NUS
National University
of Singapore

WASEDA
University

NUS WASEDA Workshop 2023

Arrival:	**March 12, 2023**
Workshop:	**March 14, 2023**
Discussions:	**March 15-17, 2023**
Departure:	**March 18, 2023**

Workshop details

Time:	**09:30 - 18:30**
Place:	**Institute of Operations Research and Analytics** **Innovation 4.0, 3 Research Link, #04-01, Singapore 117602** **I4-01-03 - SEMINAR ROOM**

09:30 — 10:00	**Keynote : In My Statistical Life** Masanobu Taniguchi | Waseda University
10:00 — 10:30	**Coffee Break**
10:30 — 11:00	**Likelihood Ratio Processes under Non-standard Settings** Yuichi Goto | Kyushu University
11:00 — 11:30	**Sampling With Constraints Using Variational Methods** Xin Tong | National University of Singapore
11:30 — 12:00	**Optimal One-pass Nonparametric Estimation Under Memory Constraint** Zhenhua Lin | National University of Singapore
12:00 — 12:30	**Q & A**
12:30 — 14:00	**Lunch Break**
14:00 — 14:30	**Factor Modelling for Clustering High-dimensional Time Series** Guangming Pan | Nanyang Technological University
14:30 — 15:00	**Sparse Principal Component Analysis for High-dimensional Stationary Time Series** Kou Fujimori | Shinshu University
15:00 — 15:30	**Asymptotic Theory for Time Series** Yujie Xue | Waseda University
15:30 — 16:00	**Q & A**
16:00 — 16:30	**Coffee Break**
16:30 — 17:00	**Integer-valued Transfer Function Models for Counts that Show Zero Inflation** Cathy W.S. Chen | Feng Chia University
17:00 — 17:30	**Log-linear Zero-inflated Generalized Poisson Autoregression for Covid-19 Pandemic Modelling** Xiaofei Xu | Wuhan University
17:30 — 18:00	**Multi-stage Stochastic Bunker Procurement Planning** Wei Li | National University of Singapore
18:00 — 18:30	**Q & A**
19:30 — 21:30	**Dinner**

Speakers

Keynote — In My Statistical Life

Abstract. This talk surveys my life research. The following topics will be delivered. (1) Introduction of spectral divergence and discussion on efficiency and robustness. (2) Development in high-order asymptotic theory of time series analysis. Beyond the simultaneous equation analysis. (3) Statistical analysis of "curved" stochastic models. (4) Foundation of time series discriminant analysis (5) Statistical theory based on integral functional of nonparametric spectral estimators. Semiparametric estimation for spectra. Introduction of high-order asymptotic theory for semiparametric time series estimators. (6) LAN based asymptotic theory for time series, including long memory ones. (7) Systematic approach for portmanteau tests. (8) Empirical likelihood approach for time series. (9) Non-regular estimation for time series, and Bartlett adjustment for nonstandard settings. (10) Asymptotic theory of shrinkage estimation for time series. (11) Asymptotic theory for portfolio estimation. (12) New look at circular distributions in view of high-order spectral distribution of stationary processes. (13) Analysis of variance for time series

Masanobu Taniguchi received the B.S. degree in mathematics and the M.S. and Dr. degrees in mathematical science from Osaka University, Japan, in 1974, 1976 and 1981, respectively. He joined the Department of Mathematics, Hiroshima University, and the Department of Mathematical Science, Osaka University, in 1983 and 1990, respectively. He was a Visiting Professor at the University of Bristol, UK, in 2000. He is currently a Professor in the Department of Applied Mathematics, Waseda University, Japan. His research interests include time series analysis, mathematical statistics, multivariate analysis, information geometry, signal processing, econometric theory and financial engineering. His main contributions in time series analysis are collected in his book : "Asymptotic Theory of Statistical Inference for Time Series" (New York : Springer-Verlag, 2000). He received the Ogawa Prize (Japan) in 1989, the Econometric Theory Award (USA) in 2000, the Japan Statistical Society Prize in 2004, and Analysis Award in 2013 (Mathematical Society of Japan), Distinguished Author Award 2020, Journal of Time Series Analysis, 文部科学大臣表彰、科学技術賞 (2022). He is a Fellow of the Institute of Mathematical Statistics (USA, 1987 –), and acted the Editor of the Journal of the Japan Statistical Society (2006 – 2009). From 2022, he is an Emeritus Professor of Waseda University, Japan.

