XXXVII International Seminar on Stability Problems for Stochastic Models

13-17 October, 2025, Shenzhen, China

Collection of abstracts

Contents

Modeling and analysis of binary objects based on group observations

Andrei G. Belov, Oksana A. Belova

Lomonosov Moscow State University

Infection rate of a population of ixodid ticks with tick-borne encephalitis virus and Borrelia burgdorferi sensu lato is simulated and estimated using the method of maximum likelihood and the method of moments and their comparative analysis is given. Methods for solving direct and inverse problems of binary object transmission by individual and group observations are reviewed.

Numerical and Analytical Solution of Fredholm Integral Equation that Appears in one Geophysical Problem

Konstantin P. Belyaev, Victor Yu. Korolev, Ksenia Romashina

Lomonosov Moscow State University

In order to correctly describe the observed time-series geophysical data, in particular heat-flux, sea surface temperature and others we apply the model of stochastic difference equations with random coefficients depending both on time and the process itself. Those coefficients are estimated (reconstructed) from the observed data set. In our work we use ERA5 gridded geophysical dataset which contains data from 1979 until 2024 yr in the North Atlantic.

The distribution of the sought characteristics of the processes can be determined from the Fredholm integral equations which needs to be solved both analytically and numerically. Our work shows that under some reasonable assumption this equation has an analytical solution which occurs to be the Laplace two-side distribution. However, in the general case this equation can be solved numerically and the result of this numerical solution has been presented and analyzed.

A queueing system with general vacations

Alexey K. Bergovin, Vladimir G. Ushakov

Lomonosov Moscow State University

A single-channel queueing system with a Poisson incoming flow, general distribution of the service, infinite number of places in the queue, and vacations of the servicing device is considered in the report. During vacations, a server cannot service requests. The server can go on vacation both at the moments when the system is idle and at the moments of completion of service of requests. The main result, which will be presented during the report, is the distribution of the number of requests in the system in a non-stationary mode for some distributions of the device's departure for vacation

An Available Bandwidth Estimation Algorithm Based on Optimal State Filtering Results

Andrei V. Borisov

Federal Research Center “Informatics and Control” of RAS, Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University

This report presents an approach to monitoring the available bandwidth of a network channel used for packet data transmission, based on optimal filtering methods in stochastic differential systems.

The channel's throughput is determined by its bottleneck, which consists of a transmission element with a buffer of known capacity. The channel handles two data flows. The first is a Poisson process with known intensity. It is observable, and available statistics include the current number of data packets in the buffer and the packet losses of the first flow due to congestion. The second flow represents the aggregate of all other traffic passing through the channel. This flow is non-stationary, non-observable, and is modeled as a Cox process, where its intensity is governed by a hidden continuous-time Markov chain.

The goal of available bandwidth monitoring is to estimate the intensity of an additional, stationary data flow that can be transmitted through the channel with the current load without violating service-level agreement (SLA) requirements. These SLA constraints include both the average transmission delay and the packet loss rate caused by congestion.

Theoretically, this applied problem falls into the class of optimal filtering in stochastic dynamic observation systems. The system's state is modeled as a semi-Markov process. Observations consist of both noise-free (perfect) functions of the state and counting

observations whose intensity depends on the state. The optimal filter is expressed as a recursive system of ordinary differential equations that describe the conditional state distribution between observation arrivals, along with algebraic updates to the distribution upon each new observation.
A numerical example is included in the report to demonstrate the performance of the proposed bandwidth estimation method.

On Population Density of Catalytic Branching Random Walk

Ekaterina Vl. Bulinskaya

Lomonosov Moscow State University

The work was supported by the Russian Science Foundation under grant no.24-11-00037 and performed at Steklov Mathematical Institute of Russian Academy of Sciences.

The concentration of particles bounded by
the population propagation front, when time tends to infinity, is studied within the model of catalytic branching random walk on Z^{d}. We assume that the regime is supercritical (thus the Malthusian parameter is positive) and the tails of the random walk jump are light, i.e. the Cramer condition is satisfied. In that case the front is known to spread asymptotically linearly in time (see, e.g., [1] and [2]), so now we consider the layers of particles behind the front, which also grow linearly in time but at a slower rate. It is established that the number of particles in a layer grows exponentially fast although with an index smaller than the Malthusian parameter. Such analysis is started in a quite new paper [3].

