

Magali Champion

Assistant professor in Statistics

✉ magali.champion@u-paris.fr

📄 <https://magalichampion.github.io/>

🐙 <https://github.com/magalichampion>

Born August 11, 1988 (France)

Experience

- Since January 2022 **Guest researcher**, Seminar for Statistics, ETH Zürich, Switzerland.
- Since September 2018 **Assistant professor in statistics**, Laboratoire MAP5, IUT de Paris - Rives de Seine, Université Paris Cité, France.
- 2016-2018 **Postdoctoral researcher**, Laboratoire MAP5, Université Paris Descartes, France.
- Identification of deregulated transcription factors in specific subtypes of bladder cancer.
 - Project LIONS (*Large-scale Integrative approach to unravel the complex relationships between differentiatiON and tumorigenesiS*) funded by the french national institut of health and medical research (INSERM).
- 2015-2016 **Postdoctoral researcher**, Stanford Center of Biomedical Informatics Research, Stanford University, USA.
- Development of statistical algorithms for the identification of co-expressed gene modules and multi-omics data integration.
 - Project funded by the US National Institute of Health (NIH).
- 2014-2015 **Teaching and research assistant**, INSA de Toulouse, France .
- 2011-2014 **PhD student with teaching responsabilites**, Institut de Mathématiques de Toulouse, Université Toulouse III and Unité Mathématiques et Informatique Appliquées, INRA de Toulouse, France.
- Contributions to gene regulatory networks modelling and inference.
 - Advisors: Sébastien Gadat, Christine Cierco-Ayrolles and Matthieu Vignes.

Education

- 2014 **PhD in Applied Mathematics**, Université Toulouse III, France.
- Defense: December 5, 2014
 - Examinators: N. Vayatis, C. Matias, E. Birmelé, F. Gamboa, C. Cierco-Ayrolles, S. Gadat et M. Vignes (PhD advisors)
- 2011 **Master Research in Applied Mathematics**, Université Toulouse III, France.
- 2009 **Bachelor in Pure Mathematics**, Université Toulouse III, France .
- 2006 **Baccalauréat (High School diploma)**, Lycée Théophile Gautier, Tarbes, France.

Research activities

Research themes

- **Statistical learning:** network inference, graphical models, clustering, high dimension, penalized multi-linear regressions, sparsity.
- **Computational biology:** cancer, multi-omics data integration, gene regulatory networks.

International journals publications

- [1] M. Champion, J. Chiquet, P. Neuvial, M. Elati, F. Radvanyi & E. Birmelé. Identification of deregulation mechanisms specific to bladder cancer subtypes. *Journal of Bioinformatics and Computational Biology*, **19**(01):2140003, 2021.
- [2] M. Champion, K. Brennan, A. Gentles, T. Croonenborghs, N. Pochet & O. Gevaert. Module analysis captures pancancer genetically and epigenetically deregulated cancer driver genes for smoking and antiviral response. *EBioMedicine*, **27**:156-166, 2018.
- [3] M. Champion, V. Picheny & M. Vignes. Inferring large graphs using ℓ_1 -penalized likelihood. *Statistics and Computing*, **28**(4):905-921, 2017.
- [4] R. Dubey, A.M. Lebensohn, Z. Bahrami-Nejad, C. Marceau, M. Champion, O. Gevaert, B.I. Sikic, J.E. Carette & R. Rohatgi. Chromatin-remodeling complex SWI/SNF controls multidrug resistance by transcriptionally regulating the drug efflux pump ABCB1. *Cancer Research* **76**(19):5810-5821, 2016.
- [5] M. Champion, G. Chastaing, S. Gadat & C. Prieur. \mathbb{L}_2 -Boosting for sensitivity analysis with dependent inputs. *Statistica Sinica* **25**:1477-1502, 2015.
- [6] M. Champion, C. Cierco-Ayrolles, S. Gadat & M. Vignes. Sparse regression and support recovery with \mathbb{L}_2 -Boosting algorithms. *Journal of Statistical Planning and Inference*, **155**(C):18-40, 2014.

