Proposition de parcours MA2 2024-2025

Contents

1 Probability and statistics **1**

2 Refresher Course **2**

3 Courses **3**

1 Probability and statistics

This program consists of two intertwined series of courses in probability (P) and statistics (S), the aim of which is to provide solid and interdisciplinary teaching in modern theoretical aspects of the mathematics of randomness. Probability and statistics are represented in a balanced way, and a number of topics at the interface between several fields are offered (EDP, mathematical physics, neural networks).

Students will choose 3 courses out of the 4 offered in the first semester (P1, P2, S1, S2) and 4 courses out of the 8 offered in the second semester (P3-P6, S3-S6). It is also possible to take courses in other courses, particularly in PDE.

Note that four of the statistics-oriented courses (one in the first semester and three in the second) are offered by the M2 *Maths in action* program.

Finally, it is strongly recommended to take refresher courses.

2 Refresher Course

url latex

Stochastic tools (Thomas Budzinski, 15h)

1. Discrete time martingales: stopping theorems and convergence. Extensions for continuous time martingales.

2. Construction of Brownian motion. Regularity of trajectories.

3. Some properties of Brownian trajectories. Connection with the heat equation.

3 Courses

3.1 First semester

P1 : Stochastic calculus (Marielle Simon, 24h)

This lecture series will present some of the most important tools allowing to build and study continuous-time stochastic processes, the central example of which is of course Brownian motion. To this end, we will have to introduce and study semimartingales, a rich class of processes for which one can develop a differential and integral calculus, and set and solve certain type of stochastic differential equations.

Just as for the familiar ordinary differential equations (or PDEs), the motivation to study such stochastic differential equations comes from the goal of understanding the global behaviour of random processes by equations describing their infinitesimal behaviour. But since we are dealing with random processes, these equations contain a random “noise”, which informally is an infinitesimal increment of Brownian motion. The main problem of their study comes from the fact that Brownian motion (and therefore the other processes of interest) have too rough trajectories (nowhere differentiable, for instance) for the usual differential and integral calculus to make sense.

In front of this obstacle, we will develop a notion of stochastic integral, due to Itô. It will give rise to a particular integral calculus, in which Itô’s formula acts as an integration by parts (or a fundamental theorem of analysis) of a new kind. This integral calculus will allow us to study the stochastic differential equations for continuous semimartingales, and will shed a new light on these processes, for instance via Lévy’s characterization of Brownian motion, or the Dubins-Schwarz theorem according to which continuous martingales are appropriate time-changes of Brownian motion. Contents:

1. Generalities on continuous-time processes
2. Continuous-time martingales, regularization. Local martingales, semimartingales. Bracket of a continuous semimartingale.
3. Stochastic integral with respect to a continuous semimartingale.
4. Itô’s formula and applications. The theorems of Lévy, Dubins-Schwarz, Girsanov. Burkholder-Davis-Gundy inequalities.
5. Stochastic differential equations. The Lipschitz case.
6. (Time allowing) Continuous-time Markov processes. Generators. Diffusions.

References:
[1] Karatzas-Shreve: Brownian motion and stochastic calculus
[2] Le Gall: Brownian motion and stochastic calculus
[3] Mörters-Peres: Brownian motion
[4] Revuz-Yor: Continuous martingales and Brownian motion
[5] Varadhan: Stochastic processes

P2 : Random walks (Thomas Budzinski, 24h)

The goal of the course will be to offer a panoramic view of techniques used to study random walks on various kinds of objects. The course will be divided into two parts. First, we will be interested in random walks (i.e. sums of i.i.d. variables) on Z or on Z^d (recurrence and transience, cyclic lemma, renewal theory. We will then focus on the simple random walk on general graphs. In particular, we will study the links between random walks and electric networks, and the interplay between geometric and probabilistic properties of a graph.
References:

[1] Lawler-Limit: Random walk, a modern introduction
[2] Lyons-Peres: Probability on trees and networks

