"Statistical Analysis for High-Dimensional, Circular or Time Series Data"

Date: March 2 and 3, 2017. (March 4: Free Discussion)

Venue: Keio University, Hiyoshi Campus,
Raiosha, 2nd Floor, large conference room
(Access map: http://www.hc.keio.ac.jp/en/about/index.html)

Organizer: Hiroshi SHIRAISHI
(Department of Mathematics, Keio University)

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

Keio International Symposium

"Statistical Analysis for High-Dimensional, Circular or Time Series Data"

Date: March 2 and 3, 2017. (March 4: Free Discussion)

Venue: Keio University, Hiyoshi Campus, Raiosha, 2nd Floor,
large conference room
(Accessmap: http://www.hc.keio.ac.jp/en/about/index.html)

Organizer: Hiroshi SHIRAISHI
(Department of Mathematics, Keio University)

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

Program (* Speaker)

March 2

Session (I): 10:00 - 11:25: chaired by Hiroshi Shiraishi

10:00 - 10:10: Hiroshi Shiraishi (Keio Univ.) Opening

10:10 - 10:40: Takaaki Koike (Keio Univ.)
Estimation of Risk Contributions with MCMC

10:40 - 11:25: Cathy W. S. Chen* (Feng Chia Univ.), Mike K. P. So, Jessica C. Li, and Songsak Sriboonchitta
Autoregressive Conditional Negative Binomial Model Applied to Over-dispersed Time Series of Counts

11:25 - 11:40: Break

Session (II): 11:40 - 13:10: chaired by Cathy W. S. Chen

11:40 - 12:25: Yan Liu (Waseda Univ.)
Statistical inference for quantiles in frequency domain

12:25 - 13:10: Tommaso Proietti* (Università di Roma), and Alessandra Luati
Generalised Linear Cepstral Models for the Spectrum of a Time Series

13:10 - 14:20: Lunch

Session (III): 14:20 - 15:50: chaired by Tommaso Proietti

14:20 - 15:05: Fumiya Akashi* (Waseda Univ.), H. Dette, and Y. Liu
Empirical likelihood approach for robust change point detection of infinite variance time series models

15:05 - 15:50: S.N. Lahiri (North Carolina State Univ.)
Time Series model fitting and regularization methods

15:50- 16:05: Break

Session (IV): 16:05 - 17:35: chaired by S.N. Lahiri

16:05 - 16:50: Xiaoling Dou (Waseda Univ.)
A generalized least squares estimator for non-linear time series model

16:50 - 17:35: Meihui Guo* (National Sun Yat-sen Univ.), L.C. Lin, R.B. Chen, and M.N.L. Huang
Robust principal expectile component analysis

18:30 - : Dinner

March 3

Session (V) 10:00 -11:30 : chaired by Meihui Guo

10:00 - 10:45: Hiroaki Ogata (Tokyo Metropolitan Univ.)
Multi-order circular Markov processes with canonical vine representations

10:45 - 11:30: Ngai Hang Chan (Chinese Univ. of Hong Kong)
Group Orthogonal Greedy Algorithm for Structural Break Time Series: From Univariate to Multivariate

11:30- 11:45: Break

Session (VI): 11:45 - 13:15: chaired by Ngai Hang Chan

11:45 - 12:30: Toshihiro Abe (Nanzan Univ.)
Symmetric circular distributions and their sine-skewed extensions

12:30 - 13:15: Alessandra Luati* (Univ. of Bologna), and Andrew Harvey
Score driven models for time varying parameters

13:15 - 14:20: Lunch

Session (VII): 14:20 - 15:50: chaired by Alessandra Luati

14:20 - 15:05: Shogo Kato (Inst. Stat. Math.)
The wrapped Cauchy distribution on the circle and its bivariate extension

15:05 - 15:50: Mohsen Pourahmadi (Texas A&M Univ.)
A Frequency Domain Approach to Stationary Subspace Analysis

15:50 - 16:05: Break

Session (VIII): 16:05 - 17:35: chaired by Mohsen Pourahmadi

16:05 – 16:50: Kunio Shimizu (Inst. Stat. Math.)
Discrete distributions on the circle