References

1. Bulinskaya E.V. Spread of a catalytic branching random walk on a multidimensional lattice // Stoch. Proc. Appl. 2018. Vol. 128, No. 7. P. 2325–2340.
2. Liu R. The spread speed of multiple catalytic branching random walks // Acta Math. Appl. Sin., Eng. Ser. 2023. Vol. 39, No. 2. P. 262–292.
3. Bulinskaya E.V. Population density inside the propagation front of a catalytic branching random walk // J. Math. Sci. 2025. Vol. 292, No. 2. P. 1–20.

Statistical Estimation of Information Characteristics

Alexander V. Bulinskii

Lomonosov Moscow State University

Various concepts of information theory such as entropy, mutual information, conditional interaction information, transfer entropy and f-divergences are widely used in many fields, including the domain of modern mathematical statistics, called the feature selection, see, e.g., Macedo et al. (2019), Bulinski and Kozhevin (2021), Bulinski and Dimitrov (2021), Mielniczuk (2022), Bulinski (2023), Pradip and Chandrashekhar (2023), Bulinski and Wang (2025) and references therein. We discuss different methods for constructing estimates of relevant information characteristics based on a set of independent identically distributed observations. Moreover, we consider the asymptotic behavior of proposed statistical estimates as the number of observations tends to infinity. Some applications to analysis of medical and biological data are indicated as well.

References:

1. Macedo F., Oliveira M.R., Pacheco A., Valadas R. (2019). Теоретические основы методов прямого отбора признаков на основе взаимной информации. *Neurocomputing*, 325, 67–89.
2. Bulinski A., Kozhevin A. (2021). Статистическая оценка взаимной информации для смешанной модели. *Methodol. Comput. Appl. Probab.*, 23, 123–142.
3. Bulinski A., Dimitrov D. (2021). Статистическая оценка дивергенции Кульбака–Лейблера. *Mathematics*, 9:5, 544, 1-36.
4. Mielniczuk J. (2022). Теоретико-информационные методы отбора переменных — обзор. *Entropy*, 24:8, 1–25.
5. Bulinski A.V. (2023). О выборе релевантных признаков на основе теории информации. *Theory Probab. Appl.*, 68:3, 392–410.
6. Pradip D., Chandrashekhar A. (2023). Комплексный обзор отбора признаков в различных областях машинного обучения. *Appl. Intell.*, 52:4, 4543–4581.
7. Bulinski A., Wang S. (2025). Статистическая оценка условной информации взаимодействия методом дельта. *Sankhyā A: The Indian Journal of Statistics*. DOI: 10.1007/s13171-025-00400-1

Probability-Informed Models for Data Processing

Andrey K. Gorshenin

Federal Research Center ``Computer Science and Control'', Russian Academy of Sciences; Lomonosov Moscow State University

The deep learning methods are effective for a wide range of practical problems. However, in real-world scenarios, we have to deal with incomplete and noisy data, which arises due to various random factors. The lack of training data, including labelled data, also creates obstacles to obtaining high-quality research results. To correctly model regularities in real data, it is natural to use the methods of probability theory and mathematical statistics. Mathematical models provide additional features for neural networks and machine learning algorithms. Therefore, we introduce a new approach called probabilistic machine learning. This is an extension of the well-known physics-informed methodology, which initially proposed the use of physics models as sources of information about objects that can be modelled using partial differential equations. We extend this approach based on probabilistic approximations for situations where physics models are not applicable due to various reasons.

Forecasting Ito-type processes using features based on dynamic Gaussian mixtures

Mikhail A. Ivanov

Lomonosov Moscow State University

Many real-world processes, such as stock price dynamics in economics, are often described by Itô-type stochastic processes. Suppose is the process under consideration. Then its dynamics is written as follows: where is the Wiener process, is the drift coefficient and is the diffusion coefficient. These coefficients are unknown and in general can be governed by arbitrarily complex stochastic processes.

An important research problem is estimation of these coefficients. One strand of research [1, 2] proposes to approximate the distribution of increments by finite Gaussian mixtures with PDF. This implies that the dynamics are governed by the random vector with the categorical distribution. Estimates of and are given by the corresponding expectations of [2]. The number of mixture components is a hyperparameter chosen by the user.

