HOKKAIDO INTERNATIONAL SYMPOSIUM

Recent Developments of Statistical Theory in Statistical Science

Leader: Masanobu Taniguchi (Waseda University)

Date: October 27 - 29, 2016
Location: Hokkaido University
Graduate School of Economics and Business Administration
3rd floor
Access Map: http://www.econ.hokudai.ac.jp/en08/access

Supported by
- **Kiban (A-15H02061)**
  M. Taniguchi
  Research Institute for Science & Engineering, Waseda University
- **Houga (26540015)**
  M. Taniguchi
  Research Institute for Science & Engineering, Waseda University

Organizers: Yoshihide Kakizawa (Hokkaido University)
Akio Suzukawa (Hokkaido University)
Fumiya Akashi (Waseda University)

Hokkaido International Symposium

Recent Developments of Statistical Theory in Statistical Science

Date: October 27 - 29, 2016
Location: Hokkaido University
Graduate School of Economics and Business Administration
3rd floor
Access Map: http://www.econ.hokudai.ac.jp/en08/access

Leader: Masanobu Taniguchi (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

Organizers: **Yoshihide Kakizawa (Hokkaido University)**
Akio Suzukawa (Hokkaido University)
Fumiya Akashi (Waseda University)

Program

(* Speaker)

October 27

Session Ⅰ　13:45 - 16:20　Chaired by Y. Kakizawa

13:45 - 13:55 Masanobu Taniguchi (Waseda Univ.)
Opening

13:55 - 14:30 Yan Liu (Waseda Univ.)
Statistical inference for quantiles in frequency domain

14:30 - 15:05 Hiroshi Yamada (Hiroshima Univ.)
Quantile Hodrick-Prescott filtering

15:20 - 15:50 Rizky Reza Fauzi* (Kyushu Univ.), Yoshihiko Maesono
Modifying gamma kernel density estimator by reducing variance

15:50 - 16:20 Taku Moriyama (Kyushu Univ.)
Improvement of kernel estimators using boundary bias reduction methods

Session Ⅱ　16:35 - 17:45　Chaired by H. Ogata

16:35 - 17:10 Minehiro Takahashi, Ryota Tanaka and Junichi Hirukawa* (Niigata Univ.)
On the extension of Lo's modified R/S statistics against to locally stationary short-range dependence

17:10 - 17:45 Fumiya Akashi* (Waseda Univ.) and Jianqing Fan
Self-normalized and random weighting approach to likelihood ratio test for the model diagnostics of stable processes

October 28

Session Ⅲ　9:30 - 12:00　Chaired by A. Suzukawa

9:30 - 10:05 Hiroaki Ogata (Tokyo Metropolitan Univ.)
Multi-step circular Markov processes with canonical vine representations

10:05 - 10:40 Nobuhiro Taneichi* (Kagoshima Univ.), Yuri Sekiya, Jun Toyama
On improvement of generalized ϕ-divergence goodness-of-fit statistics for GLIM of binary data

10:55 - 11:25 Kou Fujimori* (Waseda Univ.) and Yoichi Nishiyama
The l_q consistency of the Dantzig selector for Cox's proportional hazards model

11:25 - 12:00 Yuan-Tsung Chang* (Mejiro Univ.) and Nobuo Shinozaki
Shrinkage estimators of Poisson means based on prior information and its applications to multiplicative Poisson models

12:00 - 13:40 Lunch

Session Ⅳ　13:40 - 15:45　Chaired by J. Hirukawa

13:40 - 14:25 (Invited) Kun Chen (Southwestern Univ. of Finance and Economics)
Bartlett correction of empirical likelihood with unknown variance

14:25 - 15:00 Hiroki Masuda (Kyushu Univ.)
Locally stable regression without ergodicity and finite moments

15:00 - 15:45 (Invited) Sangyeol Lee (Seoul National Univ.)
Recent developments in inferences for conditionally heteroscedastic location-scale time series models

Session Ⅴ　16:00 - 17:30　Chaired by M. Taniguchi

16:00 - 16:45 (Invited) Arup Bose (Indian Statistical Institute)

Large sample behaviour of high dimensional autocovariance matrices when the dimension grows slower than the sample size

