Waseda International Symposium

"High Dimensional Statistical Analysis for Time Spatial Processes & Quantile Analysis for Time Series"

Date: February 27 - March 1, 2017.

Venue: Waseda University, Nishi-Waseda Campus, Building 63, 1st Floor, Math & Applied Math Meeting Room
(Access map: http://www.sci.waseda.ac.jp/eng/access/)

Organizer: Masanobu TANIGUCHI (Research Institute for Science & Engineering, Waseda University)

Supported by Kiban (A-15H02061) & Houga (26540015)

Waseda International Symposium

"High Dimensional Statistical Analysis for Time Spatial Processes & Quantile Analysis for Time Series"

Date: February 27 - March 1, 2017.

Venue: Waseda University, Nishi-Waseda Campus, Building 63,1st Floor, Math & Applied Math Meeting Room
(Access map: http://www.sci.waseda.ac.jp/eng/access/)

Organizer: Masanobu TANIGUCHI (Research Institute for Science & Engineering, Waseda University)

Supported by

(1) Kiban (A-15H02061) M. Taniguchi, Research Institute for Science & Engineering, Waseda University

(2) Houga (26540015) M. Taniguchi, Research Institute for Science & Engineering, Waseda University

Program (* Speaker)

February 27

Session (I): 13:30 -15:10: chaired by Tommaso Proietti

13:30 - 13:40: Masanobu Taniguchi (Waseda Univ.)
Opening

13:40 - 14:25: Yan Liu (Waseda Univ.)
Robust parameter estimation for irregularly observed stationary process

14:25 -15:10: Cathy W. S. Chen* (Feng Chia Univ.) and S.Lee
Bayesian Causality Test for Integer-valued Time Series Models with Applications to Climate and Crime Data

15:10 - 15:30: Break

Session (II): 15:30 - 17:00: chaired by C.W.S. Chen

15:30 - 16:15: Tommaso Proietti* (Università di Roma) and Alessandro Giovannelli
Optimal linear prediction of stochastic trends

16:15 - 17:00: S. N. Lahiri* (North Carolina State University) , Arindam Chatterjee and Debraj Das
Higher order asymptotic properties of the Bootstrap in post model selection inference in high dimensional regression models

February 28

Session (III) 9:30 -10:30: chaired by N.H. Chan

9:30 - 10:00: Yurie Tamura* (Waseda Univ.) and Masanobu Taniguchi
Asymptotic theory of sphericity test statistic for high-dimensional time series

10:00 - 10:30: Kou Fujimori* (Waseda Univ.) and Yoichi Nishiyama
The Dantzig selector for diffusion processes with covariates

10:30- 10:40: Break

Session (IV): 10:40 - 12:10: chaired by Mei Hui

10:40 - 11:25: Hiroshi Shiraishi* (Keio Univ.) and Zudi Lu
Semi-parametric estimation for optimal dividend barrier with insurance portfolio

11:25 - 12:10: Yasutaka Shimizu (Waseda Univ.)
Applications of central limit theorems for equity-linked insurance

12:10 - 13:20: Lunch

Session (V): 13:20 - 14:50: chaired by Mohsen Pourahmadi

13:20 - 14:05: Makoto Aoshima* (Univ.Tsukuba) and Kazuyoshi Yata
High-dimensional statistical analysis based on the inference of eigenstructures

14:05 - 14:50: Yosihiko Ogata (Inst. Stat. Math.)
Forecasting of a large earthquake - An outlook of the research

14:50 - 15:10: Break

Session (VI): 15:10 - 17:25: chaired by Masanobu Taniguchi

15:10 - 15:55: Alessandra Luati*(Univ. of Bologna)
Quantum information and statistical inference

15:55 - 16:40: Ngai Hang Chan*(Chinese Univ.of Hong Kong), K.C. Cheung and Sam P.S. Wong
Directed Markov Multi-graphs as Models of Networks

16:40 - 17:25: Mohsen Pourahmadi* (Texas A&M Univ.) and Raanju Ragavendar Sundararajan
Nonparametric Change Point Detection in Multivariate Nonstationary Time Series

18:00 - : Buffet Party

March 1

Session (VII): 9:30 - 10:30: chaired by Arthur Pewsey

9:30 - 10:00: Yoshihiro Suto (Waseda Univ.)
Shrinkage estimation for the autocovariance matrix of vector-valued Gaussian processes

10:00 - 10:45: Fumiya Akashi*(Waseda Univ.), S. Bai and M. Taqqu
A self-normalized block sampling method to quantile regression on time series

