## Contents

<table>
<thead>
<tr>
<th>[Keynote Speaker Session 1] Estimation and Comparison of Conditional Moment Models</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimation and Comparison of Conditional Moment Models (<em>Siddhartha Chib</em>)</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>[Session 1a] Innovative Bayesian Approaches for High-dimensional Variable Selection</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable Selection in Bayesian Nonparametric Models for High-dimensional Confounding (<em>Michael J. Daniels</em>)</td>
<td>2</td>
</tr>
<tr>
<td>Moment Matching Priors for Non-regular Models (<em>Shintaro Hashimoto</em>)</td>
<td>2</td>
</tr>
<tr>
<td>Bayesian Nonparametric Differential Analysis for Dependent Multigroup Data with Application to DNA Methylation Analyses (<em>Subharup Guha</em>)</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>[Session 1b] Developments in Small Area Estimation (Chiba University SAE Session)</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small Area Estimation for Grouped Data (<em>Yuki Kawakubo</em>)</td>
<td>4</td>
</tr>
<tr>
<td>Estimation and Inference for Area-wise Spatial Income Distributions from Grouped Data (<em>Genya Kobayashi</em>)</td>
<td>5</td>
</tr>
<tr>
<td>Bayesian Approach to Lorenz Curve from Time Series Grouped Data (<em>Yuta Yamauchi</em>)</td>
<td>5</td>
</tr>
<tr>
<td>Corrected Empirical Bayes Confidence Region in a Multivariate Fay-Herriot Model (<em>Tsubasa Ito</em>)</td>
<td>6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>[Session 1c] New Developments in Bayesian Forecasting and Finance</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bayesian Network Analysis of Systemic Risk in Financial Markets (<em>Mike K.P. So</em>)</td>
<td>7</td>
</tr>
<tr>
<td>Intraday Range-based Stochastic Volatility Models with Application to the Japanese Stock Index (<em>Toshiaki Watanabe</em> (presenting) and <em>Jouchi Nakajima</em>)</td>
<td>7</td>
</tr>
<tr>
<td>Bayesian Dynamic Fused LASSO (<em>Kaoru Irie</em>)</td>
<td>8</td>
</tr>
<tr>
<td>Bayesian Indicator Approach for Variable Selection in Gaussian Process (Ray-Bing Chen)</td>
<td>8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>[Session 2a] Bayesian Modeling of Networks and Scalable Inference</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modeling Networks in the Presence of Informative Community Structure (<em>Alexander Volfovsky</em>)</td>
<td>9</td>
</tr>
<tr>
<td>Session</td>
<td>Title</td>
</tr>
<tr>
<td>---------</td>
<td>-------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>2b</td>
<td>Advances in Dynamic Modelling and Computation</td>
</tr>
<tr>
<td></td>
<td>Bayesian Multivariate Factor Analysis for Evaluating the Causal Impact</td>
</tr>
<tr>
<td></td>
<td>Particle Learning for Stochastic Volatility with Leverage</td>
</tr>
<tr>
<td></td>
<td>Large-Scale Dynamic Predictive Regressions</td>
</tr>
<tr>
<td></td>
<td>Randomization Tests of Causal Effects Under General Interference</td>
</tr>
<tr>
<td>2c</td>
<td>Bayesian Joint Modeling</td>
</tr>
<tr>
<td></td>
<td>Bayesian Joint Analysis Using a Semiparametric Latent Variable Model with Non-ignorable Missing Covariates for CHNS Data</td>
</tr>
<tr>
<td></td>
<td>A Bayesian Joint Model of Mark and Intensity of Marked Spatial Point Processes with Application to Basketball Shot Chart</td>
</tr>
<tr>
<td></td>
<td>Electricity Price Modelling with Stochastic Volatility and Jumps: An Empirical Investigation</td>
</tr>
<tr>
<td>3a</td>
<td>Recent Development and Application of Bayesian Analysis</td>
</tr>
<tr>
<td></td>
<td>Bayesian Factor Models for Probabilistic Cause of Death Assessment with Verbal Autopsies</td>
</tr>
<tr>
<td></td>
<td>The Challenges of Analyzing Drug Safety Data with Competing Risk Events and a Bayesian Mixture Model</td>
</tr>
<tr>
<td></td>
<td>Bayesian Modeling of Graph Laplacians</td>
</tr>
<tr>
<td>3b</td>
<td>Recent Developments in High-dimensional Bayesian Inference</td>
</tr>
<tr>
<td></td>
<td>Robust Bayesian Regression with Shrinkage Priors</td>
</tr>
<tr>
<td></td>
