Journées de Statistique à Rennes, 2026

IRMAR, CREST-ENSAI

19-20 mars 2026

JSTAR

Since 2004, statisticians from IRMAR, CREST/ENSAI and INRIA Rennes/IRISA have organised the Rennes Statistics Days (JSTAR) every year. The twenty-second edition of JSTAR will be held at ENSAI in March 2026. The 2026 edition of JSTAR will be an opportunity to discuss research related to **functional data analysis** during two days of presentations by speakers who are specialists in the field.

- to develop and strengthen synergies between members of the Rennes statistical community through themed days,
- to provide an overview of recent advances in dynamic and varied areas of statistical research, presented by leading national specialists,
- to foster the emergence of new areas of research within the Rennes site.

Organizing committee

David Causeur (Institut Agro)	Jean-François Dupuy (INSA)
Mohamed El Hasnaoui (ENSAI)	Madison Giacofci (Université Rennes 2)
Marion Gousse (ENSAI)	Clément Levrard (Université de Rennes)
Anouar Meynaoui (Université Rennes 2)	Valentin Patilea (ENSAI)
Myriam Vimond (ENSAI)	

Sponsors

CREST	ENSAI	ANR FUNMATHStat	IRMAR	Centre Henri Lebesgue	Rennes Métropole

Timetable : Thursday, 19 of March

10:00–10:30	**Welcome coffee with pastries**	
10:30–11:10	**Vincent Rivoirard** Université Paris Dauphine	Minimax estimation of functional principal components from noisy discretized functional data
11:10–11:50	**Nassim Bourarach** CEREMADE (Dauphine-PSL)	Minimax-Optimal Functional Principal Components Estimation from Noisy Common-Grid Data: Smooth Processes and Spectral Identifiability
11:50–12:30	**Sophie Dabo-Niang** Université de Lille	Principal component analysis of multivariate spatial functional data
12:30–14:00	**Lunch**	
14:00–14:40	**Anna Simoni** CREST-CNRS, ENSAE and Ecole Polytechnique	Impulse Response Function and nowcast with functional approaches and mixed-frequency data
14:40–15:20	**Eva-Maria Maier** Humbolt-Universität	Additive density-on-scalar regression in Bayes Hilbert spaces with an application to gender economics
15:20–15:50	**Coffee**	
15:50–16:30	**Martin Mugnier** Paris School of Economics	Inference after discretizing time-varying unobserved heterogeneity
16:30–17:10	**Myriam Vimond** CREST-ENSAI	Data depth for unparametrized curves
19:00	**Dinner**	

Timetable : Friday, 20 of March

9:30–10:10	**Aymeric Stamm** Univ. Nantes	Domain Selection and Familywise Error Rate for Functional Data: A Unified Framework
10:10–10:40	**Coffee**	
10:40–11:20	**Lucy Attwood** ENS Lyon	A Normalised Kernel Test for Independence
11:20–12:00	**Zaineb Smida** ICJ, INSA Lyon	Spatial scan statistics for functional data : application to economic and climate data
12:00–13:30	**Lunch**	
13:30–14:10	**Céline Duval** Sorbonne Université	Nonparametric estimation for additive concurrent regression models
14:10–14:40	**Coffee**	
14:40–15:20	**Hassan Maissoro** CREST-ENSAI	Adaptive Prediction for Weakly Dependent Functional Time Series
15:20–16:00	**Angelina Roche** Univ. Paris-Cité	Adaptive nonparametric estimation in the functional linear model with functional output

Minimax estimation of functional principal components from noisy discretized functional data

19 March
10:30
Amphi

Vincent Rivoirard

Université Paris Dauphine

Functional Principal Component Analysis is a reference method for dimension reduction of curve data. Its theoretical properties are now well understood in the simplified case where the sample curves are fully observed without noise. However, functional data are noisy and necessarily observed on a finite discretization grid. Common practice consists in smoothing the data and then to compute the functional estimates, but the impact of this denoising step on the procedure's statistical performance are rarely considered. Here we prove new convergence rates for functional principal component estimators. We introduce a double asymptotic framework: one corresponding to the sampling size and a second to the size of the grid. We prove that estimates based on projection onto histograms show optimal rates in a minimax sense. Theoretical results are illustrated on simulated data and the method is applied to the visualization of genomic data.

Minimax-Optimal Functional Principal Components Estimation from Noisy Common-Grid Data: Smooth Processes and Spectral Identifiability

19 March
11:10
Amphi

Nassim Bourarach

CEREMADE (Dauphine-PSL)