10:45 - 10:55: Break

Session (VIII): 10:55 - 12:25: chaired by S. N. Lahiri

10:55 - 11:40: Ritei Shibata*(Waseda & Keio Univ.) and Yuuki Rikimaru
Non-Identifiability of Simultaneous Autoregressive Model and Its Inference

11:40 - 12:25: Arthur Pewsey (Univ. of Extremadura)
An Introduction to Directional Statistics

Abstracts

Yan Liu:
Robust parameter estimation for irregularly observed stationary process

Abstract: We define a new class of disparities for robust parameter estimation of irregularly observed stationary process. The proposed disparities of spectral densities are derived from the point of view of prediction problem. The proposed disparities are not contained in the class of either location disparities or scale disparities. We investigate asymptotic properties of the minimum contrast estimators based on the new disparities for irregularly observed stationary processes with both finite and infinite variance innovations. The method provides a new way to estimate parameters robust against irregularly observed stationary process with infinite variance innovations. The relative efficiencies under regularly observed cases and the ratio of mean squared error under irregularly observed cases for comparison of different disparities are shown in our simulation studies.

Cathy W. S. Chen:
Bayesian Causality Test for Integer-valued Time Series Models with Applications to Climate and Crime Data

Abstract: In this study we investigate the causal relationship between climate and criminal behavior. Considering the characteristics of integer-valued time series of criminal incidents, we propose a modified Granger causality test based on GARCH-type integer-valued time series models to analyze the relationship between the number of crimes and the temperature as an environmental factor. More precisely, we employ the Poisson, negative binomial, and log-linear Poisson integer-valued GARCH (INGARCH) models and particularly adopt a Bayesian method for our analysis. The Bayes factors and posterior probability of the null

hypothesis help determine the causality between the considered variables. Moreover, employing an adaptive Markov chain Monte Carlo (MCMC) sampling scheme, we estimate model parameters and initial value. As an illustration, we evaluate our test through a simulation study, and to examine whether or not temperature affects crime activities, we apply our method to datasets categorized as sexual offences, drug offences, theft of motor vehicle, and domestic violence-related assault in Ballina, New South Wales, Australia. The result reveals that more sexual offences, drug offences, and domestic violence-related assaults occur during the summer than in other seasons of the year. This evidence strongly advocates the causal relationship between crime and temperature.

*: Joint work with Sangyeol Lee

Tommaso Proietti:
Optimal linear prediction of stochastic trends

Abstract: A recent strand of the time series literature has considered the problem of estimating high-dimensional autocovariance matrices, for the purpose of out of sample prediction. For an integrated time series, the Beveridge-Nelson trend is defined as the current value of the series plus the sum of all forecastable future changes. For the optimal linear projection of all future changes into the space spanned by the past of the series, we need to solve a high-dimensional Toeplitz system involving n autocovariances, where n is the sample size. The paper proposes a non parametric estimator of the trend that relies on banding, or tapering, the sample partial autocorrelations, by a regularized Durbin-Levinson algorithm. We derive the properties of the estimator and compare it with alternative parametric estimators based on the direct and indirect finite order autoregressive predictors. (Joint work with Alessandro Giovannelli).

Keywords: High-dimensional autocovariance matrices; Toeplitz

systems; Beveridge-Nelson decomposition.

S. N. Lahiri:
Higher order asymptotic properties of the Bootstrap in post model selection inference in high dimensional regression models

Abstract: Chatterjee and Lahiri (2013) showed that under suitable conditions, the residual Bootstrap is second order correct for studentized pivots based on the ALASSO. One of the major limitations of their result is the existence of a preliminary estimator satisfying certain probabilistic bounds that are hard to verify in the p>n case. In this talk, we show that the second order correctness property holds quite generally for a number of penalized regression methods satisfying a version of the Oracle property of Fan and Li (2001). In particular, we show that under some suitable conditions, the LASSO and some popular nonconvex penalization functions including the SCAD and the MCP also enjoy second order correctness.

*Joint work with Arindam Chatterjee and Debraj Das

Yurie Tamura:
Asymptotic theory of sphericity test statistic for high-dimensional time Series

Abstract: Recently, several studies on sphericity test statistic U for multivariate data have been proposed under the condition that the dimension p of observation is comparable with the sample size n. Ledoit and Wolf (2002) proved its asymptotic normality and consistency in i.i.d. case, as n and p tend to infinity at the same rate, i.e., $p/n \to c \in (0, \infty)$. In this talk, under the same condition, the asymptotics of U are elucidated when the observations are generated from multivariate Gaussian stationary processes. Then, it is shown that an appropriately standardized

version of U is asymptotically normal even in the case of high-dimensional time series. Some interesting numerical examples are provided. (Joint work with Masanobu Taniguchi) .