<td>The Berry-Esseen Type Bound for the Bernstein-von Mises Theorem in Moderately High Dimensions</td>
</tr>
<tr>
<td></td>
<td>Singular Value Shrinkage Priors for Bayesian Prediction</td>
</tr>
<tr>
<td></td>
<td>The Laplace Approximation to a High Dimension Model</td>
</tr>
<tr>
<td>3c</td>
<td>Marginal Likelihoods: Theory, Computation, and Applications</td>
</tr>
<tr>
<td></td>
<td>On Comparing Asset Pricing Models</td>
</tr>
<tr>
<td></td>
<td>A New Monte Carlo Method for Estimating Marginal Likelihoods</td>
</tr>
<tr>
<td></td>
<td>Inflated Density Ratio and Its Variation and Generalization for Computing Marginal Likelihoods</td>
</tr>
<tr>
<td></td>
<td>Monte Carlo Methods for Computing Marginal Likelihoods with Applications to Item Response Theory Models</td>
</tr>
<tr>
<td>Tutorial Session 1</td>
<td>Learning Semiparametric Regression with Missing Covariates Using Gaussian Process Models</td>
</tr>
<tr>
<td>-------------------</td>
<td>-------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td></td>
<td>Learning Semiparametric Regression with Missing Covariates Using Gaussian Process Models (Dipak K. Dey)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Poster Session</th>
<th>Poster Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measuring Cross-country Interconnectedness with a Panel Unobservable Component Model (Gianni Amisano)</td>
<td>26</td>
</tr>
<tr>
<td>Disjunct Support Prior for Practically Significant Variable Selection in Regression (Daniel Andrade *)</td>
<td>26</td>
</tr>
<tr>
<td>Quantile Nonlinear Effects of Return and Abnormal Trading Volume for the Realized Kernel (Dong Manh Cuong *)</td>
<td>27</td>
</tr>
<tr>
<td>Bayesian Inferences of Multiple Structural Change GARCH Model with Skew Student $t$ Errors (Bonny Y.F. Lee *)</td>
<td>27</td>
</tr>
<tr>
<td>Bayesian Spatiotemporal Varying Coefficients Models for Ischemic Stroke Study (Kuo-Jung Lee)</td>
<td>28</td>
</tr>
<tr>
<td>Objective Priors in Robust Quasi-Bayesian Inference Using the Divergences (Tomoyuki Nakagawa)</td>
<td>28</td>
</tr>
<tr>
<td>Bayesian Analysis of Intraday Stochastic Volatility Models with Skew Heavy-tailed Error and Smoothing Spline Seasonality (Makoto Nakakita *)</td>
<td>29</td>
</tr>
<tr>
<td>Bayesian and Decomposition Analysis for Health Inequality in Japan (Haruhisa Nishino)</td>
<td>29</td>
</tr>
<tr>
<td>Information Geometry of Complex Autoregressive Models and Its Positive Superharmonic Priors (Hidemasa Oda *)</td>
<td>30</td>
</tr>
<tr>
<td>Novel Conjugate Analysis in Unknown Dimensional Multinomial Probabilities (Toru Ogura)</td>
<td>30</td>
</tr>
<tr>
<td>A More Stable and Scalable Posterior Computation for Bayesian Graphical Models (Sakae Oya *)</td>
<td>31</td>
</tr>
<tr>
<td>The Hedonic Regression with Bayesian Hierarchical Model for Japanese Rice Wine. What Affects Price? (Wakuo Saito *)</td>
<td>31</td>
</tr>
<tr>
<td>Hierarchical Factorial Structure of WISC-4 Data (Kazu Shigemasu)</td>
<td>32</td>
</tr>
<tr>
<td>Modeling Financial Durations with Limit Order Book Information (Tomoki Toyabe *)</td>
<td>32</td>
</tr>
<tr>
<td>Modeling Temporal Data by Using Mixture for Marked Poisson Processes (Toshiya Yoshida *)</td>
<td>32</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Keynote Speaker Session 2</th>
<th>Bayesian Learning in High-dimensional State-space Models</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bayesian Learning in High-dimensional State-space Models (Hedibert Freitas Lopes)</td>
<td>33</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Session 4a</th>
<th>Recent Developments for Statistical Models with Structures and Dependence</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Semiparametric Mixture Method for Local False Discovery Rate Estimation from Multiple Studies (Woncheol Jang)</td>
<td>34</td>
</tr>
<tr>
<td>Amortized Variational Inference Methods for Bayesian Model Criticism (Seongil Jo)</td>
<td>34</td>
</tr>
</tbody>
</table>
## Analysis of Longitudinal Binary and Survival Time Data Using Joint Models with General Random Effects Covariance Matrix (Keunbaik Lee)  
Post-processed Posteriors for Band Structured Covariances (Jaeyong Lee)  

