Speakers

Abstracts

A Tale of Rare Events — Symposium in Honour of Louis Chen on his 82nd Birthday

(17–21 Jun 2022)

1 Xiao Fang

The Chinese University of Hong Kong, China

Multivariate normal approximations by Stein's method

Abstract

In this talk, I will report some recent progress in multivariate normal approximations by Stein's method, emphasizing error bounds with optimal dependence on the sample size and improved dependence on the dimension. The distributional distances under consideration are the convex-set distance, the Kolmogorov distance, and the Wasserstein distance. This talk is based on joint works with Yuta Koike.

2 Hans Föllmer

Humboldt-Universität zu Berlin, Germany

Entropy bounds and optimal couplings for measures on Wiener space

Abstract

For a probability measure Q on Wiener space, we discuss the effect of relative entropy bounds with respect to Wiener measure P on the structure of Q. This will involve the construction of an intrinsic drift and of an intrinsic Wiener process for Q. As a result, we obtain couplings between Q and P that are optimal with respect to certain Wasserstein metrics.

3 Larry Goldstein

University of Southern California, USA
Zero biased enhanced Stein couplings

Abstract

The Stein couplings of Chen and Roellin (2010) vastly expanded the range of applications for which coupling constructions in Stein's method for normal approximation could be applied, and subsumed both Stein's classical exchangeable pair, as well as the size bias coupling. A further simple generalization includes zero bias couplings, and also allows for situations where the coupling is not exact. The zero bias versions result in bounds for which often tedious computations of a variance of a conditional expectation is not required. An example to the Lightbulb process shows that even though the method may be simple to apply, it may yield improvements over previous results that had achieved bounds with optimal rates and small, explicit constants.

The paper can be found in `https://arxiv.org/abs/2206.05413`

4 Xuming He

University of Michigan, USA
Convolution-type smoothing approach for quantile regression

Abstract

Quantile regression is a powerful tool for exploring heterogeneous effects in regression analysis. Scalable computation and inference for quantile regression with large-scale data is the focus of this presentation. We provide an in-depth analysis of a convolution-type smoothing approach that achieves adequate approximation to computation and inference for quantile regression. This method, which we refer to as conquer, turns the non-differentiable quantile loss function into a twice-differentiable, convex and locally strongly convex surrogate, which admits a fast and scalable Barzilai-Borwein gradient-based algorithm to perform optimization, and multiplier bootstrap for statistical inference. We discuss the theoretical properties and demonstrate superior empirical evidence of the proposed method, as well as extensions of the

smoothing approach to penalization-based methods for higher-dimensional problems.

5 Tailen Hsing

University of Michigan, USA
Spectral density estimation of function-valued spatial processes

Abstract

The spectral density of a stationary process characterizes the second-order properties of the process. We are interested in the estimation of the spectral density of a continuous-parameter stochastic process that takes values in an infinite-dimensional Hilbert space. We assume that the process is observed at irregularly-spaced points, which is common in spatial statistics. We consider a lag-window estimator and explore the asymptotic properties of the estimator.

6 Hsien-Kuei Hwang

Academia Sinica, Taipei
Limit theorems for Fishburn matrices

Abstract

Fishburn matrices are upper-triangular square matrices with nonnegative integers as entries such that no row and no column contains exclusively zeros. They have been found to be bijectively equivalent to several other combinatorial structures such as (2+2)-free posets, ascent sequences, certain pattern-avoiding permutations, (2-1)-avoiding inversion sequences, Stoimenow matchings, and regular linearized chord diagrams. In addition to their rich combinatorial connections and modeling capabilities, the corresponding asymptotic enumeration and the finer distributional properties are equally enriching and challenging. We present a general, powerful, direct, analytic approach for dealing with the asymptotic enumeration and stochastic properties on generalized Fishburn matrices. Many normal and Poisson limit laws are given, and they are helpful in describing the typical shape of general random Fishburn matrices.

7 Wilfrid Kendall

University of Warwick, UK

Scale-invariant Spatial Random Networks everywhere?

Abstract

Scale-invariant random spatial networks (SIRSN) are remarkable random structures providing patterns of random routes that are scale-invariant, thus modelling scale-invariance in online maps. It is not a priori apparent that such networks exist at all! Nevertheless Aldous and Ganesan (2013) constructed a SIRSN based on a dense dyadic mesh, and WSK and Kahn showed how to build a SIRSN from a dense Poisson line process. But both these constructions involve infinitely long linear paths, and in particular do not easily admit local influences. Can one do better? I shall discuss recent work showing that SIRSN can arise under far less stringent conditions.