S1 : Concentration of measure in probability and high-dimensional statistical learning (Guillaume Aubrun, Aurélien Garivier, Rémi Gribonval, 24h)

This course will introduce the notion of concentration of measure and highlight its applications, notably in high dimensional data processing and machine learning. The course will start from deviations inequalities for averages of independent variables, and illustrate their interest for the analysis of random graphs and random projections for dimension reduction. It will then be shown how other high-dimensional random functions concentrate, and what guarantees this concentration yields for randomized algorithms and machine learning procedures to learn from large training collections. This course will be based on a sample of the classical textbook “Concentration Inequalities” by Boucheron, Massart, Lugosi, and on “High-Dimensional Probability” by Roman Vershynin. Applications to machine learning will rely on “Understanding Machine Learning”, by Shalev-Schwartz and Ben-David.

S2 : Stochastic modeling and statistical learning (Aurélien Garivier and Clément Marteau, 24h)

[Also in the M2 *Math en Action* program]

- High-dimensional regression: Concrete examples and modeling, Reminders and development around the linear model (modeling, hypotheses, least squares and likelihood, Fisher/Student test, etc.)
  - Introduction to model selection (construction of Cp/AIC/BIC criteria, oracle inequalities, high-dimensional behavior)
  - Ridge method (heuristic, link with Tikhonov, property of risk)
  - Introduction to the LASSO method (construction and heuristics, link with compressed sensing, theoretical properties / oracle inequalities, compatibility conditions)

- Supervised Classification: Concrete examples and modeling, overview of some algorithms (kNN, SVM, neural networks, logistic regression, etc.), theoretical aspects (concentration inequalities, kernels, etc.)
- Unsupervised classification: PCA, Clustering (kmeans, hierarchical methods, etc.), Gaussian mixture models, Spectral clustering.
- Other risks, extremes, introduction to research problematics.

3.2 Second semester

P3 : Scaling limits of interacting particle systems (Oriane Blondel, Christophe Poquet, 18h)

This course aims to present how probabilistic particle models can be described by PDEs in the large population limit. The systems studied come from modeling of physical or biological phenomena. The course will use notions of stochastic calculation, Markov chains in continuous time and will introduce some notions of PDE.

This course will be offered in two parts, the first on medium-field type interactions and the other on network systems with short-range interactions. In the first part, we will present the classic results of convergence of mean-field diffusion models, the notion of chaos propagation as well as more recent results for dense but not complete interaction graphs. In the second part, we will focus on exclusion processes, which are Markov processes on $\{0, 1\}^N$. We will study their hydrodynamic limit and describe their fluctuations.

References:

1. L. Bertini, G. Giacomin, Stochastic Burgers and KPZ equations from particle systems. Comm. Math. Phys. 183(3): 571-607 (1997).

2. F. Coppini, H. Dietert and G. Giacomin, A law of large numbers and large deviations for interacting diffusions on Erdős–Rényi graphs. Stochastics and Dynamics 20 (2020).

3. S. Delattre, G. Giacomin and E. Luçon, A Note on Dynamical Models on Random Graphs and Fokker–Planck Equations. Journal of Statistical Physics 165 (2016).

4. J. Gärtner, On the McKean–Vlasov limit for interacting diffusions. Mathematische Nachrichten 137 (1988).

5. P. Gonçalves, Hydrodynamics for symmetric exclusion in contact with reservoirs. Stochastic Dynamics Out of Equilibrium, Institut Henri Poincaré, Paris, France, 2017, Springer Proceedings in Mathematics and Statistics book series, 137-205 (2019).

6. P. Gonçalves, M. Jara, M. Simon. Second order Boltzmann-Gibbs Principle for polynomial functions and applications, Journal of Statistical Physics, Volume 166, Issue 1, 90?113 (2017).