The goal of this research is to apply these estimates to forecasting in discrete time (as a time-series) and compare forecasting performance for different numbers of components. For the problem reduces to fitting a single normal distribution and estimating and using the moving average and the moving standard deviation respectively. Cases when require fitting mixtures using maximum likelihood or other methods.

The main idea is to validate the hypothesis that additional mixture components improve forecasting performance compared to the single Gaussian. Validation approaches include running the Diebold-Mariano and Model Confidence Set tests on both simulated and real-world data.

References:

1. Королев, В. Ю. 2011. Вероятностно-статистические методы декомпозиции волатильности хаотических процессов. Изд-во Московского университета.
2. Karpov, Kirill, Victor Korolev, and Natalia Sukhareva. 2025. “Statistical Separation of Mixtures in the Problem of Reconstructing the Coefficients of an Itô Stochastic Process-Type Model of the Interplanetary Magnetic Flux Density: ℓ2-Distance Minimization vs Likelihood Maximization.” Russian Journal of Numerical Analysis and Mathematical Modelling 40 (1): 17–31. https://doi.org/10.1515/rnam-2025-0002.

On some characterization of Marshall-Olkin distribution

Yurii S. Khokhlov, Lev B. Klebanov, Nikolay V. Torbin

Lomonosov Moscow State University

The exponential distribution plays an important role in many applied and theoretical studies. There are several variants of its multidimensional counterparts.
The most popular one is the Marshall-Olkin distribution. Very often, we would like to select this class of distributions based on a certain property.
This is the so-called characterization problem. In 1996, L.B. Klebanov proposed a very interesting characterization of the one-dimensional exponential distribution.
This characterization was then generalized to the case of a two-dimensional Marshall-Olkin distribution. In this report we propose a result that is valid
for the Marshall-Olkin distribution of arbitrary dimension. The main technical approach in the proof of this result is
some recurrent representation of the Laplace transform of the Marshall-Olkin distribution, obtained in the paper of Khokhlov Yu.S. and Korolev V.Yu. in 2021.

Limit Theorems as Characterization Problems in Probability Theory

Lev B. Klebanov

Charles University, Prague, Czech Republic

Limit theorems for sums and extremes of a random number of random variables are considered as problems of characterization of probability distributions. The main attention is paid to the study of differences between stable and limit distributions. In particular, it is shown that for positive random variables, the property of a distribution to be a limit is trivial. The property of stability is essentially significant. It leads to interesting problems of analysis, probability theory, and mathematical statistics. New examples of stable laws are given. It is noted that for sums of a random number of random variables with random normalization, the difference between the continuous and discrete cases is minimal. This applies to both limiting and stable distributions. It is interesting to note that the behavior of the extremes of a random number of random variables differs more significantly from the behavior of the sums than in the classical case.

Rates of convergence for the Functional Limit Theorems for CTRW (continuous time random walks)

Vassiliy N. Kolokotsov

Lomonosov Moscow State University

The study of the rates of convergence of the CTRWs (continuous time random walks, including Levy flights and Levy walks)) to semi-Markov processes described by fractional in time pseudo-differential equations was initiated in author's papers (1) The Rates of Convergence for Functional Limit Theorems with Stable Subordinators and for CTRW Approximations to Fractional Evolutions. Fractal Fract. (2023), 7, 335, and (2) Fractional Equations for the Scaling Limits of Levy Walks With Position Depending Jump Distributions. MDPI Mathematics 2023, 11, 2566. Developing further the ideas from these papers we improve on the rates, on the assumptions and on the applicability of this new theory. Applications include new schemes for the numeric solutions of various fractional equations

Normal variance-mean mixtures as stationary distributions of stochastic difference equations with random coefficients

Victor Yu. Korolev, Nikita R. Romanyuk

Lomonosov Moscow State University; Moscow Center of Fundamental and Applied Mathematics; Federal Research Center "Computer Science and Control", Russian Academy of Sciences