Aldous, D J."Scale-Invariant Random Spatial Networks." Electronic Journal of Probability 19, no. 15 (April 2014): 1-41. https://doi.org/10.1214/EJP.v19-2920.

Aldous, D J, and K Ganesan. "True Scale-Invariant Random Spatial Networks." Proceedings of the National Academy of Sciences of the United States of America 110, no. 22 (December 18, 2013): 8782-85. https://doi.org/10.1073/pnas.1304329110.

Kahn, J. "Improper Poisson Line Process as SIRSN in Any Dimension." Annals of Probability 44, no. 4 (July 2016): 2694-2725. https://doi.org/10.1214/15-AOP1032.

WSK "From Random Lines to Metric Spaces." Annals of Probability 45, no. 1 (2017): 469-517. https://doi.org/10.1214/14-AOP935.

8 Wei-Liem Loh

National University of Singapore, Singapore
On fixed-domain asymptotics, parameter estimation and isotropic Gaussian random fields with Matérn covariance functions

Abstract

The talk is divided into two parts.

In first part of the talk, I shall share with the audience my recollections of Louis over the years.

In the second part of the talk, I shall give an overall view of a recent paper of mine (joint with S Sun and J Wen) that is tangentially related to Louis. More precisely, I shall focus on the problem of estimating the smoothness of an isotropic Matern Gaussian random field under fixed-domain asymptotics using possibly non-lattice data. Technical details and notation will be kept to a minimum.

9 Wei-Yin Loh

University of Wisconsin, Madison, USA
Post-selection inference for regression trees

Abstract

To prevent data snooping and over-fitting, statistical analysis has traditionally been constrained by limiting the number of fitted models and tested hypotheses. Submission plans for clinical trials, for example, typically require all models and hypotheses to be prespecified before data collection. Failure to do so can lead to higher than anticipated rates of false discoveries. This has given rise recently to vocal calls to ban the use of p-values and formal statistical inference in general in scientific journals.

There is no fundamental reason for limiting data exploration except for the ineffectiveness of traditional statistical methods to properly control for multiplicity of tests. Bonferroni corrections are traditionally used, but they are very hard to apply when analysis steps are unplanned or undocumented. Besides, Bonferroni corrections tend to be highly conservative if the number of tests is large. Machine learning algorithms, such as regression trees, provide a timely solution. Once an algorithm is turned into computer code, it

can be used for virtually unlimited exploration while allowing for easy multiplicity control through bootstrap calibration. This talk shows how to achieve this.

10 Ivan Nourdin

Université du Luxembourg, Luxembourg
The Breuer-Major theorem: old and new

Abstract

Let $B = (B_x)_{x\in^d}$ be a collection of standard Gaussian random variables, forming a continuous stationary Gaussian field. Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a measurable function, square integrable with respect to the standard Gaussian density. Finally, set $Y_t = \int_{\{\|x\|\leq t\}} \varphi(B_x)dx$, where $\|\cdot\|$ denotes the Euclidean norm of ^{d}.

Since the pioneering works in the 1980s by Breuer, Dobrushin, Major, Rosenblatt, Taqqu, and others, central and noncentral limit theorems for Y_t have been constantly refined, extended and applied to an increasing number of diverse situations, to such an extent that it has become a field of research in its own right.

The common belief, representing the intuition that specialists in the subject have developed over the last four decades, is that when $t \to \infty$ the fluctuations of Y_t around its mean are, in general (i.e. except possibly in very special cases), Gaussian when B has short memory and non-Gaussian when B has long memory (and the Hermite rank of f is different from 1).

We will show in this talk that this intuition forged over the last forty years can be wrong, and not only marginally or in critical cases. We will indeed bring to light a variety of situations where Y_t admits Gaussian fluctuations in a long memory context.

This is based on a work in collaboration with my PhD student, Leonardo Maini.

