

Abstract Book
2026 IMS Annual Meeting
July 6–9, 2026, Salzburg, Austria

All abstracts received as of 23 June 2026. Alphabetical by presenting author's last name. Each entry includes Title, Author, Email, Affiliation, and Abstract.

Title: A Predictive Bootstrap Technique for Double-Censored Data

Author: Aldawsari, Abdulrahman

Email: abd.aldawsari@psau.edu.sa

Affiliation: Prince Sattam bin Abdulaziz University

Abstract: Bootstrap methods are widely used because they are simple to implement and have strong theoretical and practical performance. We introduce a novel predictive bootstrap technique for double-censored data, designed specifically for predictive inference within a fully parametric setting. The proposed method is evaluated using simulation scenarios commonly employed to assess the predictive performance of bootstrap procedures. The proposed approach is compared with Efron's classical bootstrap in terms of prediction interval coverage probabilities. In addition, a global measure of coverage accuracy is examined, and a chi-squared goodness-of-fit test is applied to provide a broader evaluation of performance. The results show that the proposed procedure achieves strong predictive performance, which is consistent with its explicitly predictive construction.

Title: Uncovering Latent Structures in Large-Scale Cognitive Conflict Data Using a Novel Probabilistic Model-based Clustering Approach

Author: Ahmed, Zaheer

Email: ahmed@ifado.de

Affiliation: Leibniz Research Center for Working Environment and Human Factors (IfADo) of TU Dortmund Germany

Abstract: Uncovering latent structures and interactions in cognitive data is essential for advancing our understanding of complex cognitive processes in humans. Despite the abundance of data from diverse studies, effective approaches for disentangling these latent patterns are yet to be developed. We employed a recently established probabilistic model-based clustering approach to uncover latent structures in an extensive dataset, the Dortmund Vital Study (DVS, ClinicalTrials.gov Identifier: NCT05155397), a prospective study on healthy cognitive aging, which includes data recorded from about 600 participants in various cognitive tasks. Beyond uncovering latent structures, this method is suitable for identifying and eliminating outlier participants, indicating that outcome variables might differ at the level of various participant groups identified by the analysis. The main advantage of the proposed analytical tool is that it can simultaneously identify latent groups of participants and tasks, offering a data-driven, more comprehensive understanding of latent structures. By capturing these latent structures and participant groups, the probabilistic model-based clustering method allows for a more sophisticated analysis of how participants' cognitive profiles relate to the tasks or conditions they encounter.

Title: Sampling-Based Regularisation for Adversarial Linear Bandits

Author: Akhavan, Arya

Email: arya.akhavan@stats.ox.ac.uk

Affiliation: University of Oxford

Abstract: We study the design of sampling-based regularisation in sequential decision-making. In the online learning setting, we show that the exploration-exploitation trade-off can be optimally controlled by sampling from carefully designed probability distributions, providing a statistical alternative to explicit penalty functions. Focusing on adversarial linear bandits, we develop a unified framework that connects Follow-the-Regularized-Leader (FTRL) and Follow-the-Perturbed-Leader (FTPL), extending their known equivalence beyond the full-information setting. Within this framework, we introduce self-concordant perturbations, a family of distributions that mirror the role of self-concordant barriers used in the FTRL-based SCRiBLE algorithm. This yields a new FTPL algorithm that combines self-concordant regularisation with efficient stochastic exploration. The resulting method achieves regret $O(dn \ln n)$ on both the d -dimensional hypercube and the Euclidean ball. On the Euclidean

ball, this matches the best known guarantees for self-concordant FTRL methods, while on the hypercube it improves them by a factor of d , matching the optimal rate up to logarithmic factors.

Title: A Predictive Approach for Classifying Future Observations into Ordered Groups

Author: Alharbi, Abdulmajeed

Email: abdulmajeed1@KSU.EDU.SA

Affiliation: Department of Statistics and Operations Research, College of Science, King Saud University, Riyadh, Saudi Arabia

Abstract: In diagnostic testing and decision-making problems, classification often relies on selecting suitable decision thresholds. The choice of these thresholds plays an important role in how observations are assigned to different groups and can substantially influence classification performance, particularly when the groups have a natural ordering. This work develops a predictive approach for selecting thresholds that focuses on classification performance for future observations while relying on minimal distributional assumptions. The proposed method explicitly incorporates multiple future observations and can be applied in a range of classification settings, including those involving more than two groups. Illustrative examples and extensive simulation studies compare the proposed method with commonly used threshold selection approaches. The results show that the proposed method consistently achieves a higher number of correctly classified future observations across groups, particularly in more challenging scenarios.

Title: Probability Distributions for Component Failures in Multi-Type Systems

Author: Alharshan, Anas

Email: af.alharshan@psau.edu.sa

Affiliation: Prince Sattam bin Abdulaziz University

Abstract: This research develops a formal probability distribution for the number of failed components in systems comprising multiple distinct types. Although predicting failure counts is essential for spare part logistics and reliability centering, the existing literature is largely confined to single-component types or restricted architectures. Furthermore, current models often neglect the “survival history” of the system: the fact that the distribution of failed components at the moment of system collapse is contingent on the system having been operational up to a specific time t . By leveraging the survival signature framework, we derive a generalized expression for the joint probability of failed components across multiple types at the exact moment of system failure. This model explicitly accounts for the system's functional status at time t and provides the distribution for failures occurring within specific future time intervals. In this work, we demonstrate the derivation of these probabilistic results and provide illustrative examples of their application in complex, multi-type system configurations.

Title: Conditional validity and a fast approximation formula of full conformal prediction sets

Author: Amann, Nicolai

Email: nicolai.amann@univie.ac.at

Affiliation: University of Vienna

Abstract: Prediction sets based on full conformal prediction have seen an increasing interest in statistical learning due to their universal marginal coverage guarantees and versatility across a wide range of applications. However, practitioners have refrained from using it in real-world implementations for two reasons: firstly, it comes at very high computational costs, exceeding even that of cross-validation. Secondly, an applicant is typically not interested in a marginal coverage guarantee which averages over all possible training data sets, but rather in a guarantee conditional on the specific training data. We tackle these problems by showing that full conformal prediction provides reliable prediction sets conditional on the training data if the conformity score is stochastically bounded and satisfies a stability condition. We also propose an approximation for the full conformal prediction set that has asymptotically the same training conditional coverage as full conformal prediction under the stability assumption, and can be computed more easily. Furthermore, we show that under the stability assumption, another variant, n -fold cross-conformal prediction, also has the same asymptotic training conditional coverage guarantees as full conformal prediction. If the conformity score is defined as the out-of-sample prediction error, our approximation of the full conformal set coincides with the symmetrized Jackknife.

Title: Parameter estimation for graphon-interacting particle systems from discrete observations

Author: Amorino, Chiara

Email: chiara.amorino@upf.edu

Affiliation: Universitat Pompeu Fabra

Abstract: In this paper, we address the problem of joint parameter estimation for the drift and diffusion coefficients of heterogeneously interacting diffusive particle systems and their associated large population limits. Unlike the homogeneous setting, the interaction here is of mean-field type with weights characterized by an underlying graphon. This heterogeneity introduces significant analytical complexity: while homogeneous particles converge to an i.i.d. limit, our system leads to particles that are independent but not identically distributed in the limit. Starting from discrete observations of the system over a fixed interval, we propose a contrast function based on a pseudo-likelihood approach. We establish that the resulting estimators are consistent as the discretization step tends to zero and the number of particles tends to infinity. Furthermore, we prove asymptotic normality under an additional regime, demonstrating that the graphon-based structure can be successfully navigated despite the lack of identical distribution in the limit.

Title: An ordering for the strength of functional dependence

Author: Ansari, Jonathan

Email: jonathan.ansari@plus.ac.at

Affiliation: University of Salzburg

Abstract: We introduce a new dependence order—the conditional convex order—whose minimal and maximal elements characterize independence and perfect dependence. Moreover, it characterizes conditional independence, satisfies information monotonicity, and exhibits several invariance properties. Consequently, it is an ordering for the strength of functional dependence of a random variable Y on a random vector X . As we show, various recently studied dependence measures—including Chatterjee’s rank correlation, Wasserstein correlations, and rearranged dependence measures—are increasing in this order and inherit their fundamental properties from it. We characterize the conditional convex order by the Schur order and by the concordance order, and we verify it in settings such as additive error models, the multivariate normal distribution, and various copula-based models. Our results offer a unified perspective on the behavior of dependence measures across statistical models.

Title: Statistical foundations of representation learning in generative models

Author: Aragam, Bryon

Email: bryon@chicagobooth.edu

Affiliation: University of Chicago

Abstract: One of the key paradigm shifts in statistical machine learning over the past decade has been the transition from handcrafted features to automated, data-driven representation learning. A crucial step in this pipeline is to identify latent and learn structured representations from data. In many applications, meaningful concepts are not directly observed, and must be learned from data, often using flexible, nonparametric models such as deep generative models. These settings present new statistical and computational challenges that will be the focus of this talk. We will revisit the statistical foundations of nonparametric latent variable models as a lens into the problem of identifying representations in deep generative models. We discuss recent work on developing methods for identifying and learning causal representations from data with rigorous guarantees, and discuss how even basic statistical properties are surprisingly subtle. Along the way, we will explore the connections between deep generative models, nonparametric latent variable models, and causal graphical models.

Title: Reoptimization in the dynamic and stochastic knapsack problem: from near-optimality to distributional equivalence

Author: Arlotto, Alessandro

Email: aa249@duke.edu

Affiliation: Duke University

Abstract: We study a dynamic and stochastic knapsack problem in which a decision maker is sequentially presented with n items with unitary rewards and independent weights that are drawn from a known continuous distribution. The decision maker seeks to maximize the expected reward she collects by including items in a knapsack while satisfying a capacity constraint, and while making terminal decisions as soon as each item weight is revealed. We propose a reoptimized heuristic and compare its total rewards with that of the optimal dynamic programming policy. We show that the two total rewards have the same asymptotic mean, the same asymptotic variance, and the same limiting distribution. In contrast, other asymptotically optimal heuristics considered in the literature have different higher moments and different limiting distributions. This is joint work with Xinchang Xie and Yun-Tung Kuo.

Title: Validating spatial-temporal separability for stationary processes

Author: Bai, Lujia

Email: lujia.bai@rub.de

Affiliation: Ruhr University Bochum

Abstract: A crucial assumption to reduce computational complexity in spatial-temporal data analysis is separability, which factors the covariance structure into a purely spatial and a purely temporal component. In this paper, we develop statistical inference tools for validating this assumption for a second-order stationary process under both domain-expanding-infill asymptotics and domain-expanding asymptotics. In contrast to previous work, the methodology does not require normal distributed data nor spectral methods. Our approach is based on nonparametric estimates of measures for the deviation between the covariance matrix and separable approximations, which vanish if and only if the assumption of separability is satisfied. We derive asymptotic distributions of appropriate estimators for these measures with non-standard limiting distributions and use these results to develop confidence intervals, tests for exact separability, and tests for whether deviation from separability is smaller than a pre-specified threshold.

Title: Percolation on Hierarchical Lattices

Author: Baldasso, Rangel

Email: rangel@puc-rio.br

Affiliation: PUC-Rio

Abstract: In this talk we examine families of hierarchical lattices, with focus on fine percolative properties. Thanks to the self-similarity of the graphs in these sequences, several properties of the model can be examined, including critical exponents, scaling relations, and noise sensitivity. Furthermore, for these models, we obtain explicit expressions for several critical exponents of percolation. The results are sharp and proofs employ a combination of tools from probability and dynamical systems. Based on joint work with Alves, Moreira, and Teixeira.

Title: Numerical Radius of Non-Hermitian Random Matrices

Author: Bao, Zhigang

Email: zgbao@hku.hk

Affiliation: The University of Hong Kong

Abstract: For a square matrix, the range of its Rayleigh quotients is known as the numerical range, a compact and convex set by the Toeplitz–Hausdorff theorem. The largest value in this convex set is known as the numerical radius, often used to study the convergence rate of iterative methods for solving linear systems. This talk introduces a recent result on the asymptotic behavior of the numerical radius of a large-dimensional, complex, non-Hermitian random matrix and its elliptic variants. For the former, the radius can be represented as the extremum of a stationary Airy-like process, which undergoes a correlation-decorrelation transition from a small to a large time scale. Based on this transition, we obtain the precise first and second order terms of the numerical radius. In the elliptic case, the fluctuation of the numerical radius reduces to the maximum of two independent Tracy-Widom variables. Based on joint work with Giorgio Cipolloni.

Title: On the role of Rademacher complexities in statistical learning

Author: Bartl, Daniel

Email: bartld@nus.edu.sg

Affiliation: National University of Singapore

Abstract: We study the problem of learning with respect to the squared loss over a convex class of functions. It has long been believed that the sample complexity in this setting is governed by localized Rademacher complexities. We show that, assuming access to coarse information on the covariance structure of the model class, the sample complexity is instead controlled by a localized complexity associated with the limiting Gaussian process. In heavy-tailed regimes, this quantity can be significantly smaller than the Rademacher complexity.

Title: Testing for Isotropy of Function-Valued Random Fields

Author: Baumhake, Julius

Email: baumhake@tugraz.at

Affiliation: Graz University of Technology

Abstract: We propose a new nonparametric approach for testing isotropy, i.e. invariance in distribution under rotations around

the origin, of function-valued random fields. The key idea is to extract local, directional-dependent characteristics of the field and analyze their distributions with appropriate adjustments to account for spatial dependence. While function-valued random fields are well studied in spatial statistics, nonparametric methods for assessing their isotropy have not yet been established, and our framework is designed to fill this gap. We outline how such a test can be used for applications in materials science, particularly to assess cylindrical isotropy of local volume fractions in paper-based materials extracted from 3D image data.

Title: Online Pandora's Box for Contextual LLM Cascading

Author: Belloni, Alexandre

Email: abn5@duke.edu

Affiliation: Duke and Amazon

Abstract: Motivated by Large Language Model (LLM) cascading, we propose an online contextual Pandora's Box model for adaptively querying and selecting LLM APIs. In each period, a decision-maker observes a request context and faces a two-phase decision problem. In the query phase, the decision-maker sequentially queries APIs, where each query reveals a generated output and the decision-maker incurs an output-dependent cost. In the selection phase, the decision-maker selects one of the generated outputs to deploy and observes only the downstream reward of the deployed output. This output-mediated feedback structure differs from classical online contextual Pandora's Box models, in which opening a box directly reveals its reward. Rather than estimating the full conditional output and cost distributions of each API, we directly model the reservation index and develop a learning approach for the query phase. Specifically, we impose a parametric structure on the contextual reservation index functions induced by the classical Weitzman's policy. Our policy combines generalized method of moments type estimation of these reservation indices with UCB-style confidence bounds for both these indices and the shared output-level reward evaluator. Under regularity conditions, we prove that the resulting policy achieves dimension-dependent cumulative regret over a finite horizon.

Title: Cutoff for independent random walks on the circle conditioned not to intersect

Author: Ben-Hamou, Anna

Email: anna.ben_hamou@sorbonne-universite.fr

Affiliation: Sorbonne University

Abstract: In this talk, we consider a class of Markov processes on the discrete circle introduced by König, O'Connell and Roch. These processes describe movements of exchangeable interacting particles and are discrete analogues of the unitary Dyson Brownian motion: a random number of particles jump together either to the left or to the right, with trajectories conditioned to never intersect. We provide asymptotic mixing times for stochastic processes in this class as the number of particles goes to infinity, under a sub-Gaussian assumption on the random number of particles moving at each step. As a consequence, we prove that a cutoff phenomenon holds independently of the transition probabilities, subject only to the sub-Gaussian assumption and a minimal aperiodicity hypothesis.

Title: Emergence of host dormancy in the presence of a persistent virus epidemic

Author: Blath, Jochen

Email: blath@math.uni-frankfurt.de

Affiliation: Goethe Universität Frankfurt

Abstract: We study a stochastic individual-based model for a microbial population challenged by a persistent virus epidemic. We assume that the resident microbial host population and the virus population are in stable coexistence upon arrival of a single new mutant host individual capable of switching into a reversible state of dormancy upon contact with virions as a means of avoiding infection. At the same time, the dormancy trait comes with a cost, namely a reduced individual reproduction rate while in the active state. We prove that the mutants can nevertheless invade the resident population with strictly positive probability in the large population limit. Given the reduced reproductive rate, such an invasion would be impossible in the absence of the virus epidemic. We characterize the parameter regime where this emergence of a host dormancy trait is possible, determine the success probability of a single invader and the typical time it takes successful mutants to reach macroscopic population size. We conclude by investigating the fate of the population after successful emergence of dormancy. This is joint work with András Tóbiás.

Title: Efficient Learning of Stationary Diffusions with Stein-type Discrepancies

Author: Bleile, Fabian

Email: f.bleile@tum.de

Affiliation: TUM, MCML

Abstract: Learning a stationary diffusion amounts to estimating the parameters of a stochastic differential equation whose stationary distribution matches a target distribution. We build on the recently introduced kernel deviation from stationarity (KDS), which enforces stationarity by evaluating expectations of the diffusion's generator in a reproducing kernel Hilbert space. Leveraging the connection between KDS and Stein discrepancies, we introduce the Stein-type KDS (SKDS) as an alternative formulation. We prove that a vanishing SKDS guarantees alignment of the learned diffusion's stationary distribution with the target. Furthermore, under broad parametrizations, SKDS is convex with an empirical version that is epsilon-quasiconvex with high probability. Empirically, learning with SKDS attains comparable accuracy to KDS while substantially reducing computational cost and yields improvements over most competitive baselines.

Title: Mixtures of Experts: Granularity, Routing, and Hidden Structure

Author: Boix, Enric

Email: eboix@wharton.upenn.edu

Affiliation: University of Pennsylvania

Abstract: Mixture-of-experts (MoE) layers are a key component of modern transformer architectures, yet remain poorly understood. We study MoEs in the regime of many experts and high granularity, seeking to understand their power. We establish the strength of MoEs in this regime, showing an exponential separation in expressivity in terms of the granularity parameter, and that many-expert, high-granularity MoEs are well suited to transformers because of their interface with sparse dictionary structure in activations. Finally, we propose a modification to the MoE architecture, called MoRE, that enables efficient scaling to many more experts.

Title: Two-Sample Hypothesis Testing for Subspace Equality in Network Data

Author: Brahma, Rajdeep

Email: rbrahma2@illinois.edu

Affiliation: University of Illinois Urbana-Champaign

Abstract: In many settings one is interested in determining whether two networks share joint structural connectivity patterns such as communities. While communities may be shared across networks, edge probabilities may differ significantly. We consider testing a general null hypothesis that two networks have the same underlying subspace, including the setting in which communities are the same for stochastic blockmodels or mixed-membership stochastic blockmodels even if edge probabilities differ. We propose a test statistic based on the Frobenius norm of the difference of the leading subspace projection matrices and prove that, after appropriate centering and scaling, it converges in distribution to a Gaussian random variable as long as the average expected degree grows at least logarithmically in the number of vertices. We provide estimators for the asymptotic mean and variance, show consistency under a stronger signal condition, and give the local power of the test when networks are sufficiently dense. We demonstrate the results through simulations and an application to U.S. flight data.

Title: Self-organized criticality under arbitrary sampling frequencies

Author: Brutsche, Johannes

Email: johannes.brutsche@stochastik.uni-freiburg.de

Affiliation: University of Freiburg

Abstract: We derive the rate of convergence and limiting distribution of the profile maximum likelihood estimator for the level of self-organized criticality of an oscillating Brownian motion under equidistant sampling schemes, including low-frequency asymptotics where the observations are realizations of a null recurrent Markov chain. The derived nonstandard Cox-type limit, caused by the discontinuity of the profile likelihood in the parameter of interest, is independent of the sampling frequency while the rate of convergence interpolates between low- and high-frequency observation schemes. The limit depends continuously on the true level of self-organized criticality and unknown volatility levels, enabling statistical inference. The proof relies on distributional self-similarity of the shifted oscillating Brownian motion and a coupling argument based on the strong Markov property.

Title: Identifying risky flight landings using multivariate vine copula based regression models

Author: Buchner, Ferdinand

Email: f.buchner@tum.de

Affiliation: Technische Universität München

Abstract: Runway overruns remain a significant safety concern in commercial aviation, particularly during landing. Identifying landings with elevated overrun risk is challenging because landing performance is driven by interacting phases and operational factors. Existing statistical models often consider only a single rollout component or classify flights into overrun versus non-overrun events, limiting their ability to capture dependencies within the landing process. This work proposes a trivariate distributional vine copula regression framework that jointly models touchdown distance, distance until controllable speed, and average deceleration. By modeling their joint distribution and incorporating meteorological, aircraft-specific, and operational covariates, the approach provides a comprehensive representation of landing dynamics and enables probabilistic assessments of overrun risk. The method is evaluated using QAR data from 711 landings of identical cargo aircraft on a single runway under dry conditions.

Title: Calibration and minimizing loss

Author: Buchweitz, Erez

Email: erezmb@gmail.com

Affiliation: UC Berkeley

Abstract: Are calibration and minimizing loss compatible goals? Is calibration desirable for minimizing loss? Does minimizing loss lead to calibration? We study these questions via tradeoff and recalibration arguments and show that, all other things equal, the answer to all three questions is yes for various calibration modes and proper losses. But, in general, miscalibrated forecasts can be superior to calibrated ones, even across all proper losses at the same time. Furthermore, excluding auto-calibration, the answer can be loss-dependent. We introduce a new family of calibration modes that generalize marginal calibration in a unique way for each proper loss and arise naturally from minimizing loss. This family enjoys explicit tradeoffs analogous to the bias-variance tradeoff in regression when bias is identified with miscalibration.

Title: BabyTransformer CODL learns latent domains

Author: Bunea, Florentina

Email: fb238@cornell.edu

Affiliation: Cornell University

Abstract: Token-level contextual embeddings for text corpora are readily available for both natural language and genetic sequences. We introduce CODL, a continuous topic model that leverages these embeddings to uncover latent domains represented in a corpus. The model is a very large p -mixture of high-dimensional Gaussian distributions, where mixture weights themselves are mixtures of K Softmax probability vectors supported on the p Gaussian means. By first learning many latent Gaussian means from the tokenized corpus, we construct a corpus-specific vocabulary that supports a discrete K -Softmax mixture ensemble whose atoms encode the corpus' latent domains in embedding space. These learned softmax atoms can then be mapped back to natural language, enabling interpretable topic discovery. We provide end-to-end theoretical guarantees for EM-based atom estimates and illustrate their interpretability with data examples.

Title: Linear spreading speed of a branching annihilating random walk

Author: Callegaro, Alice

Email: alice.callegaro@tum.de

Affiliation: Technical university of Munich

Abstract: We study a branching annihilating random walk in which particles evolve on the lattice in discrete generations. Each particle produces a Poissonian number of offspring which independently move to a uniformly chosen site within a fixed distance from the parent's position. Whenever a site is occupied by at least two particles, all particles at that site are annihilated. For certain parameter ranges, we show that the system dies out almost surely or survives with positive probability. In a more restricted parameter range, we strengthen survival results to complete convergence with a non-trivial invariant measure and show existence of a linear spreading speed on the integer line. A central tool is comparison with oriented percolation on a coarse-grained space-time lattice, using carefully tuned density profiles reminiscent of discrete travelling wave solutions. Based on joint works with Matthias Birkner, Jiří Černý, Nina Gantert and Pascal Oswald.

Title: Casewise and Cellwise Robust Multilinear Principal Component Analysis

Author: Centofanti, Fabio

Email: fabio.centofanti@kuleuven.be

Affiliation: KU Leuven

Abstract: Multilinear Principal Component Analysis (MPCA) is an important tool for analyzing tensor data. It performs dimension reduction similar to PCA for multivariate data. However, standard MPCA is sensitive to outliers. It is highly influenced by observations deviating from the bulk of the data, called casewise outliers, as well as by individual outlying cells in the tensors, so-called cellwise outliers. This latter type of outlier is highly likely to occur in tensor data, as tensors typically consist of many cells. This talk introduces a novel robust MPCA method that can handle both types of outliers simultaneously, and can cope with missing values as well. This method uses a single loss function to reduce the influence of both casewise and cellwise outliers. The solution that minimizes this loss function is computed using an iteratively reweighted least squares algorithm with a robust initialization. Graphical diagnostic tools are also proposed to identify the different types of outliers that have been found by the new robust MPCA method. The performance of the method and associated graphical displays is assessed through simulations and illustrated on two real datasets.

Title: Doubly self-normalized change point test under local stationarity via sparse subsampling and stationary pooling

Author: Chan, Kin Wai

Email: kinwaichan@cuhk.edu.hk

Affiliation: The Chinese University of Hong Kong

Abstract: This project studies change point tests for locally stationary time series based on sparse subsampling and stationary pooling. First, studentized data are sparsely subsampled in such a way that all selected observations are lagged by a divergent amount. Consequently, a studentized sparse cumulative sum (CUSUM) process is constructed for testing change points. The resulting test is proven to be asymptotically pivotal. This approach automatically transforms a change point test designed for stationary conditions into one suitable for local stationarity, while bypassing change point-sensitive bootstrapping or computationally intensive simulations. Second, the p-values computed from different studentized sparse CUSUM processes initiated at various time points are shown to be approximately stationary and satisfy a functional central limit theorem. A method called stationary pooling is proposed to efficiently aggregate these approximately stationary p-values to enhance statistical power. Stationary pooling is shown to retrieve the change point detection power as in the stationary setting. Building on sparse subsampling and stationary pooling, we propose a doubly self-normalized test that employs self-normalization twice for two layers of aggregation: across sparsifying lags and along consecutive initial time points. Simulations demonstrate improvements in power, size accuracy, and computational efficiency.

Title: Convergence of the EM algorithm beyond classical Gaussian regimes

Author: Chandak, Rajita

Email: rajita.chandak@epfl.ch

Affiliation: École Polytechnique Fédérale de Lausanne (EPFL)

Abstract: The EM algorithm has been widely studied theoretically for finite Gaussian mixture models. However, the practical use cases of the EM algorithm span beyond these specialized theoretical models. In this talk we will cover recent advances in theory of the EM algorithm beyond the classical setting in two primary regimes. First, we will show the advantages of generalizing the EM algorithm to the Federated Learning regime wherein data heterogeneity poses a significant challenge to the applicability of classical statistical methods. We characterize the convergence rate of the EM algorithm under all regimes of data heterogeneity under finite mixtures of linear regressions. We show that rather than being a bottleneck, in some cases, data heterogeneity can surprisingly accelerate the convergence of iterative federated algorithms. Secondly, we delve briefly into the behaviour of the EM algorithm under mixtures of non-Gaussian densities. We show in particular that even with slower tail decay, the convergence rates of the EM algorithm for mixture of densities or linear regressions do not suffer significantly, providing the foundations for extending many existing results to a larger class of mixture models.

Title: DiPMInd: Distance Profile based Mutual Independence testing for random objects

Author: Chen, Yaqing

Email: yqchen@stat.rutgers.edu

Affiliation: Rutgers University

Abstract: We develop a novel unified framework for testing mutual independence among random objects in possibly different

metric spaces. We introduce new measures of mutual independence and propose tests—including existing methods as special cases—that achieve minimax rate optimality and strong empirical power. The foundation is the new concept of joint distance profiles, which uniquely characterize the joint law of random objects under mild conditions. Our test statistics quantify the difference of the joint distance profiles of each data point with respect to the joint law and the product of marginal laws of the vector of random objects. To enhance power, we integrate this difference with respect to different measures and incorporate flexible data-adaptive weight profiles. We derive the limiting distribution of the test statistics under the null hypothesis of mutual independence and show that the proposed tests with certain weight profiles are asymptotically distribution-free if the marginal distance profiles are continuous. We also establish the consistency of the tests under sequences of alternative hypotheses converging to the null. For practical implementations, we use a permutation scheme to approximate p-values and prove that the permutation-based tests maintain type I error control under the null and achieve consistency under alternatives. We demonstrate the power of the proposed tests across various types of data objects through simulations and real data applications.

Title: MOSAiC: a unified framework for lossless, one-shot, federated learning algorithms

Author: Chen, Yong

Email: ychen123@penndmedicine.upenn.edu

Affiliation: University of Pennsylvania

Abstract: Multi-site studies are now central to biomedical research, but inference across sites remains challenging due to regulatory limits on sharing individual-level data, heterogeneous data distributions, sparse events in small centers, and the logistical burden of multi-round communication. In this talk, I introduce MOSAiC (Multi-site One-Shot Aggregation of Compressed Risk Functions), a unified framework for modern distributed research networks. Notably, MOSAiC reframes distributed learning as a mathematical problem of compressing and aggregating risk functions, leveraging recent advances in tensor networks—state-of-the-art tools for high-dimensional function approximation in scientific computing. Our MOSAiC enjoys four desirable properties that have never been achieved by any of the existing federated algorithms (except linear regressions): one-shot communication, lossless recovery of pooled-data estimates, inclusiveness of all sites irrespective of size or event rarity, and analytic submodel exploitability without re-querying partners. I will illustrate MOSAiC's validity and efficiency through applications in drug relabeling, drug repurposing, and post-market safety surveillance.

Title: Recent progress on Talagrand's convolution conjecture

Author: Chen, Yuansi

Email: yuansi.chen@stat.math.ethz.ch

Affiliation: ETH Zurich

Abstract: In 1989, Talagrand conjectured that under the heat semigroup on the Boolean hypercube, any nonnegative function on the cube exhibits a uniform tail bound that is better than that by Markov's inequality. We discuss recent progress on the conjecture and the reverse heat process used in the proof.

Title: Large field problem in coercive SPDE via scaling

Author: Chevyrev, Ilya

Email: ichevyrev@gmail.com

Affiliation: SISSA (International School for Advanced Studies)

Abstract: In this talk, I will show an approach to deriving a priori bounds for coercive SPDEs based on scaling. The basic idea is to first show bounds for the equation with a small noise and then rescale the bounds to a global scale. While many equations that the approach can handle have been treated recently with other methods, its advantages are that it is quite simple and allows one to state a single result that is applicable to a variety of equations, such as rough differential equations and parabolic/elliptic SPDEs. Based on joint work with Massimiliano Gubinelli.

Title: Dual-Homotopy Framework for Constrained EM Algorithm

Author: Choi, jisoo

Email: ddsy999@gmail.com

Affiliation: Seoul national university

Abstract: We propose a new constrained EM algorithm that is applicable to general constrained estimation problems. This proposed method is based on a novel framework, the 'dual homotopy framework,' which combines deterministic annealing

EM with a barrier-based optimization, enabling stable estimation under parameter constraints. Building on this framework, we further introduce an adaptive constrained EM algorithm that preserves likelihood monotonicity, regardless of the underlying distributional form or the specific structure of the constraints. Through simulation studies using two models and a real-data analysis under parameter constraints, we demonstrate that the proposed algorithm yields more stable and accurate estimates than existing methods, including the standard EM algorithm.

Title: The random XORSAT decimation process

Author: Coja-Oghlan, Amin

Email: amin.coja-oghlan@tu-dortmund.de

Affiliation: TU Dortmund

Abstract: We analyse the performance of Belief Propagation Guided Decimation, a physics-inspired message passing algorithm, on the random k -XORSAT problem. Specifically, we derive an explicit threshold up to which the algorithm succeeds with a strictly positive probability that we compute explicitly, but beyond which the algorithm with high probability fails to find a satisfying assignment. In addition, we analyse a thought experiment called the decimation process for which we identify a (non-) reconstruction and a condensation phase transition. The main results of the present work confirm physics predictions from [Ricci-Tersengh, Semerjian: J. Stat. Mech. 2009] that link the phase transitions of the decimation process with the performance of the algorithm, and improve over partial results from a recent article [Yung: Proc. ICALP 2024]. Joint work with Arnab Chatterjee, Amin Coja-Oghlan, Mihyun Kang, Lena Krieg, Maurice Rolvien, Gregory B. Sorkin.

Title: Large Deviations and Bahadur Efficiency for General Divergence-based Estimators

Author: Collamore, Jeffrey

Email: collamore@math.ku.dk

Affiliation: University of Copenhagen

Abstract: Minimum divergence estimators have long been studied as robust alternatives to maximum likelihood methods; see, for example, Beran (1977) and Lindsay (1994). In this work we investigate, from a general perspective, the rare-event behavior of such divergence-based inferential methods. Specifically, we establish large deviation estimates for the probabilities that a divergence-based estimator deviates from the true parameter values, and we characterize the associated exponential rate of decay as the sample size tends to infinity. This decay is described via a large deviation rate function whose form depends on the choice of the divergence measure. We then apply these results to study Bahadur efficiency for general divergence estimators, thereby extending and complementing earlier work of Bahadur (1967) and Arcones (2006).

Title: Assessing the Quality of Denoising Diffusion Models in Wasserstein Distance

Author: DALALYAN, Arnak

Email: arnak.dalalyan@ensae.fr

Affiliation: CREST / ENSAE / IP Paris

Abstract: Generative modeling aims to produce new random examples from an unknown target distribution, given access to a finite collection of examples. Among the leading approaches, denoising diffusion probabilistic models (DDPMs) construct such examples by mapping a Brownian motion via a diffusion process driven by an estimated score function. In this work, we first provide empirical evidence that DDPMs are robust to constant-variance noise in the score evaluations. We then establish finite-sample guarantees in Wasserstein-2 distance that exhibit two key features: (i) they characterize and quantify the robustness of DDPMs to noisy score estimates, and (ii) they achieve faster convergence rates than previously known results. Furthermore, we observe that the obtained rates match those known in the Gaussian case, implying their optimality.

Title: A non-equilibrium equivalence of ensembles result via propagation of chaos

Author: Darshan, Shiva

Email: shiva.darshan@essec.edu

Affiliation: ESSEC Business School

Abstract: We consider an alternative formulation of non-equilibrium stochastic dynamics referred to as Norton dynamics. A reference system is perturbed by an external forcing with varying intensity so that the value of a given observable is fixed. This amounts to constraining the dynamics to a level set of this observable. For observables that are linear functions of the empirical measure of the dynamics, this can be viewed as a constraint on the empirical mean of the dynamics. This leads us to consider mean-field interacting particle systems. We show via a propagation of chaos that an equivalence of ensembles

holds between the Norton formulation of non-equilibrium dynamics and the standard formulation in which the intensity of the forcing is fixed. We start by showing the well-posedness of the finite N mean-field Norton dynamics and the corresponding McKean-Vlasov dynamics. This novel non-linear diffusion has a non-linear drift term due to an oblique mean-constraint. This dynamics has a continuum of invariant probability measures which are the invariant probability measures of the McKean-Vlasov dynamics corresponding to the standard formulation of non-equilibrium dynamics. Equivalence of ensembles is then deduced from the fact that the McKean-Vlasov dynamics of the two formulations coincide when started at the same invariant probability measure.

Title: A flexible parametric regression approach to handle left-censored outcome and covariate information.

Author: De Batselier, Inez

Email: inez.debatselier@student.uhasselt.be

Affiliation: Hasselt University

Abstract: Left-censored data, where observations are only known to fall below a threshold, are common in environmental and biomedical sciences. While methods for censored responses are well-established, handling censored covariates remains challenging, often relying on biased “naive” substitutions or restrictive parametric assumptions. We propose a flexible parametric regression framework handling left censoring in both response and a covariate. To relax rigid assumptions of existing methods (e.g. Tran et al., 2021), we model the covariate distribution using a power piecewise exponential distribution with fixed cut-off points. Offering semi-parametric flexibility while maintaining parametric tractability. Under mild regularity conditions, we establish model identifiability. We introduce one-stage and two-stage maximum likelihood estimators for the regression and covariate parameters. The one-stage approach estimates all parameters simultaneously, maximising efficiency. The two-stage approach first estimates covariate distribution parameters, then plugs these into the full likelihood to estimate regression parameters, reducing dimensionality and computational burden. Consistency and asymptotic normality are established for both procedures. Finite-sample performance and robustness are assessed through extensive simulations. We apply this framework to an immunological study of transplacental antibody transfer, quantifying the relationship between maternal and infant antibody levels.