It is shown that any normal variance-mean mixture can be a stationary distribution of stochastic difference equations with random coefficients – the discrete-time autoregressive process of the first order – with random coefficients. The form of drift and diffusion coefficients is described which ensures that a given variance-mean mixture is a stationary distribution of the corresponding stochastic difference equation. It is also demonstrated that a stationary mode of the discrete-time autoregressive process of the first order with random coefficients possesses the property of stability. Both univariate and multivariate cases are considered

Analogs of the multiplication theorem for some families of distributions

Victor Yu. Korolev, Yuri K. Khomutov

Lomonosov Moscow State University; Moscow Center of Fundamental and Applied Mathematics; Federal Research Center "Computer Science and Control", Russian Academy of Sciences

The definitions of the quasi-exponentiated normal (Korolev 2023a) and generalized Student distributions (Korolev 2023b), are extended for wider ranges of parameters. Analogs of the multiplication theorem for stable laws (Theorem 3.3.1 in (Zolotarev 1986) are formulated for generalized gamma-distributions, beta-distributions, generalized Student and exponential power distributions. These analogs demonstrate that these distributions can be represented as scale mixtures of the same distributions with greater parameters. Also, a representation of a strictly stable distribution in the form of a scale mixture of a special non-stable law. This alternative representation is a complement to the multiplication theorem for stable laws.

References:

1. V. Korolev. Quasi-Exponentiated Normal Distributions: Mixture Representations and Asymmetrization // Mathematics, 2023. Vol. 11. No. 17. P. 3797. DOI: 10.3390/math11173797

2. V. Korolev. Analytic and Asymptotic Properties of the Generalized Student and Generalized Lomax Distributions // Mathematics, 2023. Vol. 11. No. 13. Art. 2890. DOI: 10.3390/math11132890
3. V. M. Zolotarev. One-dimensional Stable Distributions. – Providence, RI: American Mathematical Society, 1986.

On the method of estimating the parameters of scale mixtures of generalized gamma distributions

Alexey A. Kudryavtsev, Oleg V. Shestakov

Lomonosov Moscow State University

The report discusses the method of estimating the parameters of a new digamma distribution, which is a scale mixture of two generalized gamma distributions. The difficulties of using classical methods of statistical estimation are described. An alternative method based on the logarithmic moments of the distribution is proposed. An analytical form of estimates of unknown parameters is given

Factors of the Price Gap between Primary and Secondary Housing Markets: A Regional Analysis

Ianina A. Roshchina

Lomonosov Moscow State University and MSU-BIT University

This study examines the price gap between primary and secondary housing markets in Russian regions during 2022–2025. The analysis is based on 24.5 million apartment sale listings processed using a two-stage matching procedure to pair properties with comparable characteristics. The results reveal substantial heterogeneity: in most regions new housing is more expensive, but cases of zero or negative gaps are also observed. At the second stage, a factor analysis was conducted using panel and spatial fixed-effects models. The findings show that the magnitude of the gap depends not only on macroeconomic conditions, but also on industry structure and the availability of financial services. These results have practical implications for both homebuyers and regulators, highlighting the need to account for the institutional environment when designing mortgage support programs.

Asymptotic results for the mean-square risk when using multiple hypothesis testing methods for weakly dependent observations

Oleg V. Shestakov, Mikhail O. Vorontsov

Lomonosov Moscow State University

We consider an approach to solving the problem of noise removal in a large array of sparse data under conditions of weak dependence, based on the method of controlling the average proportion of false hypothesis rejections. The statements about the strong consistency and asymptotic normality of the SURE estimate are proved. The rates of convergence of the estimate distribution to the normal law are also obtained.

Rates of convergence of random sums to the normal variance-mean mixtures

Irina G. Shevtsova

Lomonosov Moscow State University

The class of normal variance-mean mixtures is very wide, in particular, it contains the generalized hyperbolic, the generalized variance gamma, the skew Student distribution, the skew exponential power law, the skew two-sided Weibull law (asymmetric Weibull distribution of the second kind), asymmetric Linnik distribution of the third kind, M-asymmetric quasi-exponentiated normal law, Skew logistic distribution. We present convergence rate estimates of compound Cox processes to the normal variance-mean mixtures in a wide class of mean metrics including the Kolmogorov and the Kantorovich ones.