Kou Fujimori:
The Dantzig selector for diffusion processes with covariates

Abstract: The Dantzig selector for a special parametric model of diffusion processes is studied in this paper. In our model, the diffusion coefficient is given as the exponential of the linear combination of other processes which are regarded as covariates. We propose an estimation procedure which is an adaptation of the Dantzig selector for linear regression models and prove the l_q consistency of the estimator for all $q \in [1, \infty]$. (Joint work with Yoichi Nishiyama).

Hiroshi Shiraishi:
Semi-parametric estimation for optimal dividend barrier with insurance portfolio

Abstract: How an insurance portfolio is used to provide dividend income for insurance company's shareholders is an important problem in application of risk theory, where the premium income as dividends is paid to the shareholders, whenever the surplus attains a level barrier, until the next claim occurs. Under the aggregate claims process taken as a compound Poisson model, we define optimal dividend barrier as the level of the barrier that maximizes the expectation of the discounted dividends until ruin. In the literature, the optimal dividend barrier was derived explicitly under some fixed models concerning claim distribution where parametric estimation is possible. In practice, it may often be hard to provide with the claim distribution parametric model either only from theoretical point of view or from a finite sample, and thus non-parametric estimation is preferred. We consider the semi-parametric estimation of the optimal

dividend barrier. This study would contribute in practice to decision-making on dividend barrier in the case where a new product is launched or optimality of an existing dividend barrier is tested. (Joint work with Zudi Lu).

Yasutaka Shimizu:
Applications of central limit theorems for equity-linked insurance

Abstract: In both the past literature and industrial practice, it was often implicitly used without any justification that the classical strong law of large numbers applies to the modeling of equity-linked insurance. However, as all policyholders' benefits are linked to common equity indices or funds, the classical assumption of independent claims is clearly inappropriate for equity-linked insurance. In other words, the strong law of large numbers fails to apply in the classical sense. In this paper, we investigate this fundamental question regarding the validity of strong laws of large numbers for equity-linked insurance. As a result, extensions of classical laws of large numbers and central limit theorem are presented, which are shown to apply to a great variety of equity-linked insurance products.

Makoto Aoshima:
High-dimensional statistical analysis based on the inference of eigenstructures

Abstract: Aoshima and Yata (2016, Sinica) provided two disjoint models: the strongly spiked eigenvalue (SSE) model and the non-SSE (NSSE) model, for high-dimensional data. In this talk, we focus on the SSE model and develop new statistical inference for high-dimensional data. We first consider the two-sample problem for the SSE model. We provide a general test statistic as a function of positive-semidefinite matrices. We investigate the test statistic by considering strongly spiked eigenstructures and create a new test

procedure under the SSE model. Next, we consider the classification problem for the SSE model. We create a new classifier by estimating strongly spiked eigenstructures. For both problems, we give sufficient conditions for the statistics to satisfy consistency properties and to be asymptotically normal. Finally, we discuss the performance of the procedures in actual data analyses. (Joint work with Kazuyoshi Yata).

Aoshima, M. and Yata, K. (2016). Two-sample tests for high-dimension, strongly spiked eigenvalue models. Statistica Sinica, in press. (arXiv:1602.02491)

Yosihiko Ogata:
Forecasting of a large earthquake - An outlook of the research

Abstract: Although unconditional probability of a major earthquake in an area is extremely small, the probability can be increased in the presence of anomalies as potential precursors. Precursor like anomalies of only a single type may not sufficiently enhance the probability of a large earthquake, but the probability can be substantially increased using the “multi-elements prediction formula” when independent precursor like anomalies of plural types are observed at the same time. Despite such pioneering suggestions and applications for successful predictions in later 1970’s, this idea has yet hardly developed for more than 40 years in spite of the importance. I discuss the prescriptions by means of exemplar case study of 2016 Kumamoto, Japan, Earthquake of M7.3 and provide an outlook in this study.

Alessandra Luati:
Quantum information and statistical inference

Abstract: We discuss the role of quantum information in statistical inference. Parameter estimation based on likelihood is considered

and the classical Fisher information quantity is derived in parametric quantum models. Cramér-Rao type bounds are reviewed based on quantum information quantities such as Helstrom information and Wigner-Yanase information. The issue of designing experiments based on optimal measurements is considered, where optimality is related to the maximisation of an information quantity. Recent advances on the statistical analysis of open quantum systems, interacting with external forces, will be illustrated.