### Session 4b Bayesian Methods and Their Applications  
36  
Use of Two Non-informative Priors in an Empirical Bayes Estimator of Multiple Poisson Means (Takemi Yanagimoto and Toru Ogura)  
Bayesian Methods for Accelerated Materials Discovery (Ryo Yoshida)  
High-dimensional Motion Segmentation with Semi-Markov Latent Gaussian Processes (Daichi Mochihashi)  

### Session 4c Advances in Bayesian Methods for Business Analytics Problems  
40  
Quantile Forecasting Based on a Multivariate Hysteretic Autoregressive Model with GARCH Errors and Time-varying Correlations (Cathy W.S. Chen)  
Bayesian Analysis of Realized Matrix-Exponential GARCH Models (Manabu Asai)  
A Generalized Threshold Stochastic Volatility Model Incorporating with Realized Measures (Feng-Chi Liu)  

### Session 5a Recent Advances in Bayesian Methods and Applications  
42  
Nonparametric Bayesian Two-part Random Effects Model for Longitudinal Semi-continuous Data Analysis (Yeongseung Chung)  
Use of Model Reparametrization to Improve Variational Bayes (Linda S. L. Tan)  
Nonparametric Bayesian Modeling in Government Statistics: Recent Developments in Imputation, Editing, and Data Protection (Hang J. Kim)  

### Session 5b Bayesian Regression Models of Complex Dependent Data  
44  
Estimation and Prediction for Spatial Generalized Linear Mixed Models (Vivekananda Roy)  
A Bayesian Approach for Spatial Cluster Detection of Regression Coefficients (Huiyan Sang)  
A Bayesian Approach to Assess Intervention Effects on Opiate Overdose Incidents in Space and Time (Richard Zehang Li)  
Bayesian Generalized Regression Models with Gaussian Process Priors (Xia Wang)  

### Session 5c Recent Advances in Bayesian Predictive Inference  
47  
Bayesian predictive distributions for Poisson and negative binomial models when the parameter spaces are restricted (Yasuuki Hamura)  
Asymptotic Properties of Bayes Estimators Based on Shrinkage Priors for Curved Exponential Families (Michiko Okudo)  
Dual Roles of Maximizing Likelihood and Shannon Entropy under Alpha-divergence Loss (Toshio Ohnishi)
### Session 6a  Monte Carlo Methods for Complicated Target Distributions

<table>
<thead>
<tr>
<th>Analysis of Markov Chain Monte Carlo Method with Heavy-tailed Target Distributions (Kengo Kamatani)</th>
<th>49</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bayesian Inference for Intractable Likelihood Models (Krzysztof atuszyski)</td>
<td>49</td>
</tr>
<tr>
<td>Randomized Hamiltonian Monte Carlo as Scaling Limit of the Bouncy Particle Sampler (Daniel Paulin)</td>
<td>50</td>
</tr>
</tbody>
</table>

### Session 6b  New Advance in Bayesian Theorem and Nonparametric

<table>
<thead>
<tr>
<th>Bayesian Methods</th>
<th>51</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shrinkage in Bayesian Shape Constrained Inference (Debdeep Pati)</td>
<td>51</td>
</tr>
<tr>
<td>Bayesian Nonparametric Clustering Analysis for Single Cell RNA Sequencing Data (Mengjie Chen)</td>
<td>52</td>
</tr>
<tr>
<td>Bayesian Generative Training is Robust (Chao Gao)</td>
<td>52</td>
</tr>
</tbody>
</table>