16:45 - 17:30 (Invited) Hans R. Künsch* (Seminar for Statistics, ETH Zurich), Marco Frei and Sylvain Robert

Monte Carlo filtering and data assimilation

18:30- Dinner party

October 29

Session Ⅵ　9:45 - 12:10　Chaired by F. Akashi

9:45 - 10:15 Yuma Uehara (Kyushu Univ.)
Statistical inference for misspecified ergodic Lévy driven stochastic differential equation models

10:15-10:50 Nobumichi Shutoh* (Kobe Univ.), Takahiro Nishiyama, Masashi Hyodo
Bartlett correction to the likelihood ratio test for MCAR with two-step monotone missing data

11:05 - 11:35 Yasunori Tamura (Osaka Prefecture Univ.)
Multi-group profile analysis for high-dimensional

11:35 - 12:10 Masashi Hyodo* (Osaka Prefecture Univ.), Takahiro Nishiyama
Simultaneous testing of the mean vector and the covariance matrix for high-dimensional data

Abstract

October 27

Yan Liu

Statistical inference for quantiles in frequency domain

Abstract: In this talk, we consider the estimation and testing problems of quantiles in frequency domain. For second order stationary process, the spectral distribution function is uniquely determined by the autocovariance function of the process. We first define the quantiles of the spectral distribution function. The asymptotic distribution of the naive quantile estimator is shown to be non-Gaussian. This result is different from that considered in time domain. We recover the asymptotic normality of quantile estimation by smoothing the periodogram. Besides, we consider the quantile tests in frequency domain from our estimation procedure. Strong statistical power is shown in our numerical studies. The power of our proposed statistic under local alternatives is also discussed.

Hiroshi Yamada

Quantile Hodrick-Prescott filtering

Abstract: Quantile regression was introduced in the seminal work by Koenker and Bassett (1978) and widely applied in econometrics. Hodrick-Prescott (HP) (1997) filtering is used frequently to estimate trend components of macroeconomic time series. In this paper, we contribute to the literature on macroeconometrics by introducing a filtering method that combines these two statistical tools. We refer to it as quantile HP (qHP) filtering. qHP filtering enables us to obtain not only the median trend, which is more robust to outliers than HP filtering, but also other quantile trends, which may provide a deeper understanding of time series properties. As in the case of HP filtering, it requires selection of tuning parameter. We propose a method for selecting it, which enables us to compare trends from (q)HP filtering. As an empirical example, we present some estimated quantile trends from Japan's industrial index of production (IIP).

Rizky Reza Fauzi

Modifying gamma kernel density estimator by reducing variance

Abstract: We discuss a new kernel type estimator for nonnegatively supported density function $f(x)$. Using pdf of a gamma distribution, Chen (2000, Ann. Inst. Stat. Math.) has introduced new kernel estimators of $f(x)$. However, the variances of Chen's estimators diverge when x closes to zero. In this talk, modifying the gamma kernel we propose a new estimator which improves convergence rates of variance. Applying bias reduction by transformation, we improve the mean squared error of the proposed estimator too. The estimator is free from the boundary problem of bias, and the variance of the new estimator does not diverge when x is close to zero. (Joint work with Yoshihiko Maesono)

Taku Moriyama

Improvement of kernel estimators using boundary bias reduction methods

Abstract: In this talk, I discuss improvement of kernel estimators using boundary bias reduction methods. As we know, kernel estimators are biased (as we call boundary bias) if the true density has finite or semi-finite support. When the support is known, we have various methods to improve the estimators. On the other hand, there are few ways to adjust them to the unknown support. Hall and Park (2002) introduced a way to estimate the density whose support is estimated by order statistics. I propose a new way to estimate the unknown support using a boundary bias reduction method which is used to correct the distribution estimator. The asymptotic properties and some simulation results of kernel estimators using the proposed boundary estimator are shown. Furthermore, I discuss the necessity of giving appropriate (bounded) support of the estimated density even if the information is not available.