Title: Local e-values for intersection hypotheses

Author: de Heide, Rianne

Email: r.deheide@utwente.nl

Affiliation: University of Twente and Centrum Wiskunde en Informatica, Amsterdam

Abstract: The recent e-closure framework (Xu et. al., 2025) shows that multiple testing procedures controlling an expected loss can be understood through families of local e-values assigned to intersection hypotheses, yielding a unifying route to valid and often uniformly improved procedures. In this talk I will show new FDR controlling tests for settings in which it was not known before that such a test would exist. Xu, Z., Solari, A., Fischer, L., de Heide, R., Ramdas, A., and Goeman, J. J. (2025). Bringing Closure to False Discovery Rate Control: A General Principle for Multiple Testing. arXiv:2509.02517

Title: Controlled stochastic growth of structures in the zero-temperature Ising model

Author: de Jongh, Maïke

Email: m.c.dejongh@utwente.nl

Affiliation: University of Twente

Abstract: The optimization of structure growth in spatially stochastic systems poses a compelling challenge in a wide range of applications, including biological pattern formation and solution-based crystallization. In this talk, we investigate control strategies for a two-dimensional zero-temperature Ising model on a finite square lattice evolving under Metropolis dynamics. We consider a setting in which an external controller aims to drive the system from a configuration containing one or two small droplets of +-spins toward the all-plus configuration by flipping selected spins at prescribed times. To analyze this control problem, we formulate it as a Markov decision process (MDP), a classical framework for sequential decision-making problems under uncertainty. To compute an optimal policy, we construct a simplified model by reducing the configuration space to the local minima of the Hamiltonian. Leveraging structural properties of this simplified model, we characterize the optimal policy by solving the Bellman equations in a recursive manner. Finally, we present simulation results demonstrating the performance of the optimal policy and illustrating the induced growth mechanism. Moreover, we show that its geometric features persist at low but positive temperatures.

Title: Rao-Blackwellized e-variables

Author: de Roos, Dante

Email: dante.de.roos@cwi.nl

Affiliation: Centrum Wiskunde & Informatica

Abstract: The Rao-Blackwell theorem is a cornerstone result in classical statistics which, in its earliest form, states that the mean squared error (MSE) of an (unbiased) estimator decreases after conditioning on a sufficient statistic. This conditioning has thus come to be known as "Rao-Blackwellization." In this talk we present an analogous result for e-variables, namely that any e-variable can be Rao-Blackwellized into a new e-variable that, for every concave utility function, attains greater expected concave utility than the original. The simplest version of this Rao-Blackwell theorem for e-variables holds in full generality, but its interpretation and applicability can be limited without suitable integrability assumptions. For the log-utility, we present a more sophisticated version that overcomes these limitations and is therefore fully applicable. The conceptual breadth of these Rao-Blackwell theorems for e-variables will be illustrated with an example on testing the variance in a linear regression model.

Title: Surviving from the tip of a cone

Author: Deijfen, Maria

Email: mia@math.su.se

Affiliation: Stockholm University

Abstract: In two-type first passage percolation on the square lattice, two entities compete to capture the sites of the lattice. The entities spread between nearest neighbor sites at times specified by random passage times associated with the edges. We consider the case when both types have the same passage time distribution, with one type starting at the origin and the other from an infinite cone with tip at the origin. Can the type starting at the origin grow unboundedly? We give a partial answer to this question in terms of the slope of the cone.

Title: Pairwise Comparison Metrics: AUC, Win Ratio, and the Problem of Non-Transitivity

Author: Demler, Olga

Email: olgademler@gmail.com

Affiliation: ETH, Zurich / BWH, Harvard Medical School, Boston, USA

Abstract: The area under the receiver operating characteristic curve (AUC) is a standard measure of discrimination for risk prediction models. Despite its appealing probabilistic interpretation, its limitations are often overlooked. In this talk, we examine several challenges in the use of AUC, including its sensitivity to study design, the distribution of predictors, and the presence of competing events. We also highlight pitfalls in external validation, where AUC may not be directly comparable across populations due to differences in case-mix. These considerations underscore that AUC is not an intrinsic property of a model, but depends on the underlying data-generating mechanism. We further investigate AUC and the Win Ratio as members of a pairwise comparison group of statistics. Under certain conditions, the Win Ratio can be expressed as a monotonic transformation of AUC. However, as a head-to-head comparison measure, it may exhibit non-transitivity, a phenomenon related to classical results of Trybula. We discuss the implications of this property for the interpretation of treatment effects, particularly in randomized trials and meta-analysis.

Title: Nonparametric Estimation of Archimax Copulas and Consequences for Modeling Dependence

Author: Dietrich, Nicolas

Email: nicolaspascal.dietrich@plus.ac.at

Affiliation: University of Salzburg

Abstract: Modeling the dependence structure of extreme events has been a central topic in extreme value theory. Archimax copulas provide a flexible framework capturing both, asymptotic dependence between extreme events, and intermediate-level dependence. They are characterized by an Archimedean generator ψ and a Pickands dependence function A . In this work, we first investigate identifiability in the Archimax class. Building on these results, we propose a novel nonparametric estimator for Archimax copulas C , based on the Kendall distribution function F_K of C and the CFG estimator $A_n^{\{CFG\}}$ of A . We establish consistency of the proposed estimator under mild regularity conditions and examine its performance. Finally, we demonstrate how it can be used to estimate dependence when the underlying data-generating process is assumed to follow an Archimax copula.

Title: Efficient sampling with discrete diffusion models: Sharp and adaptive guarantees

Author: Dmitriev, Daniil

Email: dmitrievdaniil97@gmail.com

Affiliation: University of Pennsylvania

Abstract: Diffusion models over discrete spaces have recently shown striking empirical success, yet their theoretical foundations remain incomplete. In this talk, we discuss the sampling efficiency of score-based discrete diffusion models under a continuous-time Markov chain (CTMC) formulation, with a focus on tau-leaping-based samplers. We establish sharp convergence guarantees for attaining small Kullback-Leibler (KL) divergence for both uniform and masking noising processes. For uniform discrete diffusion, we show that the tau-leaping algorithm achieves an almost linear iteration complexity in the ambient dimension of the target distribution, improving the quadratic dependency from the prior work and eliminating linear dependency on the vocabulary size. Moreover, we establish a matching algorithmic lower bound showing that linear dependence on the ambient dimension is unavoidable in general. For masking discrete diffusion, we introduce a modified tau-leaping sampler whose convergence rate is governed by an intrinsic information-theoretic quantity, termed the effective total correlation, which can be sublinear or even constant for structured data. As a consequence, the sampler provably adapts to low-dimensional structure without prior knowledge or algorithmic modification, yielding sublinear convergence rates for various practical examples (such as hidden Markov models, image data, and random graphs). Joint work with Zhihan Huang and Yuting Wei.

Title: Right-Sizing Communication and Recommendation Set Size in AI-Assisted Search

Author: Dong, Jing

Email: jing.dong@gsb.columbia.edu

Affiliation: Columbia University

Abstract: We model the interaction between a user and an AI-driven recommendation system. The user initiates the process by conveying preference information through a costly and noisy message. The AI assistant, acting as a Bayesian agent, interprets the user's message to form a posterior belief about their preferences and make product recommendations. In particular, it determines how many recommendations to present to maximize the user's expected utility from their final choice, while accounting for the search cost. We use a mutual information-based cost to model the two distinct costs incurred by the user: a communication cost, which increases with the precision of the preference message, and a search cost, which increases with the size of the recommendation set. We study the high-dimensional setting and characterize how optimal message precision and recommendation set size depend on the cost parameters, under two distinct distributions from which recommendations can be sampled from the product universe: Bayes' posterior belief and an optimized tilted distribution. Under Bayes' posterior sampling, we identify a hybrid regime in which an efficient interaction policy requires jointly optimizing the amount of information conveyed by the user and the number of recommendations provided by the AI assistant. In the tilted sampling, our results show that the optimal interaction policy uses only one of communication and search, favoring whichever of them is less costly.

Title: Localization in Quantum and Classical Systems

Author: Drogin, Reuben

Email: reuben.drogin@yale.edu

Affiliation: Yale University

Abstract: Random band matrices are Hermitian matrices with random entries supported in a band of width W around the diagonal. The eigenfunctions of such matrices are expected to decay exponentially at the scale W^2 in dimension one, and $\exp(CW^2)$ in dimension two. Remarkably, the same scaling is expected for the cycle lengths in various models of random permutations where points are typically displaced by distances of order W . In this talk I will discuss some recent progress on these problems and some open questions.

Title: On Non-Markovian INAR(1) Processes

Author: Droullier, Frieder

Email: frieder.droullier@gmail.com

Affiliation: Department of Mathematics and Statistics, University of Konstanz

Abstract: We study INAR(1) processes with random coefficients. INAR(1) models were originally introduced by McKenzie (1985) and Al-Osh and Alzaid (1987). They are based on the binomial thinning operator and are Markovian. Integer-valued

processes have many applications, such as high-frequency trading and epidemiology. We introduce random coefficient INAR(1) processes with long memory and examine their characteristics. Estimation and prediction is considered. Asymptotic rates of convergence are derived. The methods are illustrated by applications to trading volumes of Microsoft and Netflix. This is joint work with Jan Beran.

Title: Parameter identification in linear non-Gaussian causal models under general confounding

Author: Drton, Mathias

Email: mathias.drton@tum.de

Affiliation: Technical University of Munich

Abstract: Linear non-Gaussian causal models postulate that each random variable is a linear function of parent variables and non-Gaussian exogenous error terms. We study identification of the linear coefficients when such models contain latent variables. Our focus is on the commonly studied acyclic setting, where each model corresponds to a directed acyclic graph (DAG). For this case, prior literature has demonstrated that connections to overcomplete independent component analysis yield effective criteria to decide parameter identifiability in latent variable models. However, this connection is based on the assumption that the observed variables linearly depend on the latent variables. Departing from this assumption, we treat models that allow for arbitrary non-linear latent confounding. Our main result is a graphical criterion that is necessary and sufficient for deciding the generic identifiability of direct causal effects. Moreover, we provide an algorithmic implementation of the criterion with a run time that is polynomial in the number of observed variables. Finally, we report on estimation heuristics based on the identification result and explore a generalization to models with feedback loops.

Title: A Proof of The Changepoint Detection Threshold Conjecture in Preferential Attachment Models

Author: Du, Hang

Email: hangdu@mit.edu

Affiliation: MIT

Abstract: We investigate the problem of detecting and estimating a changepoint in the attachment function of a network evolving according to a preferential attachment model on n vertices, using only a single final snapshot of the network. Bet et al. show that a simple test based on thresholding the number of vertices with minimum degrees can detect the changepoint when the change occurs at time $n - \Omega(\sqrt{n})$. They further make the striking conjecture that detection becomes impossible for any test if the change occurs at time $n - o(\sqrt{n})$. Kaddouri et al. make a step forward by proving the detection is impossible if the change occurs at time $n - o(n^{1/3})$. In this paper, we resolve the conjecture affirmatively, proving that detection is indeed impossible if the change occurs at time $n - o(\sqrt{n})$. Furthermore, we establish that estimating the changepoint with an error smaller than $o(\sqrt{n})$ is also impossible, thereby confirming that the estimator proposed in Bhamidi et al. is order-optimal.

Title: Seeing through correlations: Disentangled feature importance

Author: Du, Jin-Hong

Email: jinhongd@hku.hk

Affiliation: University of Hong Kong

Abstract: Quantifying feature importance with valid statistical uncertainty is central to interpretable machine learning, yet classical model-agnostic methods often fail under feature correlation, producing unreliable attributions and compromising statistical inference. Existing approaches—such as Shapley values and leave-one-covariate-out—are designed for feature selection and vulnerable to correlation distortion, limiting their robustness on model interpretation. We introduce Disentangled Feature Importance (DFI), a model-agnostic framework that resolves these limitations by combining principled statistical inference with computational flexibility. DFI leverages entropic optimal transport to learn flexible disentanglement maps and provide an interpretable pathway for understanding how importance is attributed through the data's correlation structure. The framework generalizes to flow matching and differentiable loss functions, enabling statistically valid importance assessment for black-box predictors in both regression and classification. We establish statistical inference theory, which enables valid confidence intervals and hypothesis testing with Type I error control. Empirical results on synthetic and biomedical datasets show that DFI delivers substantially higher statistical power than removal-based and conditional permutation methods, while maintaining robust, interpretable attributions under severe feature interdependence.

Title: Inference-Time Alignment for Diffusion Models via Variationally Stable Doob's Matching

Author: Duan, Chenguang

Email: duan@igpm.rwth-aachen.de

Affiliation: RWTH Aachen University

Abstract: Inference-time alignment for diffusion models aims to adapt a pre-trained reference diffusion model toward a target distribution without retraining the reference score network, thereby preserving the generative capacity of the reference model while enforcing desired properties at the inference time. A central mechanism for achieving such alignment is guidance, which modifies the sampling dynamics through an additional drift term. In this work, we introduce variationally stable Doob's matching, a novel framework for provable guidance estimation grounded in Doob's h-transform. Our approach formulates guidance as the gradient of logarithm of an underlying Doob's h-function and employs gradient-regularized regression to simultaneously estimate both the h-function and its gradient, resulting in a consistent estimator of the guidance. Theoretically, we establish non-asymptotic convergence rates for the estimated guidance. Moreover, we analyze the resulting controllable diffusion processes and prove non-asymptotic convergence guarantees for the generated distributions in the 2-Wasserstein distance. Finally, we show that variationally stable guidance estimators are adaptive to unknown low dimensionality, effectively mitigating the curse of dimensionality under low-dimensional subspace assumptions.

Title: Provable Diffusion Posterior Sampling for Bayesian Inversion

Author: Duan, Chenguang

Email: duan@igpm.rwth-aachen.de

Affiliation: RWTH Aachen University

Abstract: This paper proposes a novel diffusion-based posterior sampling method within a plug-and-play (PnP) framework. Our approach constructs a probability transport from an easy-to-sample terminal distribution to the target posterior, using a warm-start strategy to initialize the particles. To approximate the posterior score, we develop a Monte Carlo estimator in which particles are generated using Langevin dynamics, avoiding the heuristic approximations commonly used in prior work. The score governing the Langevin dynamics is learned from data, enabling the model to capture rich structural features of the underlying prior distribution. On the theoretical side, we provide non-asymptotic error bounds, showing that the method converges even for complex, multi-modal target posterior distributions. Finally, we present numerical experiments demonstrating the effectiveness of the proposed method across a range of inverse problems.

Title: Online policy learning and inference by matrix completion

Author: Duan, Congyuan

Email: cduanac@connect.ust.hk

Affiliation: Hong Kong University of Science and Technology

Abstract: Is it possible to make online decisions when personalized covariates are unavailable? We take a collaborative-filtering approach for decision-making based on collective preferences. By assuming low-dimensional latent features, we formulate the covariate-free decision-making problem as a matrix completion bandit. We propose a policy learning procedure that combines an eps-greedy policy for decision-making with an online gradient descent algorithm for bandit parameter estimation. Our novel two-phase design balances policy learning accuracy and regret performance. For policy inference, we develop an online debiasing method based on inverse propensity weighting and establish its asymptotic normality. Our methods are applied to data from the San Francisco parking pricing project, revealing intriguing discoveries and outperforming the benchmark policy.

Title: Neyman-Orthogonal Comparison of Regression Functions

Author: Dubey, Paromita

Email: paromita@marshall.usc.edu

Affiliation: University of Southern California

Abstract: Testing the equality of regression functions across two populations is fundamental for assessing whether regression relationships generalize across populations (external validity), whether different populations exhibit distinct response mechanisms (treatment heterogeneity), and whether regression structures remain stable across environments or regimes (regime stability). Classical residual-based procedures deteriorate in high dimensions and are not well suited to modern machine learning estimators. We develop a framework that recasts two-sample regression comparison as a Neyman-orthogonal conditional moment restriction on pooled data, and constructs a test based on the dual norm of its representer in a reproducing kernel Hilbert space. The resulting statistic admits a closed-form quadratic representation and is calibrated using an empirically centered Gaussian multiplier bootstrap. Under the null hypothesis, we establish a finite-sample

Kolmogorov distance bound to a Gaussian limit, prove non-asymptotic validity of the bootstrap, derive lower bounds on power against local alternatives, and show minimax optimality of the proposed test. For modern overparameterized learners such as neural networks, we additionally propose a novel k-fold cross-fitted statistic inspired by the proposed framework. Simulation studies and a real data application illustrate the utility of the proposed methods. This talk is based on joint work with Dr. Ao Sun, postdoctoral scholar at USC.

Title: Optimal binning for treatment effect heterogeneity

Author: Dukes, Oliver

Email: oliver.dukes@ugent.be

Affiliation: Ghent University

Abstract: Recent work has focused on nonparametric estimation of conditional average treatment effects (CATE) and learning optimal treatment regimes. However, inference has remained relatively unexplored, and may be challenging when the CATE depends on more than a handful of variables. In this talk I will present a flexible workflow for inference on treatment effect heterogeneity in clinical trials and observational studies. This begins a nonparametric hypothesis test of homogeneity using a class of decision rules. By inverting the test, one can obtain the optimal decision rule with provable regret guarantees. I will describe on-going work on study design, how to adaptively refine the class of rules and how to obtain a confidence interval on the treatment effect in those that benefit from treatment.

Title: Generalised Bayesian model selection

Author: Dutta, Ritabrata

Email: ritabrata.dutta@warwick.ac.uk

Affiliation: University of Warwick

Abstract: Scritly proper scoring rules (SPSR), e.g. Energy score, Kernel score, are widely used as diagnostics of probabilistic forecasting. As we can derive statistical divergences from SPSRs, recently they have become popular for the purpose of inference in both frequentist (as minimum scoring rules estimator) and Bayesian framework (as generalised posterior) in the absence of tractable likelihood functions via replacement of negative log-likelihoods with SPSRs. Here we explore a generalised Bayesian model selection framework for models without tractable likelihood functions, using Bayes factor derived from generalised posteriors. To compute the generalised marginal evidence under each model, we propose a path sampling based estimator in conjunction with sequential Monte Carlo sampling scheme. We study the consistency of the proposed procedure theoretically and using simulation studies, finally illustrating their use for the choice of challenging biological models without tractable likelihood functions.

Title: Omnibus Goodness-of-Fit Testing for Distributions on Stiefel Manifolds

Author: Edelman, Dominic

Email: dominic.edelmann@dkfz.de

Affiliation: German Cancer Research Center

Abstract: In this talk, a comprehensive framework for goodness-of-fit testing for distributions on Stiefel manifolds is developed. The approach is based on integrals of the squared differences between empirical and theoretical characteristic functions, yielding test statistics that are consistent against all fixed alternatives. For the Fisher-Bingham family of distributions, explicit computable forms of the test statistic are derived. Simplified expressions for important special cases, including the matrix Fisher, matrix Bingham, and uniform distributions are provided. In the case of testing uniformity on hyperspheres, we obtain the complete asymptotic distribution of the test statistic, enabling computationally efficient asymptotic testing. For general Fisher-Bingham distributions, we establish theoretically justified Monte Carlo testing procedures for both simple and composite hypotheses. Simulation studies demonstrate accurate Type I error control and strong power across a wide range of alternatives. The practical relevance of the proposed methodology is illustrated using an application to data on the orbits of comets.

Title: Yule's "nonsense correlation": Moments, density, and tests of independence for non-stationary Gaussian processes

Author: Ernst, Philip

Email: p.ernst@imperial.ac.uk

Affiliation: Imperial College London

Abstract: In 1926, G. Udny Yule considered the following problem: given two i.i.d. random walks of length n , what is the

distribution of their empirical (Pearson) correlation coefficient? Yule observed empirically that this statistic is both frequently far away from 0 and heavily dispersed in $(-1,1)$, leading him to call it "nonsense correlation." This led to a formulation of two concrete questions: (i) Find analytically the variance of the empirical correlation as n tends towards infinity; (ii) Find analytically the higher order moments and the density of the empirical correlation as n tends towards infinity. Donsker's theorem shows that the limiting law for the empirical correlation is the natural empirical correlation of two independent Wiener processes on $[0,1]$. In 2017, we closed question (i) after 90 years by calculating the variance of this continuous-time empirical correlation as an explicit double integral, numerically equal approximately to 0.2405. This talk begins where Ernst et al. (2017) leaves off. I shall explain how Ernst, Rogers, and Zhou (2025) succeeded in closing question (ii) by calculating all moments of the continuous-time empirical correlation, explicitly up to order 16, leading, for the first time, to an approximation to the limiting density of Yule's nonsense correlation. I will conclude by explaining the key ideas of Ernst, Viens, and Yan (2025) on the first statistical tests of independence for pairs of paths of long-memory fractional Brownian motion.

Title: High-dimensional learning dynamics of multi-pass Stochastic Gradient Descent in multi-index models

Author: Fan, Zhou

Email: zhou.fan@yale.edu

Affiliation: Yale University

Abstract: We study the learning dynamics of a multi-pass, mini-batch Stochastic Gradient Descent (SGD) procedure for empirical risk minimization in high-dimensional multi-index models with isotropic random data. In an asymptotic regime where the sample size n and data dimension d increase proportionally, for any sub-linear batch size $\kappa \sim n^\alpha$ where $\alpha \in [0,1)$, and for a commensurate "critical" scaling of the learning rate, we provide an asymptotically exact characterization of the coordinate-wise dynamics of SGD. This characterization takes the form of a system of dynamical mean-field equations, driven by a scalar Poisson jump process that represents the asymptotic limit of SGD sampling noise. We develop an analogous characterization of the Stochastic Modified Equation (SME) which provides a Gaussian diffusion approximation to SGD. Our analyses imply that the limiting dynamics for SGD are the same for any batch size scaling $\alpha \in [0,1)$, and that under a commensurate scaling of the learning rate, dynamics of SGD, SME, and gradient flow are mutually distinct, with those of SGD and SME coinciding in the special case of a linear model. We recover a known dynamical mean-field characterization of gradient flow in a limit of small learning rate, and of one-pass/online SGD in a limit of increasing sample size n/d to infinity.

Title: Central limit theorem for high temperature Ising models via martingale embedding

Author: Fang, Xiao

Email: xfang@cuhk.edu.hk

Affiliation: The Chinese University of Hong Kong

Abstract: We use martingale embeddings to prove a central limit theorem (CLT) for projections of high-dimensional random vectors satisfying a Poincaré inequality. We obtain a non-asymptotic error bound for the CLT in 2-Wasserstein distance involving two-point and three-point covariances. We present three illustrative applications: Ising model with finite-range interactions, ferromagnetic Ising model under the Dobrushin condition, and the Sherrington-Kirkpatrick spin glass model at sufficiently high temperature. In all the examples, we allow heterogeneous external fields. This talk is based on joint work with Yang Xie and Yi-Kun Zhao. The manuscript is available at: <https://arxiv.org/abs/2511.06196v3>

Title: Projected Information Criteria

Author: Fedotov, Maxim

Email: maxim.fedotov@upf.edu

Affiliation: Universitat Pompeu Fabra

Abstract: We introduce Projected Information Criteria, a new class of variable selection criteria for regression and classification problems. The proposed approach offers a computationally efficient alternative to classical criteria such as AIC, BIC, and EBIC by allowing to reduce problem dimensionality and avoiding costly repeated optimization for each candidate subset. The projected criteria are based on a global estimate computed once using all available covariates, which is then being reused during inference. We show that projected criteria exhibit strong theoretical guarantees: in a broad class of models, including generalized linear models, they achieve selection consistency under mild conditions on the model and the global estimator. We also establish non-asymptotic error probability bounds for additive models with Gaussian noise, providing finite-sample guarantees and supporting their use beyond the fixed-dimensional asymptotic regime. In addition, we

show that the resulting selection problem can be formulated as a binary quadratic optimization problem, enabling practical computation with modern solvers. We argue that the computational gain compared to classical criteria is particularly pronounced when the global estimate is sparse. In simulation studies, the projected criteria demonstrate highly competitive performance relative to popular alternatives, including the regular and adaptive Lasso.

Title: Gaussian Free Field with Hard Wall Constraints From Discrete Trees to Planar Domains

Author: Fels, Maximilian

Email: hmmmwhat1@gmail.com

Affiliation: Universität zu Köln

Abstract: The Discrete Gaussian Free Field (DGFF) serves as a canonical model for random interfaces in statistical mechanics. When conditioned to stay non-negative on a domain—an event known as the "hard wall" constraint—the field undergoes entropic repulsion, lifting away from the wall to accommodate its fluctuations. The study of this phenomenon in two dimensions was pioneered by Bolthausen, Deuschel, and Giacomin (2001), who established the leading-order asymptotics for the probability of the hard wall event. In this talk, I will present two major recent advancements that significantly refine our understanding of this problem: The Binary Tree (Fels, Hartung, Luidor 2024): We first discuss the case of the DGFF on the binary tree, where the hierarchical structure permits a complete derivation of the conditioned field's behavior. We will show that the hard wall constraint leads to a repulsion profile that makes the conditioned law asymptotically mutually singular with respect to the unconditioned law. The 2D Plane (Fels, Luidor, Wu 2026): Turning to the Euclidean lattice \mathbb{Z}^2 , I will present very recent work that improves upon the classical BDG result by capturing the subleading order of the probability asymptotics. The analysis relies on a novel orthogonal decomposition of the field into a "constant" component and a conditioned remainder, handled via multi-scale analysis and double-exponential tail bounds.

Title: Learning the score under shape constraints

Author: Feng, Oliver

Email: o.feng@lse.ac.uk

Affiliation: London School of Economics

Abstract: Score estimation has recently emerged as a key statistical challenge due to its pivotal role in generative modelling via diffusion models, and as an essential ingredient in a new approach to linear regression via convex M-estimation. Motivated by these applications, we study the minimax risk of estimating the score function of a log-concave density. On its own, this shape constraint is insufficient to guarantee a finite minimax risk with respect to a density-weighted L^2 norm, so we define subclasses capturing two fundamental aspects of the problem. First, we establish the crucial impact of tail behaviour by determining the minimax rate over log-concave densities whose score function exhibits controlled growth relative to the quantile levels. Second, we explore the interplay between smoothness and log-concavity by considering densities with a scale restriction and a Hölder assumption on the log-density. When the smoothness exponent is less than 2, the minimax rate is faster than under either constraint alone. Our upper bounds are attained by a locally adaptive, multiscale estimator constructed from a uniform confidence band for the score function. This study highlights intriguing differences between score and density estimation.

Title: A Single-Observation Uniformity Test via Multiscale Lacunary Harmonics

Author: Ferrari, Davide

Email: davferrari@unibz.it

Affiliation: Free University of Bolzano

Abstract: A test of uniformity on $[0,1]$ is developed for the setting of a single observation recorded with sufficient precision. Although consistency against general alternatives is not attainable with only one draw in the classical large-sample sense, a multiscale harmonic digit expansion provides a framework for structured inference. By aggregating trigonometric components across digit scales at suitably separated frequencies, a quadratic test statistic is constructed whose null distribution converges to a chi-square law via a lacunary central limit theorem. Under contiguous alternatives, the statistic admits a noncentral chi-square limit, with noncentrality determined by a Fourier projection of the alternative density. The resulting procedure detects multiscale harmonic structure that remains invisible to classical digit-frequency methods.

Title: Selective Inference in Adaptive Trials

Author: Freidling, Tobias

Email: tobias.freidling@epfl.ch

Affiliation: EPFL

Abstract: Many clinical trials follow a design with multiple stages: After each stage, the data is provisionally analysed and – based on these results – the recruitment of participants for the next stage as well as the administered treatment is chosen adaptively. For instance, we may want to exclude poorly performing drugs early or gather more samples from a certain subpopulation that shows a potentially beneficial response. Analysing such adaptive studies is challenging as the data is used twice: (1) for selection of the design of later stages and the null hypothesis, (2) for testing the null hypothesis with the data generated under the chosen design. Since the data generating mechanism and null hypothesis are not pre-specified, classical statistical methods do not provide valid inference. Existing solutions such as data splitting often lose efficiency or are specific to a certain design. In this presentation, we advocate for the use of selective inference methods to construct more powerful tests that control the selective type-I error. Randomization- or design-based inference is particularly amenable to this approach: We can provide finite-sample valid selective inference for a wide range of trial designs without any assumptions on the law of the outcomes and covariates. We show that our proposed method improves power compared to other valid tests, construct confidence intervals and address computational aspects.

Title: Niche construction emerging from fast ecological and slow evolutionary and environmental timescales in individual-based models

Author: Fritsch, Coralie

Email: coralie.fritsch@inria.fr

Affiliation: Inria

Abstract: I will introduce an individual-based model of niche construction based on birth-death processes of interacting (sub)species immersed in a slowly-varying environment which is influenced by the population state. Extinction and/or emergence of negligible (sub)species on long time scales can be observed. I will present a convergence result showing that the joint dynamics of the species sizes on a logarithmic scale and the environment can be approximated by an explicit dynamical system in the limit of large populations. This convergence result holds true under the assumption that no "bad events" occur during the population dynamics, corresponding to situations where several species become dominant or go extinct simultaneously. No such bad event occurs for Baire-almost all initial conditions of the population sizes and the environmental resources. I will apply this framework to study the long term co-existence of two specialist species consuming two resources, with a "joint" niche construction where each species constructs the niche of the other while depleting its resources. I will also present an example of immune escape in cancer, where the environmental variable is associated to the state of the immune system and the species are associated to three types of tumor cells. This is joint work with Nicolas Champagnat, Cristóbal Quiñinao, Leonardo Videla and Nicolás Zalduendo-Vidal.

Title: Reducing correlation bias of MDI importance for tree-based estimators through local sample weighting

Author: Fröhlich, Benedikt

Email: benedikt.froehlich@informatik.uni-regensburg.de

Affiliation: University of Regensburg

Abstract: Feature importance (FI) statistics provide valuable insight into the degree to which individual features contribute to the prediction of a machine learning model. However, standard metrics such as Mean Decrease in Impurity (MDI) for tree-based estimators are notoriously biased when features are not independent, attributing inflated FI to noise features that are highly correlated with signal variables. To address this issue, we propose "local sample weighting", which can be integrated locally in the split selection of tree-based models. At each node, it reweights the sample for each candidate feature with an estimated density ratio, reducing the correlation of the candidate with the other variables, such that the split can be chosen purely based on the marginal effect of the candidate. Our approach comes with a natural tuning parameter, the minimum effective sample size of the weighted population. It introduces an interpretation-prediction-tradeoff by moderating between the degree of decorrelation and the effective sample size used for estimation. We validate our approach with an extensive simulation study and support our simulation findings with theory on consistency of our model as well as asymptotic behavior of the MDI. We also investigate the relationship between the minimal effective sample size and population MDI as a key component of the interpretation-prediction-tradeoff. Part of this work is joint work with Alison Durst and Merle Behr, see arXiv: 2508.06337.

Title: Continuity of Chatterjee's rank correlation and related dependence measures

Author: Fuchs, Sebastian

Email: sebastian.fuchs@plus.ac.at

Affiliation: University of Salzburg

Abstract: While measures of concordance such as Spearman's rho and Kendall's tau are continuous with respect to weak convergence, Chatterjee's rank correlation, recently introduced by Azadkia and Chatterjee (2021), does not share this property, causing drawbacks in statistical inference as pointed out by Bücher and Dette (2025). As we show, Chatterjee's rank correlation is instead weakly continuous with respect to conditionally independent copies, namely the Markov products. To establish weak continuity of Markov products, we provide several sufficient conditions, including copula-based criteria and conditions relying on the concept of conditional weak convergence from Sweeting (1989). As a consequence, we obtain continuity results for Chatterjee's rank correlation and related dependence measures and verify their continuity in standard models such as multivariate elliptical and l1-norm symmetric distributions.

Title: Decentralised bipartite matching

Author: Ganesh, Ayalvadi

Email: a.ganesh@bristol.ac.uk

Affiliation: University of Bristol

Abstract: Modern data centres connect tens of thousands of processors. We model the scheduling of communications between these processors as a bipartite matching problem. The scale of the problem, and the very low latencies required, render traditional centralised algorithms for maximum matching unsuitable. We present a class of probabilistic matching algorithms that are decentralised and only require knowledge each node to know its 2-neighbourhood in the bipartite graph. We also present an analysis of the performance of this algorithm on a class of random graphs.

Title: Mean-Field Analysis of an Interacting Particle System on Dynamically Coevolving Graphs

Author: Ganguly, Ankan

Email: ylugnaga@gmail.com

Affiliation: Boston University

Abstract: Interacting particle systems (IPS) model collections of stochastic processes called "particles" whose interactions are governed by an underlying graph. Such models have applications throughout the social and physical sciences. Because IPS are high-dimensional stochastic processes, their large-scale limits are of interest. When the interaction graph is dense, particles in the IPS are typically asymptotically independent. This is called "propagation of chaos." Under propagation of chaos, each particle converges to a mean-field process. Previous work on mean-field limits of IPS has primarily focused on models with exogenously defined graphs. We present a simple interacting particle system with a dynamic interaction graph subject to bidirectional feedback: particle evolution depends on the underlying graph, and the graph dynamics depend on particle states. Because the graph is endogenous to the model, the mean-field limit is no longer sufficient to completely describe the limiting behavior of the IPS. We introduce a sampling perspective that yields mean-field limits for the particles, graphon limits for the network, and other limits that capture the joint distribution of the particles and graph. We show that IPS has a rich asymptotic conditional structure, motivating a general conditional propagation of chaos result that extends a famous propagation of chaos theorem by Sznitman. We finish with some numerical illustrations of our results.

Title: Some result about opinion dynamics

Author: Gantert, Nina

Email: gantert@ma.tum.de

Affiliation: Technical University of Munich

Abstract: We explain some results about opinion dynamics, more precisely about the averaging process on infinite graphs. Vertices have real-valued opinions and when they interact along edges, both vertices take the average value of the two endpoints of the edge. No prerequisites required. Based on joint work with Timo Vilkas.

Title: Some Examples on Optimal Bayesian Estimation of the Sampling Distribution

Author: García Nogales, Agustín

Email: agnogales@gmail.com

Affiliation: Universidad de Extremadura

Abstract: This paper addresses, from a Bayesian perspective, the problem of estimating the unknown probability distribution governing a random experiment and, in the dominated case, estimating its density function. It is shown that the posterior predictive distribution is the optimal solution (Bayes estimator) to the first of these problems for the squared total variation loss function. The posterior predictive density is the Bayes estimator of the density for the L1-squared loss function. The results are presented within a very general framework encompassing continuous and discrete, uni- and multivariate, parametric and non-parametric cases, without making any explicit assumptions about the prior distribution. Some examples of applications of these results are shown, two of them of a nonparametric nature.