### Session 6c  Bayesian Statistics and Marketing

<table>
<thead>
<tr>
<th>Bayesian Statistics and Marketing</th>
<th>53</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bayesian Estimation of Population Moments and Parameters in Biased Sampling (Yuya Shimizu)</td>
<td>53</td>
</tr>
<tr>
<td>Modeling Heterogeneous Impacts of Mental Accounting and Household Stock to Consumers Inter-shopping Duration (Kazuhiro Miyatsu)</td>
<td>54</td>
</tr>
<tr>
<td>Dynamic Two-Stage Modeling for Category-Level and Brand-Level Purchases with a Bayes Inference (Kei Miyazaki)</td>
<td>54</td>
</tr>
</tbody>
</table>

### Tutorial Session 2  Discrete Random Structures and Bayesian Nonparametrics

<table>
<thead>
<tr>
<th>Discrete Random Structures and Bayesian Nonparametrics (Igor Prnster)</th>
<th>56</th>
</tr>
</thead>
</table>
Estimation and Comparison of Conditional Moment Models

Siddhartha Chib
Washington University in St. Louis

We provide a Bayesian analysis of models in which the unknown distribution of the outcomes is specified up to a set of conditional moment restrictions. The prior-posterior analysis is made possible by taking advantage of the nonparametric exponentially tilted empirical likelihood function, constructed to satisfy a sequence of unconditional moments, obtained from the conditional moments by an increasing (in sample size) vector of approximating functions (such as tensor splines based on the splines of each conditioning variable). We show that subject to a growth rate condition on the number of approximating functions, the posterior distribution satisfies the Bernstein-von Mises theorem, even when the set of conditional moments contain misspecified moment conditions. Large-sample theory for comparing different conditional moment models shows that the marginal likelihood criterion selects the model that is less misspecified, that is, the model that is closer to the unknown true distribution in terms of the Kullback-Leibler divergence. Examples to illustrate the framework and results are provided.

This is joint work with Minchul Shin and Anna Simoni.
Variable Selection in Bayesian Nonparametric Models for High-dimensional Confounding

Michael J. Daniels
University of Florida

Enriched Dirichlet processes (EDPs) provide very flexible conditional distributions that are powerful for modeling the distribution of outcomes given confounders for causal inference. However in most applications, many of the potential confounders are not true confounders. Further complicating any sort of variable selection is that such flexible models, which detect non-additivity and non-linearity in default ways, do not correspond to a single parameter (e.g., being zero) for each potential confounder. As such, we propose a backward selection algorithm that identifies covariates potentially unrelated to the outcome and then use a Bayesian model selection statistic to decide whether to remove the variable. The procedure is explored via simulations and applied to a real dataset.

Moment Matching Priors for Non-regular Models

Shintaro Hashimoto
Hiroshima University

We present objective priors with the moment matching criterion for non-regular family of distributions whose supports depend on an unknown parameter. Both one-parameter and multi-parameter cases are considered. The resulting priors are given by matching the posterior mean and bias-adjusted maximum likelihood estimator up to the higher order, and achieve Bayesian-frequentist synthesis of point estimates. The moment matching priors for regular cases were proposed by Ghosh and Liu (2011), and they showed that these priors were not invariant for one-to-one transformation of parameter. The proposed priors for non-regular cases are not also invariant and often improper. We show the posterior propriety under the proposed improper priors in some examples.
Bayesian Nonparametric Differential Analysis for Dependent Multigroup Data with Application to DNA Methylation Analyses