7. C. Kipnis and C. Landim, Scaling limits of interacting particle systems. Grundlehren der Mathematischen Wissenschaften. 320. Berlin: Springer. xvi, 442 p. (1999).

8. A.-S. Sznitman, Topics in propagation of chaos. Ecole d'été de probabilités de Saint-Flour XIX–1989. Springer (1991).

P4 : Spectral Theory of random operators and graphs (Raphaël Ducatez and Christophe Sabot, 18h)

The aim of the course is to give an introduction to several models of random operators and random graphs, focusing on the understanding of their spectrum and on the behavior of their eigenfunctions. These models are motivated by fundamental questions in physics or applications in other domains.

The course will start by a short overview of the spectral theory of self-adjoint operators. The first main part of the course will concentrate on the Anderson model that is the basic model to describe the quantum propagation of electrons in a disordered media. Mathematically, it can be expressed in terms of the asymptotic behavior (localisation or delocalization) of typical eigenfunctions of the Laplacian perturbed by a random potential. The Anderson model was introduced in the 50's, but fundamental questions still remain open on this problem.

In a second part, the course will be illustrated by two other related models for which some these questions have been addressed. The first one concerns the spectral theory of Erdös-Rényi graphs where both localization and delocalization regimes can be proved to exist. The second one concerns an apparently different problem, the edge reinforced random walk, where a related localization and delocalization phenomena appears. Some illustrative references:

1. P.W. Anderson, "Absence of Diffusion in Certain Random Lattices". Phys. Rev. 109 (5): 1492 – 1505.
2. Michael Aizenman and Simone Warzel, “Random Operators: Disorder Effects on Quantum Spectra and Dynamics”, Graduate Studies in Mathematics Volume: 168; 2015
3. Werner Kirsch, “An invitation to random Schrödinger operators”. Panorama et synthèse 25, (2008)
4. Béla Bollobás “Random graphs”, Cambridge University Press, 2nd edition, (2011)
5. Johannes Alt, Raphaël Ducatez, Antti Knowles “Extremal eigenvalues of critical Erdős-Rényi graphs”, Ann. Probab. 49(3): 1347 – 1401 (May 2021)
6. M. Disertori, T. Spencer, M.R. Zirnbauer, “Quasi-Diffusion in a 3D Supersymmetric Hyperbolic Sigma Model”, Commun. Math. Phys. 300, 435 – 486 (2010)
7. C. Sabot, X. Zeng, “A random Schrödinger operator associated with the Vertex Reinforced Jump Process on infinite graphs”, J. Amer. Math. Soc. 32 (2019), 311 - 349

P5 : Phase transitions in spin systems (Christophe Garban, 18h)

The goal of this course will be to give an introduction to spin systems defined on a d-dimensional lattice $\mathbb{Z}^d$. We will focus in particular on the intriguing phenomenon of *symmetry breaking*. When a symmetry is broken in a spin system, this usually happens only at low enough *temperature*. We can then identify a *phase transition* at some critical temperature T_c.

This course will provide the mathematical tools to apprehend such phase transitions. It will be structured as follows:

Part I. Discrete symmetry spin systems.

The most celebrated example of such systems is the **Ising model** which assigns spins $\sigma_x \in \{\pm 1\}$ to each site $x \in \mathbb{Z}^d$. (See Figure below). This part I. will focus mainly on the Ising and Potts models whose discrete underlying symmetry is *broken* at low temperature.