Asymptotics of Worst-Case Mixed Market Strategies for a Binary Option is a Stopped Geometrical Brownian Motion

Sergey N. Smirnov

Lomonosov Moscow State University

We employ a Guaranteed Deterministic Approach (GDA) to address the superhedging problem in discrete time, aiming to ensure coverage of contingent liabilities from a sold option across all admissible scenarios with minimal premium. Specifically, we analyze Kolokoltsov's multiplicative model of market price dynamics without trading constraints, focusing on a European-style binary option. The mathematical interest in this problem stems from the discontinuity of the option's payoff function. The GDA adopts a game-theoretic framework, leading to the Bellman–Isaacs equations. Unlike its probabilistic

counterpart, GDA directly characterizes the most unfavorable mixed market strategy, offering a more interpretable solution. Our 2021 paper derives key properties of the solutions to these equations and proposes an algorithm for their numerical computation, alongside formulating several hypotheses. In our 2024 paper, we establish new analytical properties of the Bellman–Isaacs equation solutions and confirm the previously proposed hypotheses. We also prove a general result for options with monotonic payoff functions, demonstrating that, under mild assumptions, the GDA yields superhedging prices equivalent to those of the traditional probabilistic approach, specifically for binary options. Furthermore, we show that the price process, under conditional distributions corresponding to the worst-case market scenarios, can be approximated on a logarithmic scale by a random walk with two absorbing barriers. Under appropriate normalization, as the number of steps in the discrete-time model approaches infinity, the price process weakly converges to a geometric Brownian motion with a single absorbing barrier at the strike price.

References

1. Smirnov, S.N., Zanochkin, A.Y.: Guaranteed deterministic approach to superhedging: case of binary European option. Abstr Appl Anal 1–18 (2021). https://doi.org/10.1155/2021/5568636
2. Smirnov, S., Sotnikov, D. & Zanochkin, A. Approximation and asymptotics in the superhedging problem for binary options. Ann Finance 20, 421–458 (2024). https://doi.org/10.1007/s10436-024-00454-5

On Hanson-Wright type deviation inequalities for α-subexponential entries

Vladimir V. Ulyanov

Lomonosov Moscow State University

We obtain a p-th moment bound for the suprema of a log-concave-tailed non-homogeneous chaos process, which is optimal in some cases. With this p-th moment bound, we get two uniform Hanson-Wright type deviation inequalities for α-subexponential entries ($1 \le \alpha \le 2$), which generalize some known results.

The talk is based on joint results with G.Dai , Z.Su (both from Zhejiang University) and H.Wang (Shandong University).

Limit theorems for branching random walks in various branching environments

Elena B. Yarovaya

Lomonosov Moscow State University

Continuous-time two-type branching random walks are considered on a multidimensional lattice. The main results are devoted to the study of the generating function and the limiting behaviour of the moments of subpopulations generated by a single particle of each type. It is assumed that particle types differ from each other not only by the laws of branching, as in multi-type branching processes, but also by the laws of walking. For a critical branching process at each lattice point and recurrent random walk of particles with a finite variance of jumps, the effect of limited spatial clustering of particles over the lattice is studied. Simulation results are presented to demonstrate the effect of limit clustering, see Makarova et al. (2022). For the local number of particles of each type at a lattice point, the moments and their limiting behaviour are studied. For a branching random walk under additional assumptions on particle transformations, the limit theorem on mean square convergence of the normalized number of particles of each type in an arbitrary fixed point of the lattice is proved by Smorodina and Yarovaya (2024). For a branching random walk with a single generation center under different assumptions on the structure of a killing medium, a condition guaranteeing the appearance of exponential growth of the average number of particles in each point of the lattice have been obtained by Filichkina and Yarovaya (2024). Limit theorems based on these results in various branching environments are presented.

An approach for bounding the rate of convergence for nonstationary queueing systems

Alexander I. Zeifman, Yacov A. Satin, Ilia A. Usov, Galina Shilova

Vologda State University, Federal Research Center ``Computer Science and Control'' of the Russian Academy of Sciences

We consider the problem of bounding the rate of convergence to the limiting characteristics for general time-nonstationary queuing systems, which are described by inhomogeneous continuous-time Markov chains. The simplest and most convenient method is to use the logarithmic norm. However, the direct approach is not always applicable, and here the c-matrix method, first used in [1], can help, see also [2]. Here we provide a general description of this approach and consider a number of classes of models that can be studied using it. In particular, these are systems with catastrophes

(disasters), including the recently considered and studied systems with several types of catastrophes [3]. The most important class of Markov chains that can be studied using the c-matrix method are quasi birth-death processes, see for example [4]. In this lecture we consider the description of the method and the models that can be studied using it.

