El Mehdi Issouani

PhD in Applied Mathematics and Mathematical Applications

Postdoctoral Researcher at the Laboratory of Applied Mathematics of Compiègne (LMAC), University of Technology of Compiègne (UTC)

Chatou, France
https://issouani.perso.math.cnrs.fr
emissouani@gmail.com

Research Areas

- Statistical learning, deep learning, predictive language models
- Generalized empirical likelihood, φ-divergences, self-normalized sums (Hotelling's T^2 statistics)
- Random matrix spectrum and covariance matrices, shrinkage and regularization in high dimensions
- Applications:
  - Linguistics: NLP, generative models (MaxEnt, Markov chains, embeddings)
  - Molecular biology: probabilistic and statistical models applied to biological systems for the *in silico* design of DNA and RNA sequences

Academic Background

Postdoctoral Researcher - University of Technology of Compiègne **LMAC** 2023 - Present

Supervisors: Ghislaine Gayraud and Miraine Dávila Felipe (LMAC, University of Technology of Compiègne)
Funding: as part of the Num4Lyme project, *AAP équipes projets* ISCD, Sorbonne University

- Development of mathematical models to detect the Lyme disease pathogen
- Application of genetic sequence representation techniques, RBM, MCMC, and Gibbs sampling
- Tools: Python (RBM, Bio (SeqIO), nltk, scipy, numba, sklearn, re, matplotlib, itertools)

Keywords: Statistical learning, bioinformatics, molecular data (DNA, RNA), probabilistic models

PhD - University Paris Nanterre **MODAL'X** 2017 – 2023

Title: *Models and Algorithms for Automatic Text Simplification* (link[1])

Supervisor: Patrice Bertail (MODAL'X, University Paris Nanterre)
Co-supervisors: Mélanie Zetlaoui and Thierry Dumont (MODAL'X, University Paris Nanterre)

Funding: Doctoral contract (MESR), Doctoral School ED 139 – University Paris Nanterre

Keywords: High-dimensional statistics, deep learning, applications (NLP), empirical likelihood

Thesis Committee: Antoine Chambaz (Full Professor, MAP5 – University Paris Cité), Amor Keziou (Associate Professor (HDR), LMR – University of Reims Champagne-Ardenne), Estelle Kuhn (Senior Researcher (HDR), INRAE Jouy-en-Josas), Delphine Battistelli (Full Professor, MODYCO – University Paris Nanterre), Jean-François Pradat Peyre (Full Professor, LIP6 – University Paris Nanterre), Marianne Clausel (Full Professor, IECL – University of Lorraine)

Defense Date: 23/06/2023

Publications

Peer-reviewed publications

E. Issouani, P. Bertail, E. Gautherat. *Exponential bounds for regularized Hotelling's T^2 statistic in high dimension.* **Journal of Multivariate Analysis (JMVA)**,[2] **2024.** DOI: `https://doi.org/10.1016/j.jmva.2024.105342`

Working papers and preprints

E. Issouani, P. Bertail, T. Dumont, M. Zetlaoui. *Classifiers based on Penalized Generalized Empirical Likelihood: Application to POS-Tagging.* Submitted to **Econometrics and Statistics**,[3] **2026.**

E. Issouani, M. D. Felipe, G. Gayraud, M. Guerin, H. Da Ponte, S. Padiolleau-Lefèvre and I. Maffucci. *Design of DNA Aptamers for Lyme disease Diagnosis Combining experimental and numerical approaches.* Preprint[4], **2025.** DOI: https://doi.org/10.64898/2026.05.13.724892.

E. Issouani, P. Bertail, N. Lysenko. *Some exponential inequalities for functional data: application to tests in Hilbert space.* Manuscript in preparation.

[1] **Thesis:** `https://bdr.parisnanterre.fr/theses/internet/2023/2023PA100049/2023PA100049.pdf`
[2] **Supplementary Material:** `https://ars.els-cdn.com/content/image/1-s2.0-S0047259X24000496-mmc1.pdf`
[3] `https://hal.science/hal-05626099/document`
[4] `https://doi.org/10.64898/2026.05.13.724892`

Teaching Experience

Lecturer (ATER) - University Paris Nanterre 2020 – 2022

192 teaching hours/year.

Courses taught: Statistical Inference (Levels 1 & 2), Descriptive Statistics, Applied Statistics for Management, Statistical Tests, Calculus, Semantic Analysis of Textual Data (Master's level).