Ngai Hang Chan:
Directed Markov Multi-graphs as Models of Networks

Abstract: Directed multi-graphs of network data are studied in this paper. By assuming that an observed network is a discrete realization of a dynamical Markov directed graph process, a parametric model is proposed to describe graph change. The proposed model is a scale-free network, where the network dynamics are based both on a preferential linking mechanism and an initial attractiveness of nodes in the network. Since real networks are usually self-organized into a common topological structure, the proposed model is found to capture many of the common properties of real networks. An asymptotic maximum likelihood estimation procedure on model parameters is also developed. Applications of the model on routers network of financial and economic data and the Web Google network are examined. Validation on fitted models is conducted based on predictions of average in-degree and average out-degree of a randomly selected node. The result confirms that the proposed model is sufficiently adequate to model financial and economic networks. (Joint work with K.C. Cheung and Sam P.S. Wong. Research supported in part by grants from HKSAR-RGC-GRF.)

Mohsen Pourahmadi:
Nonparametric Change Point Detection in Multivariate Nonstationary

Time Series

Abstract: Detecting change points in multivariate time series is an important problem with numerous applications. Much of change point literature is on tackling this problem in the univariate case or is parametric in nature. We develop a nonparametric method to detect multiple change points in multivariate piecewise stationary processes when the locations and number of change points are unknown. Based on a test statistic that measures differences in the spectral density matrices through the L^2 norm, we propose a two stage procedure. The first stage identifies potential change points using a sequential test and the second stage tests for the significance of each of the potential change points. The asymptotic properties of the test for significant change points under the null and alternative hypothesis are derived. Monte Carlo simulation of values of a stationary process given its spectral density matrix is used to obtain critical values of the test statistic under the null. We illustrate the better performance of our method in comparison to some of the recent methods through a few simulation examples and discuss an application of our method in seismology. (Joint work with Raanju Ragavendar Sundararajan, Texas A&M University).

Yoshihiro Suto:
Shrinkage estimation for the autocovariance matrix of vector-valued Gaussian processes

Abstract: In this talk, we develop improved estimation for the autocovariance structure to the case of a vector-valued Gaussian stationary process. Under a criterion of the generalized mean squares error (GMSE), we propose the two empirical Bayes estimators in the cases of the mean is zero and non-zero. It is shown that the two shrinkage estimators improve the usual autocovariance matrices, and the asymptotic risk differences are evaluated similarly to the case of scalar-valued Gaussian stationary processes. We also propose the practical estimators, and evaluate the asymptotic

differences of the GMSEs. We illustrate goodness of the proposed estimators.

Fumiya Akashi:
A self-normalized block sampling method to quantile regression on time series

Abstract: This talk develops the quantile regression-based self-normalized block sampling method for time series models. We focus on the estimation problem of the coefficient of the linear regression model, whose error terms possibly have long-range dependence or heavy tails. A robust estimation procedure for the regression parameter of the model is proposed. Remarkably, the proposed confidence region does not involve any unknown quantity of the model such as the intensity of the dependence structure or the heaviness of the tail distribution of the time series. Some simulation experiments illustrate the finite sample performance of the proposed method, and it is also observed that the proposed test has robustness in the practical situations. (Joint work with S. Bai and M.S. Taqqu)

Ritei Shibata:
Non-Identifiability of Simultaneous Autoregressive Model and Its Inference

Abstract: Simultaneous autoregressive (SAR) model is widely used for spatial data analysis, observed at a set of grid points in a space. However a problem, which is not so well known, is that there exists no unique SAR model for given autocovariances or spectral density unlike time series AR model. Such a non-identificability of the SAR model implies existence of multiple maximum likelihood estimates under Gaussianity and may cause inconsistency of the estimate, which implies the singularity of Fisher Information Matrix. Several types of necessary and sufficient conditions for the singularity are

also given. (Joint work with Yuuki Rikimaru).

Arthur Pewsey:
An Introduction to Directional Statistics

Abstract: The aim of the talk is to provide an introduction to the field of Directional Statistics. We start with an overview of some of the basic ideas underpinning Circular Statistics, before moving on to methods used to investigate the fundamental issues of isotropy, symmetry and the number of modes of a circular distribution. We then consider classical distributions used to model circular data, as well as some of the more flexible models proposed recently in the literature. Developments in the use of mixture models, regression and time series analysis for circular data will also be discussed. The final part of the talk focuses on other forms of analysis and applications involving data distributed on the cylinder, torus, sphere and extensions thereof.