16:50 - 17:35: Arthur Pewsey (Univ. of Extremadura)
Circulas: copulas for modelling data on the torus

March 4

Session (IX): 10:00 - 17:00: Free Discussion

Abstract

Takaaki Koike:
Estimation of Risk Contributions with MCMC

Abstract: Determining risk contributions by unit exposures to portfolio-wide economic capital is an important task in financial risk management. In this talk, we address the problem of estimating risk contributions when the total risk is measured by Value-at-Risk (VaR). We propose a new estimator of VaR contributions, that utilizes Markov chain Monte Carlo (MCMC) method. Unlike the existing estimators, our MCMC-based estimator is computed by samples of conditional loss distributions given the rare event of our interest. Thanks to this feature, our estimator has improved sample-efficiency compared with the crude Monte Carlo method. We show that our MCMC-based estimator has several attractive properties, such as consistency and asymptotic normality. Our experiment also shows that, in various risk models used in practice, the MCMC estimator has smaller bias and MSE compared with these of existing estimators.

Cathy W. S. Chen:
Autoregressive Conditional Negative Binomial Model Applied to Over-dispersed Time Series of Counts

Abstract: Integer-valued time series analysis offers various applications in biomedical, financial, and environmental research. However, existing works usually assume no or constant over-dispersion. In this paper, we propose a new model for time series of counts, the autoregressive conditional negative binomial model that has a time-varying conditional autoregressive mean function and heteroskedasticity. The location and scale parameters of the negative binomial distribution are flexible in the proposed set-up, inducing dynamic over-dispersion. We adopt Bayesian methods with a Markov chain Monte Carlo sampling scheme to estimate model parameters

and utilize deviance information criterion for model comparison. We conduct simulations to investigate the estimation performance of this sampling scheme for the proposed negative binomial model. To demonstrate the proposed approach in modeling time-varying over-dispersion, we consider two types of criminal incidents recorded by New South Wales (NSW) Police Force in Australia. We also fit the autoregressive conditional Poisson model to these two datasets. Our results demonstrate that the proposed negative binomial model is preferable to the Poisson model. (Joint work with Mike K. P. So, Jessica C. Li, and Songsak Sriboonchitta)

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.

Tommaso Proietti:
Generalised Linear Cepstral Models for the Spectrum of a Time Series

Abstract: The paper introduces the class of generalised linear models with Box-Cox link for the spectrum of a time series. The Box-Cox transformation of the spectral density is represented as a finite Fourier polynomial, with coefficients, that we term generalised cepstral

coefficients, providing a complete characterization of the properties of the random process. The link function depends on a power transformation parameter and encompasses the exponential model (logarithmic link), the autoregressive model (inverse link), and the moving average model (identity link). One of the merits of this model class is the possibility of nesting alternative spectral estimation methods under the same likelihood-based framework, so that the selection of a particular parametric spectrum amounts to estimating the transformation parameter. We also show that the generalised cepstral coefficients are a one to one function of the inverse partial autocorrelations of the process, which can be used to evaluate the mutual information between the past and the future of the process. (Joint work with Alessandra Luati)

Key words and phrases: Generalised Linear Models; Box-Cox link; Whittle Likelihood; Mutual Information.

Fumiya Akashi:
Empirical likelihood approach for robust change point detection of infinite variance time series models

Abstract: This talk considers the change-point detection problem of the autoregressive models, and constructs the robust empirical likelihood ratio test statistic for the heavy-tails of the model. It is shown that the proposed test statistic has a known limit distribution regardless whether the model has finite variance or not. In contrast to other works on change point tests using the empirical likelihood, we show the result without knowledge of the location of the change point or the tail behavior of the innovations. We also compare the finite sample performance of the proposed test with the classical cumulative sum-based test via simulation experiments. As a results, the proposed test is shown to have nice properties in practical situations. (Joint work with H. Dette and Y. Liu)

S. N. Lahiri:
Time Series model fitting and regularization methods

Abstract: Traditional methods of fitting time series models are often based on various information criteria, such as the AIC, BIC and their bias corrected versions. In this talk, we consider some popular penalization methods that have been recently proposed in the statistical machine learning literature as alternatives for model fitting . We derive conditions that ensure selection of the correct model with high probability and investigate large sample properties of the resulting estimators. Results from a moderate simulation study will be used to illustrate finite sample performance of the proposed methods.

