Waseda International Symposium

Introduction of General Causality to Various Data & its Innovation of the Optimal Inference

Date: October 22 – October 24, 2018.

Venue: Waseda University, Nishi-Waseda Campus

Building 55N Room 02

Organizer: Masanobu TANIGUCHI

(Research Institute for Science & Engineering, Waseda University)

Supported by: JSPS KAKENHI Kiban (S) Grand-in-Aid No. 18H05290 (M. Taniguchi)

Waseda Research Institute for Science & Engineering,

Institute for Mathematical Science

Waseda International Symposium

"Introduction of General Causality to Various Data & its Innovation of the Optimal Inference"

Date: October 22 – October 24, 2018.

Venue: Waseda University, Nishi-Waseda Campus Building 55N Room 02

(Access map: http://www.sci.waseda.ac.jp/eng/access)

Organizer: Masanobu TANIGUCHI

(Research Institute for Science & Engineering, Waseda University)

This workshop is supported by:

- JSPS KAKENHI Kiban (S) Grand-in-Aid No. 18H05290 (M. Taniguchi)
- Waseda Research Institute for Science & Engineering, Institute for Mathematical Science

Program (*speaker)

October 22

Session I (13:30 – 14:50): chaired by Konstantinos Fokianos

13:30 – 13:40: Opening
Masanobu Taniguchi

13:40 – 14:15:
Robust local polynomial regression for circular regression models with heavy-tailed error
Fumiya Akashi* (Waseda Univ.)

14:15 – 14:50:
Asymptotic properties of mildly explosive processes with locally stationary disturbance
Junichi Hirukawa* (Niigata Univ.) and Sangyeol Lee (SNU)

14:50 – 15:05: Coffee break

Session II (15:05 – 17:05): chaired by Jean-Michel Zakoian

15:05 – 15:45:
Spatio-temporal extension of GARCH models for analysis of high dimensional financial time series
Yasumasa Matsuda* (Tohoku Univ.)

15:45 – 16:25:
Identification of possibly nonfundamental Structural VARMA models using higher order moments
Carlos Velasco* (Department of Economics Universidad Carlos III de Madrid)

16:25 – 17:05:
Multivariate Count Autoregression
Konstantinos Fokianos* (Lancaster Univ.)

October 23

Session III (9:30 – 10:35): chaired by Ta-Hsin Li

9:30-10:00:
Robustness of zero crossing estimator
Yuichi Goto* (Waseda Univ.) and Masanobu Taniguchi

10:00 – 10:35:
Higher order approximation of the distribution of test statistics for high-dimensional time series ANOVA models
Hideaki Nagahata* (ISM)

10:35 – 10:50: Coffee break

Session IV (10:50-12:00): chaired by Makoto Aoshima

10:50 – 11:25:
A test for missingness in time series
William Dunsmuir and Yan Liu* (Kyoto Univ.)

11:25 – 12:00:
Loss given default estimation for corporation loans: combining a two-stage model with classification tree-based boosting and support vector regression with logistic transformation
Yuta Tanoue* (Waseda Univ.)

12:00 – 13:15: Lunch

Session V (13:15 – 15:15): chaired by Carlos Velasco
13:15 – 13:55:
Sequential estimation and unit root tests for autoregressive processes
Yoshihiko Nishiyama* (Kyoto Univ.), Kohtaro Hitomi, Keiji Nagai & Junfan Tao

13:55 – 14:35:
New techniques in high-dimensional statistical analysis
Makoto Aoshima* (Tsukuba Univ.)

14:35 – 15:15:
Noncausal heavy-tailed autoregressive process and the modeling of bubbles
Jean-Michel Zakoian* (CREST and Lille Univ.)

15:15 – 15:30: Coffee break

VI Keynote Session (15:30-17:10): chaired by Masanobu Taniguchi
15:30 – 16:20:
Extending the causality-related interdependence measures
Yuzo Hosoya* (Emeritus Professor, Tohoku Univ.)