Subharup Guha
University of Florida

Cancer omics datasets involve widely varying sizes and scales, measurement variables, and correlation structures. An overarching scientific goal in cancer research is the development of general statistical techniques that can cleanly sift the signal from the noise in identifying genomic signatures of the disease across a set of experimental or biological conditions. We propose BayesDiff, a nonparametric Bayesian approach based on a novel class of first order mixture models, called the Sticky Poisson-Dirichlet process or multicuisine restaurant franchise. The BayesDiff methodology flexibly utilizes information from all the measurements and adaptively accommodates any serial dependence in the data, accounting for the inter-probe distances, to perform simultaneous inferences on the variables. The technique is applied to analyze the motivating DNA methylation gastrointestinal cancer dataset, which displays both serial correlations and complex interaction patterns. In simulation studies, we demonstrate the effectiveness of the BayesDiff procedure relative to existing techniques for differential DNA methylation. Returning to the motivating dataset, we detect the genomic signature for four types of upper gastrointestinal cancer. The analysis results support and complement known features of DNA methylation as well as gene association with gastrointestinal cancer. This is joint work with Chiyu Gu and Veera Baladandayuthapani.
This paper proposes a new model-based approach to small area estimation for grouped data or frequency data, which is often available from sample surveys. Grouped data contains information on frequencies of some pre-specified groups in each area, for example the numbers of households in the income classes, and thus provides more detailed insight about small areas than area level aggregated data. A direct application of the widely used small area methods, such as the Fay–Herriot model for area level data and nested error regression model for unit level data, is not appropriate since they are not designed for grouped data. The newly proposed method assumes that the unobserved unit level quantity of interest follows a linear mixed model with the random intercepts and dispersions after some transformation. Then the probabilities that a unit belongs to the groups can be derived and are used to construct the likelihood function for the grouped data given the random effects, which is in the form of the multinomial likelihood. The unknown model parameters (hyperparameters) are estimated by a newly developed Monte Carlo EM algorithm using an efficient importance sampling. The empirical best predicts (empirical Bayes estimates) of small area parameters can be calculated by a simple Gibbs sampling algorithm. The numerical performance of the proposed method is illustrated based on the model-based and design-based simulations. In the application to the city level grouped income data of Japan, we complete the patchy maps of the Gini coefficient as well as mean income across the country.
Estimation and Inference for Area-wise Spatial Income Distributions from Grouped Data

Genya Kobayashi
Chiba University

Estimating income distributions plays an important role in the measurement of inequality and poverty over space. The existing literature on income distributions predominantly focuses on estimating an income distribution for a country or a region separately and the simultaneous estimation of multiple income distributions has not been discussed in spite of its practical importance. In this work, we develop an effective method for the simultaneous estimation and inference for area-wise spatial income distributions taking account of geographical information from grouped data. Based on the multinomial likelihood function for grouped data, we propose a spatial state-space model for area-wise parameters of parametric income distributions. We provide an efficient Bayesian approach to estimation and inference for area-wise latent parameters, which enables us to compute area-wise summary measures of income distributions such as mean incomes and Gini indices, not only for sampled areas but also for areas without any samples thanks to the latent spatial state-space structure. The proposed method is demonstrated using the Japanese municipality-wise grouped income data. The simulation studies show the superiority of the proposed method to a crude conventional approach which estimates the income distributions separately.

Bayesian Approach to Lorenz Curve from Time Series Grouped Data

Yuta Yamauchi
University of Tokyo

This study is concerned with estimating the inequality measures associated with the underlying hypothetical income distribution from the times series grouped data on the Lorenz curve. We adopt the Dirichlet pseudo likelihood approach where the parameters of the Dirichlet likelihood is set to the difference in the Lorenz curve of the hypothetical income distribution for the consecutive income classes and propose a state space modelling approach which combines the parameters of the income distribution through a time series structure after some transformation. Furthermore, the information on the sample size in each survey is introduced into the nuisance Dirichlet precision parameter to take into account the variability from the sampling. From the simulated data and real data on Japanese income survey, it is confirmed that the proposed model produces more stable estimates on the inequality measures.
Corrected Empirical Bayes Confidence Region in a Multivariate Fay-Herriot Model

Tsubasa Ito
Institute of Statistical Mathematics

In the small area estimation, the empirical best linear unbiased predictor (EBLUP) in the linear mixed model is useful because it gives a stable estimate for a mean of a small area. For measuring uncertainty of EBLUP, much of research is focused on second-order unbiased estimation of mean squared prediction errors in the univariate case. In this talk, I consider the multivariate Fay-Herriot model where the covariance matrix of random effects is fully unknown, and obtain a confidence region of the small area mean that is based on the Mahalanobis distance centered around EBLUP and is second order correct. A positive-definite, consistent and second-order unbiased estimator of the covariance matrix of the random effects is also suggested. The performance is investigated through simulation study.
Bayesian Network Analysis of Systemic Risk in Financial Markets