Xiaoling Dou:

A generalized least squares estimator for non-linear time series model

Abstract: Ochi(1983) proposed an estimator to estimate the autoregressive coefficient of the first order of autoregression (AR(1)) model by using two constants for the end points of the process. Classical estimators, such as, the least square estimator, Burg's estimator and Yule-Walker's estimator of the parameter in AR(1) model are special choices of the constants in Ochi's estimator. First, we provide a simulation for AR(1) model and examine the performance of Ochi's estimator. Writing the autoregressive conditional heteroskedasticity model of order 1, ARCH(1), into a similar form to AR(1), we extend Ochi's estimator to the ARCH(1) model, introduce the ideas of the least squares estimator, Burg's estimator and Yule-Walker's estimator and compare the relationships of them with Ochi's estimator for ARCH(1) model. With a simulation, we investigate Ochi's estimator for ARCH(1) with different values of parameters and different sample sizes.

Meihui Guo:

Robust principal expectile component analysis

Abstract: Principal component analysis (PCA) is widely used in dimensionality reduc- tion for high dimensional data. It finds principal

components by sequentially maximizing the component score variance around the mean. However, in many applications, one is interested in capturing the tail variables of the data rather than variation around the center. In order to capture the tail characters, Tran, Osipenko, and H¨ardle (2014), based on an asymmetric L2 norm, proposed prin- ciple expectile components (PEC). In this study, we introduce a new method called principal Huber-type expectile component (PHEC) using an asymmetric Huber norm to produce robust PECs. Statistical properties of the PHEC are derived and a derivative free optimization approach, particle swarm optimiza- tion (PSO), is used to find PHECs. As illustrations, PHEC is applied to real and simulated data with encouraging results. (Jointwork with L.C. Lin, R.B. Chen and M.N.L. Huang)

Key words: asymmetric norm, Huber norm, particle swarm optimization, principle component.

Hiroaki Ogata:
Multi-order circular Markov processes with canonical vine representations

Abstract: A way of construction of multi-order 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-order 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-order circular Markov process to real circular data is also considered.

Ngai Hang Chan:
Group Orthogonal Greedy Algorithm for Structural Break Time Series: From Univariate to Multivariate

Abstract: The problem of estimating change-points in a structural break autore- gressive (SBAR) model when the number of change-points m is unknown is considered in this paper. By reformulating the problem in a high-dimensional regression context, a modified high-dimensional variable selection method, namely, the so-called GOGA+HIDC+Trimming, is proposed to estimate the change-points $\{t_1, \ldots, t_m\}$ and the unknown parameter m. It is further shown that these estimators are consistent and the computation can be efficiently performed. This method is further extended to the multivariate case to consider the structural break vector autoregressive (SBVAR) model. It is shown that the proposed method can be adopted to integrate the informa- tion across different components even when the change-points are relatively packed across components. Simulation studies are conducted to assess the finite sample performance.

Joint work with Yuanbo Li and C.Y. Yau. Research supported in part by grants from HKSAR-RGC-GRF.

Toshihiro Abe:
Symmetric circular distributions and their sine-skewed extensions

Abstract: In this talk, we review the history of the symmetric circular distributions, then, consider their skew extentions by sine perturbation. The main focus of the talk, the sine-skewed family of circular distributions, is a special case of the construction due to Umbach and Jammalamadaka (2009). Very general results are provided for the properties of any such distribution, and the sine-skewed Jones--Pewsey (SSJP) distribution is introduced as a particularly flexible model of this type. We study its properties as well as those of three of its special cases. General results are also provided for maximum likelihood estimation of the parameters of any sine-skewed

distribution. The proposed models and inferential methods are applied in the analysis of the thunderstorms data.