References:

1. Satin Y., et al. Upper bounds on the rate of convergence for constant retrial rate queueing model with two servers // Statistical Papers. 2018. Vol. 59, No. 4. P. 1271–1282.
2. Razumchik R., Rumyantsev A.S. Some ergodicity and truncation bounds for a small scale Markovian supercomputer model // ECMS Proceedings. 2022. P. 324–330. DOI: 10.7148/2022-0324
3. Li J. Birth–death processes with two-type catastrophes // Mathematics. 2024. Vol. 12, No. 10. P. 1468.
4. Usov I., Satin Y., Zeifman A. Estimating the rate of convergence of the PH/M/1 model by reducing to quasi-birth-death processes // Mathematics. 2023. Vol. 11, No. 6. P. 1494.

Rates of Ergodicity in Lévy-Driven Storage Processes

Miha Bresar, Aleksandar Mijatovic and Nikola Sandric

Chinese University of Hong Kong

In this talk, I will present new results on the Levy-driven storage process $X(t) = x + A(t) - \int_0^t r(X(s))ds$, where A is a nondecreasing zero-drift Levy process with Levy measure ν and r is a general release rule. We give sharp criteria distinguishing transience, null recurrence, and positive recurrence, and obtain explicit convergence rates to stationarity, subgeometric, geometric, and uniform, in both total variation and Wasserstein metrics, together with precise tail asymptotics of the invariant measure. Applications to classical storage models illustrate the range of ergodic behaviors. Our proofs combine clasical Foster–Lyapunov techniques with recent L-drift conditions. The talk is based on joint work with Aleksandar Mijatovic (UNIVERSITY OF WARWICK) and Nikola Sandric (UNIVERSITY OF ZAGREB)

Modified Information Criterion for Testing Changes in the Inverse Gaussian Degradation Process

Xia Cai

Hebei University of Science and Technology

The Inverse Gaussian process is a useful stochastic process to model the monotonous degradation process of a certain component. Owing to the phenomenon that the degradation processes often exhibit multi-stage characteristics because of the internal degradation mechanisms and external environmental factors, a change-point Inverse Gaussian process is studied in this paper. A modified information criterion method is applied to illustrate the existence and estimate of the change point. A reliability function is derived based on the proposed method. The simulations are conducted to show the performance of the proposed method. As a result, the procedure outperforms the existing procedure with regard to test power and consistency. Finally, the procedure is applied to hydraulic piston pump data to demonstrate its practical application

Central limit theorem for irregular discretization scheme of multilevel Monte-Carlo method

Yi Guo

Shandong University

In this talk, we study the asymptotic error distribution for a two-level irregular discretization scheme of the solution to the stochastic differential equations (SDE for short) driven by a continuous semimartingale and obtain a central limit theorem for the error processes with the rate $\sqrt{n}$. As an application, in the spirit of the result of Ben Alaya and Kebaier, we get a central limit theorem of the Linderberg-Feller type for the irregular discretizationscheme of the multilevel Monte Carlo method.

Two-way Popularity Model for Directed and Bipartite Networks

Bingyi Jing

Southern University of Science and Technology

There has been extensive research on community detection in directed and bipartite networks. However, they often fail to consider the popularity of nodes in different communities, which is a common phenomenon in real-world networks. Here we propose a new probabilistic framework called the Two-Way Popularity Model, which offers several advantages in terms of estimation accuracy and computational efficiency

Reliability analysis based on the Wiener process integrated with historical degradation data

Wenda Kang

Anhui University

For electronic devices that require high reliability and long service life, conducting accurate reliability analysis with limited degradation data from small samples remains challenging. There has been growing interest in utilizing abundant historical degradation data to enhance such analyses. Effectively extracting and integrating information from historical records to improve current reliability assessments has thus become an important research focus. This paper proposes a novel approach based on the Wiener process to integrate historical and current degradation data, under the assumption of consistent failure mechanisms across different batches. Simulation results show that the proposed method not only improves the estimation accuracy of reliability but also remains robust against deviations in the failure mechanism consistency assumption. Finally, the method is validated using a real-world dataset of degradation measurements from metal-oxide-semiconductor field-effect transistors (MOSFETs), demonstrating its practical applicability.

