Syllabus

- Asymptotic Theory
- Introduction to graphical models and bayesian networks
- Mathematical Optimization for Statistics
- Real Analysis
- Tensorial methods
- Finite Mixture Models - I
- Insights into Probability and Stochastic Processes
- Introduction to Point Processes
- Probability for Data Science
- Clustering of complex data structures
- A gentle introduction to combinatorial stochastic processes
- Advanced Data Analysis in Environmental Biology
- Regression models for categorical and survival data

Real Analysis

Prof. Simone Creo

Tentative programme

1) Abstract integration and Borel measurability. Recalls on topology, $\sigma-$ algebras, measurable spaces, measurable functions, Riemann integration. Conti−nuity and examples of measurable functions. Borel sets, Borel measurable func−tions. Simple functions, positive measures and first properties. Recalls on infinite arithmetics. Lebesgue integral. Monotone and dominated convergence theorem. Fatou Lemma. Remarks on null sets.

2) Lebesgue measure. Topological preliminaries: compact and locally compact spaces, vector spaces, generalities about Banach and Hilbert spaces. Compact support functions. Riesz representation theorem. Lebesgue measure. Continu−ity of Lebesgue measurable functions: Lusin theorem and Vitali−Carath ́eodory theorem. Absolute continuity, Radon−Nikodym theorem.

3) L^p spaces. Definitions and properties. Convex functions. Jensen, H ̈older and Minkoswki inequalities. L^p spaces. L^p norm and completeness of L^p. Approxi−mation with continuous functions. Duality properties. The case $p = 2$: Hilbert properties of L^2, application to Fourier series.

4) Product measures. Positive and signed measures. Measurability on product sets. Fubini Theorem. Complete measures and completion of product measures. Convolutions.

5) Applications. Hints about partial differential equations. Solvability of partial differential equations using Lebesgue spaces.

Mathematical Optimization for Statistics

Prof. Lavinia Amorosi

Mathematical Programming

- Unconstrained optimization: existence and optimality conditions
- Constrained optimization:
    - General case
    - Convex case
    - Linear Programming
    - Quadratic Programming
    - Integer and Mixed Integer Linear Programming

A glimpse on solution algorithms and related computational complexity:

- Linear search methods:
    - Gradient method
    - Newton's method
- Quasi-Newton methods: DFP and BFGS
- Simplex method
- Branch and bound

Applications to statistical problems:
- Linear regression
- Classification trees
- Sparse canonical correlation analysis

References

Serafini P., Ricerca Operativa, Springer-Verlag Italia, Milano, 2009.
Bertsekas D.P., Nonlinear Programming, Athena Scientific, Massachusetts, 1995.
D. Bertsimas, J. N. Tsitsiklis, Introduction to Linear Optimization, Athena Scientific, Belmont, Massachusetts, 1997.
P. Rardin, Optimization in Operations Research, Upper Saddle River, Prentice-Hall, 1998. Conforti M., Cornujols G., Zambelli G., Integer Programming, Springer, 2014.
Bertsimas D., King A., OR Forum - An Algorithmic Approach to Linear Regression, Operations Research, 64(1): 2-16, 2016.
Bertsimas, D., Dunn, J. Optimal classification trees. Mach Learn 106: 1039 - 1082, 2017. Amorosi L., Padellini T., Puerto J., Valverde C., A Mathematical Programming Approach to Sparse Canonical Correlation Analysis, Expert Systems with Applications, 237 Part A, 2024.

Clustering of complex data structures

Prof. Maria Brigida Ferraro

Il corso si rivolge agli studenti del dottorato di Scienze Statistiche.

Lo scopo del corso **è** introdurre i concetti base della classificazione sfocata non supervisionata discutendone l' utilizzo nell' analisi di dati complessi.

Argomenti trattati

Fuzzy k−means and its variants

Extensions of Fuzzy k−Means to categorical/mixed data

Fuzzy k−Means for fuzzy data and its variants

Fuzzy k−Means for functional data

Fuzzy clustering of other complex data

Clustering of Big Data

Fuzzy Double k−Means

Regression models for categorical and survival data

Prof. Marco Alfo'

Il corso si rivolge agli studenti del dottorato di Scienze Statistiche.

Lo scopo principale del corso è di introdurre i concetti base dei lineari generalizzati, additivi generalizzati, regressione per dati di sopravvivenza

Modelli lineari generalizzati

Introduzione Geometria Inferenza Estensioni (QL) Applicazioni

Modelli additivi generalizzati

Introduzione Geometria Inferenza Applicazioni

Modelli di regressione per dati di sopravvivenza

Introduzione
Disegni
Modelli AFT
Modello a rischi proporzionali

Tensorial methods

Prof. Paolo Giordani

Il corso si rivolge agli studenti del dottorato di Scienze Statistiche.