Junichi Hirukawa

On the extension of Lo's modified R/S statistics against to locally stationary short-range dependence

Abstract: Lo (1991) developed test for long-run memory that is robust to short-range dependence. It was extension of the ``range over standard deviation'' or R/S statistic, for which the relevant asymptotic sampling theory was derived via functional central limit theory. In this talk, we consider extension of Lo's modified R/S statistics which is aimed to be robust against locally stationary short-range dependence. (Joint work with Minehiro Takahashi and Ryota Tanaka)

Fumiya Akashi

Self-normalized and random weighting approach to likelihood ratio test for the model diagnostics of stable processes

Abstract: In the second order stationary case, testing problems described by the restriction of the spectral density of the model appear in various situations involving variable selection or serial correlation of the process. More generally, many kinds of the model diagnostic problems can be grasped by the spectral restriction. Various approach for model diagnostics has been developed under the assumption of finite variance. In this talk, we extend the classical approach toward the infinite variance case. In particular, we consider the symmetric alpha-stable linear process and testing problem described by the power transfer function of the process. Since the rate of convergence and limit distributions of some fundamental statistics for stable process depend on unknown tail-index and scale parameter of innovation processes, we make use of the methods of self-normalization and bootstrap with random weighting approach. As a result, we propose a feasible testing procedure, which does not require any priori estimation of auxiliary parameters. By simulation experiments, we also check that the proposed method works well in practical situations. (Joint work with Jianqing Fan)

October 28

Hiroaki Ogata

Multi-step circular Markov processes with canonical vine representations

Abstract: A way of construction of multi-step circular Markov processes is introduced. Wehrly and Johnson (1980) proposed bivariate circular distributions with specified marginal distributions and its representation naturally induces a circular Markov process with two arbitrary circular densities. One of the densities is called a binding density and it can be regarded as a copula density. This paper extends this circular Markov process to a multi-step one by employing a pair-copula decomposition of a multivariate distribution. The way of decomposition corresponds to the graphical model denoted as the canonical vine. Fitting the multi-step circular Markov process to real circular data is also considered.

Nobuhiro Taneichi

On improvement of generalized ϕ-divergence goodness-of-fit statistics for GLIM of binary data

Abstract: We consider generalized linear models of binary data. For testing the null hypothesis that the considered model is correct, ϕ-divergence family of goodness-of-fit test statistics $C_{\phi\phi^*}$ which is based on minimum ϕ-divergence estimator is considered. A family of statistics $C_{\phi\phi^*}$ includes a power divergence family of statistics $R^{a,b}$ which is based on minimum power divergence estimator. The derivation of an expression of continuous term of asymptotic expansion for the distribution of $C_{\phi\phi^*}$ under the null hypothesis is shown. Using the expression, a transformed $C_{\phi\phi^*}$ statistic that improves the speed of convergence to the chi-square limiting distribution of $C_{\phi\phi^*}$ is obtained. In the case of $R^{a,b}$, the performance of the transformed statistic and that of the original statistic are compared numerically. (Joint work with Yuri Sekiya and Jun Toyama)

Kou Fujimori

The l_q consistency of the Dantzig selector for Cox's proportional hazards model

Abstract: The Dantzig selector proposed by Candés and Tao (2007) is an estimation procedure for linear regression models in a high-dimensional and sparse setting. We apply this method for the proportional hazards model proposed by D.R. Cox and prove the l_q consistency for all $q \in [1, \infty]$ of the estimator based on the compatibility factor, the weak cone invertibility factor, and the restricted eigenvalue for certain deterministic matrix which approximates the Hessian matrix of log partial likelihood. Our proofs for the consistency and the rate of convergence of the Dantzig selector are similar to those of the Lasso estimator for the proportional hazards model, but our matrix conditions for the three factors are weaker than those of previous researches. (Joint work with Yoichi Nishiyama)

Yuan-Tsung Chang

Shrinkage estimators of Poisson means based on prior information and its applications to multiplicative Poisson models

Abstract: In estimating $p\ (\geq 2)$ independent Poisson means, Clevenson and Zidek (1975) have given a class of estimators that shrink the unbiased estimators to the origin and have shown that the shrinkage estimators dominate the unbiased ones under the normalized squared error loss. Here we consider shrinking the unbiased estimators to the specified values only when observed values are greater than or equal to the specified ones, to their minimum, or more generally to some order statistics. We apply the proposed method to the simultaneous estimation of the means in multiplicative Poisson models and propose a class of multiple-shrinkage estimators which shrinks the MLE to the row-wise and column-wise order statistics. (Joint work with Nobuo Shinozaki)