Alessandra Luati:
(K) Score driven models for time varying parameters

Abstract: Score driven models models have been recently introduced to account for data that may be generated from heavy tail distributions. Due to their flexibility and the fact that a comprehensive asymptotic theory based on maximum likelihood can be developed in closed form, these models have immediately received a great deal of attention. The key feature is that the dynamics of the parameters that specify the distribution of the data generating process depend on the score of the conditional likelihood function. We describe an observation driven model, based on the conditional Student t-distribution, that is tractable and retains some of the desirable features of the linear Gaussian model. Illustrations and directions for future research will be highlighted. (joint with Andrew Harvey)

Shogo Kato:
The wrapped Cauchy distribution on the circle and its bivariate extension

Abstract: The wrapped Cauchy distribution is a probability distribution defined on the circle. In this talk we discuss two topics related to the wrapped Cauchy distribution. First some known results about this distribution are introduced, including basic properties, derivations, parameter estimation and an association with the transformation on the circle called the Möbius transformation. Second we provide a review of a bivariate extension of the wrapped Cauchy distribution proposed by Kato and Pewsey (2015). It is seen that this extended model displays numerous tractable properties such as closed-form expressions for the density and characteristic functions, unimodality and pointwise symmetry, marginal and conditional distributions that

are all wrapped Cauchy, and fast parameter estimation based on the method of moments and maximum likelihood.

Mohsen Pourahmadi:
A Frequency Domain Approach to Stationary Subspace Analysis

Abstract: Stationary Subspace Analysis (SSA) is a recent time domain technique for finding linear transformations of nonstationary processes that are stationary in the limited sense that the first two moments or means and lag-0 covariances are time-invariant. It finds a matrix that projects the nonstationary data onto a stationary subspace by minimizing a Kullback-Leibler divergence between Gaussian distributions measuring the non-constancy of the means and covariances across several segments. We propose a frequency domain SSA for general multivariate second-order nonstationary processes. Using the asymptotic uncorrelatedness property of the discrete Fourier transform of a stationary time series, we construct a measure of departure from stationarity and optimize it to find the stationary subspace. The dimension of the subspace, a key parameter, is estimated using a sequential testing procedure and its asymptotic properties are discussed. We illustrate the broader applicability and better performance of the frequency domain method in comparison to time domain SSA methods through simulations and discuss an application in analyzing EEG data from Brain-Computer Interface experiments. (Joint work with Raanju Ragavendar Sundararajan, Texas A&M University)

Kunio Shimizu:
Discrete distributions on the circle

Abstract: Wrapping is a fruitful method for generating circular distributions, not only continuous but also discrete. After reviewing some discrete wrapped distributions, we propose a method for constructing a discrete circular distribution (DCD) from any

continuous circular distribution whose probability density function (pdf) is represented by a Fourier series. The probability mass function is defined by taking the normalized circular pdf values at some pre-fixed equidistant points on the circumference of the circle. The normalizing constants of our generated DCDs are explicitly expressed by the cosine moments of the continuous circular distributions. Some examples are presented, along with their cumulative sums and trigonometric moments.

Arthur Pewsey:
Circulas: copulas for modelling data on the torus

Abstract: We consider the circular analogues of copulas, which we refer to as ‘circulas’, and concentrate on one particular class of bivariate circulas which is pre-existing but has not been studied in such explicit form or detail before. This class is appealing in many ways but does not necessarily result in especially attractive bivariate circular models when the marginals are not circular uniform. A major exception is an elegant bivariate wrapped Cauchy distribution proposed and developed in Kato & Pewsey (2015). We will consider properties of the circulas themselves, as well as those of distributions generated using circulas with marginals that are not circular uniform. Likelihood based inference for the latter distributions will be considered and applied in the modelling of wind directions at a Texan weather station and data on the pre-earthquake direction of steepest descent and post-earthquake direction of lateral ground movement before and after, respectively, an earthquake in Noshiro, Japan.