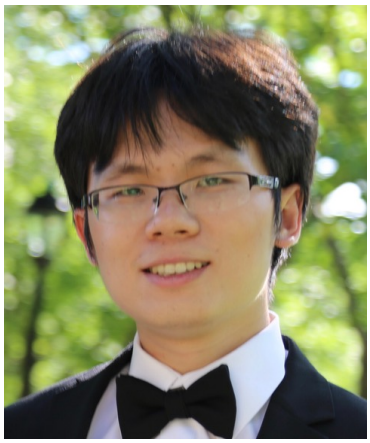
Likelihood Ratio Processes under Non-standard Settings

Abstract. In this talk, we investigate likelihood ratio (LR) processes under non-standard settings. First, a curved Gaussian family and a simultaneous equation system are discussed. We show that both models have the local asymptotic normal (LAN) property. Hence, we can construct optimal inference and testing methods based on LAN property. Second, one-way random ANOVA models are scrutinized. We elucidate that the LR process of this model has unusual limit distributions that depend on the contiguity orders. Consequently, the ordinary optimal theory based on LAN property is not available. By Neymann–Pearson framework, we show the test based on LR is asymptotically most powerful. This is a joint work with T. Kaneko (Waseda Univ.), S. Kojima (Waseda Univ.), M. Taniguchi (Waseda Univ.).

Yuichi Goto received a D.S. degree from Waseda University in March 2021. He is currently an assistant professor in Department of Mathematical Sciences, Faculty of Mathematics, Kyushu University, 744 Motooka, Nishi-ku, Fukuoka, Japan. His research interests include time series analysis, especially, binary series and spectral density. He is a member of the Japan Mathematical Society of Japan and Japan Statistical Society.

Sampling with constraints using variational methods

Abstract. Sampling-based inference and learning techniques, especially Bayesian inference, provide an essential approach to handling uncertainty in machine learning (ML). As these techniques are increasingly used in daily life, it becomes essential to safeguard the ML systems with various trustworthy-related constraints, such as fairness, safety, interpretability. We propose a family of constrained sampling algorithms which generalize Langevin Dynamics (LD) and Stein Variational Gradient Descent (SVGD) to incorporate a moment constraint or a level set specified by a general nonlinear function. By exploiting the gradient flow structure of LD and SVGD, we derive algorithms for handling constraints, including a primal-dual gradient approach and the constraint controlled gradient descent approach.We investigate the continuous-time mean-field limit of these algorithms and show that they have O(1/t) convergence under mild conditions.



Xin Tong is an associate professor at the National University of Singapore, department of mathematics. He received his Ph.D. degree from Princeton University in 2013. Prior to his position at the National University of Singapore, he was a postdoc at the Courant Institute of New York University. His recent research focuses on the analysis and derivation of stochastic algorithms.

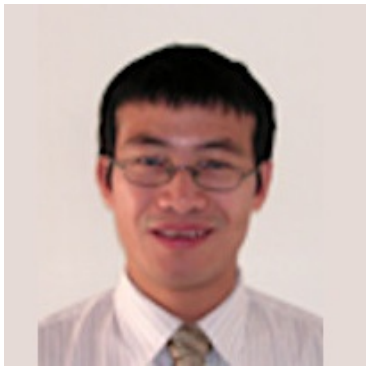
Optimal One-pass Nonparametric Estimation Under Memory Constraint

Abstract. For nonparametric regression in the streaming setting, where data constantly flow in and require real-time analysis, a main challenge is that data are cleared from the computer system once processed due to limited computer memory and storage. We tackle the challenge by proposing a novel one-pass estimator based on penalized orthogonal basis expansions and developing a general framework to study the interplay between statistical efficiency and memory consumption of estimators. We show that, the proposed estimator is statistically optimal under memory constraint, and has asymptotically minimal memory footprints among all one-pass estimators of the same estimation quality. Numerical studies demonstrate that the proposed one-pass estimator is nearly as efficient as its non-streaming counterpart that has access to all historical data.

Zhenhua Lin is currently Presidential Young Professor in Department of Statistics and Data Science, National University of Singapore. He received his Ph.D. degree in statistics from University of Toronto, master's degrees in both statistics and computing science from Simon Fraser University, and bachelor's degree in computer science from Fudan University. His research primarily focuses on non-Euclidean, high-dimensional and functional data analysis, with papers published in The Annals of Statistics, Journal of the American Statistical Association, Biometrika, Journal of the Royal Statistical Society (Series B), Biometrics, Journal of Computational and Graphical Statistics, etc.