16:20 – 17:10:
Mixed fractional Brownian motion and statistical inference
B.L.S. Prakasa Rao* (CR RAO AIMSCS, Hyderabad, India))

Special Session: chaired by Yoshihiko Nishiyama (Kyoto Univ.)
17:20 – 18:00:
Music expression and statistical consideration in opera
Yuya Harada* (Tokyo Univ. of the Arts, Baritone singer)

18:30 – Buffet Party

October 24

Session VII (9:30 – 10:40): chaired by Yoshihide Kakizawa

9:30 – 10:05:
The Dantzig selector for a linear model of diffusion processes
Kou Fujimori* (Waseda Univ.)

10:05 – 10:40:
Equality tests of high-dimensional covariance matrices on the basis of strongly spiked eigenvalues
Aki Ishii* (Tokyo Univ. of Science), Kazuyoshi Yata and Makoto Aoshima

10:40 – 10:50: Coffee break

Session VIII (10:50 – 12:50): chaired by Prakasa Rao

10:50 – 11:30:
Some boundary-bias-free density estimators and their bias-reductions
Yoshihide Kakizawa* (Hokkaido Univ.)

11:30 – 12:10:
The one-way effect causal relationships between Stock Exchange Markets of Hong Kong and others countries
Feng Yao* (Kagawa Univ.)

12:10 – 12:50:
Quantile regression for spectral analysis of time series
Ta-Hsin Li* (IBM T. J. Watson Research Center)

Abstracts

Title: Robust local polynomial regression for circular regression models with heavy-tailed error
Author: Fumiya Akashi* (Waseda Univ.)
Abstract: Statistical treatment of a "circular observation" has attracted much attention in these decades. Such datasets are observed in variety of fields involving political science, analysis of wildfire orientation, and seismic wave analysis. This talk considers nonparametric estimation problem of a nonlinear regression function of circular random variables. In particular, we use a circular kernel to approximate the regression function by a polynomial function locally. Furthermore, motivated by the concept of least absolute deviations regression, a new robust estimator is proposed. The result in Di Marzio, Panzera & Taylor (2009, Stat. Probab. Lett. 79, 2066-2075) is then nicely extended to infinite variance innovation case. Some simulation experiments illustrate the finite sample performance of the proposed method and elucidate the advantage of the robust local polynomial regression estimator.
[Back to talk list]

Title: Asymptotic properties of mildly explosive processes with locally stationary disturbance
Author: Junichi Hirukawa (Niigata Univ.)
Abstract: In this talk the limit distribution of the least squares estimator for mildly explosive autoregressive models with locally stationary disturbance is established, which is shown to be Cauchy as in the iid case. The result is then applied to identify the onset and the end of an explosive period of a financial time series. Simulations and data analysis are conducted to demonstrate the validity of the result.
[Back to talk list]

Title: Spatio-temporal extension of GARCH models for analysis of high
Author: Yasumasa Matsuda* (Tohoku Univ.)
dimensional financial time series
Abstract: Autoregressive Conditional Heteroscedasticity (ARCH) models, which were originally proposed by Engle (1982), have been playing critical roles in modelling volatilities in financial time series. This talk aims at a multivariate extension of ARCH models to evaluate volatility matrices for high dimensional multivariate financial time series. The critical difficulty in the multivariate extension is in so called "curse of dimension" caused by larger number of parameters for higher dimension of multivariate series. We introduce financial distances among components of multivariate series, which are different from the usual geographical one but are based on closeness of financial conditions, and apply dynamic panel data models by spatial weight matrices constructed by the financial distance. As a result, we propose spatio-temporal GARCH models that can identify volatility matrices for high

dimensional financial time series. We conduct comparative studies by real financial time series and show empirical features of the spatio-temporal GARCH models in terms of forecast of volatilities.
[Back to talk list]

Title: Identification of possibly nonfundamental Structural VARMA models using higher order moments
Author: Carlos Velasco* (Department of Economics Universidad Carlos III de Madrid)
Abstract: We introduce a frequency domain criterion to identify the parameters of, possibly noncausal and/or noninvertible, structural vector autoregressive moving average (VARMA) models. We use information from higher order moments to achieve identification on the location of the roots of the VAR and VMA matrix polynomials for non-Gaussian vector time series possibly non-fundamental. This information also provides identification on the rotation of the model errors leading to the structural innovations. We develop general representations of the higher order spectral density arrays of vector linear processes and describe sufficient conditions for the parameter identification that rely on both sufficiently rich (linear) dynamics and higher order dependence structure of the vector of linear innovations. These results generalize previous univariate analysis to develop more efficient estimates and relate to the predictability of Wold innovations of nonfundamental processes.
[Back to talk list]