In this talk, we consider minimax estimation of covariance eigenfunctions and eigenvalues in functional principal component analysis (FPCA) when sample trajectories are observed on a common discretization grid and corrupted by additive noise. A central theme is that covariance-kernel regularity should be understood jointly through functional smoothness and spectral structure: when no specific parametric eigenvalue decay is imposed, additional spectral identifiability—captured through weighted eigenvalue separation—becomes necessary.
To formalize this interplay, we introduce a class of stochastic processes that simultaneously controls (i) the Hölder smoothness of the covariance kernel and (ii) a weighted separation measure for the ℓ-th eigenvalue. We will first see that this class covers numerous standard stochastic processes. then over this class, we derive sharp, nonasymptotic minimax lower bounds that disentangle the respective roles of sampling and discretization. These results reveal phase transitions in the difficulty of estimating both eigenfunctions and eigenvalues, with representative rates of the form $n^{-1} + p^{-2\alpha}$ (up to a spectral factor) for eigenfunctions and $n^{-1} + p^{-4\alpha}$ for eigenvalues.
On the methodological side, we present a fully computable wavelet projection estimator, based on Coiflet scaling functions and a quadrature scheme tailored to cover low and high-order smoothness. We then show that this estimator attains the corresponding minimax upper bounds for any Hölder index $\alpha > 0$.

Principal component analysis of multivariate spatial functional data

19 March
11:50
Amphi

Sophie Dabo-Niang
Université de Lille

Impulse Response Function and nowcast with functional approaches and mixed-frequency data

19 March
16:30
Amphi

Anna Simoni
CREST-CNRS, ENSAE and Ecole Polytechnique

Prediction and causal models often involve economic time series with different sampling frequencies. The mix-frequency problem has received a lot of attention in the nowcasting/forecasting literature which has proposed solutions that work especially well for prediction when the frequency gap is small, like the gap between monthly and quarterly data. This paper proposes an alternative approach to deal with mix-frequency that is particularly designed for situations where the gap in sampling frequencies is large, like daily and quarterly. Such a large gap can be easily found when one mixes series from standard and non-standard sources, like internet or newspapers. Our approach focuses on both prediction and recovering of the impulse response function (IRF), and exhibits excellent performances. By treating the high-frequency variable as a realization of a stochastic process in continuous time, we cast the estimation problem in the class of ill-posed inverse problems and propose different regularized estimators. Importantly, for each high-frequency covariate, we recover a measure of its influence on the target variable as a function of the time-gap between the forecasting horizon and the date of the information in the past.

Additive density-on-scalar regression in Bayes Hilbert spaces with an application to gender economics

19 March
14:40
Amphi

Eva-Maria Maier
Humbolt-Universität

Motivated by research on gender identity norms and the distribution of the woman's share in a couple's total labor income, we consider additive regression models for densities as responses with scalar covariates. To preserve nonnegativity and integration to one under vector space operations, we formulate the model for densities in a Bayes Hilbert space, which allows to not only consider continuous densities but also, for example, discrete or mixed densities. Mixed ones occur in our application, as the woman's income share is a continuous variable having discrete point masses at zero and one for single-earner couples. Estimation is based on a gradient boosting algorithm, allowing for potentially numerous flexible (linear, nonlinear, categorical, interaction etc.) covariate effects and model selection. We show useful properties of Bayes Hilbert spaces related to subcompositional coherence, also yielding new

(odds-ratio) interpretations of effect functions and simplified estimation for mixed densities via an orthogonal decomposition. Applying our approach to data from the German Socio-Economic Panel Study (SOEP) shows a more symmetric distribution in East German than in West German couples after reunification and a smaller child penalty comparing couples with and without minor children. These West–East differences become smaller but are persistent over time.

Inference after discretizing time-varying unobserved heterogeneity

19 March
15:50
Amphi

Martin Mugnier
Paris School of Economics

Approximating time-varying unobserved heterogeneity by discrete types has become increasingly popular in economics. Yet, provably valid post-clustering inference for target parameters in models that do not impose an exact group structure is still lacking. This paper fills this gap in the leading case of a linear panel data model with nonseparable two-way unobserved heterogeneity. Building on insights from the double machine learning literature, we propose a simple inference procedure based on a bias-reducing moment. Asymptotic theory and simulations suggest excellent performance. In the application on fiscal policy we revisit, the novel approach yields conclusions in line with economic theory.

Data depth for unparametrized curves

19 March
16:30
Amphi

Myriam Vimond
CREST-ENSAI

In 1975, John W. Tukey defined statistical data depth as a function that determines the centrality of an arbitrary point with respect to a data cloud or to a probability measure. During the last decades, this seminal idea of data depth evolved into a powerful tool proving to be useful in various fields of science. Recently, extending the notion of data depth to the functional setting attracted a lot of attention among theoretical and applied statisticians. We go further and suggest a notion of data depth suitable for data represented as curves, or trajectories, which is independent of the parameterization. We show that our curve depth satisfies theoretical requirements of general depth functions that are meaningful for trajectories. We apply our methodology to diffusion tensor brain images and also to pattern recognition of handwritten digits and letters.