Mike K.P. So
Hong Kong University of Science and Technology

Analyzing systemic risk in financial markets has been an active research area in financial econometrics, risk management and big data analytics. This paper proposes an approach based on network analysis to study the interrelationship between financial companies. We develop statistical models to understand how the financial network, and thus systemic risk, changes over time. We adopt Bayesian inference methods to estimate the financial network, do network prediction and use listed companies in Hong Kong to illustrate our idea.

Intraday Range-based Stochastic Volatility Models with Application to the Japanese Stock Index

Toshiaki Watanabe (presenting)¹ and Jouchi Nakajima ²
¹Hitotsubashi University; ²Bank for International Settlements

Realized stochastic volatility (RSV) models, where the true volatility is modelled jointly with a realized measure (RM) of volatility taking account of the bias in the RM, are extended for the analysis of high-frequency intraday volatility. The proposed model consists of the persistent autoregressive stochastic volatility process, seasonal components of the intraday volatility patterns, and correlated jumps in prices and volatilities. The range of the logarithmic prices within each intraday time interval is used as a RM in the proposed model. A Bayesian method for the analysis of this model is developed using Markov chain Monte Carlo (MCMC) with the exact multi-move sampler for the SV process. Using this method, the proposed model is applied to the 5-minute returns of Nikkei 225 index. It is also examined whether the intraday range-based RSV model improves the predictive ability of volatility compared with the intraday SV model without the range information and commonly-used models for daily realized volatility.
Bayesian Dynamic Fused LASSO

Kaoru Irie
University of Tokyo

In the context of dynamic modeling, a special case of Bayesian fused LASSO is considered. It is crucial in the high-dimensional settings to address two types of sparsity – sparsity in predictors and dynamics of time-varying parameters – which implies the use of two penalties that shrinks the current state variable to zero and its previous value. This defines a new Markov process that is proven stationary and reversible. The transition density of this process is decomposed into two parts, the synthetic likelihood and prior, and is realized as the posterior of a conditionally dynamic linear model. As the prior for state variables in state space models, this CDLM representation enables the efficient posterior and predictive analysis by Markov chain Monte Carlo and other statistical methods. Through the posterior analysis of real and simulated data, the state space models with the new Markov process prior are extensively compared with the DLMs of Gaussian and double-exponential state evolutions to illustrate that the two different shrinkage effects can penalize the volatile behavior of time-varying state variables while aggressively shrinking the state variables toward zero to eliminate the contribution of noisy predictors.

Bayesian Indicator Approach for Variable Selection in Gaussian Process

Ray-Bing Chen
National Cheng Kung University

In this work, we are interested in the Bayesian variable selection problems in Gaussian process models. Here we do not only focus on the mean regression functions but also take the covariate structure into account. Thus a variable is active if the corresponding regression coefficient is not zero or its hyperparameter in the covariate structure is not zero. To accomplish our goal, we first treat the regression coefficient and the hyperparameter in the covariate structure as a group and an indicator is added into the model to denote the status of this group. Thus the active variables can be identified based on the posterior samples of the indicators. The performances of the proposed Bayesian selection approach are illustrated by the simulations and the real applications in Computer Experiments.
[Session 2a] Bayesian Modeling of Networks and Scalable Inference

Organizer
Leo Duan, University of Florida

Session Chair
Leo Duan, University of Florida

Modeling Networks in the Presence of Informative Community Structure

Alexander Volfovsky
Duke University

The study of network data in the social and health sciences frequently concentrates on detecting community structures among nodes and associating covariate information to edge formation. A popular class of latent space network models captures community structure by learning an embedding of the network in Euclidean space. However in much of this data, it is likely that the effects of covariates on edge formation differ between communities (e.g. age might play a different role in friendship formation in communities across a city). This information is lost by ignoring explicit community membership. In this work we introduce an extension of the additive and multiplicative effects latent space network model where coefficients associated with certain covariates can depend on the latent community membership of the nodes. We show that ignoring such structure can lead to either over- or underestimation of covariate importance to edge formation and propose a Markov Chain Monte Carlo approach for simultaneously learning the latent community structure and the community specific coefficients.
Consistent Bayesian Joint Variable and DAG Selection in High Dimensions