Lo scopo principale del corso è di introdurre i concetti base dello studio di matrice a tre vie (unità x variabilil x occasioni), come diretta estensione dell' analisi delle matrice a due vie (unita x variabili).

Analisi in Componenti Principali, richiami e notazione di base
Dati a tre (e più) vie
L' analisi dei dati a più vie

- Il modello Candecomp/Parafac
- Il modello Tucker3
- La Higher Order SVD

La stima dei parametri
La selezione del modello
Alcuni casi di studio (laboratorio con R)

Finite mixture models

Prof. Roberto Rocci (roberto.rocci@uniroma1.it)

Introduction

Motivations:
- flexible density;
- non parametric estimation of a mixing distribution;
- unsupervised classification.

Maximum likelihood estimation

EM algorithm;
"Fuzzy" interpretation of EM;
ML estimation of a mixture of Gaussians.

Mixture of linear regression models

Omitted variables;
Random effects;
Heterogeneity; EM
algorithm.

Latent class analysis

Latent variables models;
Latent class models for binary variables.

How to choose the number of components?

LR test;
Bootstrap;
Automatic selection criteria.

Principal Stratification for casual inference

Model;
ML estimation;
Examples.

References

Frangakis, C. E., Rubin, D. B. (2002). Principal stratification in causal inference. *Biometrics*, 58 21 - 29

Frühwirth-Schnatter, S. (2006). *Finite Mixture and Markov Switching Models*. Springer, New York.

McLachlan , G.J., Peel. D. (2000). *Finite Mixture Models*. New York: Wiley.

Advanced Data Analysis in Environmental Biology
Giovanna Jona Lasinio

Il corso si rivolge ai dottorandi del dottorato di Scienze Biologiche e del dottorato di Scienze Statistiche.

Lo scopo principale del corso è di creare un ponte tra le problematiche che sorgono nell' ambito delle tesi di dottorato nei diversi ambiti della biologia e le tecniche statistiche che si suppone i dottorandi in scienze statistiche padroneggino in modo più che sufficiente.

Il corso prevede

- la scelta di un numero prefissato di temi, presumibilmente quelli per i quali l' apporto statistico è più rilevante ai fini del raggiungimento del risultato.
- Sulla base dei temi scelti si fissa un numero corrispondente di incontri in cui i dottorandi di biologia presentano le diverse problematiche. Creando una prima occasione di scambio con i dottorandi di statistica.
- Al termine della serie di incontri si organizzeranno dei gruppi di lavoro misti, tra i due dottorati che avranno lo scopo di risolvere adeguatamente i problemi posti.

Nel corso ci si propone di stimolare collaborazioni tra giovani ricercatori di discipline spesso collegate. I partecipanti impareranno a collaborare e ad utilizzare un linguaggio che permetta a persone con background diversi di capire le diverse problematiche.
Ci si aspetta un trasferimento di conoscenze tra le due comunità.
I dottorandi di statistica impareranno a trattare problemi scientifici reali.
I dottorandi di biologia avranno la possibilità di approfondire la conoscenza delle tecniche statistiche più adeguate alla soluzione delle loro problematiche di ricerca.

Insights into Probability and Stochastic Processes

Prof. Costantino Ricciuti

Part 1: Some reminders and insights into probability:

Convergence in distribution. Central limit theorems. Infinitely divisible distributions.

Part 2: An overview of Markov Processes

Discrete-time and continuous-time Markov chains. Levy processes.

Part 3: Brief notes on Semi-Markov processes

Some results and examples for both discrete and continuous time.

PROBABILITY FOR DATA SCIENCE

Prof. Alessandro De Gregorio – Prof. Francesco Iafrate

Brief course description: The course will cover the main topics in high–dimensional probability theory. High–dimensional probability is an area of probability theory that studies random objects in R^n where the dimension n can be very large.

The applications in data science of the introduced theoretical tools will be discussed.
Course topics:

- Basic tail and concentration bounds: from Markov to Chernoff, sub–Gaussian variables and Hoeffding bounds, sub–exponential variables and Bernstein bounds, some one–sided results. Some applications in data science.
- Sparse linear models in high dimensions: problem formulation and applications and penalized estimators, recovery in the noiseless setting, estimation in noisy settings and LASSO estimator, bounds on prediction error, variable or subset selection.

Textbooks:

- R. Vershynin (2018) High–dimensional probability. An introduction with applications in Data Science. Cambridge University Press.
- M. J. Wainwright (2019) High–dimensional statistics: A non–asymptotic viewpoint. Cam–bridge University Press.