⋆ Course design and instruction for the Master's course *Semantic Analysis of Textual Data*.

Additional Teaching Assignments - University Paris Nanterre 2017 – 2023

Teaching in: Statistics, Calculus, Data Science, Statistical Inference.

A detailed overview of teaching activities is available **online**[5].

Research Activity

Invited Mini-Course February 2025

⋆ **Course on Predictive Language Models (total: 16 h)** Delivered as a series of mini-lectures at the Institute of Mathematical Stochastics, Universität Braunschweig (Braunschweig University, Germany, February 2025)
Title: *The Development of Predictive Language Models*
Audience: Mathematicians from Otto-Friedrich-Universität Bamberg, Fakultät Statistik - TU Dortmund, Leibniz Universität Hannover, Universität Leipzig et Universität Heidelberg, TU Braunschweig

Invitations 2023 – Present

- ProbaStatBio26 (35min) Presentation of postdoctoral research (University Paris Nanterre, May 2026).
- Parisian Statistics Seminar (1h) Presentation of postdoctoral research (Session of March 16 at IHP, Paris 2026).
- SFaMS Conf. 2025 (45 min) Student-Faculty Mathematical and Statistical Conference. *Design of DNA Aptamers for Lyme disease Diagnosis Combining experimental and numerical approaches*. (Caraga State University, Philippines, May 2025)
- Présentation (30 min) *Aptamer embedding: RBM-Based Continuous and k-mer Discrete Representations*. Laboratory Team Meeting of the Enzyme and Cell Engineering Laboratory (GEC), (University of Technology of Compiègne, February 2025)
- Journee IA (20 min) An Interdisciplinary Approach to Artificial Intelligence: Law, Economics, Management, Computer Science, Mathematics. *Mathematical Models for Textual Data Analysis and Language Modeling*. (University Paris Nanterre, November 2024)
- Elements Student Scientific Conference (40 min) AGH University of Science and Technology of Krakow. *Aptamer Embedding: using a Continuous vector representation for aptamer design, analysis and deeper understanding*. (AGH University of Science and Technology, Kraków, Poland, October 2024)
- Journées MAS 2024 (30 min) *Continuous Vector Representation for the Design and Analysis of Aptamers*. (University of Poitiers, August 2024)
- GT EMS : Groupe de Travail Entropie, Mots, Stat (40 min) Maximum d'entropie appliquée au NLP (Université de Caen Normandie, mai 2024)
- LMAC Seminar (1 h) Exponential bounds for regularized Hotelling's T^2 in high dimension (University of Technology of Compiègne, May 2023)

[5] `https://issouani.perso.math.cnrs.fr/src/files/Services_issouani.ods`.

Conferences and Symposia 2017 - Present

- ISNPS2024 (20 min) International Symposium on Nonparametric Statistics (Braga, Portugal, juin 2024)
  Exponential bounds for penalized Hotelling's T^2 statistics
- CMStatistics2023 (20 min) Computational and Methodological Statistics (Berlin, Allemagne, déc. 2023)
  Optimal penalty selection for high-dimensional covariance matrices with an application in NLP
- StatMathAppli2023 Mathematical Statistics and Applications (Fréjus, France, septembre 2023)
- ECAS - SFdS 2021 (10 min) Text Mining (course) (Fréjus, France, octobre 2021)
  Automatic Text Simplification
- EMS2019 (20 min) European Meeting of Statisticians (Palerme, Italie, juillet 2019)
  Models and Algorithms for Automatic Text Simplification
- StatMathAppli2019 (20 min) Mathematical Statistics and Applications (Fréjus, France, septembre 2019)
  Automatic Text Simplification
- ISNPS2018 (20 min) International Society for Nonparametric Statistics (Salerne, Italie, juin 2018)
  Model and Algorithm for Text Simplification