Xuan Cao
University of Cincinnati

Motivated by the eQTL analysis, we consider joint sparse estimation of the regression coefficient matrix and the error covariance matrix in a high-dimensional multivariate regression model for studying conditional independence relationships among a set of genes and discovering possible genetic effects. The error covariance matrix is modeled via Gaussian directed acyclic graph (DAG) and sparsity is introduced in the Cholesky factor of the inverse covariance matrix, while the sparsity pattern in turn corresponds to specific conditional independence assumptions on the underlying variables. In this talk, we consider a flexible and general class of these DAG-Wishart priors with multiple shape parameters on the space of Cholesky factors and a spike and slab prior on the regression coefficients. Under mild regularity assumptions, we establish the joint selection consistency for both the variable and the underlying DAG when both the number of predictors and the dimension of the covariance matrix are allowed to grow much larger than the sample size. We demonstrate our theoretical results through a marginalization-based collapsed Gibbs sampler that offers a computationally feasible and efficient solution for exploring the sample space.

Inference You Can Trust: A New Approach to Boosting

Trevor Campbell
University of British Columbia

Variational inference algorithms are well-known to be computationally tractable for large-scale models and data; but they are equally well-known to provide unreliable results and underestimate posterior uncertainty. In order for variational methods to be competitive with Markov Chain Monte Carlo (MCMC) and trusted in the statistical domain, they must come with rigorous finite-data guarantees. This talk will focus on boosting methods, i.e., those that incrementally build complex variational approximations up from simple component distributions. After reviewing some recent exciting developments in the area, the talk will introduce a new approach to variational boosting that comes with rigorous theoretical convergence guarantees. Unlike previous approaches, the method requires no ad-hoc regularization. Experiments on popular models show the practicality of the approach.
A Bayesian Model for Sparse Graphs with Flexible Degree Distribution and Overlapping Community Structure

Juho Lee
University of Oxford

We consider a non-projective class of inhomogeneous random graph models with interpretable parameters and a number of interesting asymptotic properties. Using the results of Bollobás et al. [2007], we show that i) the class of models is sparse and ii) depending on the choice of the parameters, the model is either scale-free, with power-law exponent greater than 2, or with an asymptotic degree distribution which is power-law with exponential cut-off. We propose an extension of the model that can accommodate an overlapping community structure. Scalable posterior inference can be performed due to the specific choice of the link probability. We present experiments on five different real-world networks with up to 100,000 nodes and edges, showing that the model can provide a good fit to the degree distribution and recovers well the latent community structure.
Bayesian Multivariate Factor Analysis for Evaluating the Causal Impact of Policy Interventions

Silvia Montagna
University of Turin

A problem frequently encountered in many areas of scientific research is that of estimating the impact of a non-randomised binary intervention on an outcome of interest using time-series data on units that received the intervention (treated) and units that did not (controls). One popular estimation method in this setting is based on the factor analysis (FA) model. The FA model is fitted to the pre-intervention outcome data on treated units and all the outcome data on control units, and the counterfactual treatment-free post-intervention outcomes of the former are predicted from the fitted model. Intervention effects are estimated as the observed outcomes minus these predicted counterfactual outcomes. We propose two extensions of the FA model for estimating intervention effects: 1) the joint modelling of multiple outcomes to exploit shared variability, and 2) an autoregressive structure on factors to account for temporal correlations in the outcome. We demonstrate that our approach improves the precision of the intervention-effect estimates both when the number of pre-intervention measurements is small and when the number of control units is small. We apply our method to estimate the impact of stricter alcohol licensing policies on alcohol-related harms.
Particle Learning for Stochastic Volatility with Leverage
Naoki Awaya
Duke University

New filtering methods are proposed for stochastic volatility models, a class of state space models for financial time series. Due to the non-linear/non-Gaussian nature of the stochastic volatility models, the on-line posterior and predictive analysis require the computation by particle filters. The new method is based on the preceding technique of particle filters, which approximates the non-linear/non-Gaussian component by the mixture of normal distributions and correct the approximation bias by the reweighting step added to the algorithm. The new algorithm of particle filter is also designed for the models including the leverage effect, i.e., the asymmetric relationship between returns and volatilities, which has been observed in many time series of financial assets. Numerical studies show that the new methods correctly estimate the posterior and the correction step improves the accuracy of posterior analysis.