Debiased distributed PCA under high dimensional spiked model

Zeng Li

Southern University of Science and Technology

We study distributed principal component analysis (PCA) in high-dimensional settings under the spiked model. In such regimes, sample eigenvectors can deviate significantly from population ones, introducing a persistent bias. Existing distributed PCA methods are sensitive to this bias, particularly when the number of machines is small. Their consistency typically relies on the number of machines tending to infinity. We propose a debiased distributed PCA algorithm that corrects the local bias before aggregation and incorporates a sparsity-detection step to adaptively handle sparse and non-sparse eigenvectors. Theoretically, we establish the consistency of our estimator under much weaker conditions compared to existing literature. In particular, our approach does not require symmetric innovations and only assumes a finite sixth moment. Furthermore, our method generally achieves smaller estimation error, especially when the number of machines is small. Empirically, extensive simulations and real data experiments demonstrate that our method consistently outperforms existing distributed PCA approaches. The advantage is especially prominent when the leading eigenvectors are sparse or the number of machines is limited. Our method and theoretical analysis are also applicable to the sample correlation matrix

Smoothly managing dependence and heavy tails in concentration inequalities on sums

Cosme Louart

Chinese University of Hong Kong

In this talk, we will present new tools from concentration of measure theory and financial statistics to derive concentration inequalities for sums of random variables. First, we introduce a new proof of the Fuk-Nagaev inequality that avoids truncation and instead relies on a Talagrand-type concentration inequality for convex functions of random vectors with bounded independent entries. Second, we establish an optimal concentration inequality for sums of random variables with only a bounded first moment and no independence assumption. This result is achieved using techniques involving superquantiles—a concept popularized by Rockafellar in financial statistics.

Singularity of biased discrete random matrices

Zeyan Song

Shandong University

We study the singularity probability of $n \times n$ random matrices with i.i.d. entries from highly biased discrete distributions. We obtain sharp non-asymptotic bounds for this probability and derive estimates on the least singular values. Our method combines combinatorial, geometric, and probabilistic techniques such as sphere decomposition and anticoncentration inequalities. The results extend classical invertibility theory to biased discrete settings and resolve an open problem by characterizing the dominant causes of singularity in biased discrete random matrices, namely the presence of zero columns or linearly dependent column pairs.

Statistical inference for power autoregressive conditional duration models with stable innovations

Yuxin Tao

Southern University of Science and Technology

Financial data often exhibit irregular intervals in terms of transaction time, and the study of transaction timing has become increasingly important in empirical finance. However, the durations between financial events sometimes display heavy-tailed characteristics, which current duration models fail to capture. In this paper, we propose a first-order power autoregressive conditional duration model with positive stable innovations (sPACD), which effectively addresses the excess kurtosis in the durations. Further, the power form of the model structure mitigates the issue of overpredicting short durations in the classical ACD model. We study the properties of maximum likelihood estimation (MLE) within a unified framework of stationary and explosive cases, and discuss the estimation of the asymptotic covariance matrix. Strict stationarity test statistics and a modified Kolmogorov-type test statistic are established for stationarity testing and diagnostic checking in both stationary and explosive scenarios. Monte Carlo simulation studies demonstrate the good performance of the MLE and test statistics in finite samples. An empirical example is analyzed to illustrate the usefulness of sPACD models

Sequential Design for Quantile Estimation in Multifactor Sensitivity Experiments

Dianpeng Wang

Beijing Institute of Technology

Cavitation in hydrofoil experiments is a critical physical phenomenon, which is primarily influenced by three key factors. Efficiently identifying the combination of factor levels at which cavitation occurs with 50% probability presents a significant challenge in sensitivity testing, particularly when constrained by the high cost of hydrofoil experiments. Despite its importance, sensitivity experiments with multiple factors have received scant attention in literature. In order to tackle this challenge, we propose a novel sequential design based on the upper credible bound criterion for the generalized Gaussian process models, which is used to model the hydrofoil experimental data. Some numerical simulation studies demonstrate that the proposed method effectively balances exploitation and exploration, achieving superior performance compared to existing approaches. Finally, a practical application to hydrofoil cavitation experiments validates the superiority of the proposed method.

