownian motion; multidimensional Brownian motion.

Stationary Processes: mean square distance; autoregressive processes; ergodic theory and stationary processes; Gaussian processes

Diffusion Processes: differential equations associated with some functionals of the process; backward and forward equations; stationary measures; boundary classification for regular diffusion processes; conditioned diffusion processes; spectral representation of the transition density for a diffusion; diffusion processes and stochastic differential equations; jump-diffusion processes; first passage time problems for diffusion processes.

An 8 hours module, included in the courseload, will be taught by Visiting Professor Vassili Kolokoltsov on Brownian motion.

SUGGESTED TEXTBOOKS AND READINGS

Schilling, Partzch, "Brownian Motion", De Gruyter

Karlin, Taylor. "A first Course in Stochastic Processes", Academic Press.

Karlin, Taylor. "A second Course in Stochastic Processes", Academic Press.

Mörters, Peres. "Brownian Motion", Cambridge University Press.

Kannan. "An introduction to stochastic processes", North Holland.

Course webpage: http://www.master-sds.unito.it/do/corsi.pl/Show?_id=n6jq

---