Title: Universal diffusion limit for Markovian models of evolution in structured populations with migration

Author: Garcia Pareja, Celia

Email: celiagp@kth.se

Affiliation: KTH Royal Institute of Technology

Abstract: The evolution of microbial subpopulations that migrate within spatial structures has gained interest in recent years. Questions of relevance include, for instance, the ability of a migrant mutant to take over the population (fixate). Estimating fixation probabilities is, however, usually hindered by the lack of analytical formulas and by computational complexity of simulation-based strategies when considering large populations. In this work, we study several population genetics models where the population is divided into D subpopulations (called demes) consisting of two types of individuals, mutants and wild-types, that evolve through discrete Markovian updates. We prove that under certain assumptions all the considered models converge to the same diffusion approximation, which we call universal. This diffusion approximation is amenable to simulation strategies that underly methods of statistical inference while significantly reducing computational costs. In all models, each Markovian update follows two phases: First, a local growth phase in each subpopulation, where the growth of each type of individual depends on its fitness, and then a sampling phase that implements migration between subpopulations. Our proof relies on existing diffusion approximation results for degenerate diffusions but requires further technicalities due to fact that sample sizes in each deem are not necessarily fixed but change randomly with each update. This is joint work with Alia Abbara.

Title: Variance renormalisation of parabolic stochastic PDEs

Author: Gerencsér, Máté

Email: mate.gerencser@tuwien.ac.at

Affiliation: TU Wien

Abstract: Scaling arguments give a natural guess at the regularity condition on the noise in a stochastic PDE for a local solution theory to be possible, using the machinery of regularity structures or paracontrolled distributions. This guess of "subcriticality" is often, but not always, correct. In cases when it is not, a the blowup of the variance of certain nonlinear functionals of the noise necessitates a different, multiplicative renormalisation. We present some recent progress on this variance renormalisation, and discuss the strengths and weaknesses of different approaches such as regularity structures, Itô calculus, or stochastic sewing.

Title: Signal-to-noise ratio aware minimax analysis of sparse linear regression

Author: Ghosh, Shubhangi

Email: sg4156@columbia.edu

Affiliation: Department of Statistics, Columbia University

Abstract: We consider parameter estimation under sparse linear regression— an extensively studied problem in high-dimensional statistics and compressed sensing. While the minimax framework has been one of the most fundamental approaches for studying statistical optimality in this problem, we identify two important issues that the existing minimax analyses face: (i) The signal-to-noise ratio appears to have no effect on the minimax optimality, while it shows a major impact in numerical simulations. (ii) Estimators such as best subset selection and Lasso are shown to be minimax optimal, yet they exhibit significantly different performances in simulations. In this paper, we tackle the two issues by employing a minimax framework that accounts for variations in the signal-to-noise ratio (SNR), termed the SNR-aware minimax framework. We adopt a delicate higher-order asymptotic analysis technique to obtain the SNR-aware minimax risk. Our theoretical findings determine three distinct SNR regimes: low-SNR, medium-SNR, and high-SNR, wherein minimax optimal estimators exhibit markedly different behaviors. The new theory not only offers much better elaborations for empirical results, but also brings new insights to the estimation of sparse signals in noisy data.

Title: Inference for Change Point in Marginal Distribution of High Dimensional Time Series

Author: Ghoshal, Deep

Email: deepg2@illinois.edu

Affiliation: University of Illinois Urbana-Champaign

Abstract: This work addresses the challenge of nonparametric inference for change points in the context of marginal distribution of high-dimensional, temporally dependent Euclidean or functional data. The literature on change-point testing for marginal distribution of high-dimensional data is still limited but growing. However, in our knowledge, there is currently no theoretically valid method for this problem that handles temporal dependence; applying methods designed for independent data typically results in severe size distortion. We propose a kernel-based test that embeds the data into the Cartesian product of reproducing kernel Hilbert spaces, constructing a CUSUM-statistic via sample splitting, trimming, projection, and self-normalization techniques. For stationary and absolutely regular processes, we establish that under weak cross-sectional dependence and standard moment assumptions, the test statistic possesses a pivotal limiting null distribution as both sample size and dimension diverge. We further analyze the asymptotic power of our test in the presence of a single change point. Our theoretical framework leverages Berbee's coupling result and introduces a novel martingale approximation argument to establish a functional central limit theorem in the high-dimensional setting. Finally, we validate the finite-sample size and efficiency of our method through simulation studies.

Title: Calibrated Multi-Level Quantile Forecasting

Author: Gibbs, Isaac

Email: isaac@fprime.com

Affiliation: University of Toronto

Abstract: In this talk, I will discuss an online method for calibrating quantile forecasts at multiple levels simultaneously. We say that a sequence of quantile forecasts is calibrated if its q -level predictions are greater than or equal to the target value at a q fraction of time steps, for each level q . I will introduce a lightweight procedure, multi-level quantile tracker (MultiQT), that can wrap around any baseline sequence of point or quantile predictions to produce adjusted quantile forecasts that are guaranteed to be calibrated, even against adversarial distribution shifts. Critically, this method ensures that the quantiles remain ordered, e.g., the 0.5-level quantile forecast will never be larger than the 0.6-level forecast. Moreover, the method has a no-regret guarantee, implying it will not degrade the performance of the existing forecaster (asymptotically) with respect to the quantile loss. In experiments, we find that MultiQT significantly improves the calibration of existing real-world forecasters in epidemic and energy forecasting problems, while leaving the quantile loss largely unchanged or slightly improved.

Title: Copula-based estimation of functionals of aggregated risks

Author: Gijbels, Irène

Email: irene.gijbels@wis.kuleuven.be

Affiliation: University of Leuven (KU Leuven)

Abstract: The interest in this talk is in aggregated random variables, where the dependency between the components is modelled via copulas. We are in particular interested in a general class of functionals of such an aggregated random variable, and estimation of these functionals. The separation of dependence structure and margins allows for highly adaptable estimation procedures, accommodating parametric and nonparametric methods for both components. The generality of the framework enables its application to diverse contexts. The practical use of the developed methodology is demonstrated in real-data applications.

Title: Assessing Monotone Dependence: Area Under the Curve Meets Rank Correlation

Author: Gneiting, Tilmann

Email: tilmann.gneiting@h-its.org

Affiliation: Heidelberg Institute for Theoretical Studies

Abstract: The assessment of monotone dependence between two random variables is a classical problem in statistics and a gamut of application domains. Consequently, researchers have sought measures of association that are invariant under strictly increasing transformations of the margins, with the extant literature being splintered. Rank correlation coefficients, such as Spearman's Rho and Kendall's Tau, have been studied in the statistical literature, mostly under the assumption of continuous margins. In the case of a dichotomous outcome, receiver operating characteristic (ROC) analysis and the asymmetric area under the ROC curve (AUC) measure are used to assess monotone dependence of a binary outcome on a

covariate. Here we unify and extend thus far disconnected strands of literature, by developing common population level theory, estimators, and tests that bridge continuous and dichotomous settings and apply to all linearly ordered outcomes. In particular, we introduce the asymmetric grade correlation (AGC) and coefficient of monotone association (CMA) measures, which correspond to Spearman's Rho in the continuous case and to AUC for a dichotomous outcome. We establish central limit theorems for their sample versions and develop associated tests. In case studies, we assess progress in data-driven weather prediction and evaluate methods of uncertainty quantification for large language models. Joint work with Eva-Maria Walz and Andreas Eberl.

Title: Hierarchies of Calibration: Classification and Regression

Author: Gneiting, Tilmann

Email: tilmann.gneiting@h-its.org

Affiliation: Heidelberg Institute for Theoretical Studies

Abstract: Concepts of calibration formalize the compatibility between probabilistic predictions and the respective outcomes. In a nutshell, the outcomes ought to be indistinguishable from random draws from the predictive distributions. The talk strives to review and extend notions of calibration that have been proposed for classification and regression tasks. Particular emphasis is given to hierarchical relations between the various notions, as they apply to general real-valued, continuous, nominal, and binary outcomes, respectively. Furthermore, we discuss concepts of calibration that are expressed in terms of properties or functionals of the predictive distribution, such as means, quantiles, or event probabilities. To illustrate the applied and methodological relevance of these notions, we revisit associated decompositions of proper scoring rules and consistent scoring functions into measures of miscalibration, discrimination, and uncertainty. While calibration checks apply to out-of-sample assessments of predictive performance, they relate closely to in-sample model diagnostics, and we elucidate these connections in classification and regression settings. Joint work with Johannes Resin and Lu Yang.

Title: Principles and Flexibility in Multiple Testing

Author: Goeman, Jelle

Email: j.j.goeman@lumc.nl

Affiliation: Leiden University Medical Center

Abstract: Closed testing is known to underlie many multiple testing methods. Any method controlling a tail probability of a number or proportion of errors is either equivalent to a closed testing procedure or is uniformly improved by one. We extend closed testing to a novel necessary and sufficient principle for all multiple testing methods controlling any expected loss. This e-Closure principle asserts that every multiple testing method is a special case of a generalized closed testing procedure based on e-values. It applies to a large class of error rates — in particular to the false discovery rate (FDR). We show that any procedure controlling FDR is either a e-Closure method or is uniformly improved by one. By writing existing methods as special cases of this procedure, we show uniform improvements, as we demonstrate for the e-Benjamini-Hochberg and the Benjamini-Yekutieli procedures, and the self-consistent method of Su (2018). We also show that methods derived using our novel e-Closure Principle generally control their error rate not just for one rejected set, but simultaneously over many, allowing post hoc flexibility for the researcher. This flexibility is important in many application areas including neuroimaging and genomics. Moreover, we show that because all multiple testing methods for all error metrics are derived from the same procedure, researchers may even choose the error metric post hoc, allowing post hoc switching from e.g. familywise error to FDR.

Title: Simulation-Based Methods for Valid Inference from Privacy-Protected Data

Author: Gong, Ruobin

Email: ruobin.gong@rutgers.edu

Affiliation: Rutgers University

Abstract: Data subject to disclosure control is becoming increasingly ubiquitous in the social sciences. While such protections are essential for safeguarding the privacy of individual data subjects, privacy-protected data poses substantial challenges for statistical inference. In particular, the mechanisms used to protect privacy often occlude the underlying confidential data in highly intractable ways. This talk discusses a set of approaches that leverage simulation-based methods to enable valid statistical inference in these settings. These methods are applicable to data swapping, an important technique in traditional statistical disclosure control, as well as to modern mechanisms that satisfy differential privacy. These methods

sketch a flexible framework for principled inference when the true data generating process is obscured by probabilistic privacy protection.

Title: Causal Spatial Quantile Regression

Author: Gong, Yan

Email: yan.gong@mbzuai.ac.ae

Affiliation: MBZUAI

Abstract: Treatment effects in a wide range of economic, environmental, and epidemiological applications often vary across space, and understanding the heterogeneity of causal effects across space and outcome quantiles is a critical challenge in spatial causal inference. To effectively capture spatial heterogeneity in distributional treatment effects, we propose a novel semiparametric neural network-based causal framework leveraging deep spatial quantile regression and then construct a plug-in estimator for spatial quantile treatment effects (SQTE). This framework incorporates an efficient adjustment procedure to mitigate the impact of spatial hidden confounders. Extensive simulations across various scenarios demonstrate that our methodology can accurately estimate SQTE, even with the presence of spatial hidden confounders. Additionally, the spatial confounding adjustment procedure effectively reduces neighborhood spatial patterns in the residuals. We apply this method to assess the spatially varying quantile treatment effects of maternal smoking on newborn birth weight in North Carolina, United States. Our findings consistently show negative effects across all birth weight quantiles, with particularly severe impacts observed in the lower quantile regions.

Title: The distribution of partial sums of the Steinhaus function

Author: Gorodetsky, Ofir

Email: ofir.goro@gmail.com

Affiliation: Technion - Israel Institute of Technology

Abstract: The Steinhaus function is a random, completely multiplicative function on the positive integers, whose values on primes are iid random variables uniformly distributed on the complex unit circle. Its study is motivated by the study of classical deterministic multiplicative functions such as the Möbius function and Dirichlet characters. We shall describe recent joint work with Mo Dick Wong, where the limiting distribution of the partial sums of the Steinhaus function was determined. The limiting distribution is Gaussian with random variance; the variance is given by the total mass of a random measure. This measure is an instance of critical multiplicative chaos. In the talk we shall discuss the motivation and history behind the problem, and the intuition behind the result.

Title: E-Values obtained by Conditioning on a Sufficient Statistic are Essentially Optimal

Author: Grünwald, Peter

Email: pdg@cwi.nl

Affiliation: CWI and Leiden University

Abstract: Whenever a composite null hypothesis admits a sufficient statistic T , the conditional likelihood ratio, in which both the null and a representative of the alternative are conditioned on T provides an e-variable. We provide a precise analysis of such conditional e-variables for the case of exponential family nulls and alternatives. If T is also sufficient for the alternative, as happens in applications such as 2-sample testing, this conditional e-variable performs exceedingly well: it is asymptotically indistinguishable from the much harder to compute growth, i.e. log-optimal, e-variable under any prior W_1 on the alternative with compact support: for any such prior W_1 , the corresponding reverse information prior on the null W^* converges to the same W_1 (they induce the same densities on the mean-value parameters), an intriguing result in its own right. Further, the likelihood ratio of the Bayes marginal distributions with priors W_1 and W^* converges to the conditional e-variable. As such, the conditional e-variable asymptotically achieves uniform growth optimality - a novel concept that is stronger than both absolute and relative growth optimality. We also briefly highlight how to extend the definition and analysis if the alternative is of larger dimensionality than the null (this includes GLMs) and in more general cases in which the null is not an exponential family, yet simple sufficient statistics exist. Joint work with Dante de Roos, Sebastian Arnold and Yunda Hao.

Title: Distributionally robust risk evaluation with an isotonic constraint

Author: Gui, Yu

Email: yugui@wharton.upenn.edu

Affiliation: University of Pennsylvania

Abstract: Statistical learning under distribution shift is challenging when neither prior knowledge nor fully accessible data from the target distribution is available. Distributionally robust learning (DRL) aims to control the worst-case statistical performance within an uncertainty set of candidate distributions, but how to properly specify the set remains challenging. To enable distributional robustness without being overly conservative, in this paper we propose a shape-constrained approach to DRL, which incorporates prior information about the way in which the unknown target distribution differs from its estimate. More specifically, we assume the unknown density ratio between the target distribution and its estimate is isotonic with respect to some partial order. At the population level, we provide a solution to the shape-constrained optimization problem that does not involve the isotonic constraint. At the sample level, we provide consistency results for an empirical estimator of the target in a range of different settings. Empirical studies on both synthetic and real data examples demonstrate the improved accuracy of the proposed shape-constrained approach.

Title: Mean field particles system : propagation of chaos and concentration

Author: GUILLIN, Arnaud

Email: arnaud.guillin@uca.fr

Affiliation: Université Clermont Auvergne

Abstract: We will review recent progress for sharp propagation of chaos for mean field particles system, in particular showing new results stated in terms of Fisher information. We will also prove concentration inequalities via autorefinement of Poincaré inequalities, in particular for mean field particles system. In collaboration with Jules Grass, Christophe Poquet, Boris Nectoux and Liming Wu.

Title: Rank-restricted Wald tests for goodness-of-fit testing

Author: Guo, Hanke

Email: hanke.guo@tum.de

Affiliation: Technical University of Munich / Munich Center for Machine Learning (MCML)

Abstract: We present a modified version of a Wald test that allows one to test multiple constraints in the case where the estimated restrictions have an asymptotic covariance matrix that is singular. In such settings, it is tempting to replace the matrix inverse in the construction of the test statistic by the Moore-Penrose pseudoinverse. However, this approach does not provide a valid test for problems in which the rank of the covariance matrix estimate may exceed the asymptotic rank. To address this issue, we propose the use of a rank-restricted pseudoinverse, defined for a specified fixed rank. For this construction, it is not necessary to know the exact rank of the asymptotic covariance matrix, which can be difficult to determine when testing complex sets of constraints. We then clarify under which conditions the rank-restricted Wald statistics admit an asymptotic chi-square distribution. As a motivating example, we consider goodness-of-fit testing for Independent Component Analysis (ICA) models. In ICA, one observes a collection of linear combinations of independent and non-Gaussian source variables. This setting, which also encompasses linear non-Gaussian structural equation models (LiNGAM), leads to algebraic relations among second and third moments, which we test via the proposed rank-restricted Wald tests.

Title: Statistical Analysis of Conditional Group Distributionally Robust Optimization with Cross-Entropy Loss

Author: Guo, Zijian

Email: zijguo@zju.edu.cn

Affiliation: Zhejiang University

Abstract: In multi-source learning with discrete labels, distributional heterogeneity across domains poses a central challenge to developing predictive models that transfer reliably to unseen domains. We study multi-source unsupervised domain adaptation, where labeled data are available from multiple source domains and only unlabeled data are observed from the target domain. To address potential distribution shifts, we propose a novel Conditional Group Distributionally Robust Optimization (CG-DRO) framework that learns a classifier by minimizing the worst-case cross-entropy loss over the convex combinations of the conditional outcome distributions from sources domains. We develop an efficient Mirror Prox algorithm for solving the minimax problem and employ a double machine learning procedure to estimate the risk function, ensuring that errors in nuisance estimation contribute only at higher-order rates. We establish fast statistical convergence rates for the empirical CG-DRO estimator by constructing two surrogate minimax optimization problems that serve as theoretical bridges. A distinguishing challenge for CG-DRO is the emergence of nonstandard asymptotics: the empirical CG-DRO estimator may fail to converge to a standard limiting distribution due to boundary effects and system instability. To address this, we introduce

a perturbation-based inference procedure that enables uniformly valid inference, including confidence interval construction and hypothe

Title: Goggin's Corrected Kalman Filter: Guarantees and Filtering Regimes

Author: Gurvich, Itai

Email: i-gurvich@kellogg.northwestern.edu

Affiliation: Northwestern University

Abstract: We revisit a nonlinear filter for linear state-space models with non-Gaussian signal and observation noise, introduced by Eimear Goggin in 1992. Goggin showed that applying the Kalman filter (KF) to score-transformed observations is asymptotically optimal in a specific signal-to-noise ratio (SNR) regime. We provide non-asymptotic convergence rates across a broad range of SNR regimes, recovering Goggin's setting as a special case, with bounds explicit in the SNR. Our analysis combines two ingredients: a posterior Cramér-Rao lower bound for filtering and convergence-rate bounds in the Fisher information central limit theorem. We also map the parameters space into filtering regimes, identify degenerate regimes—where simple filters are nearly optimal—and isolate a balanced regime, where Goggin's filter has the most value.

Title: Conditions for Transitivity of Win Ratios, Win Proportions, and Hazard Ratios

Author: Györfy-Kerekes, Anna

Email: akerekes@ethz.ch

Affiliation: ETHZ and Max Planck ETH Center for Learning Systems

Abstract: Head-to-head comparisons, such as win proportions (WP), win ratios (WR), and hazard ratios (HR), are fundamental for evaluating efficacy in randomized controlled trials (RCTs). However, like many pairwise comparisons, they are prone to Efron-type non-transitivity. This statistical phenomenon occurs when long-run comparisons violate transitivity—for instance, where Treatment A outperforms B, B outperforms C, yet C outperforms A. In clinical research, such paradoxes can complicate decision-making and guideline development. Our work explores the scope of Efron-type non-transitivity across common clinical metrics. Specifically, we aim to identify the exact statistical conditions necessary to preclude non-transitivity for WR, WP, and HR, thereby clarifying their appropriate utility in RCTs. We employ a variety of analytical techniques to rule out non-transitivity under different scenarios. First, we explore the impact of specific distributional assumptions on the data. Second, we apply Trybula's theorem to establish conditions where non-transitivity can be ruled out purely by examining the magnitudes of the pairwise comparisons. Finally, we evaluate how the classic proportional hazards assumption is a natural safeguard against non-transitive loops. By establishing these mathematical guardrails, we provide a clearer framework for selecting and interpreting pairwise metrics, ensuring that clinical efficacy conclusions remain consistent when comparing multiple treatments.

Title: Conditional independence testing via Azadkia-Chatterjee's graph correlation

Author: Han, Fang

Email: fanghan@uw.edu

Affiliation: University of Washington, Seattle

Abstract: Since Azadkia and Chatterjee introduced their nearest-neighbor-based graph correlation as a measure of conditional dependence, an important open question has been how to use it for statistically valid testing of conditional independence. In this talk, I will present, for the first time, a complete roadmap for accomplishing this. The key ingredients are (1) joint asymptotic normality of the graph correlation under arbitrary dependence, (2) an exact characterization of the limiting variance, and (3) a careful bias analysis and correction building on Azadkia, Chen, and Han (2025).

Title: Algorithmic inference via nonconvex gradient descent

Author: Han, Qiyang

Email: qh85@stat.rutgers.edu

Affiliation: Rutgers University

Abstract: Conventional statistical inference methods are typically developed for models simple enough to admit tractable estimators through carefully designed iterative algorithms. In contrast, modern learning architectures are enormously complex, yet are trained by simple gradient-descent-type algorithms, often without any provable guarantee of algorithmic convergence to global/local optima. Can we reconcile classical inference principles with these highly complicated, modern learning paradigms? In this talk, we will present a new inference framework addressing this question, by showing that valid

statistical inference can be performed along the entire gradient descent trajectory, iteration by iteration, without requiring convexity of the loss landscape or convergence of the algorithm. To illustrate this concept, we begin with a single-index regression model and demonstrate how gradient descent iterates can be "debiased", at each iteration, to yield valid confidence intervals for the underlying signal and consistent estimates of generalization errors. We then extend this paradigm to the much more challenging setting of learning with general multi-layer neural networks in their full complexity, where the loss landscape can be arbitrarily complex. Crucially, the proposed method remains valid without requiring either algorithmic convergence or oracle knowledge of the unknowns, and may therefore inform practical decisions such as early stopping and hyperparameter tuning.

Title: Generative Fiducial Models

Author: Hannig, Jan

Email: jan.hannig@unc.edu

Affiliation: University of North Carolina at Chapel Hill

Abstract: While generalized fiducial inference (GFI) and its variants have yielded many theoretical and practical results to parametric inference and uncertainty quantification, applying it to generative models remains challenging. We identify three key issues misspecification, metric choices, and over-parameterization hinder the direct application of the GFI to generative models. In this paper, we propose a novel method based on the framework of generalized fiducial inference, designed to construct distributional estimates over the parameter space given observed data, while also enabling uncertainty quantification for generative models. We employ a truncation-based approach and further provide a theoretical analysis of its behavior under varying truncation parameters. Both theoretical results and empirical evidence suggest that, with an appropriately chosen truncation parameter, the truncated distribution derived from generalized fiducial inference achieves valid coverage of the true parameter and leads to improved generalization performance. Joint Work with Zijie Tian, T. C. M. Lee (UC Davis).

Title: Planar percolation and the loop $O(n)$ model

Author: Harel, Matan

Email: m.harel@northeastern.edu

Affiliation: Northeastern University

Abstract: Consider a tail trivial, positively associated site percolation process such that the set of open vertices is stochastically dominated by the set of closed ones. We show that, for any planar graph G , such a process must contain zero or infinitely many infinite connected components. The assumptions cover Bernoulli site percolation at parameter p less than or equal to one half, resolving a conjecture of Benjamini and Schramm. As a corollary, we prove that p_c is greater than or equal to $1/2$ for any unimodular, invariantly amenable planar graphs. We will then apply this percolation statement to the loop $O(n)$ model on the hexagonal lattice, and show that, whenever n is between 1 and 2 and x is between $1/\sqrt{2}$ and 1, the model exhibits infinitely many loops surrounding every face of the lattice, giving strong evidence for conformally invariant behavior in the scaling limit (as conjectured by Nienhuis). This is joint work with Alexander Glazman (University of Innsbruck) and Nathan Zelesko (Northeastern University).

Title: Analysis of Multiple Outcomes in Contaminated Trials Reinforced With Validation Data

Author: Harrar, Solomon

Email: solomon.harrar@uky.edu

Affiliation: University of Kentucky

Abstract: This paper is concerned with estimation and testing for treatment effects with multivariate outcomes. It primarily focuses on the situation where imperfect diagnostic tools are used to classify subjects into different groups. Oftentimes, there are more expensive and/or invasive diagnostic tools to accurately determine the subjects' status or conditions, yielding partially validated data on a smaller number of subjects. We propose moment-based approaches for estimating and testing treatment effects. We compare our methods with maximum likelihood approach using the EM algorithm, which requires strong assumptions and bears computational burden, and with traditional methods, which ignore the diagnostic tool's imperfection. The proposed methods show advantages in terms of coverage probability, computations efficiency, and robustness. The application of the methods is illustrated with gene-expression data from the Genes-environments & Admixture in Latino Americans (GALA) II study of asthma in Hispanic/Latino children.

Title: Boosting methods for interval censored data with regression and classification

Author: He, Wenqing

Email: whe@stats.uwo.ca

Affiliation: University of Western Ontario

Abstract: Boosting has gained significant interest across both machine learning and statistical communities. Traditional boosting algorithms, designed for fully observed random samples, often struggle with real-world problems, particularly with interval censored data. This type of data is common in survival analysis where exact event times are unobserved but fall within known intervals. Effective handling of such data is crucial in fields like medical research, reliability engineering, and social sciences. In this work, we introduce novel nonparametric boosting methods for regression and classification tasks with interval censored data. Our approaches leverage censoring unbiased transformations to adjust loss functions and impute transformed responses while maintaining model accuracy. Implemented via functional gradient descent, these methods ensure scalability and adaptability. We rigorously establish their theoretical properties. Our proposed methods not only offer a robust framework for enhancing predictive accuracy in domains where interval censored data are common but also complement existing work, expanding the applicability of existing boosting techniques. Empirical studies demonstrate robust performance across various finite-sample scenarios, highlighting the practical utility of our approaches.

Title: Infinite-dimensional information geometry for semiparametric statistics

Author: Hemmi, Masayuki

Email: henmi@ism.ac.jp

Affiliation: The Institute of Statistical Mathematics

Abstract: In information geometry, a parametric model (a parametric family of probability distributions) is treated as a finite-dimensional differentiable manifold, where the Riemannian metric called Fisher metric and the pair of two dual affine connections called the exponential and mixture connections play essential roles for geometrical consideration of statistical inference. For example, the maximum likelihood estimation in an exponential family is seen as the orthogonal projection of the geodesic defined by the mixture connection and its statistical properties can be understood in terms of geometry. However, semi-parametric statistical methods cannot be understood geometrically in this framework because the statistical models behind these methods are essentially infinite-dimensional. Since Pistone and Sempi (1995) introduced an infinite-dimensional manifold structure of probability distributions using the notion of Orlicz space, the geometrical theory of infinite-dimensional statistical models has been developed from a mathematical point of view. However, it has not sufficiently provided a geometrical understanding of semi-parametric statistical methods yet. The aim of this work is to reconsider the construction of infinite-dimensional manifold structure of probability distributions and to develop geometrical tools for geometrical understanding of statistical inference so that it may provide an extension of the semiparametric estimation theory.

Title: Critical behaviour for weight-dependent random connection models

Author: Heydenreich, Markus

Email: markus.heydenreich@uni-a.de

Affiliation: University of Augsburg

Abstract: The random connection model is a versatile model encompassing various variants of continuum percolation: vertices are given as point process in a suitable metric space, and edges are drawn independently with probability depending on the distance of the endpoints and possible further vertex parameters known as weights. We introduce the model and present results for the critical percolation phase transition in high-dimensional Euclidean space and in hyperbolic space.

Title: Universal prediction with extended ranks

Author: Hoff, Peter

Email: peter.hoff@duke.edu

Affiliation: Duke University

Abstract: The accuracy of predictions derived from a regression model depends on the extent to which the model is able to mimic certain aspects of the conditional distribution of an outcome given its features. Primary among these are the distributional support of the outcome variable, and the relationship between the mean and variance of the conditional distribution. In this talk we develop a single inferential framework that can represent these aspects for a wide range of conditional distributions. Our approach is based on an unknown but arbitrary non-decreasing transformation of a linear

regression model, using a pseudo-likelihood based on a partial ordering of the observed outcomes. This transformation model can accommodate any ordinal data type, including continuous and discrete ordered data, and can represent a wide range of mean-variance relationships. Theoretical results show that the pseudo-likelihood is asymptotically efficient relative to an appropriate full likelihood at the two extremes of ordinal data – continuous and binary data. Parameter estimates and predictive inference are available via a simple Gibbs sampling algorithm and conformal calibration.

Title: Rescaled SIR epidemic processes converge to super-Brownian motion in four or more dimensions

Author: Hong, Jieliang

Email: hongjl@sustech.edu.cn

Affiliation: Southern University of Science and Technology

Abstract: In dimensions four or more, by choosing a suitable scaling parameter, we show that the rescaled spatial SIR epidemic process converges to a super-Brownian motion with drift, thus complementing the previous results by Lalley (Prob. Th. Rel. Fields, 144 (2009), 429--469) and Lalley-Zheng (Prob. Th. Rel. Fields, 148 (2010), 527--566) on the convergence of SIR epidemics in dimensions three or less. The scaling parameters we choose also agree with the corresponding asymptotics for the critical probability p_c of the range-R bond percolation on Z^d as R goes to infinity.

Title: Measuring dependence between a categorical response and a functional covariate

Author: Hörmann, Siegfried

Email: shoermann@tugraz.at

Affiliation: Graz University of Technology

Abstract: We suggest a dependence coefficient between a categorical variable and some general variable taking values in a metric space. We derive important theoretical properties and study the large sample behaviour of our suggested estimator. Moreover, we develop an independence test which has an asymptotic chi-square distribution if the variables are independent and prove that this test is consistent against any violation of independence. The test is also applicable to the classical K-sample problem with possibly high- or infinite-dimensional distributions. We discuss some extensions, including a variant of the coefficient for measuring conditional dependence.

Title: Continuous Normalizing Flows for Statistical Generative Learning

Author: Huang, Jian

Email: j.huang@polyu.edu.hk

Affiliation: The Hong Kong Polytechnic University

Abstract: Continuous Normalizing Flows (CNFs) have emerged as a powerful class of generative models, distinguished by their capacity for their high-fidelity sample generation. By leveraging neural ordinary differential equations, CNFs construct a continuous-time, invertible mapping that smoothly transforms a tractable base distribution, such as standard Gaussian, into a complex, high-dimensional target data distribution. In this talk, we will explore the theoretical foundations underpinning CNFs. Building on these principles, we will examine the versatility of this framework across a broad spectrum of modern statistical and machine learning tasks. Specifically, we will highlight how the exact invertibility and tractable density evaluation inherent to CNFs can be uniquely leveraged to characterize conditional independence, advance counterfactual estimation in causal inference, provide rigorous uncertainty quantification in conformal prediction, and enable dynamic trajectory modeling in motion generation.

Title: SGD Convergence for GLMs via Local Curvature: A Concentration of Measure Approach

Author: Hyland, Peter

Email: p.hyland24@imperial.ac.uk

Affiliation: Imperial College London

Abstract: We analyse the convergence rates of stochastic gradient descent (SGD) to a maximum likelihood estimator for the parameters of a generalised linear model. Many classical convergence results for SGD require strong convexity of the loss function, which is not satisfied for GLMs, even with the canonical link. Despite this, empirical experience suggests that SGD converges at a rate expected for a strongly convex loss close to the optimum. We provide convergence guarantees for this behaviour in wide range of settings, including logistic regression, with bounded or sub-Gaussian covariates. By exploiting concentration of measure results and exit time calculations, we show that SGD adapts to the local curvature close to the optimum with high probability.

Title: A Bayesian Dynamic Latent Space Model for Weighted Networks

Author: Iacopini, Matteo

Email: miacopini@luiss.it

Affiliation: Luiss University of Rome

Abstract: A new dynamic latent space eigenmodel (LSM) is proposed for time-series of weighted networks. The model accounts for integer-valued weights with excess zeros, time-varying node positions (features), and time-varying sparsity probabilities. The latent positions evolve according to a vector autoregressive process that accounts for possible lagged and contemporaneous dependence across nodes and features, which is neglected in the LSM literature. A Bayesian approach is used to address two of the primary sources of inference intractability in dynamic LSMs: latent feature estimation and the choice of latent space dimension. Regarding the first task, we employ an efficient auxiliary-mixture sampler that performs data augmentation and supports conditionally conjugate prior distributions. A point-process representation of the network weights and the finite-dimensional distribution of the latent processes are used to derive a multi-move sampler in which each feature trajectory is drawn in a single block, without recursions. This sampling strategy is new to the network literature and can significantly reduce computational time while improving the mixing of the chain. To avoid trans-dimensional samplers, a Laplace approximation to a partial marginal likelihood is used. Overall, our partially collapsed Gibbs sampler is general, as it can be easily extended to static and dynamic settings, as well as to discrete or continuous network weights.

Title: How does limma-trend work? An empirical partially Bayes perspective

Author: Ignatiadis, Nikolaos

Email: ignat@uchicago.edu

Affiliation: University of Chicago

Abstract: In high-throughput biology, it is common to fit thousands of linear regressions---one per gene, protein, or other unit---with very few samples per unit. Limma-trend, one of the most widely used methods in this setting, improves power by shrinking variance estimates toward an estimated mean-variance trend before computing p-values and applying the Benjamini-Hochberg procedure to control the false discovery rate (FDR). In this work, we view limma-trend through the lens of empirical partially Bayes inference, a paradigm in which a prior is posited and estimated for the nuisance parameters while parameters of interest remain fixed. From this perspective, we derive formal statistical justification for a nonparametric generalization of limma-trend. We show that this approach asymptotically controls the FDR even when the mean-variance trend is misspecified or inconsistently estimated. We also develop a new empirical partially Bayes procedure that gains power by learning the full conditional distribution of variances rather than just a trend, and illustrate our results on RNA-seq, proteomics, and ChIP-seq data.

Title: Provable Recovery of Locally Important Signed Features and Interactions from Random Forest

Author: Ihlo, Nicolas Alexander

Email: Nicolas.Ihlo@informatik.uni-regensburg.de

Affiliation: Faculty of Informatics and Data Science, University of Regensburg

Abstract: Feature and Interaction Importance (FII) methods are essential in supervised learning for assessing the relevance of input variables and their interactions in complex prediction models. In many domains, such as personalized medicine, global scores summarizing overall feature importance are insufficient; instead, local interpretations for individual predictions are essential. Random Forests (RFs) are widely used in these settings, and existing interpretability methods typically exploit tree structures and split statistics to provide model-specific insights. However, theoretical understanding of local FII methods for RF remains limited, making it unclear how to interpret high importance scores for individual predictions. We propose a novel, local, model-specific FII method that identifies frequent co-occurrences of features along decision paths, combining global patterns with those observed on paths specific to a given test point. We prove that our method consistently recovers the true local signal features and their interactions under a Locally Spike Sparse (LSS) model and also identifies whether large or small feature values drive a prediction. We illustrate the usefulness of our method and theoretical results through simulation studies and a real-world data example. This is joint work with Kata Vuk and Merle Behr, see preprint <https://arxiv.org/abs/2512.11081>.

Title: Local Fréchet Regression with Riemannian Manifold Predictors

Author: Im, Changjun

Email: icj803@snu.ac.kr

Affiliation: Seoul National University

Abstract: Fréchet regression has become a central tool in object-oriented data analysis, allowing responses to take values in general metric spaces. Existing local Fréchet regression theory has largely been developed for Euclidean predictors, although many data sets involve covariates with intrinsic geometric structure, such as directions, shapes, or manifold-valued features. We develop a local Fréchet regression framework for metric-space-valued responses with predictors on a Riemannian manifold. Our approach defines intrinsic local constant and local linear estimators through normal coordinates, logarithmic maps, and the Riemannian volume density, thereby respecting the geometry of the predictor space rather than relying on an extrinsic Euclidean representation. We establish pointwise and uniform convergence rates for the proposed estimators under regularity conditions on the predictor geometry, the local design, and the Fréchet objective. The rates are governed by the intrinsic dimension of the manifold and reduce to the standard Euclidean rates in the flat case. The theory extends local Fréchet regression to genuinely non-Euclidean predictor spaces and provides tools for analyzing complex random objects indexed by manifold-valued covariates.