Successful Couplings and Strong Ergodicity for Switching Diffusion Systems with Past-Dependent Switching

Fubao Xi

Beijing Institute of Technology

In this work we consider a class of switching diffusion systems consisting of continuous and discrete components, in which the switching rates of discrete component depend on the value of the continuous component involving past history. Motivated by the aim to study the strong ergodicity or uniform ergodicity in the sense of convergence in total variation norm, we construct a type of couplings for this class of switching diffusion systems, and give some sufficient conditions to guarantee this type of couplings to be successful. In addition, we also provide some illustrative examples.

On the empty balls of critical and subcritical super-Brownian motions with general branching mechanisms

Jie Xiong

Southern University of Science and Technology

In this talk, I will explore various limiting behavior of the radius of the largest ball around the origin which is not occupied by a super-Brownian motion and that not by a branching random walk according to the spatial dimension as time tends to infinity. This talk is based on two joint papers with Shuxiong Zhang and Jiawei Liu

A general varying terminal time structure for stochastic optimal control

Shuzhen Yang

Shandong University

In this study, we propose a general varying terminal time structure for the stochastic optimal control problem under state constraints. For this novel optimal control model, we establish the related stochastic maximum principle, and clarify the relationship between mean-filed optimal control problem and the state constraints problem.

On the passage times of self-similar Gaussian processes on curved boundaries

Cheuk Yin Lee

Chinese University of Hong Kong

In this talk, we will present some results about the passage time T_c for a continuous, α-self similar Gaussian process to cross the curved boundary $\pm ct^{\beta}$. In the super-critical case$\beta > \alpha$, T_c is infinite with positive probability. In the sub-critical case $0 < \beta < \alpha$, T_c has moments of all order. In the critical case $\beta = \alpha$, there exists a decreasing convex function $\lambda(c)$ such that $P(T_c > t) = t^{-\lambda(c)+o(1)}$ for large times, and the asymptotic behavior of λ at 0 and infinity can be determined. We will also discuss some examples and numerical challenges in computation of the critical exponent function. This is based on joint work with Davar Khoshnevisan (University of Utah)

The smallest singular value of sparse discrete random matrices

Kexin Yu

Shandong University

Let M_n be an n × n random matrix with i.i.d. discrete sparse entries. In this talk, we develop a simple framework to solve the approximate Spielman-Teng theorem for M_n, which has the following form: There exist constants C, c>0 such that for all $\eta \geq 0$, $P(s_n(M_n) \leq \eta) \lesssim n^C\eta + e^{-n^c}$. As an application, we give an approximate Spielman-Teng theorem for M_n whose entries are μ-lazy random variables, extending previous work by Tao and Vu.

Convergence rates of multivariate normal approximation with applications to stochastic approximation methods

Zhuosong Zhang, Zhijun Cai, Qi-Man Shao, Jingcai Yang

Southern University of Science and Technology

In this talk, we obtain Berry-Esseen bounds for multivariate normal approximation of nonlinear statistics and exchangeable pairs. As applications, convergence rates for SGD and LSA will be discussed. The proof is based on Stein's method, the concentration inequality approach and a recursive method. This talk is based on joint works with Zhijun Cai, Qi.-Man Shao, and Jingcai Yang.

On mean-field super-Brownian motions

Jiayu Zheng

Shenzhen MSU-BIT University

The mean-field stochastic partial differential equation (SPDE) corresponding to a mean-field super-Brownian motion (sBm) is obtained and studied. In this mean-field sBm, the branching-particle lifetime is allowed to depend upon the probability distribution of the sBm itself, producing an SPDE whose space-time white noise coefficient has, in addition to the typical sBm square root, an extra factor that is a function of the probability law of the density of the mean-field sBm. This novel mean-field SPDE is thus motivated by population models where things like overcrowding and isolation can affect growth. A two step approximation method is employed to show existence for this SPDE under general conditions. Then, mild moment conditions are imposed to get uniqueness. Finally, smoothness of the SPDE solution is established under a further simplifying condition