Proceedings conference publications

- [7] M. Champion, J. Chiquet, P. Neuvial, M. Elati, F. Radvanyi & E. Birmelé. Identification of deregulated transcription factors in specific bladder cancer subtypes. *Proceedings of the 12th International Conference on Bioinformatics and Computational Biology*, **70**:1-10, 2020.

Preprints

- [8] C. Champion, M. Champion, M. Blazère, R. Burcelin & J.M. Loubes. ℓ_1 -spectral clustering algorithm: a spectral clustering method using ℓ_1 -regularization .
- [9] M. Pacault, C. Verebi, M. Champion, L. Orhant, A. Perrier, C. Férec, T. Bienvenu, R. Daveau & J. Nectoux. Non-Invasive Prenatal Diagnosis of Single Gene Disorders by Relative Haplotype Dosage Analysis: Adapted workflow for diagnosis implementation.

R-packages

- [9] ℓ_1 -spectral clustering algorithm: a spectral clustering method using ℓ_1 -regularization (M. Champion, C. Champion, M. Blazère & J.M. Loubes). Available at the CRAN platform.
- [10] LIONS: identification of deregulated transcription factors involved in specific cancer subtypes (M. Champion, J. Chiquet, P. Neuvial, M. Elati & E. Birmelé). Available on GitHub.
- [11] AMARETTO: a multi-omics data fusion for cancer data (M. Champion, C. Planey & O. Gevaert). Available at the Bitbucket platform.
- [12] GADAG: a hybrid Genetic Algorithm for learning DAGs (M. Champion, V. Picheny & M. Vignes). Available at the CRAN platform.

Scientific talks in international conferences

- 2022 “Non-invasive prenatal diagnosis of monogenic diseases by enhanced relative haplotype dosage analysis” in the European Human Genetics Conference (ESHG), Vienne, Austria. Presented by J. Nectoux.

- 2020 “*Identification of deregulated transcription factors in specific bladder cancer subtypes*” in the 12th International Conference on Bioinformatics and Computational Biology (InCoB), San Francisco, United States. *Canceled due to Coronavirus.*
- 2018 “*Module Analysis Captures Pancancer Genetically and Epigenetically Deregulated Cancer Driver Genes for Smoking and Antiviral Response*” in Intelligent Systems for Molecular Biology (ISMB), Boston, United States. *Presented by O. Gevaert.*
- 2017 “*Identification of deregulated transcription factors in specific subtypes of cancer*” (poster) in the 16th International Conference on Bioinformatics (InCoB), Shenzhen, China.
- 2016 “*Pancancer module analysis captures major oncogenic pathways and identifies master regulator of immune response*” in the Keystone symposia on Molecular and Cellular Biology: The Cancer Genome, Banff, Canada.
- 2015 “*Multi-omics data fusion for cancer data*” in the 14th Annual International Conference on Critical Assessment of Massive Data Analysis (CAMDA), Dublin, Ireland. *Presented by O. Gevaert.*
- 2014 “ *\mathbb{L}_2 -Boosting on Generalized Hoeffding Decomposition for Dependent Variables*” in the SIAM Conference on Uncertainty Quantification, Savannah, United States.
- 2011 “*An \mathbb{L}_2 -Boosting algorithm for sparse multivariate regression: application to gene network recovery*” (poster) in the NIPS Workshop Machine Learning for Computational Biology, Sierra Nevada, Spain.