11 Giovanni Peccati

University of Luxembourg, Luxembourg

The fourth-moment theorem for U-statistics: Stein's bounds and functional fluctuations

Abstract

One striking demonstration of the power and flexibility of Stein's method is its ability to quantitatively capture "fourth-moment phenomena" for sequences of random variables belonging to eigenspaces of Markov operators - like e.g. multiple stochastic integrals with respect to independently scattered random measures. In my talk, I will focus on the so-called "de Jong Theorem" for degenerate U-statistics (1990), which is one of the ancestors of such a fruitful line of research. My aim is twofold

(i) On the one hand, I will show how the classical method of exchangeable pairs allows one to recover quantitative multivariate versions of de Jong's findings;

(ii) On the other hand, I will illustrate several new functional extensions of de Jong's theorem, yielding invariance principles for sequential U-processes, both in the symmetric and non-symmetric cases.

The results discussed at Point (ii) - which for the time being lie outside the scope of functional versions of Stein's method - yield in particular substantial generalizations of the Miller-Sen invariance principle for U-processes (1972), as well as functional CLTs for random graph statistics and functional versions of the "Universality of Wiener chaos phenomenon" first detected by Nourdin, Peccati, and Reinert (2010).

My presentation is mainly based on the following references:

1. Ch. Döbler and G. Peccati: Quantitative de Jong Theorems in any dimension. EJP, 2016.

2. Ch. Döbler, M. Kasprzak and G. Peccati: Weak convergence of U-processes with size-dependent kernels. Ann. App. Prob., 2022

3. Ch. Döbler, M. Kasprzak and G. Peccati. The multivariate functional de Jong CLT. Probab. Th. Rel. Fields, 2022+

12 Shige Peng

Shandong University, China
Convergence rate of CLT by Stein method under uncertainty

Abstract

How to obtain the convergence rate of Central Limit Theorem, in the framework of sublinear expectation, has been a challenging problem for a quit long period. A main obstacle is the nonlinearity of the OU-operator in the corresponding Stein characterization.

13 Gesine Reinert

University of Oxford, UK
Stein's method for exponential random graph models and assessing goodness of fit

Abstract

Exponential random graph models are a key tool in network analysis but due to an intractable normalising constant are difficult to manipulate. In this talk we shall use Stein's method to approximate these models by Bernoulli random graphs in "high temperature" regimes.

For assessing the goodness of fit of a model, often independent replicas are assumed. When the data are given in the form of a network, usually there is only one network available. If the data are hypothesised to come from an exponential random graph model, the likelihood cannot be calculated explicitly. Using a Stein operator for these models we introduce a kernelized goodness of fit test and illustrate its performance.

Finally, we extend the ideas of this goodness of fit test to provide an approximate goodness of fit test for potentially black-box graph generators.

This talk is based on joint work with Nathan Ross and with Wenkai Xu.

14 Qi-Man Shao

Southern University of Science and Technology, China
Berry-Esseen bounds for general non-normal approximation with unbounded exchangeable pairs

Abstract

Let W be a random variable of interest. Construct W' so that (W, W') is an exchangeable pair. Let $\Delta = W - W'$. Assume that

$$E(\Delta \mid W) = \lambda(g(W) + r_1), \ E(\Delta^2 \mid W) = 2\lambda(v(W) + r_2).$$

Let Y be a random variable with the probability density function

$$p(y) = \frac{c}{v(y)} \exp\Big(- \int_0^y g(t)/v(t)dt\Big).$$

In this talk, a general Berry-Esseen bound will be given for W approximated by Y. Applications to Pearson's χ^2-test, Polya urn model and isotropic mean-field Heisenberg model at the critical value will be discussed. The talk is based on a join work with Songhao Liu and Hao Shi.

15 David Siegmund

Stanford University, USA
Changepoints and probability

Abstract

Dedicated to Louis Chen, Friend and Colleague, For the Celebration of his 82nd Birthday A large number of changepoint problems lead to probability approximations involving the crossing of nonlinear boundaries by random walks and related processes. I will describe (now) classical and new approximations and applications. A few of these approximations have been addressed by Stein's method. Are there others that might be?

16 Weng Kee Wong

UCLA Fielding School of Public Health, USA
Metaheuristics for designing efficient large scale complex experiments with theoretical convergence

Abstract

Metaheuristics seems greatly under-utilized in the statistical community and this talk provides a brief overview of metaheuristics as a general-purpose optimization tool. As illustrations, we demonstrate its usefulness in finding a variety of hard-to-find optimal designs for complex models or high-dimensional models and discuss biomedical applications. We also discuss some recent mathematical advances in metaheuristics, which was made possible from a workshop on "Particle Swarm Optimization and Evolutionary Computation" sponsored by IMS at NUS a few years ago.

Joint talk with Kwok Pui Choi, National University of Singapore