Factor Modelling for Clustering High-dimensional Time Series

Abstract. We propose a new unsupervised learning method for clustering a large number of time series based on a latent factor structure. Each cluster is characterized by its own cluster-specific factors in addition to some common factors which impact on all the time series concerned. Our setting also offers the flexibility that some time series may not belong to any clusters. The consistency with explicit convergence rates is established for the estimation of the common factors, the cluster-specific factors, the latent clusters. Numerical illustration with both simulated data as well as a real data example is also reported. As a spin-off, the proposed new approach also advances significantly the statistical inference for the factor model of Lam and Yao (2012).
This is a joint work with Zhang Bo, Yao Qiwei and Zhou Wang.



Guangming Pan is currently a Professor in the School of Physical and Mathematical Sciences since 2008. He received his Ph.D. degree from University of Science and Technology of China. His research interests include random matrices theory, multivariate analysis and applications of probablity. He has published a couple of top quality international journal papers.

Sparse principal component analysis for high-dimensional stationary time series

Abstract. We consider the sparse principal component analysis for high-dimensional stationary processes. The standard principal component analysis performs poorly when the dimension of the process is large. We establish oracle inequalities for penalized principal component estimators for the large class of processes including heavy-tailed time series. The rate of convergence of the estimators is established. We also elucidate the theoretical rate for choosing the tuning parameter in penalized estimators. The performance of the sparse principal component analysis is demonstrated by numerical simulations. The utility of the sparse principal component analysis for time series data is exemplified by the application to average temperature data.

Kou Fujimori received a D.S. degree from Waseda University in March 2019. He is currently an assistant professor of Department of Economics, Faculty of Economics and Law, Shinshu University. His research interests are statistical inference for stochastic processes, time series analysis, sparse estimation and high-dimensional statistics.

Asymptotic Theory for Time Series

Abstract. This talk consists of two parts. In the first part, we introduce a time series Hellinger distance for spectra f and g : $T(f,g) = \int \log\{1/2\sqrt{g/f} + 1/2\sqrt{f/g}\}\mathrm{d}\lambda$ for a Gaussian stationary process. By evaluating $T(f_\theta, f_{\theta+h})$ of the form $O(h^\alpha)$, we elucidate the $n^{1/2}$-consistent asymptotics of the maximum likelihood estimator of θ for non-regular spectra. For regular spectra, we introduce the minimum Hellinger distance estimator $\theta^T(\hat{g}_n) = \mathrm{argmin}_\theta T(f_\theta, \hat{g}_n)$ where $\hat{g}_n$ is a nonparametric spectral density estimator, and as a benchmark, we introduce the Whittle divergence estimator $\theta^W(\hat{g}_n)$. It can be shown that both $\theta^T(\hat{g}_n)$ and $\theta^W(\hat{g}_n)$ are asymptotically efficient, and that the former is more robust than the latter. Besides, small numerical studies will be provided. In the second part, we introduce a new generalized estimator for AR(1) model which improves MLE uniformly. For the first order autoregressive model, Ochi (1983) introduced a generalized estimator $\hat{\alpha}_n(c_1, c_2)$ of the coefficient α with two constants c_1 and c_2, which includes Daniels' estimator, least-squares estimator and Durbin's estimator. From Ochi(1984), it was shown that the MLE is better than the Ochi's estimator in the third-order sense if we modify the both estimators to be "third-order asymptotically median unbiased". In this talk, we propose a new estimator when the c_1 and c_2 depend on α, i.e., $c_1(\alpha)$ and $c_2(\alpha)$. Then we show that it improves the MLE uniformly in the sense of the third-order mean square error "without bias-adjustments". Because α is unknown, the feasible estimator with $c_1(\hat{\alpha})$ and $c_2(\hat{\alpha})$ is proposed, and it is shown that it is third-order equivalent. where $\hat{\alpha}$ is a consistent estimator of α. (Joint work with Masanobu Taniguchi) .

Yujie Xue is a junior researcher at Waseda University. She mainly focuses on researches in frequency domain for stationary time series, like the estimation of spectral density, variable selection problem with long memory disturbances.

Integer-valued transfer function models for counts that show zero inflation

Abstract. Large proportions of zero counts frequently appear in many applications, which is called zero inflation. This study proposes integer-valued transfer function models with zero-inflated generalized Poisson and negative binomial distributions to help describe overdispersion, a large proportion of zeros, and the influence of exogenous variables. We provide effective Bayesian estimation and model selection for weekly dengue cases with two meteorological covariates.