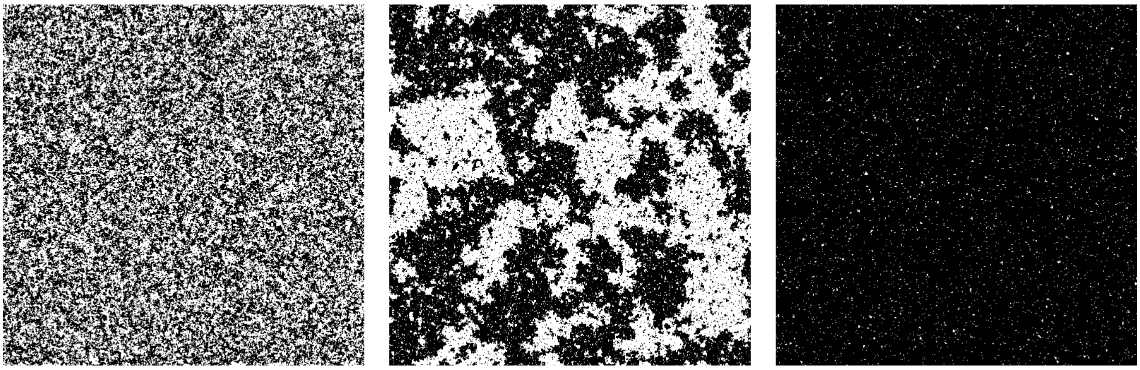

Figure 1: The Ising model in its three phases: $T > T_c, T = T_c$ and $T < T_c$.

Part II. Continuous symmetry spin systems.

Such spin systems include the following examples:

- The Gaussian Free Field whose 'spins' are $\mathbb{R}$ valued
- The XY model whose spins are $\mathbb{S}^1$ valued (i.e in the unit circle).
- The classical Heisenberg model with values in $\mathbb{S}^2$
- Lattice gauge theory with continuous gauge group G.

Some of the classical techniques which are very powerful when dealing with a discrete symmetry spin system (for example the so-called Peierls argument for the Ising model) do not apply for continuous symmetry spin systems. This Part II. will explain how such continuous symmetries affect the fluctuations in the system and will introduce some of the main relevant techniques, among which:

- Mermin-Wagner theorem (on the absence of symmetry breaking in $2d$)
- Reflection positivity
- techniques from Bayesian statistics
- Berezinskii-Kosterlitz-Thouless topological phase transition (Nobel prize in Physics 2016).

References

[1] Yvan Velenik. *Le modèle d'Ising.* `https://www.unige.ch/math/folks/velenik/Cours/2008-2009/Ising/Ising.pdf`

[2] Hugo Duminil-Copin. *Lectures on the Ising and Potts models on the hypercubic lattice.* `https://arxiv.org/abs/1707.00520`

[3] Sacha Friedli and Yvan Velenik. *Statistical Mechanics of Lattice Systems: a Concrete Mathematical Introduction.* `https://www.unige.ch/math/folks/velenik/smbook/index.html`

P6 : Scaling limits for stochastic processes: Application to Biology (Hélène Leman and Céline Bonnet, 18h)

The aim of the course is to present Markovian models for the study of biological phenomena. Following biologically motivated assumptions, we will present convergences of these stochastic processes.

We will start by constructing càd-làg (*continus à droite avec des limites à gauche*) processes, based on Poisson Point Processes and their associated martingales. These processes will model the dynamics of some key characteristics of biological populations. We will be particularly interested in structured populations (age, position, phenotypic trait...).

Then, on the one hand we will present some convergences of sequences of processes under different assumptions, starting with a limit of large population. More precisely, we will consider stochastic processes versions of the law of large numbers and the central limit theorem. On the other hand, we will study the long-time behavior of these processes. Some illustrative references:

1. N. Ikeda and S. Watanabe, «Stochastic differential equations and diffusion processes», 2nd ed. North-Holland, 1989.

2. V. Bansaye and S. Méléard, «Stochastic models for structured populations», Berlin: Springer, Vol. 16., 2015.

3. P. Billingsley, «Convergence of probability measures», John Wiley & Sons, 2013.

4. S.N. Ethier and T.G. Kurtz, «Markov processes: characterization and convergence», John Wiley & Sons, 2009.

5. T. Britton, E. Pardoux, F. Ball, C. Laredo, D. Sirl, V. Tran. «Stochastic epidemic models with inference» (Vol. 2255). Berlin: Springer, 2019.