INTRODUCTION TO POINT PROCESSES
Prof. Valentina Cammarota

Brief description of the course: this course will give a first introduction to several families of point processes, focussing in particular on the properties of repulsion and rigidity of point processes.

Course topics:

- Poisson point process
- Cox process
- Zeros of random polynomials
- Determinantal point process
- Critical points of Gaussian fields
- Point processes and continuum percolation
- Point processes in Machine Learning

Suggested books:

- Geoffrey Grimmett, *Probability on Graphs, Random Processes on Graphs and Lattices*. Sta−tistical Laboratory University of Cambridge (2018)
- Hough, Krishnapur, Peres, Vir ́ag, *Zeros of Gaussian Analytic Functions and Determinantal Point Processes*. American Mathematical Society (2009)
- Antoine Maillard, G ́arard Ben Arous, Giulio Biroli, *Landscape Complexity for the Empirical Risk of Generalized Linear Models* https://arxiv.org/abs/1912.02143 (2023)

Asymptotic theory
Prof. P.L. Conti

NOTE: Starred points* will be actually covered according to available time

1. Basic asymptotic approximation tools from probability theory (2h)
- Modes of convergence and their basic properties
- Convergence of moments and uniform integrability *
- Representation theorem: elementary version *
- Convergence of transformed/perturbed random variables
- Laws of large numbers
- Central limit theorems
- Approximation error in the central limit theorem *

2. Basic classification of statistical models: parametric, nonparametric, semiparametric

3. Empirical distribution function (EDF) and its properties (2 h)
- Definition and elementary properties
- Glivenko–Cantelli theorem
- Large deviations for empirical distribution functions: the Sanov theorem *
- The empirical process: basic weak convergence results *

4. Asymptotics for the Maximum Likelihood Estimators (MLEs) (6 h)
- Consistency of the global maximizer of the likelihood function: the Wald approach
- Consistency and asymptotic normality of roots of the likelihood equations
- Asymptotic efficiency issues
- Multidimensional extensions

5. Multiple roots of likelihood equation(s) (4 h)
- The problem of multiple roots
- One–step Newton–Raphson method
- Multidimensional extensions

6. M–estimators (6 h) *
- Basic definitions and examples
- Consistency of M–estimators
- Asymptotic normality of M–estimators

Introduction to graphical models and bayesian networks
Introduzione ai modelli grafici e reti bayesiane

Prof. Paola Vicard – Prof. Lorenzo Giammei

Richiami sull'analisi della dipendenza
Odds e odds ratio
Test di indipendenza
Modelli log-lineari
L' indipendenza condizionata - definizione e proprietà
I modelli grafici
Campi applicativi ed esempi di applicazioni reali
I grafi e le loro proprietà
Le proprietà di Markov
Decomponibilità e collassabilità
I grafi diretti aciclici DAG
Proprietà di Markov per DAG
Le reti Bayesiane
Trasmissione dell' informazione nelle reti bayesiane
Esempi di applicazione
L' apprendimento di una rete Bayesiana
Algoritmi constraint-based
Algoritmi score-based
Algoritmi ibridi
Esempi di applicazione
Script R
L' inferenza causale e i Potential Outcomes
Le assunzioni del framework
Le principali strategie di identificazione
Modelli grafici causali
Reti Bayesiane causali
Do-operator
Back-door criterion
Front-door criterion e do-calculus
Relazione con i Potential Outcomes
Esempi di applicazione
Script R

A gentle introduction to combinatorial stochastic processes (with applications to Physics, Finance and Economics)

Prof. Enrico Scalas

The theme of this course is the allocation of n objects (or elements) into g categories classes), discussed from several viewpoints. This approach can be traced back at least the early work of 24−year−old Ludwig Boltzmann in his first attempt to derive distribution of velocity for a perfect gas in probabilistic terms. We shall start from of the world as facts (taking place or not), and events as propositions (true or not) facts (taking place or not). Not everything in the world is known, and what remains is a of possibilities. For this reason, events can be probabilized and probability theory plays fundamental, but often underestimated, role in our scientific theories. Indeed, it turns out that many important problems in statistical physics and some problems in economics and finance can be formulated and solved using these methods.

Syllabus: The following topics will be addressed

Individual and statistical descriptions
The Pólya urn process
The Ehrenfest - Brillouin model
Applications to statistical physics
Applications to stylized models in economics and finance
The Ewens sampling formula
The Zipf - Simon - Yule process

The theoretical material will be complemented by computer−based sessions on Monte simulations of the processes and models introduced in the course.