Title: Multivariate Count Autoregression
Author: Konstantinos Fokianos* (Lancaster University)
Abstract: We are studying the problems of modeling and inference for multivariate count time series data with Poisson marginals. The focus is on linear and log-linear models. For studying the properties of such processes we develop a novel conceptual framework which is based on copulas. However, our approach does not impose the copula on a vector of counts; instead the joint distribution is determined by imposing a copula function on a vector of associated continuous random variables. This specific construction avoids conceptual difficulties resulting from the joint distribution of discrete random variables yet it keeps the properties of the Poisson process marginally. We employ Markov chain theory and the notion of weak dependence to study ergodicity and stationarity of the models we consider. We obtain easily verifiable conditions for both linear and log-linear models under both theoretical frameworks. Suitable estimating equations are suggested for estimating unknown model parameters. The large sample properties of the resulting estimators are studied in detail. The work concludes with some simulations and a real data example. (joint work with D. Tjostheim, B. Stove and P. Doukhan)
[Back to talk list]

Title: Robustness of zero crossing estimator
Author: Yuichi Goto* (Waseda University)
Abstract: Zero Crossing (ZC) statistic is the number of zero crossings observed in a time series. In this talk, a strictly stationary ellipsoidal ϕ-mixing processes with mean zero, finite variance are discussed. We consider the estimation problems of the autocorrelation of a time series by using zero crossing. First, we will elucidate the joint asymptotic distribution of the ZC estimator. Next, we show that the ZC estimator has a good robustness when the spectral density of the process is contaminated by a sharp peak.
[Back to talk list]

Title: Higher order approximation of the distribution of test statistics for high-dimensional time series ANOVA models
Author: Hideaki Nagahata* (Risk Analysis Research Center, The Institute of Statistical Mathematics)
Abstract: Analysis of variance (ANOVA) is tailored for independent observations. Recently, there has been considerable demand for ANOVA of high-dimensional and dependent observations in many fields. Now it is important to analyze differences among big data's averages of any areas of all over the world, for example, the financial industry, the manufacturing one, and so on. However, the numerical accuracy of ANOVA for these types of observations has been inadequately developed. In this paper, we thus present a study on Edgeworth expansion of distribution of ANOVA tests for high-dimensional and dependent observations. Specifically, we present the second-order approximation of classical test statistics proposed for independent observations. We also give Numerical examples for simulated high-dimensional time series data.
[Back to talk list]

Title: A test for missingness in time series
Author: William Dunsmuir and Yan Liu* (Kyoto Univ.)
Abstract: We consider a test for missingness in time series. Suppose we observe a time series with missing values, which is generated by a regression model with dependent disturbances. The mechanism for the missing values is supposed to be generated as Bernoulli responses from a generalized linear ARMA model. A score type test is proposed to test the null hypothesis that the data are missing completely at random. For this testing problem, we use the Laplace approximation to obtain the likelihood of the process. We investigate the performance of our proposed test statistic in several numerical simulations. The method is also applied to real data of pollution levels containing some missing observations.
[Back to talk list]