S3 : Graphs and ecological networks (Clément Marteau and Thibault Espinasse, 18h)

[Also in the M2 *Math en Action* program]

A graph, whose first use are mentionned in the 16th century, is a mathematical object widely used from the first appearance of network investigations, namely investigation of relationship between individual in wide sense. Ranging from social network to the internet, graphs are leading objects for the analysis of several data sets. Ecosystem relationships, from species relationship (prédation, interaction between plants and pollinating insects , etc...) social relationship between individuals (sociality between primates, etc...), offers several different possible applications of graphs modelling and network investigation.

In this course, we will investigate the framework of graph theory and network science. We will provide an introduction to modern research problems regarding ecosystems studies. We will use alternatively discrete mathematics, statistics and machine learning.

We will adress both theoretical and practical (case studies in ecology) questions.

Theretical keywords: Bases / definitions (graphs, path, etc...) ; Metrics ; Clustering methods ; Spectral methods ; Random graphs models ; Graphical models (graphs inference) ; Signal processing on graphs ; Multi-level graphs (time, space, link types) ; Embedding methods (optional)

Case studies : Contact network between animals. Interaction network between species in a marine and/or alpine environment. Consideration about the relevance of a graph for biodiversity support.

S4: Neural networks (Aurélien Garivier, Rémi Gribonval, Julian Tachella, 18h)

[Also in the M2 *Math en Action* program]

The goal of this course is twofold:

- To present the principles of modern deep neural networks, as well as the technical ways to implement them for solving classification and regression problems.
- To provide a detailed overview of the mathematical foundations of modern learning techniques based on deep neural networks.

Starting with the universal approximation property of neural networks, we will then see why depth improves the capacity of networks to provide accurate function approximations for a given computational budget. Tools to address the optimization problems appearing when training networks on large collections will then be covered, and their convergence properties will be reviewed. Finally, statistical results on the generalization guarantees of deep neural networks will be presented, both in the classical underfitting scenario and in the overfitting scenario leading to the so-called "double descent" phenomenon.

S5: Optimal transport and learning (Filippo Santambrigio, Ivan Gentil, Yohann de Castro, Ievgen Reedko, Julie Digne, Nicolas Bonnel, 18h)

[Also in the M2 *Math en Action* program]

The aim of the course is to present the broad outlines of optimal transport theory and some of its applications in data sciences.

A first part of the course will detail the Monge-Kantorovich problem, its formulation as a linear programming problem and the use of convex duality, as well as the distances (called Wasserstein distances) that optimal transport makes it possible to define in space. probability measures. Geodesics and barycenters in Wasserstein space, of great importance in the interpolation and comparison of data, will also be introduced.

A second part of the course will focus on numerical methods for solving optimal transport problems, with particular attention to the methods best suited to high dimension and unstructured data, in particular the Sinkhorn algorithm.

Finally, the third part of the course will present a choice of applications of transport and Wasserstein distances in learning, of which we cite as examples Wasserstein GANs, transfer learning, data generation models, etc.

S6: Inverse problems and parsimony (Yohann de Castro and Rémi Gribonval, 18h)

[Also in the M2 *Math en Action* program]

Sparsity and convexity are ubiquitous notions in Machine Learning and Statistics. In this course, we study the mathematical foundations of some powerful methods based on convex relaxation: L1-regularisation techniques in Statistics and Signal Processing; Nuclear Norm minimization in Matrix Completion; K-means and Graph Clustering.

These approaches turn out to be Semi-Definite representable (SDP) and hence tractable in practice. The theoretical part of the course will focus on the performance guarantees for theses approaches and for the corresponding algorithms under the sparsity assumption. The practical part of this course will present the standard SDP solvers for these learning problems.
Keywords: L1-regularisation; Matrix Completion; K-Means; Graph Clustering; Semi-Definite Programming.