Title: Maximum information gain in kernel ridge regression

Author: Janz, David

Email: david.janz93@gmail.com

Affiliation: University of Oxford

Abstract: We establish matching lower and upper bounds for the maximum information gain (log-determinant of the regularised Gram matrix) for kernel ridge regression with covariates in the unit hypercube for families of kernels including the Matérn and squared exponential. Our results are based on elementary methods and avoid the strong and unverified assumption of uniformly bounded Mercer eigenfunctions used in earlier work.

Title: On the number of particles in Activated Random Walk and the Abelian sandpile.

Author: Jarai, Antal A.

Email: A.Jarai@bath.ac.uk

Affiliation: University of Bath

Abstract: We consider the number of particles in two different models of self-organised criticality: the Activated Random Walk (ARW) and the Abelian sandpile model (ASM). For the driven-dissipative ARW on the complete graph with N vertices and one sink and sleep rate λ , we show that the stationary number of particles is $\rho(\lambda) N + a(\lambda) \sqrt{N \log N} (1 + o(1))$, for simple explicit functions ρ and a . For the driven-dissipative ASM on a lattice approximation of a 2D domain D , with lattice scaled by ϵ , we show that the stationary probability that a site at z in D is vacant scales as $p(0) + c f(z) \epsilon^2 + o(\epsilon^2)$, where $p(0)$ is the vacancy probability on the infinite lattice, $c > 0$ is a lattice dependent constant, and $f(z) > 0$ is a conformally covariant function on D of scale dimension 2. We also show that the symmetry assumptions we impose on the lattice cannot be dropped, that is, without them the formula does not necessarily hold. Joint works with C. Moench & L. Taggi, and M.W. Elvidge, respectively.

Title: Dynamic Networks with Node Heterogeneity and Homophily

Author: Jiang, Binyan

Email: by.jiang@polyu.edu.hk

Affiliation: The Hong Kong Polytechnic University

Abstract: The goal of this work is to model node heterogeneity and link homophily for dynamic networks. The proposed framework brings new insights on how networks evolve over time. It also provides more sophisticated tools for the prediction of future networks with statistical guarantees. The new model accounts for the link homophily associated with both observed traits and latent traits. The joint modeling of node heterogeneity and both observed and latent homophily effects also poses the significant challenge in statistical inference, resulted from the large number of confounding parameters in the model. To overcome this, we propose a novel normalized squared loss, paving the way for estimating parameters stably in a high-dimensional setting. We provide a rigorous theoretical analysis of the estimation method, and demonstrate its effectiveness through extensive simulations and the illustration with some real-world network data.

Title: Functional Calibration for Survival Prediction under Right Censoring

Author: Jonkers, Jef

Email: jef.jonkers@ugent.be

Affiliation: Ghent University

Abstract: Many survival models produce full predictive distributions, yet calibration is often assessed only through fixed-horizon summaries. We develop a general framework for functional calibration under right censoring, focusing on whether decision-relevant summaries of a predicted survival distribution, such as horizon-specific risks, quantiles, expectiles, or moments, are reliable conditional on their issued values. Using inverse-probability-of-censoring weighting, we show how these calibration targets can be identified from censored data. We then introduce censoring-aware functional reliability diagrams based on a weighted H-pool-adjacent-violators algorithm, yielding monotone empirical recalibration maps for identifiable forecast functionals. The same construction induces score decompositions into miscalibration, discrimination, and uncertainty, extending Murphy-type decompositions to the right-censored setting. We also study large-sample behavior in discrete- and continuous-support regimes, providing a basis for uncertainty bands around the reliability diagrams. The framework unifies several existing survival calibration tools and provides a flexible way to diagnose where and how predictive survival distributions are miscalibrated under censoring.

Title: Nonparametric hazard rate estimation with associated kernels: theory and application to a two-phase aging model

Author: Kaakai, Sarah

Email: kaakai@math.univ-paris13.fr

Affiliation: LAGA, Université Sorbonne Paris Nord

Abstract: In this talk, I will present a general theoretical framework for the nonparametric estimation of the intensity of random events, using so-called associated kernels, whose shape adapts locally to the estimation point. One of the main advantages of these kernels is that they correct the well-known boundary bias of classical kernel estimators. I will present several results on the convergence of the associated kernel hazard rate estimator, as well as theoretical guarantees on a minimax bandwidth selection method in the spirit of Goldenshluger and Lepski (2011). I will then apply these results to a two-phase aging model, using experimental data on *Drosophila*, obtained via the so-called "Smurf" phenotype. This phenotype, which measures intestinal permeability, is an important marker of aging in several species. I will show how the estimation of transition and mortality rates using associated kernels sheds new light on the biological mechanisms underlying the aging process, and lay the groundwork for the study of aging in natural populations. Joint work with L. Breuil, M. Doumic and M. Rera.

Title: BSNMANI: Bayesian scalar-on-network regression with manifold learning

Author: Kang, Jian

Email: jiankang@umich.edu

Affiliation: University of Michigan

Abstract: Brain connectivity analysis is crucial for understanding brain structure and neurological function, shedding light on the mechanisms of mental illness. To study the association between individual brain connectivity networks and the clinical characteristics, we develop BSNMani: a Bayesian scalar-on-network regression model with manifold learning. BSNMani comprises two components: the network manifold learning model for brain connectivity networks, which extracts shared connectivity structures and subject-specific network features, and the joint predictive model for clinical outcomes, which studies the association between clinical phenotypes and subject-specific network features while adjusting for potential confounding covariates. For posterior computation, we develop a novel two-stage hybrid algorithm combining Metropolis-Adjusted Langevin Algorithm (MALA) and Gibbs sampling. Our method is not only able to extract meaningful subnetwork features that reveal shared connectivity patterns but can also reveal their association with clinical phenotypes, further enabling clinical outcome prediction. We demonstrate our method through simulations and through its application to real restingstate fMRI data from a study focusing on Major Depressive Disorder (MDD). Our approach sheds light on the intricate interplay between brain connectivity and clinical features, offering insights that can contribute to our understanding of psychiatric and neurological disorders as well as mental health.

Title: Multiview Functional Clustering Using Latent Representations of Phase and Amplitude Components

Author: Kang, Seungwoo

Email: kangsw@skku.edu

Affiliation: Sungkyunkwan University

Abstract: Functional data often exhibit two distinct sources of variability: amplitude variation in function values and phase variation arising from domain misalignment. While many preprocessing methods have been proposed to separate these two types of variation, their role in functional clustering has not been fully understood from a probabilistic perspective. Moreover, modeling the phase component within the usual functional data geometry fails to capture its intrinsic structural

characteristics. In this work, we propose multiview functional k-means and k-medians clustering methods that explicitly incorporate both amplitude and phase variation into the clustering procedure. Within this framework, clustering is performed on latent representations consisting of both variations, allowing different views of the data to be explored by varying the relative contributions of phase and amplitude variation. With particular attention to the geometric structure of the phase component, the latent representations are defined in a separable Hilbert space, which enables theoretical guarantees under a principled probabilistic model. Simulation studies and real data analyses demonstrate that the proposed methods achieve improved clustering performance and provide multiview insights into the structure of functional data.

Title: Optimal Inference with Black-box Predictions

Author: Kania, Lucas

Email: lucaskania@cmu.edu

Affiliation: Carnegie Mellon University

Abstract: Powerful black-box predictive models have motivated many proposals for combining observed data with predictions to perform valid statistical inference, including prediction-powered inference, dimension-agnostic inference, and universal inference. Despite this progress, the field lacks a unifying principle that explains how hypothesis tests should integrate data and predictions in a way that is both valid and efficient. In this talk, I address this gap by characterizing the information-theoretic limits of inference with black-box predictions in the high-dimensional Gaussian sequence model. Building on this characterization, I develop practical hypothesis tests that adapt optimally to the unknown accuracy of the predictions. Finally, I discuss how these results provide insight into the strengths and limitations of existing methodologies, and how the proposed methodology extends to nonparametric settings. This is joint work with Abhinav Chakraborty, Edward Kennedy, Larry Wasserman, and Sivaraman Balakrishnan.

Title: Robust and Efficient Domain Adaptation via ϕ -Divergences

Author: Kepplinger, David

Email: dkepplin@gmu.edu

Affiliation: George Mason University

Abstract: Machine learning models often perform unreliably when the distribution of the training (source) data differs from the test (target) data. Domain adaptation (DA) addresses this challenge by augmenting the training process to transfer knowledge across mismatched domains. However, most conventional DA approaches rely on strict, unverifiable assumptions—such as label or covariate shift—and fail to account for distributional irregularities like outliers and model misspecification, which can severely impact performance. We propose a principled and robust DA framework based on Distributionally Robust Optimization (DRO) using ϕ -divergences. Our framework bypasses restrictive assumptions and naturally handles misspecification and outliers in both the source and target domains. We establish strong theoretical guarantees on the learned model's performance, deriving generalization bounds under mild conditions. We further show how varying levels of information from the target domain (e.g., no data, or labeled/unlabeled data) seamlessly integrate into our framework and theory. Finally, we demonstrate the advantages of our method through numerical experiments and real-world applications.

Title: Temporal connectivity of random geometric graphs

Author: Kerriou, Céline

Email: ckerriou@math.uni-koeln.de

Affiliation: University of Cologne

Abstract: A temporal random geometric graph is a random geometric graph in which all edges are endowed with a uniformly random time-stamp, representing the time of interaction between vertices. In such graphs, paths with increasing time stamps indicate the propagation of information. We determine a threshold for the existence of monotone increasing paths between all pairs of vertices in temporal random geometric graphs. The results reveal that temporal connectivity appears at a significantly larger edge density than simple connectivity of the underlying random geometric graph. This is in contrast with Erdős-Rényi random graphs in which the thresholds for temporal connectivity and simple connectivity are of the same order of magnitude. This talk is based on joint work with Anna Brandenberger, Serte Donderwinkel, Gábor Lugosi and Rivka Mitchell.

Title: Avoiding the Price of Adaptivity: Inference in Linear Contextual Bandits via Stability

Author: Khamaru, Koulik

Email: kk1241@stat.rutgers.edu

Affiliation: Assistant Professor

Abstract: Statistical inference under adaptive data collection is challenging due to the non-i.i.d. nature of the data. In particular, classical least-squares inference can fail under contextual bandit sampling, and constructing valid confidence intervals for linear functionals typically requires an unavoidable inflation of order square-root of $d \log T$ —where d is the dimension and T is the number of samples—the so-called price of adaptivity. The Lai–Wei stability condition, which requires the empirical feature covariance to concentrate around a deterministic limit, is known to circumvent this limitation. When it holds, the OLS estimator satisfies a CLT and standard Wald-type confidence intervals remain asymptotically valid without the need to inflate the confidence intervals. In this talk, we will discuss a regularized EXP4 algorithm for linear contextual bandits that provably satisfies the Lai–Wei stability condition, thereby enabling valid Wald-type inference with quantitative CLT rates. We will also show that the same algorithm achieves minimax-optimal regret up to logarithmic factors, demonstrating that stability and statistical efficiency can coexist. As an application, we will discuss confidence intervals for the conditional average treatment effect (CATE) under adaptively collected data, and present simulations illustrating the empirical normality of the resulting estimators.

Title: The weak-feature-impact phase transition of the NPMLE in monotone binary regression

Author: Kieffer, Dario

Email: dario.kieffer@stochastik.uni-freiburg.de

Affiliation: Albert-Ludwigs-Universität Freiburg

Abstract: The nonparametric maximum likelihood estimator (NPMLE) in monotone binary regression models is considered here when the impact of the features on the labels is weak, where weakness is colloquially understood as "close to flatness" of the feature-label relationship $x \rightarrow P(Y=1|X=x)$. Statistical literature provides limiting distributions of the NPMLE for the two extremal cases: If the feature-label relation is strictly monotone and sufficiently smooth, then it converges at a nonparametric rate pointwise and in L^1 with scaled Chernoff-type and Gaussian limiting distribution, respectively, and it converges at the parametric $n^{1/2}$ -rate if the underlying relation is flat. To explore the distributional transition of the NPMLE from the nonparametric to the parametric regime, we introduce a novel mathematical scenario. New restricted minimax lower bounds and matching pointwise and L^1 -rates of convergence of the NPMLE in the weak-feature-impact scenario together with corresponding limiting distributions are derived. They are shown to exhibit an elbow and a phase transition, solely characterized by the level of feature impact.

Title: Extended Fiducial Inference for Individual Treatment Effects via Deep Neural Networks

Author: Kim, Sehwan

Email: sehwankim@ewha.ac.kr

Affiliation: Ewha Womans University

Abstract: Fiducial inference was long regarded as a failure, yet Fisher’s original goal—quantifying parameter uncertainty directly from observations—remains fundamental in statistics. We develop extended fiducial inference (EFI), a scalable framework that revisits this goal using modern statistical computing. EFI jointly imputes latent random errors via stochastic gradient Markov chain Monte Carlo and estimates the inverse map from observations to parameters using a sparse deep neural network (DNN). The consistency of the sparse DNN ensures proper propagation of observational uncertainty to parameter uncertainty, yielding downstream inference. Compared with frequentist and Bayesian approaches, EFI is more robust in parameter estimation, especially under outliers, and avoids reference distributions in hypothesis testing, moving toward automated inference. EFI also provides a framework for semi-supervised learning. As an application, we propose a Double Neural Network (Double-NN) method for individual treatment effect estimation within EFI. Separate DNNs model treatment and control response functions, while an additional network estimates their parameters. The method is broadly applicable by universal approximation theory and empirically outperforms conformal quantile regression. Theoretically, it enables uncertainty quantification in models whose complexity grows with sample size, and provides a rigorous framework for uncertainty quantification of DNNs under neural scaling laws.

Title: Long-Range Order in the Monomer Double-Dimer Model with Long-Range Interactions

Author: Klippel, Andreas

Email: klippel@mathematik.tu-darmstadt.de

Affiliation: TU Darmstadt

Abstract: The dimer model and its associated double-dimer model are fundamental objects in probability theory, statistical

mechanics, and combinatorics. While their behavior in planar settings is by now well understood, much less is known in higher dimensions and in the presence of a positive density of monomers, leading to the so-called monomer double-dimer model. We study these models on Z^d -like graphs ($d \geq 1$) that allow long-range edges whose weights decay with distance. For a large class of such interactions, we prove that the monomer double-dimer model exhibits long-range order. As a consequence, monomer correlations in the dimer model remain uniformly positive, and loops in the double-dimer model become macroscopic. In this talk, I will introduce the models and outline the main ideas of the proof. We will see that the model admits a natural correspondence with a spin system, which allows us to transfer results obtained via reflection positivity and thereby establish long-range order. This is joint work with Lorenzo Taggi and Wei Wu.

Title: Random graphs as models of quantum disorder

Author: Knowles, Antti

Email: antti.knowles@unige.ch

Affiliation: University of Geneva

Abstract: Abstract: A disordered quantum system is mathematically described by a large Hermitian random matrix. One of the most remarkable phenomena expected to occur in such systems is a localization-delocalization transition for the eigenvectors. Originally proposed in the 1950s to model conduction in semiconductors with random impurities, the phenomenon is now recognized as a general feature of wave transport in disordered media, and is one of the most influential ideas in modern condensed matter physics. A simple and natural model of such a system is given by the adjacency matrix of a random graph. I report on recent progress in analysing the phase diagram for the Erdős-Renyi model of random graphs. In particular, I explain the emergence of fully localized and fully delocalized phases, which are separated by a mobility edge. I also explain how to obtain optimal delocalization bounds using a new Bernoulli flow method. Based on joint work with Johannes Alt, Raphael Ducatez, and Joscha Henheik.

Title: An algorithmic procedure for solving the generalized minimum information checkerboard copula problem

Author: Kojadinovic, Ivan

Email: ivan.kojadinovic@univ-pau.fr

Affiliation: University of Pau

Abstract: The minimum information copula principle (see Meeuwissen and Bedford, 1997) is a maximum entropy-like approach for finding the least informative copula that satisfies a certain number of expectation constraints specified either from expert knowledge or the available limited data. In this presentation, we first propose a generalization of this principle allowing the inclusion of additional constraints fixing certain higher-order margins of the copula. We next prove that the associated optimization problem has a unique solution under a natural condition. As the latter problem is intractable in general, following the existing literature, we consider its version with all the probability measures involved in its formulation replaced by checkerboard approximations. This amounts to attempting to solve a so-called discrete I-projection linear problem. We then use the seminal results of Csiszar (1975) to derive an IPFP-like procedure for solving the latter and provide theoretical guarantees for its convergence. We conclude the presentation with numerical experiments in dimensions up to four with substantially finer discretizations than those encountered in the literature. This is joint work with Tommaso Martini, University of Torino, Italy.

Title: Privacy-Utility Tradeoffs of Noisy Stochastic Gradient Descent in High Dimensions

Author: Kolaczyk, Eric

Email: eric.kolaczyk@mcgill.ca

Affiliation: McGill University

Abstract: The interplay between optimization and privacy has become a central theme in privacy-preserving machine learning. Noisy stochastic gradient descent (SGD) has emerged as a cornerstone algorithm, particularly in large-scale settings. These variants of gradient methods inject carefully calibrated noise into each update to achieve differential privacy, which is the gold standard for rigorous privacy guarantees. Prior works primarily provide various bounds on statistical risks (utility) and privacy loss for noisy SGD, yet the exact behavior of the process remain unclear, particularly in high-dimensional settings. This work leverages a diffusion approach to analyze noisy SGD precisely, providing a continuous-time perspective that captures both statistical risk evolution and privacy loss dynamics in high dimensions. Moreover, we study a variant of noisy SGD that does not require explicit knowledge of gradient sensitivity, unlike existing approaches that assume or enforce

sensitivity through gradient clipping. Specifically, we focus on the least squares problem with l_2 regularization. Joint work with Shurong Lin and Elliot Paquette.

Title: Recent progress on graph bootstrap percolation

Author: Kolesnik, Brett

Email: brett.kolesnik@warwick.ac.uk

Affiliation: University of Warwick

Abstract: In H -percolation, edges in the random graph $G(n,p)$ are initially active. Other edges of the complete graph K_n become active if they are the only inactive edge in some copy of H in K_n . If all edges of K_n become active, the process H -percolates. In this talk, we will identify the critical threshold $p_c(n,H)$ when H is the clique K_r , and we will discuss recent progress for general H . These results resolve open problems of Balogh, Bollobás, and Morris (RS&A 2012). The talk is based on joint work with Bartha, Kronenberg, and Peled, and with Makai, Nenadov, Pérez-Giménez, Prałat, and Zhukovskii.

Title: The eigenvalues of iid matrices are hyperuniform

Author: Kolupaiev, Oleksii

Email: Oleksii.Kolupaiev@ist.ac.at

Affiliation: Institute of Science and Technology Austria

Abstract: We consider the point process of the eigenvalues of real or complex non-Hermitian matrices X with independent, identically distributed entries. In this talk, I will present a recent result obtained in a joint work with L. Erdos and G. Cipolloni, which establishes that this point process is hyperuniform: the variance of the number of eigenvalues in a subdomain of the spectrum is much smaller than the volume of this subdomain. Our main technical novelty is a very precise computation of the covariance between the resolvents of the Hermitization of $X-z_1, X-z_2$, for two distinct complex parameters z_1, z_2 .

Title: Equivalence testing with data-dependent and post-hoc equivalence margins

Author: Koning, Nick

Email: n.w.koning@ese.eur.nl

Affiliation: Erasmus University Rotterdam

Abstract: Equivalence testing compares the hypothesis that an effect is large against the alternative that it is negligible. Here, 'large' is classically expressed as being larger than some 'equivalence margin'. A longstanding problem is that this margin must be specified but can rarely be objectively justified in practice. We lay the foundation for an alternative paradigm, arguing to instead report a data-dependent margin that bounds the true effect with high probability. Our key argument is that this data-dependent margin is more useful than a test outcome at a fixed margin, as measured by the guarantees it offers to decision makers. We generalize this to a curve of margins, uniformly valid under the post-hoc selection of the margin. These ideas rely on e -values, which we derive for models that are strictly totally positive of order 3, nesting the classical z -test and t -test settings.

Title: Flexible skew models via coordinatewise Gaussian mixtures

Author: Kozubowski, Tomasz

Email: tkozubow@unr.edu

Affiliation: University of Nevada, Reno

Abstract: The normal mean - variance mixture model is a versatile tool in probability and statistics, with applications in hierarchical modeling, Bayesian inference, stochastic processes, and machine learning. In this talk, we propose an extension that allows for coordinatewise mean - variance mixing in a multivariate Gaussian framework. A particular case based on gamma mixing leads to a new class of multivariate marginally generalized asymmetric Laplace (MGAL) distributions, extending the classical multivariate generalized asymmetric Laplace (GAL) model. The proposed class offers enhanced flexibility for modeling multivariate data with heterogeneous skewness and tail behavior across dimensions. We present key distributional properties of the MGAL family and discuss parameter estimation via the Expectation - Maximization (EM) algorithm. An application to real data illustrates the practical usefulness of the proposed approach. This is joint work with Amos Natido.

Title: From Theory to Practice: inference beyond independence and CLT

Author: Kuchibhotla, Arun

Email: arunku@cmu.edu

Affiliation: Carnegie Mellon University

Abstract: Construction of confidence intervals and hypothesis tests have long been based on central limit theorems for the estimators and variance estimation. Such methods require the knowledge of the rate of convergence and estimation of the asymptotic variance, and even when these are satisfied, the CLT based methods can fail to retain validity uniformly over large classes of data-generating processes and when the dimension increases too fast with the sample size. Furthermore, dependence between the observations only worsens the issues. In this talk, I will present a framework that builds on HulC (Kuchibhotla et al. 2024, JRSSB) and significantly expands its scope. In addition to general results of uniform validity under independence, dependence, and diverging dimensions, we present specific examples for constructing confidence bands for empirical processes, and high-dimensional mean parameters. This talk constitutes joint work with Sivaraman Balakrishnan, Larry Wasserman, Manit Paul, and Abhimanyu Choudhary.

Title: Universality of first-order methods on random and deterministic matrices

Author: Kunisky, Dmitriy

Email: kunisky@jhu.edu

Affiliation: Johns Hopkins University

Abstract: I will present recent results on the universal behavior of general first-order methods (GFOM), a flexible class of iterative algorithms which update a state vector by matrix-vector multiplications and entrywise nonlinearities, thereby generalizing the much-studied class of approximate message passing (AMP) algorithms. Our approach will study the dynamics of GFOM through a moment method, reducing them to the traffic distribution of a sequence of matrices, an object studied in connection with free probability that consists of the collection of limiting values of certain diagrammatic permutation-invariant polynomials in the matrix entries. I will show how this approach gives a simple diagrammatic explanation for the asymptotic Gaussianity of AMP that distinguishes these algorithms from other GFOM, and how it unifies known limit theorems for AMP and GFOM on Wigner and rotationally invariant matrices and extends them to matrices with less randomness, more complicated variance profiles, and even some fully deterministic settings, in the latter case making progress towards resolving longstanding conjectures in the statistical physics and AMP literature. Based on joint work with Nicola Gorini, Chris Jones, and Lucas Pesenti.

Title: Minimum Norm Interpolation Under Gaussian Covariates

Author: Kur, Gil

Email: gilkur1990@gmail.com

Affiliation: ETH Zurich

Abstract: Minimum-Norm Interpolators (MNI) in overparameterized linear models have gained attention as a tractable framework for studying interpolation phenomena that resemble empirical observations in neural networks. Most prior work on these interpolators either exploits closed-form solutions when available or relies heavily on Gaussian comparison results, such as the convex Gaussian Min-Max Theorem (CGMT). In this paper, we introduce a new perspective on MNI under Gaussian covariates by leveraging tools from high-dimensional geometry. First, we obtain a "localized" bound on how much the MNI shrinks the original ground truth under isotropic Gaussian covariates when the norm is in isotropic position. Next, we prove sharp rates for the Mean Squared Error (MSE) of the l_1 -MNI, as obtained by Wang et al. via a purely geometric proof; crucially, our approach does not rely on CGMT.

Title: M-Estimation for Non-stationary Spatial Data

Author: Kurisu, Daisuke

Email: dkurisu.mathstat@gmail.com

Affiliation: The University of Tokyo

Abstract: While non-stationary spatial data are increasingly prevalent in various scientific applications, developing robust estimation and inference methods remains theoretically challenging due to complex spatial dependencies and potentially non-smooth loss functions. In this talk, we investigate the local linear quantile estimation for non-stationary spatial random fields. As our primary theoretical contribution, we derive the spatial Bahadur representation for the proposed estimator. This foundational result provides a unified asymptotic framework that bypasses traditional restrictive conditions. Furthermore, based on this representation, we explore its applications in facilitating further statistical inferences, such as the construction of Simultaneous Confidence Bands (SCBs) and spatial hypothesis testing.

Title: A Central Limit Theorem for sums of spatial block variables and some applications

Author: Lahiri, Soumendra

Email: s.lahiri@wustl.edu

Affiliation: Washington University in St. Louis

Abstract: Many statistical inference methods for spatial data are based on sums of functions of blocks of spatial variables instead of sums of functions of individual observations. Common examples include Spatial Block Bootstrap, Spatial Block Empirical Likelihood, Log-likelihood functions under different Markov Random Field models, among others. Asymptotic distribution theory for these methods is complicated by nontrivial overlaps that lead to strong dependence among the summands and is often derived on a case by case basis. This work proves a general Central Limit Theorem for sums of block variables using a version of Stein's method. allowing both pure- and mixed-increasing domain spatial asymptotic structures. Applications to some important inference problems in Spatial Statistics are also given.

Title: A nonparametric test for survival curves combining the Mann-Whitney effect and an overlap index.

Author: Langthaler, Patrick

Email: p.langthaler@plus.ac.at

Affiliation: Paris-Lodron University Salzburg

Abstract: In survival analysis several methods are available for testing the identity of two survival functions. A prevalent model is that of proportional hazards, which is violated in many cases occurring in practice, such as the treatment arm conferring late or early benefits compared to the placebo arm. Here we provide a method for comparing two survival functions in a right-censoring framework, which is particularly useful for crossing survival curves, while still retaining good power in a proportional hazards setting. The method combines two stochastic functionals: the well known Mann-Whitney effect, and an overlap index which considers distributional overlap. We construct our test by adapting a joint test methodology for the two functionals to the case of right-censored data. We present essential theoretical results regarding our method as well as results from a simulation study and an application to a real dataset.

Title: Trustworthy scientific inference with neural density estimators in intractable likelihood settings

Author: Lee, Ann

Email: annlee@andrew.cmu.edu

Affiliation: Carnegie Mellon University

Abstract: Scientific inference often involves inferring internal key parameters that determine the outcome of a complex physical phenomenon. The data themselves may come in the form of a labeled set (from e.g. a simulator or cross-matched studies) that implicitly encodes an often intractable likelihood function. We refer to inference in such a setting as "Likelihood-Free Inference" (LFI). The application of neural density estimators and generative models to scientific LFI settings is becoming increasingly widespread. However, high-posterior density (HPD) regions derived from these density estimators do not necessarily have a high probability of including the true parameter of interest, even if the posterior is well-estimated and the labeled data have the same distribution as the target distribution. Furthermore, if the prior distribution is poorly specified, then the HPD regions could severely undercover and/or be biased, thereby leading to misleading scientific conclusions. In this talk, I will present new LFI methodology and algorithms for leveraging neural density estimators to produce confidence regions that have (i) nominal frequentist coverage for any value of the (unknown) parameter, even with just one observation (sample size $n=1$), and (ii) smaller average area (yielding higher constraining power) if the prior is well-specified. I will illustrate our methods on examples from astronomy and high-energy physics, and discuss where we stand and what challenges still remain.

Title: Sato--Tate random multiplicative function

Author: Leung, Sun Kai

Email: sunkaileung@gmail.com

Affiliation: University of Oxford

Abstract: Random multiplicative functions have become a central theme in modern probabilistic number theory. In this talk, I will introduce the Sato--Tate random multiplicative function and outline its connections to arithmetic quantum chaos, with emphasis on the statistical behaviour of Hecke eigenforms. I will conclude with a discussion of ongoing joint work with Jad Hamdan and Mo Dick Wong.

Title: Testing for correlation between network structure and high-dimensional node covariates

Author: Levin, Keith

Email: kdlevin@wisc.edu

Affiliation: University of Wisconsin--Madison

Abstract: In many application domains, networks are observed with node-level features. In such settings, a common problem is to assess whether or not nodal covariates are correlated with the network structure itself. Here, we present four novel methods for addressing this problem. Two of these are based on a linear model relating node-level covariates to latent node-level variables that drive network structure. The other two are based on applying canonical correlation analysis to the node features and network structure, avoiding the linear modeling assumptions. We provide theoretical guarantees for all four methods when the observed network is generated according to a low-rank latent space model endowed with node-level covariates, which we allow to be high-dimensional. Our methods are computationally cheaper and require fewer modeling assumptions than previous approaches to network dependency testing. We demonstrate and compare the performance of our novel methods on both simulated and real-world data.

Title: Comparing groups of networks

Author: Levina, Liza

Email: elevina@umich.edu

Affiliation: University of Michigan

Abstract: Network-structured data arise in many applications; joint analysis of multiple networks is especially relevant to applications like neuroimaging, where each network corresponds to a patient's brain connectome. Models for multiple networks have typically focused on estimating their shared structure. Two-sample tests have also been developed, testing some version of the hypothesis of two samples of networks being indistinguishable across all node pairs (the global null). However, scientifically relevant hypotheses rarely take this form: for example, in neuroimaging it is rarely of interest to compare whole brains of patients and healthy controls, and the focus is more often on a particular brain region. Beyond comparisons, it is also of interest to estimate structures that correspond to a specific subgroup of networks, for example, for patients with a certain trait. One could always do that using just the subgroup of interest, but using all available samples allows us to better estimate structures that are shared by all, which in turn helps separate out the structure associated with a trait. This talk will introduce two methods that help address these challenges: mesoscale testing on networks, which conducts formal hypothesis testing on a subset of edges (like a brain region) while leveraging the rest of the network to increase power; and group MultiNeSS, a method that takes a sample of networks and estimates structures

Title: Graph Release with Assured Node Differential Privacy

Author: Li, Tianxi

Email: tianxili@umn.edu

Affiliation: University of Minnesota

Abstract: Differential privacy is a well-established framework for safeguarding sensitive information in data. While extensively applied across various domains, its application to network data -- particularly at the node level -- remains underexplored. Existing methods for node-level privacy either focus exclusively on query-based approaches, which restrict output to pre-specified network statistics, or fail to preserve key structural properties of the network. In this talk, I will introduce GRAND (Graph Release with Assured Node Differential privacy), which is, to the best of our knowledge, the first network release mechanism that releases entire networks while ensuring node-level differential privacy and preserving structural properties. Under a broad class of latent space models, we show that the released network asymptotically follows the same distribution as the original network. The effectiveness of the approach is evaluated through extensive experiments on both synthetic and real-world datasets.

Title: Likelihood-Based Nonparametric Causal Discovery under Latent Confounding

Author: Liang, Yurou

Email: yurou.liang@tum.de

Affiliation: Technical University of Munich, Chair of Mathematical Statistics

Abstract: Causal discovery with latent confounding amounts to learning an acyclic directed mixed graph (ADMG) over observed variables and unmeasured confounders. Existing approaches often rely on discrete combinatorial search, which becomes computationally prohibitive for large-scale problems. Recent methods alleviate this challenge by introducing a

differentiable acyclicity and bow-freeness constraint. However, these methods either assume linear structural equations or model nonlinear causal relationships involving both observed variables and confounders. The latter approximates intractable posteriors via variational inference, resulting in complex objectives and making performance more challenging. In this work, we propose a nonlinear causal model with correlated errors that encode latent confounding. We establish structural identifiability of the proposed model under bow-free graphs and parameter identifiability under ancestral graphs. This model yields a simple maximum-likelihood objective. We further develop a differentiable optimization scheme incorporating constraints for ADMG discovery. Experiments on synthetic and real-world datasets demonstrate that our method achieves competitive performance compared to relevant baselines.

Title: Harnessing Synthetic Data from Generative AI for Statistical Inference

Author: Lin, Xihong

Email: xlin@hsph.harvard.edu

Affiliation: Department of Biostatistics and Department of Statistics, Harvard University

Abstract: Integration of statistics and generative AI plays a pivotal role for accelerating trustworthy cross-domain scientific discovery. Recent advances in generative models have dramatically increased the availability and use of synthetic data across scientific domains. While these developments create exciting opportunities for empowering data analysis, they also raise fundamental statistical challenges regarding how synthetic data can be used in a valid, reliable, and principled manner. In this talk, we first discuss the current landscape of synthetic data generation using generative AI models such as transformer- and diffusion- based models. More importantly, we present a principled framework for incorporating synthetic data in downstream statistical analysis that ensures valid statistical inference even when generative AI models are misspecified. We show that the proposed synthetic data assisted methods integrating observed and synthetic data are robust to misspecified black-box generative models and can improve statistical inferential power when the generative AI models are informative. We demonstrate the utility of these synthetic data assisted methods to the analysis of the UK biobank data, by performing genome-wide association studies (GWAS) of proteomic data and whole-genome sequencing (WGS) analyses of brain imaging phenotypes, both characterized by substantial missingness (about 90%).

Title: Simple and sharp generalization bounds via lifting

Author: Liu, Jingbo

Email: jingbol@illinois.edu

Affiliation: University of Illinois

Abstract: We develop an information-theoretic framework for bounding the supremum of stochastic processes, offering a simpler and sharper alternative to classical chaining and slicing arguments for generalization bounds. The key idea is a lifting argument that produces information-theoretic analogues of empirical process bounds, such as Dudley's entropy integral. Lifting introduces permutation symmetry, yielding sharp bounds when the classical Dudley integral is loose. This gives a simple proof of the majorizing measure theorem via the sharpness of Dudley's entropy integral for stationary processes, a result known well before the proof of the majorizing measure theorem. Furthermore, the information-theoretic formulation provides soft versions of classical localized complexity bounds in generalization theory, but is simpler and does not require the slicing argument. We apply this approach to empirical risk minimization over Sobolev ellipsoids, obtaining sharp convergence rates in settings where previous methods are suboptimal. (arXiv:2508.18682)

Title: Global dynamics and weak universality for the fractional hyperbolic Φ^4_3 -model

Author: Liu, Ruoyuan

Email: ruoyuanl@math.uni-bonn.de

Affiliation: University of Bonn

Abstract: In this talk, I consider the 3-dimensional stochastic damped fractional nonlinear wave equation (with order $\alpha > 1$) with a cubic nonlinearity, also known as the fractional hyperbolic Φ^4_3 -model. The construction of the fractional Φ^4_3 -measure (Gibbs measure for the fractional hyperbolic Φ^4_3 -model) exhibits a phase transition: when $\alpha > 9/8$, the Gibbs measure is equivalent with the base Gaussian measure; when $1 < \alpha \leq 9/8$, the Gibbs measure is mutually singular with the base Gaussian measure. I will first talk about global well-posedness of the fractional hyperbolic Φ^4_3 -model and invariance of the fractional Φ^4_3 -measure for the entire range $\alpha > 1$. Then, I will discuss a weak universality result for the fractional hyperbolic Φ^4_3 -model. Some parts of the talk are based on a joint work with Nikolay Tzvetkov (ENS Lyon) and Yuzhao Wang (University of Birmingham).