Title: Loss given default estimation for corporation loans: combining a two-stage model with classification tree-based boosting and support vector regression with logistic transformation
Author: Yuta Tanoue (Waseda Univ.)
Abstract: The Basel Accords set out a framework that regulates the minimum amount of capital banks are required to hold. To calculate the minimum amount of capital, banking institutions have to calculate risk weighted-assets(RWA). The Basel Accords allow banking institutions to calculate RWA for credit risk according to either of two approaches: a standardized approach that uses agency ratings for calculating RWA for credit risk and an internal ratings-based approach that allows a bank to use internal estimates of components of credit risk—that is, the probability of default (PD), the loss given default (LGD), and exposure at default (EAD). In past decades, significant attention has been devoted to the analysis of PD and the development of a PD estimation model. However, recently, more attention has been dedicated to the analysis of LGD and the development of an LGD estimation model. In this study we will develop LGD estimation model. Recently, various studies have pointed out that the predictive performance of non-linear or non-parametric models ([2] and so on), ensemble medhod ([3] and so on) and two-step model of [1] is superior to traditional models. Then, in this study, using a dataset composed of five Japanese regional banks, we propose a LGD estimation model using a two-stage model, classification tree-based boosting, and support vector regression. We compare the proposed model's predictive performance with existing models by performing cross validation and out-of-time validation. As a result, we find that incorporation of non-linearity into the LGD estimation model by classification and support vector regression improves the predictive performance. Furthermore, we confirm that the boosting method also improves the model's predictive performance.

References:
[1] Marc G¨urtler and Martin Hibbeln. Improvements in loss given default forecasts for bank loans. Journal of Banking & Finance, 37(7):2354–2366, 2013.
[2] Thomas Hartmann-Wendels, Patrick Miller, and Eugen T¨ows. Loss given default for leasing: Parametric and nonparametric estimations. Journal of Banking & Finance, 40:364–375, 2014.
[3] Eugen T¨ows. Advanced Methods for Loss Given Default Estimation. PhD thesis, Universit¨at zu K¨oln, 2016.
[Back to talk list]

Title: Sequential estimation and unit root tests for autoregressive processes
Author: Yoshihiko Nishiyama* (Institute of Economic Research, Kyoto University), Kohtaro Hitomi (Kyoto Institute of Technology), Keiji Nagai (Yokohama National University) & Junfan Tao (Yokohama National University)
Abstract: Estimation and testing for autoregressive processes have a long history, where the literature mostly assume that off-line sampling data is available. Lai and Siegmund (1983) consider the case of online sampling, where they propose to stop sampling using the observed Fisher information and show

that the OLS estimator of the AR coefficient is asymptotically normally distributed for both stationary and nonstationary AR(1) processes. Motivated by the unit root testing problem, we consider the case of nearly nonstationary processes where we employ a unifying approach by diffusion approximation for both nonstationary and nearly nonstationary cases to examine the statistical properties. It allows us to scrutinize the properties of the stopping times we consider, which are shown to be characterized by Bessel processes of dimension 3/2. We also consider the stationary autoregressive processes and prove the asymptotic normality of the coefficient estimator and the stopping times.
[Back to talk list]

Title: New techniques in high-dimensional statistical analysis
Author: Makoto Aoshima* (Tsukuba Univ.)
Abstract: Any high-dimensional data is classified into two disjoint models: the strongly spiked eigenvalue (SSE) model and the non-SSE (NSSE) model. In actual high-dimensional data, a non-sparse and low-rank structure which contains strongly spiked eigenvalues is often found; a structure which fits the SSE model. Under the SSE model, it may be noted that the asymptotic normality of high-dimensional statistics is not valid because it is heavily influenced by strongly spiked eigenvalues. To enable a unified treatment of both the SSE models and non-SSE models, data transformation techniques that transform the SSE models to the non-SSE models were developed by Professor Aoshima and Dr. Yata. Following this novel methodology, strongly spiked eigenvalues are accurately detected by using new PCA-type techniques. With the transformed data, one can create a new statistic which can ensure high accuracy for inferences by using asymptotic normality even under the SSE models. Professor Aoshima will demonstrate the new techniques to handle high-dimensional data to solve two-sample problems and classification problems.
[Back to talk list]