Domain Selection and Familywise Error Rate for Functional Data: A Unified Framework

20 March
10:00
Amphi

Aymeric Stamm
Univ. Nantes

Functional data are smooth, often continuous, random curves, which can be seen as an extreme case of multivariate data with infinite dimensionality. Just as componentwise inference for multivariate data naturally performs feature selection, subsetwise inference for functional data performs domain selection. In this paper, we present a unified testing framework for domain selection on populations of functional data. In detail, p-values of hypothesis tests performed on pointwise evaluations of functional data are suitably adjusted for providing control of the familywise error rate (FWER) over a family of subsets of the domain. We show that several state-of-the-art domain selection methods fit within this framework and differ from each other by the choice of the family over which the control of the FWER is provided. In the existing literature, these families are always defined a priori. In this work, we also propose a novel approach, coined thresholdwise testing, in which the family of subsets is instead built in a data-driven fashion. The method seamlessly generalizes to multidimensional domains in contrast to methods based on a priori defined families. We provide theoretical results with respect to consistency and control of the FWER for the methods within the unified framework. We illustrate the performance of the methods within the unified framework on simulated and real data examples and compare their performance with other existing methods.

A Normalised Kernel Test for Independence

20 March
10:40
Amphi

Lucy Attwood
ENS Lyon

Kernel methods are a powerful framework for non parametric statistical testing. They allow the embedding of complex probability distributions into functional spaces, known as reproducing kernel Hilbert spaces (RKHS), on which statistical properties can be studied. In particular, covariance operators can be precisely defined using kernels from which independence tests can be derived. The reference kernel test for independence uses the HSIC (Hibert-Schmidt Independence Criterion) measure. This tests for null cross-covariance between two variables embedded in distinct RKHSs. Despite it's popularity, HSIC has limitations. First an intractable null distribution demands intensive permutation strategies to compute p-values. Second, HSIC lacks an explicit underlying model and hence doesn't account for measurement errors. In this work, we show that HSIC can be described with respect to a functional linear model between response and regressor variables embedded into distinct RKHSs. Thanks to this model, we propose a new statistic on the expectation of the response variable conditional to the predictors normalised with respect to its variance, in essence a kernelised version of the Hotelling Lawley test. We use a spectral regularisation to invert covariance operators which gives a simple and

tractable asymptotic distribution under null. This regularisation is also advantageous for the model, providing a visual representations of variables in these functional RKHSs as well as the relationship between them. In addition, it draws connections between kernel based independence testing and dimension reduction. We benchmark our test on simulated and experimental biological data and show that it can be used to investigate the spatial variability of gene expression.

Spatial scan statistics for functional data : application to economic and climate data

20 March
11:20
Amphi

Zaineb Smida

ICJ, INSA Lyon

The detection of spatial clusters has been an active research area in statistics for several decades. Among the existing approaches, spatial scan statistics make it possible to identify significant clusters within a geographic region, either through a parametric method (using the likelihood ratio test) or a nonparametric one (using the Wilcoxon test statistic). In this talk, we introduce two scan statistics specifically designed for functional data (i.e., infinite-dimensional data): a nonparametric version based on the Wilcoxon–Mann–Whitney statistic, and a parametric version derived from Hotelling's test statistic. These two approaches extend classical univariate and multivariate scan statistics to the functional setting. Our simulation studies show that the Hotelling-based method outperforms several existing approaches, including the nonparametric one, in terms of both cluster detection and localization. Finally, the proposed methods are applied to two real-world datasets: economic data on unemployment rates in Spain and climate data in Great Britain, with the aim of identifying regions where climate change exhibits different patterns.

Nonparametric estimation for additive concurrent regression models

20 March
13:30
Amphi

Céline Duval

Sorbonne Université

Adaptive Prediction for Weakly Dependent Functional Time Series

20 March
14:40
Amphi

Hassan Maissoro
CREST-ENSAI

We propose an adaptive curve prediction method for stationary functional time series (FTS) with irregular sample paths that are observed with error at discrete domain points. Our approach is based on the best linear unbiased predictor and requires knowledge of the mean, covariance and autocovariance functions of the FTS, as well as the conditional variance of the measurement errors. We introduce new adaptive nonparametric estimators for the mean, covariance and autocovariance functions, using adaptive local bandwidth rules. These estimators adapt to the local regularity of the FTS, resulting in improved prediction accuracy. As a first step, we introduce a local regularity estimator and derive a non-asymptotic concentration bound for it. We also derive pointwise and uniform convergence rates for our adaptive estimators. A simulation study and a real-world data application on daily wind and solar photovoltaic electricity production illustrate the good performance of the new predictor.

Adaptive nonparametric estimation in the functional linear model with functional output

20 March
15:20
Amphi

Angelina Roche
Univ Paris-Cité

We study adaptive estimation in the functional linear regression model with functional output, in which both the covariate and the response variable are functional random variables. First, we introduce a collection of projection estimators over finite-dimensional subspaces. We then provide a non-asymptotic bias-variance decomposition of the mean square prediction error when these subspaces are generated using the empirical PCA functional basis. This automatic trade-off is realised thanks to a model selection device that selects the optimal projection dimensions: the penalised contrast estimator satisfies an oracle-type inequality and is therefore optimal from an adaptive perspective. These upper bounds allow us to derive convergence rates over ellipsoidal smoothness spaces. These rates are shown to be minimax optimal: they match a lower bound of the minimax risk, which is also proven. A numerical study based on simulated and real data sets will also be presented.