Title: Flexible Selective Inference with Flow-based Transport Maps

Author: Liu, Sifan

Email: sifan.liu@duke.edu

Affiliation: Duke University

Abstract: Data-carving methods conduct selective inference by conditioning on the observed selection event and basing inference on the distribution of the full data given that event. A key limitation of existing approaches is their reliance on an analytically tractable characterization of the selection mechanism, which excludes many modern adaptive procedures, including data-driven hyperparameter tuning and complex model selection. We propose a new framework that enables selective inference without requiring an exact characterization of the selection event or the algorithm that produces it. The core idea is to learn a transport map that pushes forward a simple reference distribution to the conditional distribution given selection. The map is parameterized using a normalizing flow and trained directly from simulated data, leveraging the flexibility of flow-based generative modeling to approximate otherwise intractable distributions. Through extensive numerical experiments on both simulated and real data, we demonstrate that this method enables selective inference by providing: (i) valid p-values and confidence intervals for adaptively selected hypotheses and parameters, (ii) a closed-form expression for the conditional density, enabling likelihood-based and quantile-based inference, and (iii) a modular adjustment for intractable selection steps that can be easily combined with existing methods for the tractable steps in a multi-step selection procedure.

Title: Preferential attachment trees with vertex death: The super-linear regime

Author: Lodewijks, Bas

Email: bas.lodewijks@sheffield.ac.uk

Affiliation: University of Sheffield

Abstract: We investigate an evolving random tree model in which, at every step, either a new alive vertex is added which connects to an existing alive vertex preferentially according to a function b , or an existing alive vertex is selected preferentially according to a function d and killed. We consider the super-linear regime, in which $1/(b(i)+d(i))$ is summable in i . Here, we investigate structural properties of the limiting infinite tree, in particular whether the limiting tree contains (i) A unique vertex with infinite degree and no infinite path, (ii) A unique infinite path and no vertex with infinite degree, (iii) Infinitely many infinite paths and no vertex with infinite degree. We shall discuss how and why each case (i)-(iii) arises, and will look at a broad range of examples. Joint work with Markus Heydenreich.

Title: Interacting point processes : How does the mean field limit help us for statistical inference on them?

Author: Loecherbach, Eva

Email: eva.loecherbach@polytechnique.edu

Affiliation: Ecole Polytechnique

Abstract: We consider mean field interacting Hawkes processes composed of N components. We suppose that we observe the whole system during some fixed time interval. Asymptotics will be taken as N tends to infinity (and sometimes also as time grows). Such models are widely used, for example in neuroscience to describe neuron's spiking behavior, an application that I will focus on. Three statistical questions are particularly important for neuroscientists in this context: Firstly, by observing only the jumps of the system, is it possible to estimate the underlying interaction graph or at least some important features of it? Secondly, is it possible to detect synchrony in the spiking behavior of large assemblies of neurons? Thirdly, is it possible to understand how a neuron reacts to incoming stimuli; that is, to estimate the underlying jump rate function? I will present some attempts of answers that have been given to these points in some recent works in which I had the chance to collaborate. In particular I will also discuss how passing to the mean field limit can help us to deal with these questions when we suppose that the underlying graph is the complete graph. My talk is based on collaborations with Julien Chevallier (France), Aline Duarte (Brazil), Kadmo Laxa (Brazil), Dasha Loukianova (France), Guilherme Ost (Brazil), Patricia Reynaud-Bouret (France), Etienne Tanré (France) and Josué Tchouanti (France).

Title: Computationally efficient reductions between some statistical models

Author: Lou, Mengqi

Email: mlou30@gatech.edu

Affiliation: Georgia Institute of Technology

Abstract: Average-case reductions establish rigorous connections between different statistical models, allowing us to show that if one problem is computationally hard, then another must be as well. Reductions from the planted clique problem have

revealed statistical-to-computational gaps in many statistical problems with combinatorial structure. However, several important models remain beyond the reach of existing reduction techniques—for example, no reduction-based hardness results are currently known for sparse phase retrieval. In this talk, we introduce a computationally efficient procedure that approximately transforms a single observation from certain source models with continuous-valued sample and parameter spaces into a single observation from a broad class of target models. I will present several such reductions and highlight their applications in computational lower bounds, including universality results and hardness in sparse generalized linear models. I will also discuss a potential application in transforming one differentially private mechanism into another.

Title: A conversion theorem and minimax optimality for continuum contextual bandits

Author: Lounici, Karim

Email: karim.lounici@polytechnique.edu

Affiliation: École polytechnique

Abstract: We study contextual continuum bandits, where a learner sequentially observes a context vector and selects an action in a convex set to minimize a context-dependent function. The objective is to minimize all functions across contexts, leading to contextual regret, a stronger notion than static regret. Assuming the objective functions are γ -Hölder continuous in the context ($0 < \gamma \leq 1$), we show that any algorithm with sublinear static regret can be extended to achieve sublinear contextual regret. We establish a general static-to-contextual regret conversion theorem bounding contextual regret in terms of static regret. We then derive implications for three settings: Lipschitz, convex, and strongly convex smooth bandits. For Lipschitz bandits with $\gamma=1$, we obtain a minimax optimal contextual regret rate scaling as $T^{\frac{d_c+d_x+1}{d_c+d_x+2}}$ (up to logarithmic factors), where d_c and d_x are the context and action dimensions. For noisy evaluations, convex and strongly convex smooth bandits share the same rate, scaling as $T^{\frac{d_c+\gamma}{d_c+2\gamma}}$ up to logs. Finally, we prove a minimax lower bound showing that sublinear contextual regret may be impossible without continuity in context, and that our algorithms achieve near-optimal rates (up to logarithmic factors) for convex and strongly convex settings.

Title: Versatile Differentially Private Learning for General Loss Functions

Author: LU, QILONG

Email: lu_qilong@stu.pku.edu.cn

Affiliation: Peking University

Abstract: We provide a versatile privacy-preserving release mechanism along with a unified approach for subsequent parameter estimation and statistical inference. We propose a privacy mechanism based on zero-inflated symmetric multivariate Laplace (ZIL) noise, which requires no prior specification of subsequent analysis tasks, allows for general loss functions under minimal conditions, imposes no limit on the number of analyses, and is adaptable to increasing data volume in online scenarios. We derive the trade-off function for the proposed ZIL mechanism, which characterizes its privacy protection level. Furthermore, to formalize the local differential privacy (LDP) property of the ZIL mechanism, we extend the classical epsilon-LDP to a more general f -LDP framework. To address scenarios where only individual attribute values require protection, we propose attribute-level differential privacy (ADP) and its local version. Within the M-estimation framework, we introduce a novel doubly random (DR) corrected loss for the ZIL mechanism, which yields consistent and asymptotically normal M-estimates under differential privacy constraints. The proposed approach is computationally efficient and does not require numerical integration or differentiation for noisy data. It applies to a broad class of loss functions, including non-smooth ones. In addition, our method can handle unbounded data, while ensuring that discrete data remain discrete and bounded data stay within their original range.

Title: Robust, sub-Gaussian mean estimators in metric spaces

Author: Lugosi, Gabor

Email: gabor.lugosi@gmail.com

Affiliation: ICREA & Pompeu Fabra University

Abstract: The optimal accuracy one may achieve with a given confidence for the problem of estimating the mean of a random vector from i.i.d. data is well understood. When the data take values in more general metric spaces, an appropriate extension of the notion of the mean is the Fréchet mean. We study the performance of estimators of the Fréchet mean in general metric spaces under possibly heavy-tailed and contaminated data. In such cases, the empirical Fréchet mean is a poor estimator. We propose a general estimator based on high-dimensional extensions of trimmed means and prove general performance bounds. Unlike all previously established bounds, ours generalize the optimal bounds known for Euclidean data. We apply our

results for metric spaces with curvature bounded from below, such as Wasserstein spaces, and for uniformly convex Banach spaces. The talk is based on joint work with Daniel Bartl, Roberto Imbuzeiro Oliveira, and Zoraida F. Rico.

Title: Identifiability in continuous Lyapunov models with heterogeneous noise

Author: Lumpp, Sarah

Email: sarah.lumpp@tum.de

Affiliation: Technical University of Munich

Abstract: Stationary distributions of multivariate diffusion processes provide a new approach to probabilistic modelling of causal systems in statistics and machine learning. By assuming each observation to arise as a one-time cross-sectional snapshot of a temporal process in equilibrium, they allow one to naturally model dependence structures that may include feedback loops. Specifically, the graphical continuous Lyapunov model consists of Gaussian stationary distributions of multivariate Ornstein-Uhlenbeck processes. Their covariance matrices are parametrized as solutions to the continuous Lyapunov equation, with sparsity assumptions on the processes' drift matrices represented by a directed graph. Allowing for coordinate-wise different scaling of the driving noise process, we show generic parameter identifiability for certain classes of directed graphs, including tree-structured graphs. Additionally, we prove sufficient conditions for generic non-identifiability and characterize the generically identifiable structures in small acyclic graphs up to 6 nodes.

Title: Conformal Prediction for Dyadic Regression

Author: Lunde, Robert

Email: lunde@wustl.edu

Affiliation: Washington University in St. Louis

Abstract: Dyadic regression, which involves modeling a relational matrix given covariate information, is an important task in statistical network analysis. We consider uncertainty quantification for dyadic regression models using conformal prediction. We establish finite-sample validity of our procedures for various sampling mechanisms under a joint exchangeability assumption. Our proof uses new results related to the validity of conformal prediction beyond exchangeability, which may be of independent interest. We also show that, under certain conditions, it is possible to construct asymptotically valid prediction sets for a missing entry under a structured missingness assumption.

Title: On the satisfaction frequency of spectral characterization conditions

Author: Lvov, Nikita

Email: nikita.lvov@mail.mcgill.ca

Affiliation: McGill University

Abstract: We give the first specific conjectures on how frequently graphs satisfy sufficient conditions for being uniquely characterized by spectral information. These conjectures arise from a theoretical framework that we develop based on abstract-algebraic random matrix statistics. Specifically, we rephrase conditions from the literature in terms of $\mathbb{Z}[x]$ -modules associated to the adjacency matrix, and study the distribution of those modules in analytically tractable profinite random matrix ensembles. We apply this new method to two distinct conditions. The first requires square-freeness of the determinant of the walk matrix, and the second uses the discriminant of the characteristic polynomial.

Title: HeteroJIVE: Weighted Spectral Estimation for Shared Subspace Recovery under Multi-View Heteroskedasticity

Author: Lyu, Zhongyuan

Email: zhongyuan.lyu@sydney.edu.au

Affiliation: The University of Sydney

Abstract: Many modern datasets consist of multiple related matrices measured on a common set of units, where the goal is to recover the shared low-dimensional subspace. While the Angle-based Joint and Individual Variation Explained (AJIVE) framework provides a solution, it relies on equal-weight aggregation, which can be strictly suboptimal when views exhibit significant statistical heterogeneity (arising from varying SNR and dimensions) and structural heterogeneity (arising from individual components). We propose HeteroJIVE, a weighted two-stage spectral algorithm tailored to such heterogeneity. Theoretically, we first revisit the "non-diminishing" error barrier with respect to the number of views K identified in recent literature for the equal-weight case. We demonstrate that this barrier is not universal: under generic geometric conditions, the bias term vanishes and our estimator achieves the $O(K^{-1/2})$ rate without the need for iterative refinement. Extending this to the general-weight case, we establish error bounds that explicitly disentangle the two layers of heterogeneity. Based on this,

we derive an oracle-optimal weighting scheme implemented via a data-driven procedure. Extensive simulations corroborate our theoretical findings, and an application to TCGA-BRCA multi-omics data validates the superiority of HeteroJIVE in practice.

Title: Chatterjee's coefficient is optimal for detecting regular oscillating signals

Author: Lyubarskaja, Anna

Email: annalyu@stanford.edu

Affiliation: Stanford University

Abstract: We investigate the power of rank-based independence tests in detecting deterministic dependence for regular k -oscillating functions. We establish that tests based on integrated CDF distances, such as Hoeffding's coefficient, Blum–Kiefer–Rosenblatt's coefficient, and Bergsma–Dassios–Yanagimoto's coefficient, asymptotically require a sample size of $n=k^2$ for detection. In contrast, Chatterjee's coefficient succeeds asymptotically with n only on the order of $Ck^{4/5}$ data points. Finally, by bounding the total variation distance between rank distributions, we show this rate is optimal: no rank-based test can detect regular k -oscillation when n is of order less than $k^{4/5}$.

Title: Multivariate Archimedean copulas may have fractal support

Author: Maislinger, Lea

Email: lea.maislinger@plus.ac.at

Affiliation: University of Salzburg

Abstract: We study the family of all d -dimensional Archimedean copulas with d greater than or equal to three, a well-known class with broad applicability. Induced by a continuous univariate function, the so-called Archimedean generator, Archimedean copulas might be expected to distribute their mass in a rather regular manner on the unit cube. Motivated by a recently shown result stating that bivariate Archimedean copulas may have fractal support (in the sense of having non-integer Hausdorff dimension) we investigate the general multivariate case. In particular, we show that for every d and arbitrary s between $d-1$ and d , there exists some d -dimensional Archimedean copula whose support has Hausdorff dimension s . Finally, considering absolute continuity of the corresponding marginal copula our results are used to point out that ignoring at least one coordinate can obscure this irregularity, thereby creating a misleading impression of regularity.

Title: Glassiness in Quantum Systems

Author: Manai, Chokri

Email: cm7411@nyu.edu

Affiliation: New York University

Abstract: Spin glasses occupy a central place in theoretical physics, mathematics, and theoretical computer science. Landmark achievements include Parisi's celebrated solution of the Sherrington–Kirkpatrick model, recognized by the 2021 Nobel Prize in Physics, and Talagrand's rigorous mathematical proof of the Parisi formula, for which he was awarded the 2024 Abel Prize. In this talk, I will discuss a natural next step: disordered mean-field quantum systems. Quantum systems are governed by noncommutative interactions, and the resulting quantum glasses may be viewed as structured random matrices. They are substantially more intricate than their classical spin-glass counterparts, while at the same time lying largely beyond the reach of standard random matrix methods. I will focus on a paradigmatic model of a quantum glass, the quantum Sherrington–Kirkpatrick model. This model already exhibits a range of novel physical phenomena, and in a series of works with Simone Warzel we have been able to place several of these features on a rigorous mathematical foundation. Time permitting, I will also briefly discuss further challenging models, such as the Sachdev–Ye–Kitaev model and the quantum Heisenberg model. These systems remain, to a large extent, mathematically unexplored and offer a rich source of difficult and important questions for mathematical physics and probability theory.

Title: Edgeworth expansion on Wiener chaos

Author: Mansanarez, Paul

Email: paul.mansanarez@ulb.be

Affiliation: Nantes Université/ULB

Abstract: Edgeworth expansions approximate the distribution of normalized sums of i.i.d. random variables as the Gaussian measure corrected by polynomial perturbations, with coefficients expressed in terms of the cumulants of the underlying distribution. In this talk, we investigate Edgeworth expansions for functionals of a Gaussian field, which are not necessarily i.i.d. sums. Using the Malliavin–Stein method and semigroup analysis, we derive a bound on the total variation distance

between the law of F and its so-called Edgeworth expansion: a modified Gaussian measure. The bounds depend only on p and a power of the fourth cumulant of F , which quantifies the distance of the law of F to the normal distribution.

Title: Cost-Aware Optimized Front-Door Experimental Design

Author: Mareis, Leopold

Email: leopold.mareis@tum.de

Affiliation: Technical University of Munich

Abstract: Causal effect estimation often succeeds cost-constrained sequential data collection. This work considers multivariate linear front-door models with arbitrary unobserved confounding on treatment and response. We optimize the experimental design by balancing the statistical efficiency and measurement costs through partial data. The full-data efficient influence function for the causal effect is derived, together with the geometry of all observed-data influence functions. This characterization yields a closed-form optimal sampling policy and an estimator to minimize the asymptotic variance of regular asymptotically linear (RAL) estimators within a class of augmented full-data influence functions. The resulting design also covers back-door estimation. In simulations and applications to biological, medical, and industrial datasets, the optimized designs achieve substantial efficiency gains (5.3% to 31.9%) over naive full-sampling strategies.

Title: The Functional Graphical Lasso

Author: Masak, Tomas

Email: tomas.masak@wu.ac.at

Affiliation: Wirtschaftsuniversitaet Wien

Abstract: We consider the problem of recovering conditional independence relationships between p jointly distributed Hilbertian random elements given n realizations thereof. We operate in the sparse high-dimensional regime, where $n \ll p$ and no element is related to more than $d \ll p$ other elements. In this context, we propose an infinite-dimensional generalization of the graphical lasso. We prove model selection consistency under natural assumptions and extend many classical results to infinite dimensions, doing away with unnecessary structural assumptions. The plug-in nature of our method makes it applicable to heterogeneous data measured under any observational regime, whether sparse or dense, and indifferent to serial dependence between samples. Moreover, the method can be understood as naturally arising from a coherent maximum likelihood philosophy.

Title: Analyzing the relationship between universal and classical likelihood ratio inference: Almost sure-relations and asymptotic rates

Author: Matz, Lorenz

Email: lorenz.matz@wu.ac.at

Affiliation: Vienna University of Economics and Business

Abstract: Universal Inference is a general method for constructing confidence sets and tests in statistical models, proposed by Wasserman et al. (2020). It is based on a combination of the likelihood ratio approach and sample splitting. Unlike procedures built on asymptotic results or most resampling methods, the resulting confidence sets/tests have finite-sample guarantees and require virtually no regularity conditions on the model or the set of null distributions. However, for many inference problems of interest, universal confidence sets and tests tend to be conservative in comparison to established methods with finite sample or asymptotic coverage/type I error guarantees. Currently, the relation between the universal method and classical approaches to inference in terms of power and confidence set diameters is still poorly understood. To shed some light on this, we investigate the relationship between the split LR and the classical LR confidence set in general settings. For one- and two-dimensional models, we establish that the classical LR set is almost surely contained in the split LR set when the maximum likelihood estimator is used for the latter. Furthermore, we find conditions under which the diameters of both confidence sets shrink at the same rate by proving some general results about these rates for 'M-type' sets.

Title: Trimming and constraints based robust clustering for matrix-variate data

Author: Mayo-Iscar, Agustín

Email: agustin.mayo.iscar@uva.es

Affiliation: Universidad de Valladolid

Abstract: Authors: Luis Ángel García-Escudero (Universidad de Valladolid); Marcus Mayrhofer (TU Wien); Agustín Mayo-Íscar (Universidad de Valladolid). Abstract: We have designed a robust clustering estimator for data coming from a mixture of matrix

normal distributions. On the one hand, we have used the Kronecker-product covariance structure based on row and column covariances. On the other hand, the mentioned robustness of this estimator is based on the joint application of trimming, in order to avoid the influence of contaminating sources, and eigenvalue constraints, in order to get a well posed estimation problem and to avoid spurious solutions. Our proposal, MTCLUST, generalizes the Matrix Minimum Covariance Determinant, which is a robust mean and covariance estimator of matrix-valued data. We will provide empirical evidence about how this methodology works based on artificial and real data.

Title: Robustness and Explainability in Multivariate Functional Data Analysis

Author: Mayrhofer, Marcus

Email: mayrhofer.m95@gmail.com

Affiliation: TU Wien

Abstract: We propose a method for robust covariance estimation for multivariate functional data by establishing a connection between a stochastic process with a separable covariance structure and the corresponding matrix-variate distribution of its basis representation. Based on this connection we use the Matrix Minimum Covariance Determinant (MMCD) approach in combination with a truncated multivariate functional Mahalanobis semi-distance for robust estimation of mean and covariance. To go beyond outlier detection, we generalize multivariate outlier explanations based on Shapley values to decompose the truncated multivariate functional Mahalanobis semi-distance of individual observations into time-coordinate-specific contributions. Our approach provides a framework for the generalization of other distance-based functional outlier detection approaches and lays the groundwork for future extensions to robust covariance estimation and explainable outlier detection for random surfaces.

Title: What we can know from complex functional data

Author: McKeague, Ian

Email: im2131@columbia.edu

Affiliation: Columbia University

Abstract: This talk discusses a nonparametric inference framework for functional data having sample paths of bounded variation, with applications in a variety of complex settings. The main application will be to wearable device data collected in a Columbia-based study of an experimental therapy for mitochondrial disease, a group of disorders that affect the body's ability to produce energy. Specifically, we provide the first clinical application of a novel, bias-adjusted outcome measure of acceleration across a range of subjects' activities to assess nucleoside therapy for thymidine kinase 2 deficiency, an ultra-rare autosomal recessive mitochondrial disease. In addition, the talk will briefly touch on the potential (in terms of "what we can know") of having access to functional data collected along causal paths in a possibly non-smooth metric spacetime, which relaxes the smooth Lorentzian framework of Einstein's field equations to allow rougher settings.

Title: Universality for the global spectrum of random inner-product kernel matrices

Author: McKenna, Benjamin

Email: mckenna@math.gatech.edu

Affiliation: Georgia Institute of Technology

Abstract: In recent years, machine learning has motivated the study of what one might call "nonlinear random matrices." This broad term includes various random matrices whose construction involves the entrywise application of some deterministic nonlinear function, such as ReLU. We study one such model, an entrywise nonlinear function of a sample covariance matrix $f(X^*X)$, typically called a "random inner-product kernel matrix" in the literature. A priori, entrywise modifications of a matrix can affect the eigenvalues in complicated ways. However, a long line of work has established that the global spectrum of such matrices actually behaves in a simple way, described by free probability, when either the randomness in X is from a restricted family (such as Gaussian) or X has proportional sidelengths. We show that this description is universal, holding both for much more general randomness in X and for nonclassical aspect ratios. Joint work with Sofiia Dubova, Yue M. Lu, and Horng-Tzer Yau.

Title: Bayesian Nonparametric Privacy-Preserving Synthetic Data Generation

Author: Menicucci, Maria Chiara

Email: mariachiara.menicucci@carloalberto.org

Affiliation: University of Torino and Collegio Carlo Alberto

Abstract: Synthetic data generation is a well-established method for releasing data while protecting individual privacy. From a statistical perspective, a release mechanism should satisfy differential privacy (DP) guarantees, and preserve statistical utility of the data. We propose two new release mechanisms based on Bayesian nonparametric models: modelling the private data as a sample from a random probability distribution, endowed with a nonparametric prior, we generate synthetic data sampling from the posterior predictive distribution. Specifically, when the private data have ties, we consider the Pitman-Yor process (PYP) prior and show that, in general, only instance-level (ϵ, δ) -DP is achievable. However, if the discount parameter σ of the PYP is zero, which corresponds to the Dirichlet process, then global (ϵ, δ) -DP can be obtained depending on the size of the released sample. When the private data have no ties, we consider the Pólya tree prior and show that global (ϵ, δ) -DP is achievable, with constraints on the size of the released sample whose stringency depends on the rate of growth of the tree coefficients. In both cases, we quantify the statistical utility of the mechanism by deriving rates for the expected 1-Wasserstein distance between the empirical measure of the released data and the data generating distribution. For the Dirichlet process, such a result is based on a novel concentration result for the posterior of a Dirichlet process, which is of independent interest.

Title: A phase transition for the corank of random band matrices over F_p

Author: Mészáros, András

Email: meszaros@renyi.hu

Affiliation: HUN-REN Alfréd Rényi Institute of Mathematics

Abstract: For the corank of random matrices over the p element field, there is a widespread universality phenomenon. Namely, for a large class of random matrices over the p element field, their cokrank has the same limiting distribution as the corank of uniform random $n \times n$ matrices. In this talk, we consider uniform random band matrices over the p element field, and determine the critical band width where the above mentioned universality breaks down. This question is motivated by the analogous phase transition for the spectrum of Gaussian band matrices, which is a well studied problem in classical random matrix theory.

Title: Moments of Exponential Functionals of Lévy Processes and Bernstein-Gamma Functions

Author: Minchev, Martin

Email: martin.minchev@math.uzh.ch

Affiliation: University of Zurich and Sofia University

Abstract: Exponential functionals of Levy processes are a classical object in probability, which appear, for example, through the Lamperti time change linking self-similar Markov processes to Levy processes, and also in mathematical finance in connection with Asian options. A standard argument shows that their moments satisfy a recurrence relation of period one, and this leads naturally to the question whether the law is determined by its moments. In this talk we first recall the early results of Bertoin and Yor from the early 2000s. We then turn to the extension from integer moments to complex exponents, that is, to the Mellin transform. Patie and Savov expressed this transform in terms of a new class of Gamma-type functions, called Bernstein-Gamma functions. We present some basic properties of these functions and explain how they arise in the study of exponential functionals. We also discuss several open questions on moment determinacy for exponential functionals, as well as possible matrix-valued extensions of Bernstein-Gamma functions. Joint work with Mladen Savov.

Title: Probabilistic forecasting via continuously weighted, proper score-optimal ensembles

Author: Minin, Volodymyr

Email: vminin@uci.edu

Affiliation: University of California, Irvine

Abstract: When forecasting with parametric models, one usually employs the classical statistical approach of selecting for parameter values that best explain observed data, generally either by maximizing the likelihood function or approximating the Bayesian posterior distribution derived from it. A class of novel procedures has arisen to shift away from this paradigm, selecting instead for parameter values that induce a predictive distribution best matching the empirical distribution. Recent work, such as Focused Bayesian Prediction and Predictive Variational Inference, modifies the Bayesian update to achieve this idea in a generalized Bayesian setting. Building on these developments but not limiting ourselves to the Bayesian framework, we introduce a new paradigm for probabilistic forecasting with an arbitrary family of parametric time series models. Each member of the family generates its own probabilistic forecast, and the forecasts are combined into a continuously-weighted ensemble. We find optimal weights by minimizing a proper score (e.g., continuous ranked probability score) of the ensemble

forecast over the training period. Our preliminary research compares the predictive performance of this new framework with the Bayesian posterior predictive distribution—the state of the art in many applications, such as infectious disease forecasting. Specifically, we apply our technique to the forecasting of healthcare utilization during peaks of respiratory virus activity.

Title: Learning Multi-Index Models in High-Dimensions

Author: Misiakiewicz, Theodor

Email: theodor.misiakiewicz@yale.edu

Affiliation: Yale University

Abstract: We study the problem of learning multi-index models (MIMs), where the label depends on the high-dimensional input vector only through an unknown low dimensional projection. Exploiting the equivariance of this problem under the orthogonal group, we obtain a sharp harmonic-analytic characterization of the learning complexity for MIMs with spherically symmetric inputs. Specifically, we derive statistical and computational complexity lower bounds within the Statistical Query (SQ) and Low-Degree Polynomial (LDP) frameworks. These bounds decompose naturally across spherical harmonic subspaces. Guided by this decomposition, we construct a family of spectral algorithms based on harmonic tensor unfolding that sequentially recover the latent directions and nearly achieve these SQ and LDP lower bounds. Depending on the choice of harmonic degree sequence, these estimators can realize a broad range of trade-offs between sample and runtime complexity.

Title: Statistics for regularized optimal transport with decreasing regularization

Author: Mordant, Gilles

Email: mordantgilles@gmail.com

Affiliation: Yale University

Abstract: In this talk, we will discuss the statistical fluctuations of entropy-regularized optimal transport between empirical measures when the regularization parameter is a decreasing function of the sample size and the latter goes to infinity. We shall discuss how the dual potentials relates to the density of the data and how entropy-regularized optimal transport is related to diffusions.

Title: Statistics for regularized optimal transport with decreasing regularization

Author: Mordant, Gilles

Email: mordantgilles@gmail.com

Affiliation: Yale University

Abstract: In this talk, we will discuss the statistical fluctuations of entropy-regularized optimal transport between empirical measures when the regularization parameter is a decreasing function of the sample size and the latter goes to infinity. We shall discuss how the dual potentials relates to the density of the data and how entropy-regularized optimal transport is related to diffusions.

Title: Stochastic control and reinforcement learning for greedy dynamic matching models

Author: Moyal, Pascal

Email: pascal.moyal@univ-lorraine.fr

Affiliation: Université de Lorraine / Inria

Abstract: We consider a dynamic matching model on a general (non-bipartite) graph: items arrive one by one into a system, and depart by pairs of two, as soon as they are matched, following compatibility rules that are given by the underlying graph structure. We assume that the matching policy is greedy, in the sense that an incoming item that finds at least one compatible item in the queue, must be matched right away. Using the tools of dynamic programming and reinforcement learning, we address the optimal control of such models following two different lines: First, we show that, for a simple graph structure, a threshold-type policy is optimal for minimizing holding costs. Second, we show how to optimize the access control to the queue, by using a Policy-Gradient Reinforcement Learning (PGRL) algorithm.

Title: Surrogate-powered Causal Inference

Author: Mukhin, Yaroslav

Email: yaroslav@mukhin.net

Affiliation: Cornell University

Abstract: Missing information is intrinsic to causal inference with survival outcomes: Experiments are constrained by budgets

and timelines; observational studies face drop-out. In addition to the unobserved counterfactual outcome, the factual event time is subject to right censoring. This degrades the effective sample size even when identification is secured by design or by assumption. We study whether an intermediate post-treatment event, e.g., recurrence or progression, can recover the power for the primary outcome estimand that is lost to censoring. With full-data and absent semiparametric restrictions, surrogate outcomes cannot deliver a first-order efficiency gain. With missing outcomes, surrogates and baseline covariates become informative for marginal causal parameters that, absent missing data, do not depend on the surrogate. The size of the gain depends on: (i) the balance of the post-treatment stratification induced by the surrogate, (ii) the prognostic separation of the strata, (iii) the temporal proximity of the stratification with the censoring.

Title: Geodesic Causal Inference

Author: Müller, Hans-Georg

Email: hgmueller@ucdavis.edu

Affiliation: UC Davis

Abstract: Causal inference methods have emerged as a popular tool to adjust for confounding and imbalance in regression models with scalar responses and also distributional responses. We introduce a general framework for causal inference when responses reside in general geodesic metric spaces, drawing on a novel geodesic calculus, which facilitates the quantification of treatment effects through the concept of geodesic average treatment effect (GATE). Employing Fréchet regression, we obtain doubly robust estimation of GATE with consistency and rates of convergence. Theoretical results are complemented with various data illustrations. The geodesic framework can also be employed for other causal designs such as difference-in-differences and discontinuity designs. This talk is based on joint work with Daisuke Kuriso, Yidong Zhou and Taisuke Otsu.

Title: ANOVA-type Wasserstein tests for factorial designs

Author: Munk, Axel

Email: munk@math.uni-goettingen.de

Affiliation: Göttingen University

Abstract: We introduce a general framework for testing statistical hypotheses for probability measures supported on finite spaces, which is based on optimal transport (OT). These tests are inspired by the analysis of variance (ANOVA) and its nonparametric counterparts. They allow for testing linear relationships in factorial designs between discrete probability measures and are based on pairwise comparisons of the OT distance and corresponding barycenters. To this end, we derive under the null hypotheses and (local) alternatives the asymptotic distribution of empirical OT costs and the empirical OT barycenter cost functional as the optimal value of linear programs with random objective function. In particular, we extend existing techniques for probability to signed measures and show directional Hadamard differentiability and the validity of the functional delta method. We discuss computational issues, permutation and bootstrap tests, and back up our findings with simulations. We illustrate our methodology on a data set from biometric fingerprint identification. This is joint work with Michel Groppe, and Shayan Hundrieser and Linus Niemöller.

Title: The Central Role of Proper Scoring Rules in Modern Statistics: Theory, Robustness, and Applications to Complex Data

Author: Musio, Monica

Email: mmusio@unica.it

Affiliation: University of Cagliari

Abstract: Proper scoring rules (PSRs) provide a foundational framework for both the evaluation of probabilistic forecasts and the construction of tailored inferential procedures. A SR is proper if it incentivizes honesty, ensuring that the expected score is optimized when the quoted distribution matches the true underlying process. While the logarithmic score is the standard choice due to its link with likelihood-based inference, its application can be limited by sensitivity to anomalous observations or the presence of intractable normalizing constants. This has motivated the development of alternative rules characterized by specific properties, such as robustness or homogeneity. In this talk, we review the theoretical foundations and practical relevance of PSRs, highlighting their deep connections with M-estimation and subjective probability. We discuss how these rules induce unbiased estimating equations, leading to consistent estimators whose asymptotic precision is governed by the Godambe information. Special emphasis is placed on applications involving complex data structures. For compositional data, we demonstrate how SRs can be integrated with isometric log-ratio transformations to respect the geometry of the simplex. Furthermore, we explore extreme value analysis, where the choice of the SR is critical. We compare the use of rules such as

the Tsallis for handling contamination and the Hyvärinen for its homogeneity, which allows bypassing complex normalizing constant.

Title: Random polynomials: old and new

Author: Nguyen, Hoi

Email: nguyen.1261@osu.edu

Affiliation: The Ohio State University

Abstract: The study of zeros of random polynomials dates back to the early twentieth century and remains an active area of research. In this talk, we discuss recent developments, focusing on robust frameworks for establishing universality. Topics include global and local correlations of zeros, the variance in the number of real zeros, concentration phenomena, central limit theorem fluctuations, and related questions.

Title: Random polynomials with dependent coefficients

Author: Nguyen, Oanh

Email: oanh_nguyen1@brown.edu

Affiliation: Brown University

Abstract: Classical results on random polynomials largely focus on the idealized setting of independent coefficients. But what happens when this assumption is relaxed? In this talk, we explore a range of random polynomials with dependent structures, highlighting both the challenges they pose and the techniques used to analyze them. We will discuss models with weakly dependent coefficients, as well as those arising from Ising interactions. This talk is based on joint work with Jürgen Angst, Guillaume Poly, and Oren Yakir.

Title: The blessing of refitting in doubly-robust conditional independence testing procedures

Author: Niu, Ziang

Email: ziangniu@wharton.upenn.edu

Affiliation: University of Pennsylvania

Abstract: Conditional randomization tests (CRTs) is a powerful framework for conditional independence (CI) testing. Two practical drawbacks are often cited: (i) they require exact knowledge of the predictor's conditional distribution given covariates, and (ii) they involve refitting a model for each resampled statistic. In practice, the conditional distribution is rarely known and must be estimated from data. We show that when this distribution is learned in-sample, the refitting in CRT is not merely a computational burden; it can improve Type-I error control in finite samples. This improvement can be understood through the lens of data sharpening—a classical bias-correction technique in nonparametric regression. Building on this insight, we first generalize the data sharpening to accommodate broad learners and then propose new CI testing procedures based on the generalized sharpening approach. These procedures require at most two additional refits and enjoy provably more robust Type-I error control compared to the vanilla unsharpened version. Our theory accommodates a range of learners, including kernel ridge regression, lasso, and local linear regression. In particular, we extend our approach to CI testing based on doubly robust functional and show that data sharpening reduces second-order bias, leading to improved Type-I error control.

Title: Functional AMUSE for multivariate functional separable time series

Author: Nordhausen, Klaus

Email: klaus.nordhausen@helsinki.fi

Affiliation: University of Helsinki

Abstract: Second-order source separation for vector-valued time series assumes that an observable stationary p -variate process can be represented as a linear mixture of p independent components. We extend this framework to the setting of multivariate functional stationary time series. The corresponding model is introduced and its identifiability properties are discussed. Furthermore, we generalize the classical AMUSE estimator to the functional setting under a separable covariance operator assumption. The performance of the proposed method is investigated through a simulation study.