Teaching activities

Teaching

Data Mining, professional public from Université de Paris (14h course) *Classification, clustering and regression technics.*

Machine Learning in high dimension, Master students from Université de Paris (20h course) *Multiple tests, dimensionality reduction (PCA, PLS), penalized linear regressions (Ridge, Lasso, Elastic Net).*

Parametric Tests, Undergraduate students from IUT de Paris (12h course, 18h exercices) *Introduction to parametric tests (notion of risks, pvalues, power of a test), one-sample test (t-test, proportion, variance).*

Linear model, Undergraduate students from IUT de Paris (28h course + exercices) *Gaussian linear models (simple and multiple), selection of variables methods.*

Statistics with R, Undergraduate students from IUT de Paris (36h exercices) *Introduction to the R software.*

Statistical survey, Undergraduate students from IUT de Paris (20h) *Project in which students analyse the results of surveys using statistical data analysis tools.*

Introduction to big data with R, professional public from Université de Paris (7h exercices) *Introduction to the R software.*

Parametric statistics, Undergraduate students from ENAC Toulouse (10h course) *Parametric statistics: statistical models, identifiability, estimation, statistical sufficiency, confidence intervals and tests.*

Random simulation technics, Master students from Université Toulouse III (30h exercices) *Stochastic, probabilistic and statistic methods and algorithms: Markov models, stochastic algorithms, complexity and high dimension (Scilab).*

Statistical software SAS, Undergraduate students from Université Toulouse III (30h exercices)

Introduction to the SAS software.

Statistics, Master students from Université Toulouse III (10h course)

Linear models, maximum likelihood estimation, non-parametric tests.

Analysis, Undergraduate students from Université Toulouse III (42,5h exercices)

Derivability, limits, integrals, real numeric and recurrent sequences.

Mathematics, Undergraduate students from Université Toulouse III (24h course + exercices)

Functions, derivability, limits, continuity, integrals, differential equations, introduction to probabilities and statistics.

Analysis, Undergraduate students from Université Toulouse III (60h course + exercices)

Complex numbers, polynomials, functions, derivability, limits, continuity, integrals, real numeric and recurrent sequences, differential equations.

Probabilities and statistics, Undergraduate students from IUT de Toulouse (28h exercices)

Descriptive statistics, combinatory, discrete probabilities, gaussian law, confidence intervals and statistical tests.

Supervised students

- High-school student 2015 Nabeel Mamoon, Stanford University (2 months), *Analysis of statistical signatures in methylation-guided automated carcinoma diagnosis*. Winner of the *Stanford Institutes of Medicine Summer Research* program.
- Undergraduate 2020 Marina Atangana and Stellan Wea, IUT de Paris (2 months), *Analyses statistiques de données AirBnB*.
- 2021 Gauthier Pervieux, IUT de Paris (2 months), *Etude du cancer du sein récepteur-positif d'hormone ER+ .*
- 2022 Michaël Vollenweider, ETH Zürich (4 months), *Benchmark of gene regulatory network inference methods*.
- Riccardo Fumagalli, ETH Zürich (4 months), *Identification of genes involved in the development of ER+ breast cancer*.
- Master 2015 Teun de Planque and Christopher Elamri, Stanford University (2 months), *Identifying genes with prognostic DNA methylation rates for breast cancer survival*.
- Reyna Zhang, Stanford University (2 months), *Data fusion for predicting cancer survival*.
- 2022 Julia Netzel, ETH Zürich (4 months), *Handling gender bias in NLP models*.

Responsabilities

- 2019-2021 Head of the 2nd year students of IUT de Paris

Research enhancement

Seminars

- Since 2020 Co-organiser of the statistics seminar MAP5 with J. El Methni

Workgroups

- Since 2011 Member of the MIA *NetBio* research group (inference and analysis of gene regulatory networks).
- 2015-2016 Member of the *Cancer Target Discovery and Development CTD2* group and of the *The Cancer Genome Atlas* project.

Editorial activities

- Associate editor of The International Journal of Biostatistics.

Paper reports for the Journal of the Royal Statistical Society, Frontiers in Public Health and Annals of Applied Statistics.

Informatic and languages

Informatic

- Mathematic: R, Matlab, Scilab, SAS, Maple.
- L^AT_EX, Open Office, html.

Languages

French (Native tongue), English (Fluent), German and Spanish (Intermediate).