Kun Chen

Bartlett correction of empirical likelihood with unknown variance

Abstract: Bartlett correction is one of the desirable features of likelihood inference, which allows constructions of confidence regions for parameters with improved coverage probabilities. In this paper we study Bartlett correction for frequency domain empirical likelihood based on the Whittle likelihood of linear time series models. Previous studies demonstrated the Bartlett correction of EL for independent observations, Gaussian short- and long-memory time series with known innovation variance. Nordman and Lahiri (2006) showed that frequency domain empirical log-likelihood ratio statistics does not have an ordinary χ^2-limit when the innovation is non-Gaussian with unknown variance, which restricts the use of empirical likelihood inference in time series. By profiling out the innovation variance from the Whittle likelihood function, we show that the empirical log-likelihood ratio statistic is χ^2-distributed and is Bartlett correctable. In particular, orders of the coverage error of confidence regions can be reduced from $1/n$ to $1/n^2$.

Hiroki Masuda

Locally stable regression without ergodicity and finite moments

Abstract: We present some recent findings on (semi) parametric estimation based on high-frequency sample for a class of continuous-time stochastic processes driven by a pure-jump noise. The underlying setup can cover not only stochastic differential equations but also some semimartingale regression models. Under the locally stable property of the noise process, asymptotic mixed normality can be proved in a unified way through a non-Gaussian strategy. What is most notable here is that the result holds true without any moment condition and any ergodicity and/or stationarity, which cannot be achieved by the conventional Gaussian counterpart; the latter most often puts limitations on data-dependence and driving-noise structures in compensation for providing an explicit quasi-likelihood. Numerical experiments are shown to illustrate effectiveness of the proposed strategies. One concern is that, except for a special case, our non-Gaussian strategy involves time-consuming numerical optimization.

Sangyeol Lee

Recent developments in inferences for conditionally heteroscedastic location-scale time series models

Abstract: This study examines the asymptotic properties of a class of conditionally heteroscedastic location-scale time series models with innovations following a generalized asymmetric Student t distribution (ASTD) or an asymmetric exponential power distribution (AEPD). We first show the consistency and asymptotic normality of the conditional maximum likelihood estimator of the model parameters under certain regularity conditions. Then, based on the maximum likelihood estimator, we estimate conditional value-at-risk (VaR) and expected shortfall (ES) by using their closed forms induced from the model. Their performance is finally compared with that of conditional autoregressive VaR and expectile methods. To ensure the adequacy of the model in advance of the VaR and ES calculation, we develop an entropy-type goodness-of-fit test based on residuals and a residual-based cumulative sum test to conduct a parameter change test. To handle the former, we also investigate the asymptotic behavior of the residual empirical process.

Arup Bose

Large sample behaviour of high dimensional autocovariance matrices when the dimension grows slower than the sample size

Abstract: Consider a sample of size n from a linear process of dimension p where $n,\ p \to \infty,\ p/n \to 0$. Let $\hat{\Gamma}_u$ be the sample autocovariance of order u. The existence of the limiting spectral distribution (LSD) of $\hat{\Gamma}_u + \hat{\Gamma}_u^*$, the symmetric sum of the sample autocovariance matrix $\hat{\Gamma}_u$ of order u, after appropriate centering and scaling, has been considered in the literature in exactly one article under appropriate (strong) assumptions on the coefficient matrices. Under significantly weaker conditions, we prove, in a unified way, that the LSD of any symmetric polynomial in these matrices such as $\hat{\Gamma}_u + \hat{\Gamma}_u^*$, $\hat{\Gamma}_u\hat{\Gamma}_u^*$, $\hat{\Gamma}_u\hat{\Gamma}_u^* + \hat{\Gamma}_k\hat{\Gamma}_k^*$, after suitable centering and scaling, exists and is non-degenerate. We use methods from free probability in conjunction with the method of moments to establish our results but unlike in the

case $p/n \to y \in (0, \infty)$, the embedding technique does not work in this scenario. In addition, we are able to provide a general description for the limits in terms of some freely independent variables. The earlier result follows as a special case. We also establish asymptotic normality results for the traces of these matrices. We suggest statistical uses of these results in problems such as order determination of high-dimensional MA and AR processes and testing of hypotheses for coefficient matrices of such processes.