Title: Functional Metric Learning for Classification Problems

Author: Ogden, Todd

Email: to166@columbia.edu

Affiliation: Columbia University

Abstract: Metric-based classification methods, such as K-nearest neighbors, are widely used, and their success often depends on the choice of distance measure. While metric learning has been shown to improve classification in multivariate settings, it faces challenges in high-dimensional spaces and when data are noisy. In many modern applications, including electroencephalograph (EEG) signals and brain images, data are more naturally represented as functional data, which exhibit inherent smoothness, continuity, and structural dependence. Motivated by this, we propose a functional metric learning framework that integrates functional principal component analysis (FPCA) with Large Margin Nearest Neighbors (LMNN) to define and learn an optimal semimetric that is tailored for functional data. Rather than relying on prespecified distances, our approach accounts for the structure of both one-dimensional and multidimensional functional data, accounts for noise, and adapts the semimetric to the classification task. Through simulations and an application to EEG data for alcoholism classification, we demonstrate that our method outperforms alternative approaches and demonstrates that frontal and parietal brain activity plays a crucial role in classification. These results establish functional metric learning as a flexible and powerful tool for improving classification in complex functional datasets.

Title: Modelling multiplex networks

Author: Olhede, Sofia

Email: sofia.olhede@epfl.ch

Affiliation: EPFL

Abstract: This talk will discuss how to model and make inferences of multiplex networks. One point-of-view starts from a special class of models based on multivariate graph limits (see Skeja and Olhede (2024)). These can be represented as a limiting object known as a decorated graphon (see Dufour and Olhede (2024)), where a more complex object is represented by a generalization of a graph limit. The choice of object will depend on the particular application at hand, and we will discuss the particular case of coma patients (Verdeyme et al (2025)). We will discuss choices of parameterization, and choices of summary statistics for particular models, highlighting both the benefits and challenges of multiplex observations. This is joint work with Charles Dufour, Jake Grainger, Anda Skeja, Karl Sawaya and Arthur Verdeyme.

Title: On the convex hull and coverage threshold of Ginibre eigenvalues

Author: Osman, Mohammed

Email: mogokzam@gmail.com

Affiliation: Queen Mary University of London

Abstract: We consider two geometric properties of Ginibre eigenvalues: the convex hull and the coverage threshold. The latter is the radius for which the union of disks of this radius centred at each eigenvalue covers a given region. We obtain limit theorems for the perimeter, area and number of vertices of the convex hull and the coverage threshold of disks.

Title: Inference on Generalized Latent Variable Models with High-Dimensional Responses and Covariates

Author: Ouyang, Jing

Email: jingoy@hku.hk

Affiliation: University of Hong Kong

Abstract: Regression models with both high-dimensional responses and covariates have attracted growing attention. Standard multivariate regression models become inadequate when the response variables depend not only on observed covariates but also on latent variables that capture key unobserved characteristics. To draw statistical inferences on covariate effects while accounting for latent variables, we consider a high-dimensional generalized latent variable model that accommodates mixed-type responses and allows for flexible dependence between covariates and latent variables, which is more suitable for many real-world applications than traditional methods that either rely on a linear regression form or restricted assumptions on the dependence between covariates and latent variables. We develop an alternating algorithm that iteratively updates the regression parameters and the latent variables, transforming an intractable nonconvex problem into a sequence of tractable convex subproblems. Theoretically, we provide algorithmic guarantees by deriving error bounds for the resulting estimator, proving statistical consistency, and characterizing its convergence rate. Further, building on this estimator, we construct debiased estimates for the covariate effects and establish their asymptotic normality.

Title: Stopping on the last success with unknown odds: Impossibility barriers and quantitative oracle bounds

Author: Paindaveine, Davy

Email: Davy.Paindaveine@ulb.be

Affiliation: Université libre de Bruxelles

Abstract: Optimal stopping problems lie at the interface of probability, statistics, and sequential decision theory. They model situations in which decisions must be made online, using only the information revealed so far and without access to future outcomes. We focus on the classical last-success problem: one observes a sequence of Bernoulli trials and seeks to stop exactly on the final success. When the success probabilities are known, the problem admits an elegant and complete solution via Bruss' sum-the-odds theorem (Bruss, 2000, Annals of Probability), which yields a simple threshold rule and an explicit formula for the optimal win probability. In most applications, however, these probabilities are unknown, so the oracle rule cannot be implemented. We therefore study the last-success problem under a minimal information structure, where the decision maker observes only the sequential Bernoulli outcomes. Our contribution provides a quantitative theory of oracle-freeness: we analyze the performance of the natural plug-in odds rule and, more importantly, characterize what is possible---and what is fundamentally impossible---when the success probabilities are unknown.

Title: A weak transport approach to the Schrödinger-Bass problem

Author: Pammer, Gudmund

Email: gudmund.pammer@tugraz.at

Affiliation: TU Graz

Abstract: The Schrödinger-Bass problem is a semimartingale optimal transport problem connecting the celebrated Schrödinger bridge with the Bass martingale problem. In this talk, I will explain how the dynamic Schrödinger-Bass problem admits an equivalent static formulation as a weak optimal transport problem with an explicit cost. This reformulation yields existence and duality results, provides a structural characterization of the optimal semimartingales, and naturally leads to a Sinkhorn-type iteration for computation.

Title: State evolution beyond first order methods

Author: Pananjady, Ashwin

Email: ashwinpm@gatech.edu

Affiliation: Georgia Tech

Abstract: We consider the dynamics of iterative optimization algorithms when applied to instances with high-dimensional, random data. When the algorithm of choice is a first-order method, it is known that the dynamics of the method are well approximated by a low-dimensional deterministic recursion known as state evolution. In this paper, we move beyond first-order methods and develop a rigorous state evolution for a far larger set of algorithms. We show that this state evolution can be written in a “canonical form”, allowing us to argue existence and uniqueness of the deterministic updates. Along the way, we develop a variant of Bolthausen’s conditioning method that relies on a sequential variant of Gordon’s Gaussian comparison inequality and enables a fully non-asymptotic analysis.

Title: Hierarchical Clustering with Confidence

Author: Panigrahi, Snigdha

Email: psnigdha@umich.edu

Affiliation: University of Michigan

Abstract: Agglomerative hierarchical clustering is one of the most widely used approaches for exploring how observations in a dataset relate to each other. However, its greedy nature makes it highly sensitive to small perturbations in the data, often producing different clustering results and making it difficult to separate genuine structure from spurious patterns. In this talk, I show how randomizing hierarchical clustering can be used not only to assess stability, but also to enable valid hypothesis testing based on the clustering results. We introduce a simple randomization scheme together with a test at each node of the hierarchical clustering dendrogram that controls the Type I error rate. Our test works with any hierarchical linkage without case-specific derivations, is more powerful than existing selective inference methods, and can be used to estimate the number of clusters with a guarantee against overestimation.

Title: From Atmospheric Rivers to Flood Risk: A Multivariate Model for Extreme Precipitation

Author: Panorska, Anna

Email: ania@unr.edu

Affiliation: University of Nevada Reno, USA

Abstract: Water resources and flood risk in the western United States are commonly driven by large scale precipitation events

lasting several consecutive days. Some of the largest precipitation events are driven by the Atmospheric Rivers, the “rivers in the sky”. We propose a multivariate model for the extreme precipitation events describing their duration, intensity, and maximum. We also link it to the atmospheric flow patterns, like those producing the Atmospheric Rivers. As the most important questions related to the impact of extreme precipitation are about the risk or probability of these extreme events, our ultimate goal is to employ the statistical model along with the meteorological and statistical methods leading to estimation of these probabilities.

Title: Tail-oriented Factor Models Using Expectiles

Author: Park, Seeun

Email: jo9373@snu.ac.kr

Affiliation: Seoul National University

Abstract: We propose a new factor model that captures tail-sensitive dependence structures in high-dimensional panel data, called the expectile factor model (EFM). While conventional factor models primarily focus on central co-movements, EFM targets conditional expectiles and extracts latent factors that characterize common dynamics in the tails. Importantly, the use of expectiles allows the model to exploit their smoothness and tail sensitivity through an asymmetric least squares criterion, which offers two key advantages. First, it yields numerically stable estimation even at very low or high expectile levels. Second, it enhances the ability to capture extreme events. We develop an iterative estimation algorithm and investigate the consistency and convergence rates of the proposed estimators. We further propose two consistent procedures for determining the number of factors, based on an information criterion and a rank-based approach. Simulation studies demonstrate that EFM remains numerically stable in extreme regions and better captures tail-driven latent components compared to existing quantile-based approaches. Empirical analyses of U.S. stock return data show that expectile factors provide substantial explanatory power for downside market risk, especially during periods of market stress. An additional application to macroeconomic forecasting demonstrates that tail-sensitive expectile factors capture information overlooked by mean-based factors and improve predictive performance.

Title: Application of control theory to online parameter estimation

Author: Parry, Matthew

Email: matthew.parry@otago.ac.nz

Affiliation: University of Otago

Abstract: Control theory studies feedback-driven decision-making for dynamical systems. In engineering, a great deal is known about how to design adaptive control processes to stabilise, regulate, or optimise system behaviour over time. These ideas can be applied in statistics to the problem of online parameter estimation. In this framework, the prediction-generating process is treated as the system, and the prediction error as the feedback signal used to update parameters from sequentially observed data. Prediction-error methods, for example, estimate parameters by minimising a loss function based on the one-step-ahead prediction error. We develop a learning framework based on proportional–integral–derivative (PID) controllers for online parameter estimation. We show that the integral component removes persistent error, while the derivative component helps to damp over-correction. In addition to connecting PID learning to momentum-based and adaptive optimisation methods, we demonstrate that PID learning can remain stable with a constant learning rate and is robust to time-varying data-generating processes.

Title: Revisiting M-estimation: risk bounds, optimality results, and probabilistic connections

Author: Pathak, Reese

Email: pathak.reese@gmail.com

Affiliation: UC Berkeley / Cornell University

Abstract: M-estimation procedures, such as empirical risk minimization (ERM) and least squares estimation (LSE), are pervasive in statistics and machine learning. Surprisingly, however, a complete understanding of their statistical properties is lacking, even in the most basic settings. In this talk we discuss recent results on the LSE and ERM, including recently obtained risk bounds, deviation inequalities, and optimality and suboptimality results that improve upon the classical ones. We also discuss some new connections between the analysis of these estimators and important probabilistic and geometric quantities such as the Gaussian width. Based on a series of joint works with: L. Aolaritei (UC Berkeley), M. I. Jordan (UC Berkeley / INRIA), A. Ulichney (UC Berkeley), N. Zhivotovskiy (UC Berkeley).

Title: On a Property of Conditional Expectation for Markov Kernels

Author: Pérez Fernández, Paloma

Email: palomaperezfernandez@gmail.com

Affiliation: Universidad de Extremadura

Abstract: A well-known property of conditional expectation states that, given three random variables such that the third is independent of the other two, then the conditional expectation of the first given the last two coincides with the conditional expectation of the first given the second. A Markov kernel can be considered as an extension of the concept of a random variable (and of a sub-sigma-field), and the authors have studied in previous works extensions to Markov kernels of certain probabilistic or statistical concepts (including independence and conditional expectation) and their properties. The main result of this communication considers the extension of the aforementioned result. A version in terms of densities is also presented.

Title: DNA replication and cell division cycles coordination in *E. coli*: a mathematical approach to cell physiology

Author: Perrin, Alexandre

Email: alexandre.perrin@polytechnique.edu

Affiliation: CMAP Ecole Polytechnique

Abstract: Despite extensive research, the quantitative principles that govern the coordination between DNA replication and cell division in bacteria remain debated. Multiple theoretical models have been proposed, some postulating that a single regulatory process is sufficient to ensure replication–division coordination, while others argue that two concurrent processes are required for robust control. To enable the comparison of these approaches, we developed a unifying mathematical framework within which models can be consistently formulated and quantitatively compared. Through theoretical analysis, this talk will present the necessary and sufficient conditions under which independent replication and division cycles recover physiological cell behaviours. Beyond the correlation-based analyses extensively used to date, this talk will discuss within a comprehensive statistical framework that double-process models more accurately recapitulate experimental data across all growth conditions. Finally, a novel model will be introduced that robustly captures the replication-division coordination in every growth regime, thereby providing a foundation for future mechanistic studies.

Title: Improved estimators for functions of scale parameters in mixture models

Author: Petropoulos, Constantinos

Email: costas@math.upatras.gr

Affiliation: Department of Mathematics, University of Patras, Greece

Abstract: Estimation of the scale parameter of the scale mixture of a location–scale family under the scale-invariant loss function is considered. The technique of Strawderman (Ann Stat 2(1):190–198, 1974) is used to obtain a class of estimators improving upon the best affine equivariant estimator of the scale parameter under certain conditions. Further, integral expressions of risk difference approach of Kubokawa (Ann Stat 22(1):290–299, 1994) is used to derive similar improvements for the reciprocal of the scale parameter. Using the improved estimator of the scale parameter and the improved estimator of the reciprocal of the scale parameter, classes of improved estimators for the ratio of scale parameters of two populations have been derived. In particular, Stein type and Brewster–Zidek type estimators are provided for the ratio of scale parameters of two mixture models. These results are applied to the scale mixture of exponential distributions, this includes the multivariate Lomax and the modified Lomax distributions.

Title: A class of continuous-time Gaussian mean-variance mixture state-space models simultaneously accounting for asymmetry, tailedness, and peakedness

Author: Podgorski, Krzysztof

Email: krys.podgorski@gmail.com

Affiliation: Lund University

Abstract: The generalized inverse Gaussian (GIG) distribution underpins generalized hyperbolic (GH) models used in Gaussian mean–variance mixtures to capture asymmetry, peakedness, and heavy tails. However, although GIG distributions are infinitely divisible (ID), they are not closed under convolution, limiting their use in continuous-time and Lévy-based frameworks. This paper introduces a related class of ID mixing distributions, termed IGG, defined as a convex combination of inverse Gaussian and gamma distributions. The IGG class retains the same number of parameters as GIG and, when used in Gaussian mean–variance mixtures, captures location, scale, skewness, and kurtosis, with an additional parameter governing

the mixing proportion, enhancing flexibility. A key advantage is closure under convolution, ensuring consistency under time scaling and irregular sampling. This enables direct use in continuous-time modeling via subordination of Brownian motion with drift to an IGG Lévy process, yielding Gaussian–IGG (NIGG) distributions. The class includes the normal–inverse Gaussian and generalized asymmetric Laplace as special cases. With intuitive parameter interpretation, the model avoids ambiguities common in GH-based time models and is well-suited for state-space formulations, as illustrated by stochastic volatility applications.

Title: An Alternative Copula Construction for Samples Exhibiting Duplicate Component Values

Author: Provost, Serge

Email: provost@stats.uwo.ca

Affiliation: The University of Western Ontario

Abstract: After a brief introductory overview, the presentator will introduce a construction that maps pseudo-observations obtained from a sample onto the cells of a grid partitioning the unit square, producing a random vector having uniformly distributed margins. This formalism is then refined into a coherent framework for representing stochastic dependence in settings where coordinate values may coincide. It will be shown that Sklar’s theorem continues to hold within this extended context, thereby situating the approach within a familiar theoretical foundation. A smooth version of the resulting distribution is subsequently achieved through Bernstein polynomial approximation. The presentation will feature illustrative examples supported by graphical displays, and a connection to a notable work of art will be highlighted.

Title: SYNTAX-GUIDED DIFFUSION LANGUAGE MODELS WITH USER-INTEGRATED PERSONALIZATION

Author: Qu, Annie

Email: aqu2@ucsb.edu

Affiliation: University of California Santa Barbara

Abstract: Large language models have achieved remarkable progress in generating human-like text, yet their outputs often lack structural diversity, limiting personalized expression. Recent advances in diffusion models offer new opportunities for text generation, surpassing the limitations of autoregressive paradigms. Building on the idea of hierarchical modeling, we develop a syntax-guided diffusion language model that incorporates latent syntactic variables to enhance text quality, diversity, and personalized conditioning. The proposed framework decomposes the text distribution into a syntactic prior and a conditional text distribution, enabling interpretable control over both syntactic and lexical components. We design a cascaded diffusion process for syntax-to-text generation and extend it to a generalized noncascaded formulation with a unified attention mechanism. To achieve fine-grained personalization, we introduce shared-latent representations that integrate information across users to capture style-specific lexical and syntactic patterns, supporting zero-shot inference. Extensive experiments demonstrate that the proposed approach substantially improves fluency, diversity, and stylistic fidelity. Further qualitative analyses highlight its interpretability and flexibility in learning personalized patterns.

Title: Community Detection and Modularity in the Preferential Attachment Block Model

Author: Racz, Miklos

Email: miklos.racz@northwestern.edu

Affiliation: Northwestern University

Abstract: I will discuss community detection in the preferential attachment block model (PABM). The key to the results is a detailed understanding of modularity of partitions in PABM, which may be of independent interest. This is based on joint work with Shuwen Chai.

Title: A Minimum Covariance Determinant Approach for Matrix and Tensor Data

Author: Radojicic, Una

Email: una.radojicic@tuwien.ac.at

Affiliation: TU Wien

Abstract: This work introduces the Matrix Minimum Covariance Determinant (MMCD), a robust estimator for location and covariance tailored to matrix-valued data. Unlike classical approaches that require vectorization—often leading to high-dimensional problems—MMCD preserves the inherent matrix structure and consistently estimates the mean matrix along with rowwise and columnwise covariance matrices within matrix-variate elliptical distributions. The estimators are matrix-affine equivariant and achieve a higher breakdown point than any affine equivariant multivariate estimator applied to

vectorized data. An efficient algorithm with convergence guarantees is proposed, enabling practical implementation. MMCD-based robust Mahalanobis distances provide a reliable tool for outlier detection. Furthermore, Shapley values are extended to the matrix setting, allowing decomposition of squared Mahalanobis distances into contributions from rows, columns, or individual cells. The framework naturally extends to tensor-valued data, enabling analogous robust estimation and interpretability across higher-order structures. Both theoretical results and simulations demonstrate that MMCD outperforms vectorization-based methods in robustness and computational efficiency, supported by real-world applications.

Title: Elements of Interactive Decision Making

Author: Rakhlin, Alexander

Email: rakhlin@mit.edu

Affiliation: MIT

Abstract: Machine learning methods are increasingly deployed in interactive environments, ranging from dynamic treatment strategies in medicine to fine-tuning of LLMs using reinforcement learning. In these settings, the learning agent interacts with the environment to collect data and necessarily faces an exploration-exploitation dilemma. We present a general framework for interactive decision making that subsumes multi-armed bandits, contextual bandits, structured bandits, and reinforcement learning. We focus on both the statistical aspect of learning---aiming to develop a tight characterization of sample complexity in terms of properties of the class of models---and on the basic algorithmic primitives.

Title: Quantitative Unimodular Propagation of Chaos and its Consequences

Author: Ramanan, Kavita

Email: Kavita_Ramanan@brown.edu

Affiliation: Brown University

Abstract: Interacting particle systems, or large collections of randomly evolving particles that interact locally with respect to an underlying network or graph, model a plethora of phenomena in physics, engineering, neuroscience, epidemiology and machine learning. When the underlying network is sparse, as is the case of many real-world networks, neighboring particles remain strongly correlated and there is no propagation of chaos in the classical sense. However, recent work has shown that there is a certain unimodular propagation of chaos and that the limit of the empirical measure can be characterized as the law of a local-field equation. We establish quantitative versions of unimodular propagation of chaos and describe some consequences, including showing how the local-field equation naturally interpolates to the mean-field equation as the graph size goes to infinity. The proofs combines techniques from stochastic analysis and random graph theory.

Title: Assumption-lean weak limits and tests for two-stage adaptive experiments

Author: Ren, Zhimei

Email: zren@wharton.upenn.edu

Affiliation: University of Pennsylvania

Abstract: Adaptive experiments are becoming increasingly popular in real-world applications for effectively maximizing in-sample welfare and efficiency by data-driven sampling. Despite their growing prevalence, however, the statistical foundations for valid inference in such settings remain underdeveloped. Focusing on two-stage adaptive experimental designs, we address this gap by deriving new weak convergence results for mean outcomes and their differences. In particular, our results apply to a broad class of estimators, the weighted inverse probability weighted (WIPW) estimators. In contrast to prior works, our results require significantly weaker assumptions and sharply characterize phase transitions in limiting behavior across different signal regimes. Through this common lens, our general results unify previously fragmented results under the two-stage setup. We further establish quantitative convergence rates in bounded-Lipschitz distance that reveal the fundamental trade-off between exploitation and inferential stability. To address the challenge of potential non-normal limits in conducting inference, we propose a computationally efficient and provably valid simulation-based method for obtaining critical values of the non-normal limiting distributions under the null, enabling practical hypothesis testing. Our results and approaches are sufficiently general to accommodate various adaptive experimental designs, including batched bandit and subgroup enrichment experiments. Simulatio

Title: Debiasing Score Decompositions for Reliable Inference

Author: Resin, Johannes

Email: resin@econ.uni-frankfurt.de

Affiliation: Goethe University Frankfurt

Abstract: Proper scoring rules are widely used to quantify the predictive performance of probabilistic predictions. Score decompositions uncover contributions of miscalibration (reliability) and discrimination (resolution) on the overall predictive performance. Recent work has proposed the use of isotonic regression to estimate score decompositions nonparametrically, allowing for consistent estimation under a natural monotonicity assumption. We demonstrate that these estimators are susceptible to bias and show empirically that cross-validation can be used to effectively debias the decomposition in simulations. We then discuss how to leverage large sample theory for cross-validated generalization errors to enable inference for the debiased score components. Finally, we illustrate the debiased decomposition in an application to common image classifiers from the machine learning literature.

Title: Robust and Sub-Gaussian Estimation of Means and Covariances

Author: Rico, Zoraida

Email: zoraida.rico@unibocconi.com

Affiliation: Bocconi University

Abstract: Estimating the mean and covariance under heavy tails and adversarial contamination remains a central challenge in robust statistics. In this talk, we revisit the classical trimmed mean estimator for one-dimensional mean estimation, providing new finite-sample insights. We show that the trimmed mean achieves optimal performance in this setting and satisfies a strong form of the CLT. We then turn to covariance estimation for a d -dimensional random vector from an i.i.d. sample. We show that, even in the presence of relatively heavy tails and adversarial contamination, this estimator achieves the optimal dimension-free rate of convergence.

Title: A stochastic continuous-space model and its large population limit for the dynamic of actin polymerisation and depolymerisation in the cell

Author: Ringeard, Timothé

Email: timothe.ringeard@gmail.com

Affiliation: Université Paris Cité

Abstract: The network of actin proteins in the cell is responsible for many structural characteristics and behaviour. In particular, during cell division, actin monomers form very long polymers around the equator, creating a ring that will contract itself to separate the two new cells. In order to better understand this phenomenon, we model the situation by a stochastic process where actin monomers follow Brownian diffusion, and can agglomerate into polymers at a prescribed location, corresponding to the contractile ring. The macroscopic behaviour of the process can be understood by a large population limit, in which the number of actin monomers goes to infinity. With the right scaling of the parameters, the limit is deterministic and follows a PDE with a singularity at the location of the polymers.

Title: Asymptotic equivalence for nonparametric additive regression

Author: Rohde, Angelika

Email: angelika.rohde@stochastik.uni-freiburg.de

Affiliation: University of Freiburg

Abstract: We prove asymptotic equivalence of nonparametric additive regression and an appropriate Gaussian white noise experiment in which a multidimensional shifted Wiener process is observed, whose dimension equals the number of additive components. The shift depends on the additive components of the regression function and solely the one- and two-dimensional marginal distributions of the covariates via an explicitly specified bounded but non-compact linear operator Γ . The number of additive components d is allowed to increase moderately with respect to the sample size. In the special case of pairwise independent components of the covariates, the white noise model decomposes into d independent univariate processes. Moreover, we study approximation in some semiparametric setting where Γ splits into a multiplication operator and an asymptotically negligible Hilbert-Schmidt operator.

Title: The distributionally robust prediction error of the square root LASSO and related estimators

Author: Rush, Cynthia

Email: cynthia.rush@columbia.edu

Affiliation: Columbia University

Abstract: We study the classical problem of predicting an outcome variable using a linear combination of a d -dimensional

covariate vector. We are interested in linear predictors whose coefficients are generated by the square root LASSO and related estimators. We provide conditions under which linear predictors based on these estimators minimize the worst-case prediction error over a ball of distributions determined by a type of max-sliced Wasserstein metric. A detailed analysis of the statistical properties of this metric yields a simple recommendation for the choice of regularization parameter. The suggested order of the regularization parameter, after a suitable normalization of the covariates, is typically given by the dimension d divided by the sample size, up to logarithmic factors. Our recommendation is computationally straightforward to implement, pivotal, has provable out-of-sample performance guarantees, and does not rely on sparsity assumptions about the true data generating process. This is joint work with Jose Montiel Olea, Amilcar Velez and Johannes Wiesel.

Title: Chaos and Riemann zeta function on the critical line

Author: Saksman, Eero

Email: eero.saksman@helsinki.fi

Affiliation: University of Helsinki

Abstract: Abstract: We will discuss some results, both old and new (in preparation), on the statistics of the Riemann zeta function on short intervals on the critical line. The talk is based on collaboration with Adam Harper (Warwick) and Christian Webb (Helsinki).

Title: Learn the score

Author: Samworth, Richard

Email: r.samworth@statslab.cam.ac.uk

Affiliation: University of Cambridge

Abstract: Score estimation has recently emerged as a key modern statistical challenge, due to its pivotal role in generative modelling via diffusion models. Moreover, it is an essential ingredient in a new approach to linear regression via convex M-estimation, where the corresponding error densities are projected onto the log-concave class. Motivated by these applications, we study the minimax risk of score estimation with respect to squared squared error loss, weighted by the true underlying log-concave distribution. Such distributions have decreasing score functions, but on its own, this shape constraint is insufficient to guarantee a finite minimax risk. We therefore define subclasses of log-concave densities that capture two fundamental aspects of the estimation problem. First, we establish the crucial impact of tail behaviour on score estimation by determining the minimax rate over a class of log-concave densities whose score function exhibits controlled growth relative to the quantile levels. Second, we explore the interplay between smoothness and log-concavity, and determine the minimax rate up to poly-logarithmic factors. Our upper bounds are attained by a locally adaptive, multiscale estimator constructed from a uniform confidence band for the score function. This study highlights intriguing differences between the score estimation and density estimation problems over this shape-constrained class.

Title: Statistical Analysis of Forensic Black Box Studies

Author: Saunders, Christopher

Email: christopher.saunders@sdstate.edu

Affiliation: South Dakota State University

Abstract: Following the recommendations of the President's Council of Advisors on Science and Technology, researchers in forensic science have focused on the foundational validity of forensic source identification disciplines. A commonly performed study focuses on treating the forensic examiner as a "black box," providing the examiner with a pair of items to compare and recording the result of the forensic examiner's analysis on an ordinal point scale. The results of the study are then summarized in a manner that measures either the error rates or accuracy of the forensic examiners when applying the standard techniques to a given population of sources. A typical black box study design has the following components: a sample of forensic examiners, a sample of sources, and a sample of objects from each source. These different components of variability make it difficult to approximate the standard error of the examiner accuracy. This difficulty has led to numerous methods being proposed for constructing confidence intervals that all lead to different results with the degree of the difference being dependent on the number of examiners and sources used in the study. In this presentation, we will review the experimental design of these studies, the various proposed methods for analyzing the experimental results, and use a functional data based Monte Carlo experiment to compare the different statistical approaches to constructing confidence intervals in these studies.

Title: Provable false discovery rate control for deep feature selection

Author: Sawaya, Kazuma

Email: sawaya@g.ecc.u-tokyo.ac.jp

Affiliation: University of Tokyo

Abstract: We develop a flexible feature selection framework based on deep neural networks that approximately controls the false discovery rate (FDR), a measure of Type-I error. The method applies to architectures whose first layer is fully connected. From the second layer onward, it accommodates multilayer perceptrons (MLPs) of arbitrary width and depth, convolutional and recurrent networks, attention mechanisms, residual connections, and dropout. The procedure also accommodates stochastic gradient descent with data-independent initializations and learning rates. To the best of our knowledge, this is the first work to provide a theoretical guarantee of FDR control for feature selection within such a general deep learning setting. Our analysis is built upon a multi-index data-generating model and an asymptotic regime in which the feature dimension n diverges faster than the latent dimension q , while the sample size, the number of training iterations, the network depth, and hidden layer widths are left unrestricted. Under this setting, we show that each coordinate of the gradient-based feature-importance vector admits a marginal normal approximation, thereby supporting the validity of asymptotic FDR control. As a theoretical limitation, we assume B -right orthogonal invariance of the design matrix, and we discuss broader generalizations. We also present numerical experiments that underscore the theoretical findings.

Title: Non-Parametric Distributional Regression via Sklar's Theorem and Empirical Copula Approximations

Author: Schärer, Kai

Email: kai.schaerer@plus.ac.at

Affiliation: University of Salzburg

Abstract: Nonparametric regression is a powerful tool for estimating unknown regression functions, yet most existing methods treat each functional, such as conditional means, quantiles, or expectiles, separately. We answer the natural question of whether there is a unified framework for regression on a broad class of functionals, using nonparametric estimation of the conditional distribution functions. We show that this can be done by exploiting Sklar's theorem, which separates the joint distribution into a copula and its marginal components. Under suitably chosen resolution and mild regularity assumptions, we establish uniform convergence to the conditional distributions of the underlying copula using the empirical checkerboard and Bernstein approximations. By plugging in the univariate empirical marginal distribution functions, we establish uniform convergence to the joint conditional distribution function. By simple aggregation of the estimator, we obtain consistent non-parametric estimators for a broad class of regression functionals, including those mentioned earlier and many others. The method provides a clear computational advantage over traditional kernel-based regressors. Simulation studies and a real-data application illustrate the finite-sample performance and flexibility of the proposed approach.

Title: Critical Decoration of Branching Brownian Motion via Time-Inhomogeneous Additive Penalization Methods

Author: Schickentanz, Dominic T.

Email: mail@dominic-schickentanz.com

Affiliation: Technion and Paderborn University

Abstract: Our aim is to provide a construction of the decoration point measure appearing in the extremal process of the branching Brownian motion, by constructing the law of this process seen from its topmost particle. This construction involves the law of a Brownian motion penalized by a solution of the F-KPP equation. We therefore give a general description of the law of a Brownian motion penalized by a time-inhomogeneous potential, which is of independent interest. This is joint work with Bastien Mallein (Toulouse).

Title: Optimal Transport for Inverse Problems

Author: Schönlieb, Carola-Bibiane

Email: cbs31@cam.ac.uk

Affiliation: cbs31@cam.ac.uk

Abstract: Optimal transport has emerged as a powerful mathematical framework for comparing structured data distributions and has recently shown significant promise in inverse problems. In this talk, I will discuss how transport-based discrepancies, priors, and geometries can be leveraged to improve robustness, reconstruction quality, and interpretability in inverse imaging problems, particularly in settings involving sparse, noisy, or distribution-shifted data. I will further highlight recent developments connecting Wasserstein metrics with variational regularisation and modern data-driven approaches.

Title: Power Fréchet Means for Robust Statistics in Metric Spaces

Author: Schötz, Christof

Email: christof.schoetz@tum.de

Affiliation: Technical University of Munich

Abstract: The 2-Fréchet mean generalizes the expectation to metric spaces by minimizing the expected squared distance to a random object, while the 1-Fréchet mean generalizes the median. Interpolating between these two notions, the alpha-Fréchet means (or power Fréchet means), for alpha between one and two, inherit robustness properties that likewise interpolate between those of the expectation and the median. We establish finite-sample bounds in expectation for their empirical counterparts in Hadamard spaces (complete metric spaces of non-positive curvature) under the sole assumption of a finite moment of order alpha. Notably, the moment of order $(2\alpha - 2)$ emerges as the dominant term in the error bound, making power Fréchet means particularly suitable for heavy-tailed distributions. These results hold in infinite-dimensional settings and are new even in Euclidean spaces.

Title: Non-parametric recovery of causal diffusion mechanisms from equilibrium observations

Author: Schwank, Richard

Email: richard.schwank@tum.de

Affiliation: Technical University of Munich

Abstract: When studying a causal system, it can be beneficial to consider its time evolution, particularly if there are feedback loops that are expected to unfold at small timescales. However, in many practical scenarios (e.g., for single-cell gene expression data), observations can only be obtained at a single point in time. We present methodology to recover the system's time-infinitesimal transition mechanism from such cross-sectional data. Precisely, we assume the system follows a time-homogeneous diffusion process that has reached an equilibrium distribution at observation time. Further, we assume the causal mechanism is fully described by the diffusion drift, is acyclic, and its causal graph is known. In this setting, we show that the full causal mechanism, i.e., the drift function, can be non-parametrically identified under mild conditions, and we provide a practical algorithm to solve this challenging inverse problem.

Title: Consistency of fraction of variance estimation from summary statistics in high-dimensional linear models

Author: Schwartzman, Armin

Email: armins@ucsd.edu

Affiliation: University of California, San Diego

Abstract: In Genome-Wide Association Studies (GWAS), heritability is defined as the fraction of variance of an outcome explained by a large number of genetic predictors in a high-dimensional polygenic linear model. This work studies the asymptotic properties of the most common estimator of heritability from summary statistics called linkage disequilibrium score (LDSC) regression, together with a simpler and closely related estimator called GWAS heritability (GWASH). We show that, with some variations, two conditions which we call weak dependence (WD) and bounded-kurtosis effects (BKE) are sufficient for consistency of both the basic LDSC with fixed intercept and GWASH estimators, for both Gaussian and non-Gaussian predictors. For Gaussian predictors it is shown that these conditions are also necessary for consistency of GWASH (with truncation) and simulations suggest that necessity holds too when the predictors are non-Gaussian. We also show that, with properly truncated weights, weighting does not change the consistency results, but standardization of the predictors and outcome, as done in practice, introduces bias in both LDSC and GWASH if the two essential conditions are violated. Finally, we show that, when population stratification is present, all the estimators considered are biased, and the bias is not remedied by using the LDSC regression estimator with free intercept, as originally suggested by the authors of that estimator.

Title: The contact process on dynamical random trees with degree dependence

Author: Seiler, Marco

Email: seiler@math.uni-frankfurt.de

Affiliation: Goethe University Frankfurt

Abstract: The contact process models the spread of an infection in a structured population given through a graph, where individuals are infected with a certain rate and recover with rate 1. If this graph is a Bienaymé-Galton-Watson (BGW) tree, the contact process exhibits a well-known dichotomy in its survival behaviour. If the offspring distribution has no exponential moments, then the phase transition of survival is trivial, which means that for every infection rate there is a positive probability to survive. On the other hand, if some exponential moments exist, then the process dies out almost surely for small infection

rates. In this talk we study the contact process on a dynamically evolving BGW tree. Initially edges are assigned an open/closed state with a degree-dependent connection probability. Additionally, edges update at a degree-dependent rate, after which each edge is independently reopened or closed with the same degree-dependent probability as before. Finally, the contact process is only allowed to spread via open edges. We determine the impact of the update speed and the connection probability on the existence of a non-trivial phase transition. In particular, we provide conditions depending on the connection probability and update rate which guarantee a trivial phase transition. This talk is based on joint work with Natalia Cardona-Tobón, Marcel Ortgiese and Anja Sturm.