Cathy W.S. Chen is a distinguished Professor at the Department of Statistics, Feng Chia University, Taiwan. She has made contributions to the areas of Bayesian inference, diagnostics, and model comparison techniques for time series, as well as forecasting. Professor Chen has an internationally recognized record of research and scholarship as is demonstrated by the 124 papers published or accepted for publication in academic journals. She became an Elected Member of the International Statistical Institute (ISI) in 2008. In 2010, She is a fellow of the American Statistical Association, the Royal Statistical Society, and ISBA (the International Society for Bayesian Analysis). She has served as an Editor for Computational Statistics (2021-) and an Associate Editor of a number of prominent journals, including Journal of Business and Economic Statistics (2013-2018), Computational Statistics & Data Analysis (2009-2016), the Australian and New Zealand Journal of Statistics (2008-), PLOS ONE (2014-), Computational Statistics (2008-2020), and Entropy (2020-).

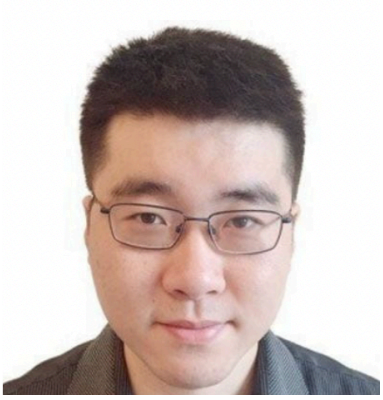
Long-memory Log-linear Zero-inflated Generalized Poisson Autoregression for Covid-19 Pandemic Modelling

Abstract. This paper describes the dynamics of daily new cases arising from the Covid-19 pandemic using a long-range dependent model. A new long memory model, LFIGX (Log-linear zero-inflated generalized Poisson integer-valued Fractionally Integrated GARCH process with eXogenous covariates), is proposed to account for count time series data with long-run dependent effect. It provides a novel unified framework for integer-valued processes with serial and long-range dependence (positive or negative), over-dispersion, zero-inflation, nonlinearity, and exogenous variables effects. We adopt an adaptive Bayesian Markov Chain Monte Carlo (MCMC) sampling scheme for parameter estimation. This new modeling is applied to the daily new confirmed cases of Covid-19 pandemic in six countries including Japan, Vietnam, Italy, the United Kingdom, Brazil, and the United States. The LFIGX model provides insightful interpretations on the impacts of policy index and temperature, and delivers good forecasting performance to the dynamics of daily new cases in different countries. This is a joint work with Ying Chen, Yan Liu, Yuichi Goto, and Masanobu Taniguchi.

Xu Xiaofei is currently an assistant professor in the Department of Probability and Statistics at Wuhan University since June 2022. Before that, she was an assistant professor at Research Institute for Science and Engineering, Waseda University from Jan 2021, and was research fellow in RMI at NUS from June 2020 to Dec 2020. Her main research interests include functional data analysis, count time series analysis, forecasting, and spectral density.

Multi-stage stochastic bunker procurement planning

Abstract. The bunker procurement decisions have a considerable impact on the international shipping industry and are closely related to the operating expenses. Obtaining optimum decisions is, however, hampered by a number of obstacles, such as requiring foresight into the future bunker price, including numerous feature variables, and accounting for their complex nonlinear relationships. This study proposed a machine learning-based multi-stage bunker procurement planning framework for the maritime sector to tackle the difficulties and investigate practical knowledge and aware- ness. We show that ensemble forecasting can effectively capture the characteristics of the multivariate, nonlinear time series, consequently facilitating operation optimization. We find that the kernel Shapley value is a useful tool for ex- plaining the prediction results of non-linear machine learning models by disentangling the effect of the multiple input variables. The eXplainable AI builds a bridge of confidence between algorithms and the conservative business sector. Compared to the present operational approach followed by the liner shipping sector, our empirical analysis demonstrates that the proposed framework can cut operating expenses by a total of USD 272,578.23 for a fleet of six vessels. To evaluate the robustness of the proposed framework, we investigate its performance after the Russia-Ukrainian war. The result shows that the framework can still save 127,941.18 dollars.



Wei Li is currently Research Fellow at the Institute of Operations Research and Analytics, National University of Singapore. He received his Ph.D. degree in computational finance from Norwegian University of Science and Technology, master's degrees in both financial forecasting and investment and financial computing from University of Glasgow and Queen Mary University of London, respectively, and bachelor's degree in business management from Jilin University. His research focuses on eXplainable AI, data science and energy analytics, with papers published in Quantitative Finance, Energy, and Journal of Commodity Market, etc.