Hans R. Künsch

Monte Carlo filtering and data assimilation

Abstract: The problem of estimating the current state of a latent Markov process based on a sequence of partial and noisy observations up to the same time is called filtering in engineering and in statistics, and data assimilation in the geosciences. In the linear Gaussian case, the Kalman filter provides the exact solution in a recursive form, but for nonlinear and non-Gaussian cases one has to rely on approximations by recursive Monte Carlo algorithms. The two most widely used such algorithms are the particle filter and the ensemble Kalman filter. The former originated in statistics and engineering, the latter in the geosciences, and until recently there was little exchange of ideas between these two areas. In this talk, I will describe the basics of both algorithms, discuss their strengths and weaknesses and present some recent proposals that aim to combine their strengths. (Joint work with Marco Frei and Sylvain Robert)

October 29

Yuma Uehara

Statistical inference for misspecified ergodic Lévy driven stochastic differential equation models

Abstract: Currently, we can obtain high-frequency data stemming from time varying phenomena, such as log returns, spike noise of neurons and so on. To better describe non-Gaussian random behavior, Lévy driven stochastic differential equation (SDE) models may be appropriate. In this study, we consider the parametric estimation problem in the case that the coefficients of the model are partly or fully misspecified. Our approach is based on the tractable Gaussian quasi-likelihood focusing on the conditional mean and variance structure. We will present the rate of convergence of the corresponding estimators and its explicit asymptotic distribution in a special case.

Nobumichi Shutoh

Bartlett correction to the likelihood ratio test for MCAR with two-step monotone missing data

Abstract: In this talk, we give Bartlett type correction to the likelihood ratio test for the satisfaction of missing completely at random (MCAR) proposed by Little (1988) under two-step monotone missing data drawn from a multivariate normal distribution. This test plays an important role in the statistical analysis of incomplete datasets. We observed that the attained significant level of Little's test is larger than the nominal significant level, especially for the case of small sample size. Therefore, by using the idea of Nagao (1973), we derive an asymptotic expansion for the distribution of the test statistic, and have Bartlett type correction to that as the main result. Finally, the advantages of our approach are confirmed in Monte Carlo simulations. Our correction drastically improves the accuracy of the significant level in Little's test for MCAR, and performs well even on moderate sample sizes. (Joint work with Takahiro Nishiyama and Masashi Hyodo)

Yasunori Tamura

Multi-group profile analysis for high-dimensional

Abstract: In this talk, we propose approximate tests for three hypotheses in profile analysis under high-dimensional elliptical populations with unequal covariance matrices. To construct these tests, we derive the limiting distribution of the quadratic form in a set of sample mean vectors defined by the specific matrix when dimension is substantially larger than sample size. This quadratic form can also be used in the linear hypotheses of a set of mean vectors not only used in the profile analysis. We also investigate asymptotic sizes and powers of proposed tests using these asymptotic results. Finally, we study the finite sample performance of the proposed test via Monte Carlo simulations. We demonstrate the relevance and benefits of the proposed approach for a number of alternative settings.

Masashi Hyodo

Simultaneous testing of the mean vector and the covariance matrix for high-dimensional data

Abstract: We treat two-sample simultaneous test procedures on high-dimensional mean vectors and covariance matrices under non-normal populations. When $p < \min\{n_1, n_2\}$, the likelihood ratio test can not be applicable for this problem. In this talk, we propose L^2-norm based test for simultaneous testing of the mean vector and the covariance matrix under high-dimensional non-normal populations. To construct this, we derive asymptotic distribution of test statistic based on both the differences mean vectors and covariance matrices. We also investigate asymptotic sizes and powers of proposed tests using this result. Finally, we study the finite sample and dimension performance of this test via Monte Carlo simulations, and apply our methods to leukemia data sets analyzed by Dudoit et al. (2002). (Joint work with Takahiro Nishiyama)