Title: Wasserstein-Cramer-Rao Theory of Unbiased Estimation

Author: Sen, Bodhisattva

Email: b.sen@columbia.edu

Affiliation: Columbia University

Abstract: The quantity of interest in the classical Cramer-Rao theory of unbiased estimation (i.e., the Cramer-Rao lower bound, exact efficiency in exponential families, and asymptotic efficiency of maximum likelihood estimation) is the variance, which represents the instability of an estimator when its value is compared to the value for an independently-sampled data set from the same distribution. In this talk we are interested in a quantity which represents the instability of an estimator when its value is compared to the value for an infinitesimal additive perturbation of the original data set; we refer to this as the "sensitivity" of an estimator. The resulting theory of sensitivity is based on the Wasserstein geometry in the same way that the classical theory of variance is based on the Fisher-Rao (equivalently, Hellinger) geometry, and this insight allows us to determine a collection of results which are analogous to the classical case: a Wasserstein-Cramer-Rao lower bound for the sensitivity of any unbiased estimator, a characterization of models in which there exist unbiased estimators achieving the lower bound exactly, and a guarantee that Wasserstein projection estimators achieve the lower bound asymptotically. We use these results to treat many statistical examples, sometimes revealing new optimality properties for existing estimators and other times revealing new estimators. This is joint work with Nicolas Garcia Trillos and Adam Jaffe.

Title: Spectral algorithms for signal recovery from multi-modal data

Author: Sen, Subhabrata

Email: subhabratasen@fas.harvard.edu

Affiliation: Harvard University

Abstract: We will study spiked Wigner matrices with correlated spikes. The model is motivated by multimodal Principal Component Analysis (PCA). We derive and analyze a spectral algorithm for signal recovery in this problem. The spectral operator is obtained by linearization of Approximate Message Passing (AMP). We discover a sharp phase transition in the behavior of this spectral operator ---above a critical SNR threshold, the spectrum has an outlier, and the corresponding top eigenvector is correlated with the latent signals. This generalizes well-known phase transitions for spectral methods used in unimodal data analysis. We will discuss inference for the latent signals based on the top eigenvector of this new spectral operator.

Title: Goodness-of-Fit Tests for the Survival Copula and Cross-Ratio Function based on Nonparametric Bernstein-based Estimators

Author: Sercik, Ömer

Email: omer.sercik@uhasselt.be

Affiliation: Hasselt University

Abstract: In multivariate statistics, and particularly in bivariate time-to-event data analysis, inference regarding the underlying dependence structure is of crucial importance. In recent years, copulas have been widely used for modeling dependence between random variables. Alongside the application of the copula framework, several global and local association measures have been developed. In this work, we study the dependence structure between two non-negative random variables focusing on the cross ratio function (CRF) as a measure of local association. As time-to-event data are often subject to (right) censoring, leading to partial information about these event times for some subjects, association measures such as the CRF should be estimated explicitly accommodating censoring. Most existing approaches for copula modeling and CRF estimation are parametric in nature, leading to the need of reliable methods to assess parametric assumptions. In this work, we propose Bernstein-based estimators for the survival copula and the CRF under both uncensored and univariate right-censored data. Specifically, we 1) introduce the proposed estimators for the survival copula and CRF, 2) develop goodness-of-fit tests for

survival copula and CRF estimation, and 3) evaluate the performance of the proposed tests through an in-depth simulation study. Finally, we illustrate the performance of the proposed tests based on real-life data-applications.

Title: Generalization via distributional learning

Author: Shen, Xinwei

Email: xwshen@uw.edu

Affiliation: University of Washington

Abstract: This talk focuses on generalization beyond the observed data distribution, including problems such as extrapolation, distribution shifts, and causal inference. These tasks require generalizing beyond what has been directly observed. We propose tackling such problems through a distributional perspective: instead of fitting only low-dimensional summaries like conditional means, we estimate the entire distribution of the observed data. While natural from an identifiability standpoint, this approach has been underexplored in estimation. We first introduce engression, a distributional learning method that blends the flexibility of generative models with proper scoring rules. In various settings, we demonstrate how distributional learning can help with better generalization better than non-distributional approaches. Applications span from single-cell prediction to large language model training.

Title: Treatment effects on Hadamard spaces

Author: Shin, Ha-Young

Email: hayoung.shin@gmail.com

Affiliation: Department of Statistics and Actuarial Science, Soongsil University

Abstract: We propose a new framework for causal inference with Hadamard space-valued response variables, in which treatment effects are defined by a non-negative real number representing magnitude and a point at infinity representing direction. In particular, the average treatment effect (ATE) is defined as a magnitude-direction pair satisfying the following: when the the control group is transported a distance equal to that magnitude in that direction, its Frechet mean coincides with that of the treatment group. We define an ATE estimator, explore properties including strong consistency, and detail algorithms for computation of the ATE, using Broyden's method, and confidence regions. Simulations in which we investigate consistency and confidence region coverage are included. We perform causal inference on real diffusion tensor imaging (DTI) data and compare with Euclidean alternatives. Interpretability, and the testability of the null hypothesis of zero ATE, are major advantages over existing attempts to generalize the ATE to non-linear spaces; this is also illustrated with the DTI data.

Title: Calibration in multiple hypothesis testing

Author: Soloff, Jake

Email: jakesoloff@gmail.com

Affiliation: University of Michigan

Abstract: Among error criteria for large-scale hypothesis testing, the local false discovery rate (lfdr) is the one most directly tied to the reliability of an individual finding. Its interpretation, however, traditionally relies on a Bayesian two-groups model. This talk develops new perspectives on empirical Bayes and compound-decision inference in this context. The guiding idea is to seek procedures that, without prior knowledge, attain in expectation the same local guarantees that an oracle Bayes rule attains almost surely. First, I introduce a simple, nonparametric procedure that exactly controls the expected maximum lfdr among the discoveries; equivalently, the probability that the least promising discovery is false. I will then discuss recent and ongoing work connecting multiple testing to probabilistic forecasting, giving a prior-free interpretation of lfdr as a calibrated conditional error probability and suggesting a broader framework for interpretation of compound decisions.

Title: Doubly regularized generalized linear models for spatial observations with high-dimensional covariates

Author: Sondhi, Arjun

Email: a2sondhi@gmail.com

Affiliation: Northwell Health

Abstract: A discrete spatial lattice can be cast as a network structure over which spatially correlated outcomes are observed. A second network structure may also capture similarities among measured features, when such information is available. Incorporating the network structures when analysing such doubly structured data can improve predictive power, and lead to better identification of important features in the data-generating process. Motivated by applications in spatial disease mapping, we develop a new doubly regularized regression framework to incorporate these network structures for analysing

high-dimensional datasets. Our estimators can be easily implemented with standard convex optimization algorithms. In addition, we describe a procedure to obtain asymptotically valid confidence intervals and hypothesis tests for our model parameters. We show empirically that our framework provides improved predictive accuracy and inferential power compared with existing high-dimensional spatial methods. These advantages hold given fully accurate network information, and also with networks which are partially misspecified or uninformative. The application of the proposed method to modelling COVID-19 mortality data suggests that it can improve the prediction of deaths beyond standard spatial models, and that it selects relevant covariates more often.

Title: Local inference for functional time series

Author: Sørensen, Helle

Email: helle@math.ku.dk

Affiliation: University of Copenhagen

Abstract: Local inference is a technique in functional data analysis where hypotheses are tested along the domain of the functional data - in principle involving infinitely many tests. The multiple testing problem is addressed by performing tests over intervals and exploiting the continuity of the data. In this talk we consider local inference for the situation where data is a time series with each data unit being a curve. We approach two questions: (1) In which parts of the functional domain is there dependence in the time series, and (2) how can local inference for hypotheses concerning fixed effects be carried out in the presence of dependence? Our tests are based on permutations and are thus non-parametric of nature. As an application we consider data consisting of daily ice cover in the Baffin Bay in the period from 1979 to 2023.

Title: An IDEAL algorithm for stream-based selective sampling in regression scenarios

Author: Spangl, Bernhard

Email: bernhard.spangl@boku.ac.at

Affiliation: BOKU University, Vienna

Abstract: In statistical learning and machine learning, the efficiency of data acquisition plays a crucial role in model performance, especially when labeled data is expensive or time-consuming to obtain. This challenge has given rise to the field of active learning, a paradigm where the learning algorithm actively selects the most informative data points for labeling to optimize the training process. While many active learning strategies consider query synthesis or pool-based sampling we focus on the problem of online active learning in regression scenarios. We propose a new algorithm to iteratively select new data points from a data stream adapting ideas of the Inverse Distance based Exploration for Active Learning (IDEAL) algorithm using thresholding of a properly chosen acquisition function. The IDEAL algorithm was originally developed for the pool-based sampling scenario and balances between model-based learning and learning based on the pure exploration of the feature-vector space to promote diversity. We compare the proposed active selection strategy to other state-of-the-art methods and to the passive selection strategy that selects the next input points from the unlabelled data stream at random with a given probability and serves as a baseline. The performance of our method is evaluated in numerical tests on illustrative synthetic problems and real-world datasets. This talk is based on a joint ongoing work with Xiaojian Xu.

Title: Tight generalization bounds in Inverse Optimization

Author: Sra, Suvrit

Email: s.sra@tum.de

Affiliation: TU Munich

Abstract: Inverse optimization (IO) seeks to infer the parameters of a decision-maker's objective from observed context-action data. We study noiseless IO, where demonstrations are generated by a ground-truth objective. We provide a high-probability $O(d/T)$ generalization bound for the induced action set, where d is the number of unknown parameters and T is the size of the training dataset. We strengthen these guarantees under additional conditions that ensure uniqueness of the chosen action, bringing our IO guarantees in line with best-arm identification results in the bandit literature. We further show that the $O(d/T)$ rate is tight over all consistent estimators considered here, and extend the result to both instantaneous and cumulative regret. Notably, the resulting regret lower bound matches the corresponding upper bounds in the adversarial setting, indicating that the stochastic IO setting is effectively adversarial for the class of estimators studied here. Finally, we propose a parameter-free algorithm with lower per-iteration complexity than generic solvers. Experiments validate the predicted rates and illustrate the tightness of our bounds.

Title: Spectral regularized kernel two-sample tests

Author: Sriperumbudur, Bharath

Email: bharathsv.ucsd@gmail.com

Affiliation: The Pennsylvania State University

Abstract: Over the last decade, an approach that has gained a lot of popularity in tackling non-parametric testing problems on general (i.e., non-Euclidean) domains is based on the notion of reproducing kernel Hilbert space (RKHS) embedding of probability distributions. The main goal of our work is to understand the optimality of two-sample tests constructed based on this approach. First, we show that the popular MMD (maximum mean discrepancy) two-sample test is not optimal in terms of the separation boundary measured in Hellinger distance. Second, we propose a modification to the MMD test based on spectral regularization by taking into account the covariance information (which is not captured by the MMD test) and prove the proposed test to be minimax optimal with a smaller separation boundary than that achieved by the MMD test. Third, we propose an adaptive version of the above test, which involves a data-driven strategy to choose the regularization parameter and show the adaptive test to be almost minimax optimal up to a logarithmic factor. Moreover, our results hold for the permutation variant of the test, where the test threshold is chosen elegantly through the permutation of the samples. Through numerical experiments on synthetic and real-world data, we demonstrate the superior performance of the proposed test in comparison to many popular two-sample tests.

Title: Spectral methods for graph inference: the case of community detection

Author: Stephan, Ludovic

Email: ludovic.stephan@ensai.fr

Affiliation: ENSAI - CREST

Abstract: Community detection is the task of recovering a hidden cluster structure from network information. Spectral methods, especially those based on the graph adjacency matrix, are a very popular family of algorithms but they can fail when the average graph degree is too low. I will present several remediations to this phenomenon, from the celebrated non-backtracking method and its variants to more efficient but less theoretically tractable methods such as the Bethe Hessian.

Title: Non-convex matrix sensing: Breaking the quadratic rank barrier in the sample complexity

Author: Stöger, Dominik

Email: dominik.stoeger@ku.de

Affiliation: KU Eichstätt-Ingolstadt

Abstract: For the problem of reconstructing a low-rank matrix from a few linear measurements, two classes of algorithms have been widely studied in the literature: convex approaches based on nuclear norm minimization, and non-convex approaches that use factorized gradient descent. Under certain statistical model assumptions, it is known that nuclear norm minimization recovers the ground truth as soon as the number of samples scales linearly with the number of degrees of freedom of the ground-truth. In contrast, while non-convex approaches are computationally less expensive, existing recovery guarantees assume that the number of samples scales at least quadratically with the rank r of the ground-truth matrix. In this talk, we close this gap by showing that the non-convex approaches can be as efficient as nuclear norm minimization in terms of sample complexity. Namely, we consider the problem of reconstructing a positive semidefinite matrix from a few Gaussian measurements. We show that factorized gradient descent with spectral initialization converges to the ground truth at a linear rate as soon as the number of samples scales linearly in the degrees of freedom of the ground truth. This improves the previous rank-dependence in the sample complexity of non-convex matrix factorization from quadratic to linear. Furthermore, we extend our theory to the noisy setting where we achieve minimax optimal error rates.

Title: Discussion on Open Questions Involving Moments and Cumulants

Author: Stoyanov, Jordan

Email: stoyanovj@gmail.com

Affiliation: Bulgarian Academy of Sciences

Abstract: The IMS meeting is an excellent opportunity for a comprehensive discussion on outstanding questions related to moments and cumulants. For non-linear Box-Cox transformations of random data, it is not clear whether the M- or C-determinacy of the transformed data is preserved or lost. To study determinacy in conjunction with infinite divisibility poses extra challenges. Numerous questions remain unanswered. More questions are related to the important qualitative result by C. Berg and co.: If a distribution F possesses finite moments and is M-indeterminate, there exist infinitely many distributions of

any kind, continuous, discrete, singular, all sharing the same moments as F . There are several checkable conditions that enable us to establish M -(in)determinacy for absolutely continuous and discrete distributions. Only a few instances are known where infinite families of both continuous and discrete distributions are explicitly identified. The Log-Normal distribution serves as a notable example. The M -indeterminacy of power 3 of a r.v. with either a normal, exponential or IG distribution can be shown in a few ways, we still do not know any 'discrete' solution in these and many other cases. I will present several open questions arising from my research. Attendees of the discussion will have the opportunity to share their own open questions. Additionally, colleagues unable to attend the meeting have graciously provided their open questions to be addressed publicly.

Title: Optimality of Self-distillation in Spiked Covariance Models

Author: Sur, Pragya

Email: pragya@fas.harvard.edu

Affiliation: Harvard University

Abstract: In modern artificial intelligence pipelines, self-distillation has emerged as a promising technique for improving model performance without requiring additional data. Yet it remains unclear when self-distillation yields improved out-of-sample risk or better estimation compared to commonly used statistical estimators. This talk will present a rigorous framework for analyzing the performance of a broad class of estimators in high dimensions, which we term spectral-shrinkage estimators. We show that for spiked covariance models, self-distillation achieves uniquely optimal prediction and estimation performance, outperforming well-known methods such as ridge regression, minimum norm interpolation, and principal components regression. This stands in contrast to the behavior observed under isotropic covariances, where suitably tuned ridge regression is optimal among these shrinkage estimators. Time permitting, we will also present a federated learning setting in which multiple data centers share their self-distilled estimators and a central server seeks to aggregate them optimally. We characterize optimal aggregation strategies in this context. Together, these results elucidate why self-distillation improves predictive performance and provide a broader statistical framework that connects it to other shrinkage-based methods.

Title: Consensus tree estimation with false discovery control via partially ordered set

Author: Taeb, Armeen

Email: ataeb@uw.edu

Affiliation: University of Washington

Abstract: Connected acyclic graphs (trees) are data objects that hierarchically organize categories. Collections of trees arise in a diverse variety of fields, including evolutionary biology, public health, machine learning, social sciences and anatomy. Summarizing a collection of trees by a single representative is challenging, in part due to the dimension of both the sample and parameter space. We frame consensus tree estimation as a structured feature-selection problem, where leaves and edges are the features. We introduce a partial order on leaf-labeled trees, use it to define true and false discoveries for a candidate summary tree, and develop novel estimation algorithms that control the family-wise error rate and the false discovery rate at a nominal level for a broad class of non-parametric generative models. Importantly, our method accommodates unequal leaf sets and non-binary trees, allowing the estimator to reflect uncertainty by collapsing poorly supported structure instead of forcing full resolution. We apply the method to study the archaelal origin of eukaryotic cells and to quantify uncertainty in deep branching orders. While consensus tree construction has historically been viewed as an estimation task, reframing it as feature selection over a partially ordered set allows us to obtain the first estimator with finite-sample and model-free guarantees.

Title: Model-oriented distances between causal graphs via posets

Author: Taeb, Armeen

Email: ataeb@uw.edu

Affiliation: University of Washington

Abstract: A well-defined distance on the parameter space is key to evaluating estimators, ensuring consistency, and building confidence sets. While there are typically standard distances to adopt in a continuous space, this is not the case for combinatorial parameters such as graphs that represent statistical models. Defined on the graphs alone, existing proposals like the structural Hamming distance ignore the structure of the model space and can thus exhibit undesirable behaviors. We propose a model-oriented framework for defining the distance between graphs that is applicable across different graph classes. Our approach treats each graph as a statistical model and organizes the graphs in a partially ordered set based on

model inclusion. This induces a neighborhood structure, from which we define the model-oriented distance as the length of a shortest path through neighbors, yielding a metric in the space of graphs. We apply this framework to probabilistic undirected graphs, causal directed acyclic graphs, probabilistic completed partially directed acyclic graphs, and causal maximally oriented partially directed acyclic graphs. We analyze theoretical and empirical behaviors of the model-oriented distance and draw comparison with existing distances. By exploiting the underlying poset structures, we develop algorithms for computing and bounding the proposed distance that scale to moderate-sized graphs.

Title: Selection of the fittest or selection of the luckiest: the emergence of Goodhart's law in evolution

Author: Talyigas, Zsofia

Email: zsofia.talyigas@boku.ac.at

Affiliation: BOKU University

Abstract: Natural selection is commonly assumed to become more effective as it becomes stronger. However, selection acts on phenotypes rather than directly on genotypes, and phenotypic success is inherently noisy. Here we study how this mismatch shapes long-term evolutionary dynamics. Using a minimal stochastic model in which individuals inherit genetic fitness while selection acts on noisy phenotypic expressions, we show that increasing selection strength accelerates adaptation only up to a critical threshold. Beyond this point, stronger selection paradoxically slows evolution and erodes genetic diversity by favoring the luckiest individuals rather than the genetically fittest. We identify two distinct evolutionary regimes—selection of the fittest and selection of the luckiest—separated by a sharp transition. This transition corresponds to a previously unrecognized change in the structure of traveling fitness waves, from semipulled to fully pulled fronts, with profound consequences for adaptation speed and genealogical structure. Our results reveal a biological instance of Goodhart's law: when phenotypic measures become overly optimized targets, they cease to reliably promote genetic improvement. These findings highlight intrinsic limits to the effectiveness of strong selection and suggest that optimal evolutionary outcomes require intermediate selection strength in noisy environments. Joint work with Colin Desmarais Bastien Mallein, Francesco Paparella and Emmanuel Schertzer.

Title: Consensus dimension reduction via multi-view learning

Author: Tang, Tiffany

Email: ttang4@nd.edu

Affiliation: University of Notre Dame

Abstract: A plethora of dimension reduction methods have been developed to visualize high-dimensional data in low dimensions. However, different dimension reduction methods often output different and possibly conflicting visualizations of the same data. This problem is further exacerbated by the choice of hyperparameters, which may substantially impact the resulting visualization. To obtain a more robust and trustworthy dimension reduction output, we advocate for a consensus approach, which summarizes multiple visualizations into a single consensus dimension reduction visualization. Here, we leverage ideas from multi-view learning in order to identify the patterns that are most stable or shared across the many different dimension reduction visualizations, or views, and subsequently visualize this shared structure in a single low-dimensional plot. We demonstrate that this consensus visualization effectively identifies and preserves the shared low-dimensional data structure through both simulated and real-world case studies. We further highlight our method's robustness to the choice of dimension reduction method and hyperparameters---a highly-desirable property when working towards trustworthy and reproducible data science.

Title: Post-reduction inference for confidence sets of models

Author: Tang, Yanbo

Email: yanbo.tang@imperial.ac.uk

Affiliation: Imperial College London

Abstract: In high-dimensional regression problems, especially in areas like genomics, we often face thousands of potential explanatory variables but only a limited number of observations. In these settings, it is often more informative to construct a set of plausible explanatory models on which follow-up studies can be based, rather than selecting a single model. Previous approaches to constructing such confidence sets of models have relied on sample splitting, but here we present an approach based on Fisher's ideas of conditional inference to avoid this. We illustrate the method with the normal linear regression model and show how it extends to time-to-event data and more general regression models.

Title: Return probabilities on Bienaymé-Galton-Watson trees

Author: Terveer, Sara

Email: saraterveer@outlook.de

Affiliation: LMU Munich

Abstract: Consider a simple random walk on a Bienaymé-Galton-Watson tree. For its annealed return probability, an approach by Piau (1998) established upper bounds with subexponential decay in time. The bounds obtained this way for offspring distributions resulting in leaf-free trees are optimal and it has been conjectured that these are also the optimal bounds for general offspring distributions. In this talk, we derive such optimal upper bounds for all offspring distributions with a finite first moment. We discuss the main techniques of the proof, specifically how to control regions of the tree with 'bad' isoperimetric properties, as well as consequences of the result, including an application to the spectral theory of Erdős-Rényi random graphs.

Title: Gradient equilibrium, Blackwell approachability, and ergodicity

Author: Tibshirani, Ryan

Email: ryantibs@berkeley.edu

Affiliation: UC Berkeley

Abstract: Gradient equilibrium is a property of a sequence of iterates in online learning such that the average of gradients along the sequence converges to zero. In general, this property is achievable by standard online learning methods, such as gradient descent, under weak conditions; furthermore, in various problem settings it translates into statistically meaningful properties such as sequential notions of calibration and unbiasedness. This talk will review some of the basic theory of gradient equilibrium, and describe new connections to Blackwell approachability, and ergodic theory.

Title: Selective and marginal selective inference for exceptional groups

Author: Tokdar, Surya

Email: surya.tokdar@duke.edu

Affiliation: Duke University

Abstract: Statistical analyses of multipopulation studies often use the data to select a particular population as the target of inference. For example, a confidence interval may be constructed for a population only in the event that its sample mean is larger than that of the other populations. We show that for the normal means model, confidence interval procedures that maintain strict coverage control conditional on such a selection event will have infinite expected width. For applications where such selective coverage control is of interest, this result motivates the development of procedures with finite expected width and approximate selective coverage control over a range of plausible parameter values. To this end, we develop selection-adjusted empirical Bayes confidence procedures that use information from the data to approximate an oracle confidence procedure that has exact selective coverage control and finite expected width. In numerical comparisons of the oracle and empirical Bayes procedures to procedures that only guarantee selective coverage control marginally over selection events, we find that improved selective coverage control comes at the cost of increased expected interval width.

Title: Optimized Inference for High-Dimensional Vector Autoregressions

Author: Trindade, Alex

Email: alex.trindade@ttu.edu

Affiliation: Texas Tech University

Abstract: Two computationally feasible estimators for high-dimensional sparse vector autoregressive (VAR) models are proposed, in which its dimension (d) or its lag-order p , or both increase with the sample size n . Both methods rely on the p -values derived from the least-squares estimator (LSE) as the basis for sparsification. The thresholding method (TLSE) sets to zero the coefficients with p -values exceeding a cutoff which depends on n , and the constrained model is refitted. The data-adaptive method (BLSE) first ranks the p -values and then fits increasingly larger models, selecting the model with the smallest information criterion from this sequence. Asymptotic normality of LSE is established, and under additional regularity conditions it is shown that both TLSE and BLSE have the oracle property, whereby the true sparsity structure can be recovered asymptotically. Weaker conditions result in the sure-screening property, guaranteeing that the true nonzero coefficients are included in the selected active set. Simulations provide a comparison with existing methods, and suggest that TLSE and BLSE are generally better at sparsity pattern recovery while maintaining computational feasibility. The methodology is applied to

high-dimensional neuroscience and econometrics real datasets, illustrating two types of Granger-causality recently proposed in the literature.

Title: d-stochastic measures with uniform (d-1)-marginals and fractal support

Author: Trutschnig, Wolfgang

Email: wolfgang.trutschnig@plus.ac.at

Affiliation: University of Salzburg

Abstract: We study the family P_d of d-stochastic measures on $[0, 1]^d$, whose (d-1)-dimensional marginals coincide with the uniform distribution on $[0, 1]^{d-1}$, and show that this family contains elements with fractal support (in the sense that the support is self-similar with non-integer Hausdorff dimension). In fact, working with Iterated Function Systems with Probabilities (IFSP) it is possible to prove the following surprising theorem: for every dimension $d \geq 3$ the set of Hausdorff dimensions of supports of elements in P_d is dense in $[d-1, d]$. Complementing this result with some additional observations on the interplay between complete dependence and uniform (d-1)-marginals allows to conclude that the family P_d is much richer than one might conjecture.

Title: Smooth Zero-Inflated Modeling on Counting Tensors

Author: Tuzhilina, Elena

Email: elena.tuzhilina@utoronto.ca

Affiliation: University of Toronto

Abstract: Inferring the three-dimensional organization of the genome from single-cell Hi-C data presents a major statistical challenge due to the extreme sparsity and noise inherent in these measurements. Each single-cell Hi-C experiment yields a contact matrix recording the frequency of spatial proximity between genomic loci. However, many entries are zeros, corresponding either to true biological absence of contact or to technical dropouts from incomplete ligation or limited sequencing depth. Distinguishing these sources is essential for accurate characterization of chromatin architecture at the single-cell level. We represent single-cell Hi-C data as a three-dimensional tensor encoding loci-by-loci contact maps across multiple cells. Building on this, we develop a tensor-based imputation framework using a zero-inflated Poisson model to distinguish true and false zeros and address data sparsity. Efficient parameter estimation procedures are derived, and the method is applied to real datasets to recover missing contacts and improve data quality across cell types.

Title: Change-point detection in preferential attachment models

Author: van der Hofstad, Remco

Email: rhofstad@win.tue.nl

Affiliation: Eindhoven University of Technology

Abstract: Many phenomena in the real world can be phrased in terms of networks. Examples include the World-Wide Web, social interactions and Internet, but also the interaction patterns between proteins, food webs and citation networks. A very popular model for such networks is the preferential attachment model, in which vertices arrive in the network sequentially, and attach to the older vertices in the network such that higher-degree vertices are more likely chosen. In this talk, the focus is on the so-called affine preferential attachment model, where vertices attach according to the degree plus a constant. Such models have power-law degree distributions, alike those observed in real-world networks. We discuss the problem of identifying a change point in the attachment mechanism, where the constant in the attachment rule changes at some time in the evolution of the network. We give conditions for when such a change can be detected, the precise condition being close to optimal. [This joint work with Gianmarco Bet, Kay Bogerd and Rui Castro.]

Title: One-arm exponents of the Ising model

Author: van Engelenburg, Diederik

Email: diederikvanengelenburg@gmail.com

Affiliation: Technische Universität Wien

Abstract: I will focus on behavior of the Ising model in high dimensions ($d \geq 4$). Widom proposed that thermodynamic quantities follow power laws governed by critical exponents, and above the upper critical dimension $d_c = 4$, these exponents reduce to the mean-field values (matching those on trees or complete graphs). I will talk about a recent work about the so-called one-arm event (the origin connects to distance n) in the FK-Ising model. We observe that this exponent, as well as the upper critical dimension, depends on the boundary condition: for wired boundary conditions, we prove that this probability

decays up to constants as n^{-1} for $d \geq 4$ (by the Edwards-Sokal coupling, this probability is the magnetization at the origin), whereas in infinite volume we prove that it decays as n^{-2} for $d \geq 6$. Based on joint work with Christophe Garban, Romain Panis and Franco Severo.

Title: On an extension of the Cox model for time-varying covariates under dependent censoring with unknown association

Author: Van Keilegom, Ingrid

Email: INGRID.VANKEILEGOM@KULEUVEN.BE

Affiliation: KU Leuven

Abstract: This paper proposes a copula-based Cox model to take dependent censoring into account in the presence of time-varying covariates. Unlike many existing approaches, the copula parameter defining the copula function is assumed to be unknown and is estimated from the observed survival data without requiring prior information. Although the standard partial likelihood method remains valid for estimating regression coefficients in Cox models with time-varying covariates, it is not suitable when the survival time and censoring time are dependent. To overcome this issue, we expand the nonparametric hazard function in the Cox model using M-spline basis functions. Then, the proposed model is estimated using both maximum likelihood and penalized likelihood methods. The latter method employs a penalty function to regularize the baseline hazard estimates, giving a numerically stable estimator of the variance-covariance matrix. Another advantage of our penalized likelihood approach is its reduced sensitivity to both the number and placement of knots. The finite-sample performance of the proposed estimators is evaluated via simulation studies, and the methodology is further illustrated using a real data example.

Title: On the spectral determinacy of random graphs

Author: Van Werde, Alexander

Email: a.van.werde@uni-muenster.de

Affiliation: University of Muenster

Abstract: Consider a network. Do the eigenvalues of the adjacency matrix characterize the graph up to isomorphism? Both positive and negative examples are known, but it remains poorly understood what happens in the "typical case" of an Erdős-Rényi random graph. I will discuss old and new conjectures as well as new results. The rigorous results concern local obstructions for spectral characterization (or the lack of such obstructions) in the giant component and its 2-core. Based on work with Nils Van de Berg, available at arXiv:2602.22757.

Title: Real-time anomaly detection in space-time processes via conformal prediction and functional data analysis

Author: Vantini, Simone

Email: simone.vantini@polimi.it

Affiliation: MOX - Dept of Mathematics, Politecnico di Milano

Abstract: Temporally variant data observed on two-dimensional domains arise naturally across several disciplines. Functional data analysis proves inherently suitable for representing and modelling this kind of data. Within this framework, discrete-time evolving surfaces can be effectively modelled as functional time series of random real-valued functions defined on a two-dimensional domain. Building upon this approach, an anomaly detection method for handling such data is here developed. The central focus revolves around conformal prediction for functional time series. The proposal extends a functional autoregressive process of order one to this context and incorporating conformal prediction bands for functional data recently introduced in Diqigiovanni et al (2025, Statistica Sinica) and Ajroldi et al (2023, CSDA). This methodology allows the real-time construction of a spatially and temporally varying prediction range for each point of the spatial domain ensuring control of the joint coverage probability. An anomaly is identified every time the observed surface deviates from the prediction range at a particular point in the domain. This approach inherently guarantees the exact control of the probability of encountering one or more false warnings across the spatial domain. Finally the procedure is applied to weekly satellite interferometric measures of land elevation speed in the Phlegraean Fields and Etna Volcano (Bortolotti et al 2024, International Astronautical Conference 2024).

Title: A "piecewise Gaussian" limit for the fluctuations of a stochastic spatial chemical reaction network

Author: Véber, Amandine

Email: amandine.veber@parisdescartes.fr

Affiliation: Université Paris Cité

Abstract: In this presentation, we will introduce a measure-valued Markov process modelling a finite set of biochemical reactions taking place in a compact, continuous space. In general, molecules react when they are close to one another; certain reactions or chemical species may be localised in space; some chemical species are abundant (with an $O(N)$ number of molecules) whilst others may be rare (with $O(1)$ molecules). We shall begin by showing that, as N tends to infinity, a properly normalised version of the counting measure describing the state of the system converges to a piecewise deterministic Markov process with measure values. We shall then describe a central limit theorem associated with this convergence, in which the limiting process is a measure-valued semi-martingale whose noise part can be seen as a conditionally "piecewise" Gaussian process. Work in collaboration with Lea Popovic (Concordia University).

Title: How Differential Privacy Reshapes Optimal Statistical Design

Author: Vuursteen, Lasse

Email: lv121@duke.edu

Affiliation: Duke

Abstract: Much of the differential privacy literature focuses on optimal inference within a fixed statistical model and design. But researchers typically face a choice among multiple study designs that address the same scientific question. For instance, to estimate the half-life of a drug, one must decide how many patients to enroll, how many measurements to take per patient, and when to take them. I will present results showing that designs which are statistically equivalent in classical settings can perform drastically differently under differential privacy: some designs permit accurate private inference, while others fundamentally do not. This distinction is especially consequential for functional data, which arises naturally in EEG monitoring, wearable biosensors, and GPS tracking. We characterize the trade-offs between the number of individuals, the number and placement of measurements per individual, and whether the design is random or fixed. Our results reveal that privacy constraints fundamentally alter the optimal allocation of sampling resources — sometimes amplifying phenomena already present in the non-private setting, and sometimes breaking entirely from the classical picture.

Title: Fast and Optimal changepoint detection and localization with Bonferroni triplets

Author: Walther, Guenther

Email: gwalther@stanford.edu

Affiliation: Stanford University

Abstract: The talk concerns the problem of detecting and localizing changepoints in a sequence of independent observations. We propose to evaluate a local test statistic on a triplet of time points, for each such triplet in a particular collection. This collection is sparse enough so that the results of the local tests can simply be combined with a weighted Bonferroni correction. This results in a simple and fast method, Lean Bonferroni Changepoint detection (LBD), that provides finite sample guarantees for the existence of changepoints as well as simultaneous confidence intervals for their locations. LBD is free of tuning parameters, and we show that LBD allows optimal inference for the detection of changepoints. Joint work with Jayoon Jang.

Title: Distribution-Free and Model-Agnostic Changepoint Detection with Finite-Sample Guarantees

Author: Wang, Guanghui

Email: ghwang.nk@gmail.com

Affiliation: Nankai University

Abstract: We introduce ART, a distribution-free and model-agnostic framework for changepoint analysis with finite-sample guarantees. ART transforms independent observations into real-valued scores via a symmetric function; under the null hypothesis of no changepoint these scores are exchangeable. Ranking and aggregating the scores yields test statistics whose null distribution is known exactly from the permutation law of ranks, enabling exact finite-sample Type I error control without repeated refitting under permutations. ART extends naturally to a multi-scale setting: by locally ranking scores over a family of intervals and aggregating them, it supports multiple changepoint testing, localization with inference, and post-detection inference, while retaining distribution-free calibration. The approach is model-agnostic: it imposes minimal structural or distributional assumptions and accommodates diverse score constructions, including features learned by statistical or machine-learning models. Across simulations and real-data applications, ART delivers valid error control and competitive power across a range of models and distributions. These properties make ART a reliable and versatile tool for modern changepoint analysis.

Title: A Bayesian Monitoring Approach Based on the Win Ratio for Randomized Phase II Trials with Multiple Time-to-Event Outcomes

Author: Wang, Jian

Email: jianwang@mdanderson.org

Affiliation: The University of Texas MD Anderson Cancer Center

Abstract: Futility monitoring in Phase II randomized clinical trials with multiple time-to-event outcomes is typically based on a single survival endpoint or a composite time-to-first-event outcome. However, these approaches are inadequate, because a single endpoint may not adequately capture treatment effects, while composite endpoints treat all events as equally important, ignoring differences in clinical priority. We develop a Bayesian monitoring design for two-arm randomized Phase II trials using the win ratio framework to account for the clinical importance of multiple time-to-event endpoints. Operating characteristics of the design are evaluated via simulation and compared with conventional monitoring approaches. The proposed design stops early when deterioration occurs in higher-priority outcomes (e.g., death) while avoiding premature stopping driven by lower-priority outcomes (e.g., recurrence). This proposed approach provides a more tailored approach to decision-making in clinical trials with multiple time-to-event endpoints.