Title: Noncausal heavy-tailed autoregressive process and the modeling of bubbles
Author: Jean-Michel Zakoian* (based on joint works with Christian Gouriéroux, CREST and Toronto University & Sébastien Fries, CREST)
Abstract: Noncausal autoregressive processes with heavy-tailed errors possesses a nonlinear causal dynamics, which allows for local explosion or asymmetric cycles often observed in economic and financial time series. They provide a new modeling for multiple local explosions in a strictly stationary framework. The causal predictive distribution displays surprising features, such as the existence of higher moments than for the marginal distribution, or the presence of a unit root in the Cauchy case. Aggregating such models can yield complex dynamics with local and global explosion as well as variation in the rate of explosion. The asymptotic behavior of a vector of sample autocorrelations is studied in a semi-parametric noncausal autoregressive framework with Pareto-like tails, and diagnostic tests are proposed. Empirical results based on the Nasdaq composite price index are provided.
[Back to talk list]

Title: Extending the causality-related interdependence measures

Author: Yuzo Hosoya* (Tohoku University (Emeritus Professor))

Abstract: The measures of one-way effect, reciprocity, and association are defined respectively as overall as well as frequency-wise quantities in the frequency domain typically for vector stationary time series ([3]). The basic idea is the elicitation of a one-way effect component of a supposedly causing series and the representation of its contribution for reduction of the prediction error by means of the Szegö representation formula of the one-step ahead prediction error. To define a partial version of the measures of interdependence, the paper [4] suggested the third-effect elimination of the one-way effect component of the third series from a pair of subject-matter series to preserve the inherent feedback structure of the pair. In this talk, I suggest extension of those measures to the processes as follows:

1. Nonstationary cointegrated processes ([5]),
2. Locally stationary processes introduced by [1],
3. Doubly-parametrized stationary processes and more generally to stationary fields ([2], [6]).

The extension is enabled by introduction of the concept of reproducible processes for the case (i), by the asymptotic approximation of the Szegö formula for the case (ii) and by extension of the Szegö formula to the prediction error by the half plane for the case (iii).

References:

[1] Dahlhaus, R. (1996). *Stochastic processes and their Applications*, **62**, 139-164.[2] Helson, H. and Lowdenslager, D. (1958). *Acta Mathematics*, **99**, 165-202.

[3] Hosoya, Y. (1991). *Probability Theory and Related Fields*, **88**, 429-444.

[4] Hosoya, Y. (2001). *Journal of Time Series Analysis*, **22**, 537-554.

[5] Hosoya, Y. *et al.* (2017). *Characterizing Interdependencies of Multiple Time Series: Theory and Application*, Springer.

[6] Whittle, P. (1954). *Biometrika*, **41**, 434-449.

[Back to talk list]

Title: Mixed fractional Brownian motion and statistical inference

Author: B.L.S. Prakasa Rao* (CR RAO Advanced Institute of Mathematics, Statistics and Computer Science, Hyderabad, India (CR RAO AIMSCS, Hyderabad, India))

Abstract: We give an overview of some recent results concerning properties of mixed fractional Brownian motion and discuss methods of statistical inference for parametric and non-parametric estimation of the trend for processes driven by a mixed fractional Brownian motion. In particular, we present methods of maximum likelihood estimation for the parameters involved in the trend and study their asymptotic properties and rates of convergence. Analogues of the Bernstein-von Mises theorem and application to Bayes estimation of the parameter will be discussed. Non-parametric estimation of

the trend and non-parametric estimation of the linear multiplier through kernel methods will be investigated.

[Back to talk list]

Title: Music expression and statistical consideration in opera

Author: Yuya Harada* (Tokyo Univ. of the Arts, Baritone singer)

Abstract:

(i) Play multiple operas with different age of different performance expressions.

(ii)Listen to the performance and consider whether statistical consideration can be done.

[Back to talk list]

Title: The Dantzig selector for a linear model of diffusion processes

Author: Kou Fujimori* (Waseda Univ.)

Abstract: In this talk, a linear model of diffusion processes with unknown drift and diagonal diffusion matrices is discussed. We will consider the estimation problems for unknown parameters based on the discrete time observation in high-dimensional and sparse settings. To estimate drift matrices, the Dantzig selector which was proposed by Candès and Tao in 2007 will be applied. We will prove two types of consistency of the Dantzig selector for the drift matrix; one is the ordinal consistency in the sense of norms and another is the variable selection consistency. Moreover, we will construct an asymptotically normal estimator for the drift matrix by using the variable selection consistency of the Dantzig selector.