Working Groups and Collaborations 2017 - Present

- WG - "Num4Lyme" between mathematicians from LMAC and members of the GEC Laboratory (Enzyme and Cell Engineering) in biology and bioinformatics (Compiègne, 2023 – present)
- LMAC Doctoral & Postdoctoral Seminar (1 h) *Aptamer design & analysis with continuous and sparse embeddings*. Talk at the monthly seminar for non-permanent researchers at LMAC (Compiègne, December 2024)
- WG - Project under development with the Paris SAMU, University Paris Nanterre, University of Technology of Compiègne, and University Paris Descartes (Paris SAMU, 2023)
- MODAL'X – MODYCO Meeting (1 h) Meeting and working group with mathematicians and linguists from University Paris Nanterre. *Complexity Measure*. (Nanterre, September 2022)
- DAP2018 Seminar (> 2 h) *Maximum Entropy for POS-Tagging* - Presentation at the PhD Students, A.T.E.R, and Postdocs Seminar, MODAL'X (Nanterre, December 2018)
- WG "Entropie, mots, statistique" (1 h) *Maximum Entropy for Translation or Text Simplification* – Organized by Valérie Girardin (Nanterre, November 2017)

Organization

Member of the Local Organizing Committee - University Paris Nanterre June 5–7, 2023

Conference: *Non-stationarity, Cyclostationarity and Applications* (**link**[6])

Main Organizer - University of Technology of Compiègne 2024 – 2025

Monthly seminar for PhD students and postdoctoral researchers at LMAC

emissouani@gmail.com · **https://issouani.perso.math.cnrs.fr** · github.com/El-Mehdi-42

[6] `https://nonstationarity.sciencesconf.org/`

Professional Experience

Project Manager - **Dalink Conseil** Feb 2017 - June 2017

Digital Transformation Project Management: Integration of innovative products
Market Analysis: Assessment of economic impact and feasibility. Meeting reports with investors, suppliers, and clients. Governance: Project segmentation based on workload; task allocation. Monitoring & Reporting: Budget tracking and performance dashboards. **Tools:** *Excel, VBA, R, InDesign*

Actuarial Analyst - **La Mutuelle Générale** Apr 2016 - Sep 2016

Technical Department – Individual and Group Insurance Division

Cost Analysis: Study of average healthcare cost deviations on a key account. Data Handling: Database construction, processing, and management. Statistics: Data analysis, exploration, and data mining. Probabilistic modeling of claims. Automation: Reporting, performance measurement, and process improvement. **Tools:** *BO, SQL, SAS Guide, Excel, VBA*

Education

Master 1 & 2 ISÉFAR - University Paris Nanterre [With Honors] 2014 – 2016

Bachelor's in Applied Mathematics - University Paris Nanterre [With Merit] 2012 – 2013

Preparatory Class – MP* - Lycée George Clémenceau (Reims) [Admitted] 2011 – 2012

Preparatory Class – MPSI - Lycée Franklin Roosevelt (Reims) [With Highest Honors] 2010 – 2011

Languages & Technical Skills

Languages - (written and spoken): French, English, Arabic (Fluent). Spanish (Beginner)

Python & R - Packages: NUMPY, NLTK, SKLEARN, PANDAS, TORCH, GENSIM, TENSORFLOW

Machine learning - Extraction, Analysis, Modeling, Parallel Programming

- Extracting, cleaning, and transforming data into vectors or other numerical formats
- Data analysis, dimensionality reduction, descriptive summarization
- Modeling: supervised and unsupervised learning, Deep learning (RNN, CNN, LSTM, GAN), RBM

NLP - Parsing, Text Classification, Embedding, Word2Vec, Language Modeling

Tools - Linux, Office, VBA, SAS, Maple, SQL, Oracle, BO, XML, LaTeX, HTML, CSS

Molecular Biology - SELEX, Aptamer Embedding & Generation, Structural Modeling

Interests

Travel - Germany, Spain, Portugal, Belgium, Italy, Poland, Morocco

Music - 10 years in conservatory, 7 years of piano, 20 years of guitar, 1.5 year of violin

Other - Club football (6 years), Running (half marathon), Theatre (1-year training)

References

Research Contact: Patrice Bertail (MODAL'X, University Paris Nanterre, France), Jens-Peter Kreiß (University of Technology of Braunschweig, Germany), Ghislaine Gayraud (LMAC, University of Technology of Compiègne, France). Jean-François Pradat Peyre (LIP6, University Paris Nanterre, France), Emmanuelle Gautherat (REGARDS, University of Reims Champagne Ardenne, France)

Teaching Contact: Olivier Couronné (MODAL'X, University Paris Nanterre, France), Marc Baudry (EconomiX, University Paris Nanterre, France)

emissouani@gmail.com · **https://issouani.perso.math.cnrs.fr** · github.com/El-Mehdi-42