Title: Drift Estimation and Donsker Theorems for SDEs with Non-smooth Drift

Author: Wang, Jixin

Email: jixin.wang23@imperial.ac.uk

Affiliation: Imperial College London

Abstract: Ergodic diffusions with rough drift arise naturally in stochastic dynamics and nonparametric inference, but obtaining variance asymptotics and concentration inequality in this setting remains difficult. We study uniformly elliptic SDEs in \mathbb{R}^d with Hölder diffusion coefficient and non-smooth, possibly discontinuous drift. Under a weak dissipativity condition guaranteeing an invariant measure and polynomial beta-mixing, we establish a linear asymptotic variance formula for additive functionals via a Clark–Ocone representation, combining Aronson-type gradient estimates with long-time mixing control. We further prove a concentration inequality for kernel estimators of the invariant density indexed by translated kernel classes, as well as a Donsker theorem for the corresponding empirical process. Together, these results extend empirical-process methods to a rough drift diffusion setting and provide the probabilistic foundation for uniform analysis of Nadaraya–Watson type drift estimators under only local drift regularity.

Title: Adaptive Lq-regularized Estimation of High-dimensional Sparse Covariance Matrix

Author: Wang, Liqun

Email: liqun.wang@umanitoba.ca

Affiliation: University of Manitoba

Abstract: The estimation of high-dimensional covariance matrix plays an important role in many application fields such as economics, biology, social and health sciences. A mainstream structural assumption for enhancing estimator accuracy is that the covariance matrix is sparse or approximately sparse. This paper proposes an adaptive Lq-regularized estimator with minimum eigenvalue constraint for high-dimensional sparse covariance matrix. This method eliminates the need for the conventional two-stage framework of sequential correlation and covariance matrix estimation. Under appropriate regularity conditions, we analyze its asymptotic and finite sample properties. The proposed iterative reweighted minimization method and its inexact variant can be employed to find a desired estimate. Simulation studies confirm that the proposed estimator performs better than some other state-of-the-art methods. The method is applied to a cancer RNA sequence data analysis. This is a joint work with X Wang, H Zhao, Z Zhou, and L Kong.

Title: High-dimensional Change-point Detection Using Generalized Homogeneity Metrics

Author: Wang, Runmin

Email: runminw@tamu.edu

Affiliation: Texas A&M University

Abstract: In this talk, we study the problem of detecting abrupt changes in the data-generating distributions of a sequence of high-dimensional observations beyond the first two moments. This problem has remained substantially less explored in the existing literature, especially in the high-dimensional context, compared to detecting changes in the mean or the covariance structure. We develop a nonparametric methodology to (i) test the existence of a change-point, and (ii) identify the change-point locations in an independent sequence of high-dimensional observations. Our approach rests upon recent

nonparametric tests for the homogeneity of two high-dimensional distributions. We construct a single change-point test statistic based on a cumulative sum process in an embedded Hilbert space. We shall derive its limiting null distribution and present the asymptotic consistency under the high dimension medium sample size framework. We also combine our statistics with wild binary segmentation to recursively estimate and test for multiple change-point locations. The superior performance of our methodology compared to other existing procedures will be illustrated via extensive simulation studies and the application to the stock return data observed during the period of the global financial crisis in the United States.

Title: On polynomial-time computation in PDE models

Author: Wang, Sven

Email: sven.wang@epfl.ch

Affiliation: EPFL

Abstract: In complex statistical models involving PDEs or stochastic process, likelihood-based methods may suffer from non-convex and possibly multimodal loss landscapes. Thus, optimization and Markov chain Monte Carlo (MCMC) methods may mix exponentially slowly. We discuss recent progress in devising statistical and polynomial-time computational guarantees in such settings. In particular, we will discuss a class of computationally tractable estimators which can be deployed in models where the likelihood function is implicitly defined via a PDE. For some prototypical non-linear inverse problems arising from elliptic PDEs, we prove that these estimators attain the best currently known statistical convergence rates while being globally computable in polynomial time. Our analysis is based on new generalized stability estimates, extending classical stability beyond the range of the forward operator, combined with tools from nonparametric M-estimation. Our estimators also provide principled warm-start initializations for polynomial-time Bayesian computation.

Title: Two-Sample Comparison in the Presence of Long-Term Survivors

Author: Wang, Weijing

Email: wjwang@stat.nycu.edu.tw

Affiliation: National Yang Ming Chiao Tung University

Abstract: Clinical trials of modern cancer therapies frequently violate the proportional hazards assumption because of delayed treatment effects, crossing survival curves, and the presence of long-term survivors. In these settings, hazard ratio-based inference can be difficult to interpret. We introduce the tau process, an intuitive and robust measure of treatment effect under nonproportional hazards that can also identify when the hazards change their ordering over time. We further extend this framework to accommodate long-term survivors by adopting a mixture (cure) modeling perspective and proposing estimands that compare both cure fractions and treatment effects within the susceptible (non-cured) subpopulation. In particular, we develop the susceptible tau process to quantify treatment effects among patients who remain at risk. A case study of the CheckMate 067 melanoma trial illustrates how these tools provide clinically meaningful insights into long-term treatment benefit. This talk is based on joint work with Professor Yi-Cheng Tai (National Chengchi University) and Professor Martin T. Wells (Cornell University).

Title: Parametric and Nonparametric Approaches for Phase II Monitoring of Nonlinear Profiles

Author: Wang, Yihua

Email: wangyi@yuntech.edu.tw

Affiliation: Department of Industrial Engineering and Management, National Yunlin University of Science and Technology

Abstract: In many manufacturing and continuous process applications, product or process quality can be characterized by a functional relationship between a response variable and one or more explanatory variables. Such data are commonly referred to as profile data, and profile monitoring aims to detect changes in this functional relationship. In practice, the underlying profile is sometimes nonlinear, making parameter estimation and monitoring more challenging in Phase II applications. To address this issue, we propose two Phase II monitoring approaches for nonlinear profiles. The first approach approximates the nonlinear function using a Taylor series expansion, transforming it into a linear form, and monitors the resulting parameters using a multivariate exponentially weighted moving average (MEWMA) control chart. The second approach is based on a nonparametric area-based statistic, which is incorporated into exponentially weighted moving average (EWMA) and homogeneously weighted moving average (HWMA) control charts. Both approaches are designed to detect changes in profile parameters, and error variance. The performance of the proposed methods is evaluated through Monte Carlo simulations using the average run length (ARL) as the performance measure. Comparisons with the existing EWMA-based (NEWMA)

monitoring method show that the proposed nonparametric schemes are more effective in detecting moderate to large shifts in profile parameters as well as increases in the error variance.

Title: A New Regression Lens on Multi-Class Classification

Author: Wegkamp, Marten

Email: marten.wegkamp@cornell.edu

Affiliation: Cornell University

Abstract: Linear Discriminant Analysis (LDA) is a foundational method for classification. Its linear structure makes it easy to interpret, and it naturally accommodates multi-class problems. LDA is also closely connected to classical multivariate techniques, including Fisher's discriminant analysis, canonical correlation analysis, and linear regression. In this talk, we deepen LDA's link to multivariate response regression by deriving an explicit relationship between discriminant directions and regression coefficients. This insight leads to a new regression-based procedure for multi-class classification that can flexibly incorporate structured, regularized, or even non-parametric regression methods. Unlike existing regression-based approaches, our formulation is particularly well-suited for establishing provable performance guarantees.

Title: Can large language models boost the power of randomized experiments without statistical bias?

Author: Wei, Waverly

Email: waverly@marshall.usc.edu

Affiliation: University of Southern California

Abstract: Randomized experiments or randomized controlled trials (RCTs) are gold standards for causal inference, yet cost and sample-size constraints limit power. Meanwhile, modern RCTs routinely collect rich, unstructured data that are highly prognostic of outcomes but rarely used in causal analyses. We introduce CALM (Causal Analysis leveraging Language Models), a statistical framework that integrates large language models (LLMs) predictions with established causal estimators to increase precision while preserving statistical validity. CALM treats LLM outputs as auxiliary prognostic information and corrects their potential bias via a heterogeneous calibration step that residualizes and optimally reweights predictions. We prove that CALM remains consistent even when LLM predictions are biased and achieves efficiency gains over augmented inverse probability weighting estimators for various causal effects. In particular, CALM develops a few-shot variant that aggregates predictions across randomly sampled demonstration sets. The resulting U-statistic-like predictor restores i.i.d. structure and also mitigates prompt-selection variability. In simulations, CALM delivers lower variance relative to other benchmarking methods, is effective in zero- and few-shot settings, and remains stable across prompt designs. By principled use of LLMs to harness unstructured data and pretraining, CALM provides a practical path to more precise causal analyses in RCTs.

Title: Statistical Inference under Adaptive Sampling with LinUCB

Author: Wei, Yuting

Email: ytwei@wharton.upenn.edu

Affiliation: University of Pennsylvania

Abstract: Adaptively collected data has become ubiquitous in modern practice. Yet even seemingly benign adaptive sampling schemes can introduce severe biases, rendering traditional statistical inference tools inapplicable. Focusing on the linear bandit problem, a fundamental and influential framework in reinforcement learning and the bandit literature, we characterize the performance of LinUCB, a canonical upper-confidence-bound algorithm that balances exploration and exploitation, and derive inferential procedures that remain valid despite the challenges posed by adaptive data collection.

Title: Asymptotic Inference for Rank Correlations

Author: Wermuth, Jan-Lukas

Email: wermuth@econ.uni-frankfurt.de

Affiliation: Goethe University Frankfurt

Abstract: Kendall's tau and Spearman's rho are widely used tools for measuring dependence. Surprisingly, when it comes to asymptotic inference for these rank correlations, some fundamental results and methods have not yet been developed, in particular for discrete random variables and in the time series case, and concerning variance estimation in general. Consequently, asymptotic confidence intervals are not available. We provide a comprehensive treatment of asymptotic inference for classical rank correlations, including Kendall's tau, Spearman's rho, Goodman-Kruskal's gamma, Kendall's tau-b,

and grade correlation. We derive asymptotic distributions for both iid and time series data, resorting to asymptotic results for U-statistics, and introduce consistent variance estimators. This enables the construction of confidence intervals and tests, generalizes classical results for continuous random variables and leads to corrected versions of widely used tests of independence. We analyze the finite-sample performance of our variance estimators, confidence intervals, and tests in simulations and illustrate their use in case studies.

Title: Testing the calibration of fixed-event probability forecast sequences

Author: Wilkinson, Thomas

Email: tswilkinson@protonmail.com

Affiliation: University of Exeter

Abstract: The martingale property for a sequence of real-valued random variables can also be stated for a sequence of random measures on a common measurable space. The martingale property for random measures has been studied in Bayesian statistics, where the random measures in the sequence are predictive probability distributions for a sequence of observations. We show that for a sequence of random probability measures representing a sequence of forecasts made for a common observation, the martingale property is equivalent to a basic requirement that the forecasts are calibrated for the observation. We present a simple statistical test of the martingale property for a finite sequence of random probability measures on \mathbb{R} , and interpret it as a test of forecast calibration. As part of the test, we construct 'synthetic PIT values', based on the popular Probability Integral Transform method for testing the calibration of a rolling-event sequence of forecasts. We also show how to extend our method to test the martingale property of any finite sequence of random probability measures on a common measurable space which is a Borel space.

Title: How do simple rotations affect the implicit bias of Adam?

Author: Willett, Rebecca

Email: willett@uchicago.edu

Affiliation: University of Chicago

Abstract: Adaptive gradient methods such as Adam and Adagrad are widely used in machine learning, yet their effect on the generalization of learned models -- relative to methods like gradient descent -- remains poorly understood. Prior work on binary classification suggests that Adam exhibits a "richness bias," which can help it learn nonlinear decision boundaries closer to the Bayes-optimal decision boundary relative to gradient descent. However, the coordinate-wise preconditioning scheme employed by Adam renders the overall method sensitive to orthogonal transformations of feature space. We show that this sensitivity can manifest as a reversal of Adam's competitive advantage: even small rotations of the underlying data distribution can make Adam forfeit its richness bias and converge to a linear decision boundary that is farther from the Bayes-optimal decision boundary than the one learned by gradient descent. To alleviate this issue, we show that a recently proposed reparameterization method -- which applies an orthogonal transformation to the optimization objective -- endows any first-order method with equivariance to data rotations, and we empirically demonstrate its ability to restore Adam's bias towards rich decision boundaries.

Title: Underreporting in count data models: some new results and an application

Author: Winkelmann, Rainer

Email: rainer.winkelmann@econ.uzh.ch

Affiliation: University of Zurich

Abstract: Winkelmann and Zimmermann (1993), although only published as a working paper, has initiated a small but persistent literature on underreporting in count data models. Their Poisson-logistic (Pogit) model was initially developed to estimate the determinants of the number of job offers during a given time interval, when the researcher only observes the number of accepted offers. Later Pogit applications can be found in epidemiology, engineering, public health, and criminology, among others. I review some of the recent developments in the literature. I also consider an extension of the original model where the reporting process can be estimated separately from the event generating process due to the availability of a validation dataset.

Title: Gaussian Approximation and Concentration of Constant Learning-Rate Stochastic Gradient Descent

Author: Wu, Wei

Email: wbwu@uchicago.edu

Affiliation: University of Chicago

Abstract: We establish a comprehensive finite-sample and asymptotic theory for stochastic gradient descent (SGD) with constant learning rates. First, we propose a novel linear approximation technique to provide a quenched central limit theorem (CLT) for SGD iterates with refined tail properties, showing that regardless of the chosen initialization, the fluctuations of the algorithm around its target point converge to a multivariate normal distribution. Our conditions are substantially milder than those required in the classical CLTs for SGD, yet offering a stronger convergence result. Furthermore, we derive the first Berry-Esseen bound – the Gaussian approximation error – for the constant learning-rate SGD, which is sharp compared to the decaying learning-rate schemes in the literature. Beyond the moment convergence, we also provide the Nagaev-type inequality for the SGD tail probabilities by adopting the autoregressive approximation techniques, which entails non-asymptotic large deviation guarantees. These results are verified via numerical simulations, paving the way for theoretically grounded uncertainty quantification, especially with non-asymptotic validity.

Title: Provably Reliable Classifier Guidance via Cross-Entropy Control

Author: Wu, Yuchen

Email: yuchen.wu@cornell.edu

Affiliation: Cornell University

Abstract: Classifier-guided diffusion models generate conditional samples by augmenting the reverse-time score with the gradient of the log-probability predicted by a probabilistic classifier. In practice, this classifier is usually obtained by minimizing an empirical loss function. While existing statistical theory guarantees good generalization performance when the sample size is sufficiently large, it remains unclear whether such training yields an effective guidance mechanism. We study this question in the context of cross-entropy loss, which is widely used for classifier training. Under mild smoothness assumptions on the classifier, we show that controlling the cross-entropy at each diffusion model step is sufficient to control the corresponding guidance error. In particular, probabilistic classifiers achieving conditional KL divergence ε^2 induce guidance vectors with mean squared error $O(\varepsilon)$, up to constant and logarithmic factors. Our result yields an upper bound on the sampling error of classifier-guided diffusion models and bears resemblance to a reverse log-Sobolev-type inequality. To the best of our knowledge, this is the first result that quantitatively links classifier training to guidance alignment in diffusion models, providing both a theoretical explanation for the empirical success of classifier guidance, and principled guidelines for selecting classifiers that induce effective guidance.

Title: Weak convergence of Volterra-type stochastic integrals in Skorokhod space

Author: Wunderlich, Fabrice

Email: fabrice.wunderlich@uni.lu

Affiliation: University of Luxembourg

Abstract: We study weak convergence of Volterra-type stochastic integrals whose integrands depend on both the integration variable and the terminal time. Such stochastic convolutions arise naturally in non-Markovian limit problems. We obtain weak convergence results in both the J_1 and M_1 Skorokhod topologies under natural conditions on the semimartingale integrators, the integrands, and their interaction. The framework applies in particular to stochastic convolutions with singular kernels and is motivated by scaling limits arising in nearly unstable Hawkes models and rough volatility. This is joint work with Giulia Livieri and Andreas Søjmark.

Title: A modified Greenwood statistic for goodness-of-fit testing in heavy-tailed multivariate samples

Author: Wylomanska, Agnieszka

Email: agnieszka.wylomanska@pwr.edu.pl

Affiliation: Wroclaw University of Science and Technology

Abstract: This presentation details an enhanced version of the Greenwood statistic, specifically tailored for analyzing complex, multivariate datasets. The study addresses the challenges of "heavy-tailed" data, i.e. where extreme outliers are frequent, common in distributions such as Pareto, Student's t , and alpha-stable. By exploring the stochastic properties of this modified statistic, the research demonstrates its superior performance in goodness-of-fit testing. A major breakthrough of this work is the statistic's ability to determine if a sample originates from a distribution with infinite variance, a recurring hurdle in modern data science. The methodology's reliability is confirmed via Monte Carlo simulations and further proven through the analysis of real data sets.

Title: Is the F-test doubly robust?

Author: Xia, Lucy

Email: lucyxia@ust.hk

Affiliation: The Hong Kong University of Science and Technology

Abstract: We study the robustness of the F-test in random design linear models, and reach a somewhat nuanced conclusion. On the positive side, one of our main results is that the size of the test is close to its nominal level as soon as either the distribution of the normalised error vector is close to uniform on the unit sphere, or the distribution of the design matrix, after applying any column space-preserving orthogonalisation scheme (e.g. Gram-Schmidt), is close to a Haar distribution. This provides a sense in which the F-test is doubly robust. Our conclusion is reached by establishing a Hölder continuity property of the Kolmogorov distance between the distribution of the F-statistic and its notional F-distribution under the null. Writing n , p and p_0 for the sample size and the dimensions of the full and null models respectively, we prove that the Hölder exponent is $1/3$ when $\min(n-p, p-p_0) = 1$ and $1/2$ when $\min(n-p, p-p_0) \geq 2$. On the other hand, these exponents are relatively small and cannot be improved in general, revealing that the size of the test may depart from its nominal level quite quickly as we move away from settings where the test is exact. In particular, the regression t-test for the significance of a single predictor corresponds to the case $p - p_0 = 1$, so this test may be especially vulnerable to model misspecification.

Title: Model-Free Inference for High-Dimensional Binary Classification Using Repro Samples

Author: Xie, Minge

Email: mxie@rutegrs.edu

Affiliation: Rutgers University

Abstract: Performing inference is particularly challenging in high-dimensional settings when the underlying data-generating regression structure is completely unknown and unspecified. In this talk, we develop a model-free inference approach for high-dimensional binary classification, leveraging a class of (likely misspecified) sparsity GLMs. Our method builds on the repro samples framework, which generates artificial samples that mimic the true data or noises. The proposed approach facilitates inference by targeting both the model support and the regression coefficients of the oracle model, defined as the working model closest to the unknown true model. The method offers three key advantages: (1) it is fully model-free, requiring neither correct model specification nor sparsity of the true model; (2) it constructs a candidate set of influential covariates with guaranteed coverage under weak signal conditions; and (3) it provides confidence sets for any linear transformation of the oracle coefficients. When the oracle model coincides with the true underlying model, the inference results directly apply to the true model, allowing simultaneous quantification of uncertainty in discrete model selection and continuous parameter estimation. Simulation studies illustrate the effective performance of the proposed method. An application to single-cell RNA-seq immune response data demonstrates that it identifies key genes, offering new insights into cellular immune response mechanisms.

Title: Identifiability and Inference for Generalized Latent Factor Models

Author: Xu, Gongjun

Email: gongjun@umich.edu

Affiliation: University of Michigan

Abstract: Generalized latent factor analysis not only provides a useful latent embedding approach in statistics and machine learning, but also serves as a widely used tool across various scientific fields, such as psychometrics, econometrics, and social sciences. Ensuring the identifiability of latent factors and the loading matrix is essential for the model's estimability and interpretability, and various identifiability conditions have been employed by practitioners. However, fundamental statistical inference issues for latent factors and factor loadings under commonly used identifiability conditions remain largely unaddressed, especially for correlated factors and/or non-orthogonal loading matrix. In this work, we focus on the maximum likelihood estimation for generalized factor models and establish statistical inference properties under popularly used identifiability conditions. The developed theory is further illustrated through numerical simulations and an application to a personality assessment dataset.

Title: Detecting High-Dimensional Geometry in Random Graphs

Author: Xu, Jiaming

Email: jx77@duke.edu

Affiliation: Duke University

Abstract: Random geometric graphs hide geometric structure inside an otherwise random-looking network: each vertex has a latent position in high-dimensional space, and edges are generated with probabilities given by a kernel applied to the inner products of latent positions. When can this hidden geometry be detected from the graph alone? I will discuss recent results on detection thresholds for high-dimensional random geometric graphs. For models with smooth kernels, we show that the critical dimension is $d^*=n^{\{3/4\}}$, far below the $d^*=n^3$ threshold known for the classical hard-threshold model. I will also describe our new progress on proving the conjectured threshold in the sparse regime of the hard-threshold model. More broadly, we formulate a unifying conjecture that the critical dimension is determined by $n^3 \text{Tr}(K^2)=1$, where K is the standardized kernel operator.

Title: Optimal Convex M-Estimation via Score Matching

Author: Xu, Min

Email: mx76@stat.rutgers.edu

Affiliation: Rutgers University

Abstract: In the context of linear regression, we construct a data-driven convex loss function with respect to which empirical risk minimisation yields optimal asymptotic variance in the downstream estimation of the regression coefficients. At the population level, the negative derivative of the optimal convex loss is the best decreasing approximation of the derivative of the log-density of the noise distribution. This motivates a fitting process via a nonparametric extension of score matching, corresponding to a log-concave projection of the noise distribution with respect to the Fisher divergence. At the sample level, our semiparametric estimator is computationally efficient, and we prove that it attains the minimal asymptotic covariance among all convex M-estimators. As an example of a non-log-concave setting, the optimal convex loss function for Cauchy errors is Huber-like, and our procedure yields asymptotic efficiency greater than 0.87 relative to the maximum likelihood estimator of the regression coefficients that uses oracle knowledge of this error distribution. In this sense, we provide robustness and facilitate computation without sacrificing much statistical efficiency. Numerical experiments using our accompanying R package "asm" confirm the practical merits of our proposal.

Title: Wasserstein Spatial Depth

Author: Yao, Yisha

Email: yy3381@columbia.edu

Affiliation: Columbia University

Abstract: Modeling observations as random distributions embedded within Wasserstein spaces is becoming increasingly popular across scientific fields, as it captures the variability and geometric structure of the data more effectively. However, the distinct geometry and unique properties of Wasserstein spaces pose challenges to the application of conventional statistical tools designed for Euclidean spaces. The space of distributions on \mathbb{R}^d with $d>1$ is not linear, and "mimic" the geometry of a Riemannian manifold. Developing new methodologies for analysis within Wasserstein spaces has become essential. In this paper, we extend the concept of statistical depth to distribution-valued data, introducing the notion of Wasserstein spatial depth (WSD). This new measure provides a way to rank and order distributions, enabling the development of order-based clustering techniques and inferential tools. We show that the WSD preserves critical properties of conventional statistical depths, notably, ranging within $[0, 1]$, transformation and geodesic invariance, vanishing at infinity, reaching a maximum at the geometric median, and continuity. Its robustness is established via the breakdown points of the empirical depth regions and its influence function. Additionally, the population WSD has a straightforward plug-in estimator based on sampled empirical distributions. We establish the estimator's consistency and asymptotic normality, and develop a two-sample test based on WSD.

Title: Causal Learning of Paired Vectors with Label Noise: Impact and Correction Methods

Author: Yi, Grace

Email: gyi5@uwo.ca

Affiliation: University of Western Ontario

Abstract: Causal inference involves determining whether a cause-effect relationship exists between two sets of interest, a task that can be framed as a binary classification problem. When dealing with a sequence of independent and identically distributed paired vectors, the kernel mean embedding of the probability distribution can be utilized to map the empirical distribution to a feature space. Subsequently, a classifier is trained in this feature space to predict causation for future vector pairs. However, this approach is susceptible to mislabeling of causal relationships, a common challenge in causation studies.

In this talk, I will discuss the impact of mislabeled outputs on the training results. Moreover, I will present robust learning methods that take into account the mislabeling effects and offer theoretical justifications for the validity of these proposed methods.

Title: Preventive Use of Korean Medicine Among Children and Adolescents in Korea: A Statistical Analysis of National Real-World Data

Author: YIM, MI HONG

Email: mh.yim@kiom.re.kr

Affiliation: KOREA INSTITUTE OF ORIENTAL MEDICINE

Abstract: Objectives: This study identified determinants of Korean Medicine (KM) use for preventive purposes among children and adolescents. Using Andersen's behavioral model, we investigated how socioeconomic characteristics influence KM utilization within a nationally representative dataset. Methods: We analyzed the 2021 Korea Health Panel Survey (n=732). To account for the multi-stage stratified cluster sampling design, all statistical analyses were performed using the survey package in R. We employed multivariable logistic regression incorporating compensatory weights and Taylor series linearization for accurate variance estimation. Generalized Variance Inflation Factors (GVIF) were used to ensure model stability by addressing multicollinearity among predictors. Results: Multivariable analyses showed KM utilization was significantly lower in the Busan/Gyeongsang region (aOR 0.28, 95% CI 0.10–0.83) versus the capital area and in larger households (≥ 5 members; 0.30, 0.09–0.98). Conversely, those in the highest income quartile exhibited substantially higher odds (8.80, 1.80–43.03). KM visits primarily involved herbal preparations (88.25%), while conventional visits were dominated by immunizations (97.61%). Conclusion: Findings suggest KM plays a unique role in pediatric wellness through herbal strategies. By employing rigorous complex-sample inference, this study provides robust evidence for Health Technology Assessments and integrated preventive care policies.

Title: Outtrigger local polynomial regression

Author: Young, Elliot

Email: ey244@cam.ac.uk

Affiliation: University of Cambridge

Abstract: Standard local polynomial estimators of a nonparametric regression function employ a weighted least squares loss function that is tailored to the setting of homoscedastic Gaussian errors. We introduce the outtrigger local polynomial estimator, which is designed to achieve distributional adaptivity across different conditional error distributions. It modifies a standard local polynomial estimator by employing an estimate of the conditional score function of the errors and an 'outtrigger' that draws on the data in a broader local window to stabilise the influence of the conditional score estimate. Subject to smoothness conditions, and only requiring consistency of the conditional score estimate, we establish that even under the least favourable settings for the outtrigger estimator, the asymptotic ratio of the worst-case local risks of the two estimators is at most 1, with equality if and only if the conditional error distribution is Gaussian. Moreover, we prove that the outtrigger estimator is minimax optimal over Hölder classes up to a multiplicative factor depending only on the smoothness of the regression function and the dimension of the covariates. For Lipschitz functions this factor is shown to be at most 1.69, and in low smoothness or high dimension regimes we approach exact constants. A further attraction of our proposal is that we do not require structural assumptions such as independence of errors and covariates, or symmetry of the conditional error distribution.

Title: On the hardness of group-conditional distribution-free predictive inference

Author: Zaffran, Margaux

Email: mzaffran@protonmail.com

Affiliation: Inria, Laboratoire de Mathématiques d'Orsay

Abstract: Predictive uncertainty quantification is crucial in decision-making problems. In this talk, we will focus on distribution-free uncertainty quantification by considering predictive intervals for the target Y enjoying validity (i.e. nominal coverage) with no assumptions on the underlying data generating process nor the sample size. After introducing the framework, we will detail the nuance between marginal validity and conditional---on the test point---validity. We will review the existing (impossibility) results on conditional validity. This will lead us to our main question: how can we relax the goal of conditional validity to make it achievable? We will present new hardness results, that characterize the limits of group conditional coverage (e.g., achieving nominal coverage not only on average but also among women on the one hand, and

among men on the other hand), a weaker goal extensively used in the literature in place of the impossible perfect conditional validity. Finally, we will dive into applying these results in the context of prediction with missing values. There, one wants to obtain not only marginally valid intervals despite missing values, but also intervals that achieve the nominal coverage regardless of which values are missing at test time. We provide an algorithm reaching this goal by constructive informative predictive intervals in the light of our hardness results.

Title: Consistent Variable Selection for GARCH-X Models

Author: Zambom, Adriano

Email: adriano.zambom@csun.edu

Affiliation: California State University Northridge

Abstract: In this talk, we present a consistent variable selection procedure for GARCH-X models designed to identify which exogenous variables truly drive volatility dynamics. While GARCH-X models allow many potential covariates to enter the volatility equation, determining which ones genuinely matter remains a challenging problem. The proposed approach combines Wald-type test statistics with a multiple testing framework based on the Benjamini–Yekutieli False Discovery Rate procedure, allowing us to systematically control false discoveries while retaining relevant predictors. We outline the intuition behind the method and discuss the theory of consistency — meaning that as the sample size grows, the procedure correctly recovers the true set of volatility drivers. Simulation studies demonstrate that the method performs well across different distributions and dependence structures. We conclude with an empirical application to S&P 500 volatility, incorporating macroeconomic and commodity indicators, illustrating how the procedure can uncover meaningful volatility determinants in practice.

Title: Empirical Bayes for Dependent Data

Author: Zhang, Cun-Hui

Email: czhang@stat.rutgers.edu

Affiliation: Rutgers University

Abstract: Empirical Bayes is built on the idea that pooling information across related decision problems can outperform procedures that treat each problem in isolation. While this principle has had a profound impact, most of its development has focused on independent observations. In contrast, modern applications often involve measuring many attributes on the same sample, leading to compound decision problems with dependent data. In such settings, the benefit of pooling is clear, but the challenge is how to do so with limited knowledge of the dependence structure. A key observation is that the fundamental theorem connecting compound decision problems to empirical Bayes does not require independence. Building on this, we develop a nonparametric marginal likelihood approach under dependence. The effect of dependence is captured by the largest eigenvalue of the correlation matrix, which acts as a discount factor on the effective number of problems that can be pooled. The required dependence conditions are no stronger than for linear estimators. For estimating the density, score function, component means, and compound risk of the mean estimator, the proposed methods achieve nearly parametric rates of convergence to the oracle and are nearly minimax optimal in rate. They also provide consistent estimation of the prior, posterior, and credible intervals and adapt to the atomic sparsity of the oracle prior.

Title: Moments and Cumulants in Probability Theory: Recent Results and Perspectives

Author: Zhang, Jiechen

Email: jiechen.zhang@epfl.ch

Affiliation: EPFL

Abstract: Moments and cumulants are two fundamental sequences of a random variable, closely related but often playing different roles. They are central in probability theory, especially in questions of determinacy and in limit theorems. Classical topics include moment determinacy versus indeterminacy, the Gaussian law characterized by vanishing higher cumulants, the Fréchet–Shohat theorem, and Wigner's semicircle law. There are also parallel cumulant-based methods, which have proved very effective. I will present two recent and somewhat unexpected results. The first is a nonconventional limit theorem which is genuinely cumulant-based and cannot be accessed effectively by moment methods alone. The second says that a moment-indeterminate random variable can be decomposed as a sum of two random variables each satisfying Carleman's condition, hence each moment-determinate. My own work focuses on what more, beyond the classical formulas, can be said when one passes from moments to cumulants, and vice versa. In direction moments \leftrightarrow cumulants, I will present recent explicit and distribution-free bounds with factorial rather than superexponential growth. In direction cumulants \leftrightarrow moments, I

will explain both what is possible and what is not, and discuss a certain 'gap'. I will also mention another perspective, related to M-indeterminate distributions and their Stieltjes classes, showing that moment equivalence may be invisible to polynomial tests yet still matter for other natural functionals.

Title: Convex Mixed-Integer Programming for Causal Additive Models with Optimization and Statistical Guarantees

Author: Zhang, Xiaozhu

Email: xzzhang@uw.edu

Affiliation: University of Washington

Abstract: Learning causal structure from nonlinear data remains computationally and statistically challenging. We address this problem for additive, nonlinear structural equation models with Gaussian noise. By expressing each non-linear function through a basis expansion, we formulate a maximum likelihood estimator with a group l0 penalty that directly controls graph sparsity. Despite the combinatorial nature of the problem, the estimator can be solved efficiently via a convex mixed-integer program combined with a novel outer approximation scheme, yielding globally optimal solutions within seconds. Our framework can incorporate prior knowledge such as edge constraints and partial orders. We prove consistency for graph recovery when the number of variables grows with sample size, and connect optimization guarantees to statistical error, yielding an early stopping rule that preserves consistency while reducing computation. Compared to methods based on linearity, equal noise variances, or heuristics, our approach provides a unified framework with provable optimality, efficiency, and statistical guarantees. Experiments demonstrate substantial improvements in graph recovery on both simulated and real data. This is joint work with Nir Keret, Ali Shojaie, and Armeen Taeb.

Title: Statistical Inference in an Interactive Learning Paradigm

Author: Zhou, Kangjie

Email: kz2326@columbia.edu

Affiliation: Columbia University

Abstract: The proliferation of generative artificial intelligence has given rise to an interactive learning paradigm, where model parameters are continuously updated using not only data generated by natural processes, but also synthetic outputs produced by other models. This paradigm introduces two major challenges: (1) training data are no longer drawn exclusively from the target population, undermining a core assumption of classical statistical learning theory, and (2) model training processes become inherently correlated, as models influence one another through repeated exposure to each other's synthetic outputs. Establishing reliable statistical inference in such interactive environments therefore remains an important open problem. In particular, there is growing concern about model collapse, a phenomenon in which the performance of generative models progressively degrades as they are trained on synthetic data produced by earlier model generations. In this talk, we study the behavior of generative models under general interaction patterns. We formalize these interactions using directed graphs and show that the occurrence of model collapse depends critically on the graph's topology. Within this framework, we derive an explicit necessary and sufficient condition characterizing when model collapse occurs.

Title: Minimax optimal differentially private synthetic data for smooth queries

Author: Zhu, Yizhe

Email: yizhezhu@usc.edu

Affiliation: University of Southern California

Abstract: Differentially private synthetic data enables the sharing and analysis of sensitive datasets while providing rigorous privacy guarantees for individual contributors. A central challenge is to achieve strong utility guarantees for meaningful downstream analysis. Many existing methods ensure uniform accuracy over broad query classes, such as all Lipschitz functions, but this level of generality often leads to suboptimal rates for statistics of practical interest. Since many common data analysis queries exhibit smoothness beyond what worst-case Lipschitz bounds capture, we ask whether exploiting this additional structure can yield improved utility. We study the problem of generating (ϵ, δ) -differentially private synthetic data from a dataset of size n supported on the hypercube $[-1, 1]^d$, with utility guarantees uniformly for all smooth queries having bounded derivatives up to order k . We propose a polynomial-time algorithm that achieves a minimax error rate of $n^{-\min\{1, k/d\}}$, up to a $\log(n)$ factor. This characterization uncovers a phase transition at $k=d$. Our results generalize the Chebyshev moment matching framework in the literature and strictly improve the error rates for k -smooth queries. Moreover, we establish the first minimax lower bound for the utility of (ϵ, δ) -differentially private synthetic data with respect to k -smooth queries.

Title: Priority queues with differing customer selection rules

Author: Ziedins, Ilze

Email: i.ziedins@auckland.ac.nz

Affiliation: University of Auckland

Abstract: We consider priority queues where each class may have its own selection rule (such as first in first out, or last in first out). We show that the customer waiting time distribution for a given class is insensitive to the customer selection rules used for the other priority classes, and depends only on the selection rule for that class. The sample path argument used to show this also yields a general approach for finding waiting time distributions in priority queues. This is joint work with Peter Taylor (University of Melbourne) and David Stanford (Western University).