[Back to talk list]

Title: Equality tests of high-dimensional covariance matrices on the basis of strongly spiked eigenvalues

Author: Aki Ishii* (Tokyo Univ. of Science), Kazuyoshi Yata and Makoto Aoshima

Abstract: Recently, Aoshima and Yata (2018, Stat. Sinica) proposed two types of eigenvalue models for high-dimensional data. One is called the strongly spiked eigenvalue (SSE) model and the other one is called the non-SSE model. In this talk, we focus on the SSE model that is often seen when we analyze microarray data sets. We consider the equality test of high-dimensional covariance matrices under the SSE model. We find the difference of covariance matrices by dividing high-dimensional eigenspaces into the first eigenspace and the others. We create a new test procedure on the basis of those high-dimensional eigenstructures. We precisely study the influence of spiked eigenvalues on a test statistic and consider a bias correction so that the proposed test procedure has a consistency property for the size. We also show that the test procedure has preferable properties for the power. We discuss its performance in simulations and real data analyses using microarray data sets.

[Back to talk list]

Title: Some boundary-bias-free density estimators and their bias-reductions
Author: Yoshihide Kakizawa* (Hokkaido Univ.)
Abstract: We consider estimation of the probability density for nonnegative data (the case [0,1] can be similarly discussed). In that case, the standard kernel density estimator (KDE) is, in general, inconsistent near the boundary, due to the so-called boundary bias. Many authors have suggested some remedies for removing the boundary bias. Among them, we focus on asymmetric kernel methods. This talk is divided into two parts. In the first part, we discuss the recent developments of asymmetric KDEs. For the nonnegative data, Chen (2000) first suggested the boundary-bias-free nonparametric density estimator using gamma kernel. This type of research since Chen has received considerable attention in the literature, using inverse Gaussian (IG)/reciprocal inverse Gaussian (RIG)/Birnbaum-Saunders (BS)/LN, and so on. The resulting estimators have the MISEs of order O($n^{-4/5}$) under suitable conditions. We stress that some existing asymmetric KDEs should be re-formulated, because of the ``bad" formulation yielding an unrealistic constraint $\hat{f}(0)=0$. In the second part, we show that some ``good" asymmetric KDEs can be bias-corrected in various ways (i.e., additive/multiplicative bias-corrections), to achieve the faster MISE-convergence rates. This is a series of joint works with Dr. Igarashi, G. (University of Tsukuba).
[Back to talk list]

Title: The one-way effect causal relationships between Stock Exchange Markets of Hong Kong and others countries
Author: Feng Yao* (Kagawa Univ.)
Abstract: The transfer of sovereignty over Hong Kong from the United Kingdom to China, referred to as "the Handover" internationally or "the Return" in China, took place on 1 July 1997. Since then, discussions of the causal characteristics between Stock Exchange Markets of Hong Kong and others countries becomes one of the most important topics about Hong Kong financial system. We apply the one-way effect causal measure and also its Wald test to investigate the causal relationships between HANG SENG Index and Nikkei 225, Dow Jones Industrial Average, FTSE 100, and Shanghai Composite. The detail discussions are focused on six periods we concerned before and after the time point of "the Return" in 1990/1/4 – 2010/4/30. The empirical results show us that Hong Kong Stock Exchange Market is not internationally integrated as we generally think.
[Back to talk list]

Title: Quantile regression for spectral analysis of time series
Author: Ta-Hsin Li* (IBM T. J. Watson Research Center)
Abstract: Quantile regression is a powerful tool that extends the capability of the traditional least-squares apparatus by focusing on the behavior of the data at different quantiles instead of the mean. This talk gives an overview of some recent advances in quantile regression for spectral analysis of time-series data. In particular, it discusses a new type of periodogram, called the quantile periodogram,

which is constructed from quantile regression with harmonic (trigonometric) regressors; it explains the theoretical underpinning of the quantile periodogram in relation with the spectrum of level-crossing processes, it demonstrates, with both simulated and real data, the capability of the quantile periodogram in offering a richer view than the one provided by the ordinary periodogram; and finally, it discusses possible extentions of the methodology.

[Back to talk list]