Large-Scale Dynamic Predictive Regressions
Kenichiro McAlinn
The University of Chicago

We develop a novel “decouple-recouple” dynamic predictive strategy and contribute to the literature on forecasting and economic decision making in a data-rich environment. Under this framework, clusters of predictors generate different latent states in the form of predictive densities that are later synthesized within an implied time-varying latent factor model. As a result, the latent inter-dependencies across predictive densities and biases are sequentially learned and corrected. Unlike sparse modeling and variable selection procedures, we do not assume a priori that there is a given subset of active predictors, which characterize the predictive density of a quantity of interest. We test our procedure by investigating the predictive content of a large set of financial ratios and macroeconomic variables on both the equity premium across different industries and the inflation rate in the U.S., two contexts of topical interest in finance and macroeconomics. We find that our predictive synthesis framework generates both statistically and economically significant out-of-sample benefits while maintaining interpretability of the forecasting variables. In addition, the main empirical results highlight that our proposed framework outperforms both LASSO-type shrinkage regressions, factor based dimension reduction, sequential variable selection, and equal-weighted linear pooling methodologies.
Randomization Tests of Causal Effects Under General Interference

David Puelz
The University of Chicago

We develop an approach for estimating causal effects under general interference structures between units that is built upon the classical Fisher test (1935). The null hypothesis is viewed as implying a bipartite graphical structure between units and assignments, and we find bicliques within this structure to impute potential outcomes and run valid tests. The method is applied to a unique randomized experiment for combating crime in Medellín, Colombia. We also discuss extending the Fisher test to settings where potential outcomes are modeled with Bayesian additive regression trees.
Bayesian Joint Analysis Using a Semiparametric Latent Variable Model with Non-ignorable Missing Covariates for CHNS Data

Zhihua Ma
Jinan University

Motivated by the China Health and Nutrition Survey (CHNS) data, a semiparametric latent variable model with a Dirichlet Process (DP) prior on the latent variable is proposed to jointly analyze mixed binary and continuous responses. Non-ignorable missing covariates are considered through a missing covariate model and a missing data mechanism model. The logarithm of the pseudo-marginal likelihood (LPML) is applied for selecting the priors, and the deviance information criterion (DIC) measure focusing on the missing data mechanism model is used for choosing different missing data mechanisms. A Bayesian index of local sensitivity (ISNI) is extended to explore the local sensitivity of the parameters in our model. A simulation study is carried out to examine the empirical performance of the proposed methodology. Finally, the proposed model and the ISNI index are applied to analyze the CHNS data in the motivating example.

A Bayesian Joint Model of Mark and Intensity of Marked Spatial Point Processes with Application to Basketball Shot Chart

Guanyu Hu
University of Connecticut

The success rate of a basketball shot may be higher at locations in the court where a player makes more shots. In a marked spatial point process model, this means that the marks are dependent on the intensity of the process. We develop a Bayesian joint model of the mark and the intensity of marked spatial point processes, where the intensity is incorporated in the model of the mark as a covariate.
Further, we allow variable selection through the spike-slab prior. Inferences are developed with a Markov chain Monte Carlo algorithm to sample from the posterior distribution. Two Bayesian model comparison criteria, the modified Deviance Information Criterion and the modified Logarithm of the Pseudo-Marginal Likelihood, are developed to assess the fit of different joint models. The empirical performance of the proposed methods are examined in extensive simulation studies. We apply the proposed methodology to the 2017–2018 regular season shot data of four professional basketball players in the NBA to analyze the spatial structure of shot selection and field goal percentage. The results suggest that the field goal percentages of three players are significantly positively dependent on their shot intensities, and that different players have different predictors for their field goal percentages.

Electricity Price Modelling with Stochastic Volatility and Jumps: An Empirical Investigation

Katja Ignatieva
University of New South Wales

Over the past few years, the electricity derivatives market has experienced a substantial growth in the volume of trade and the diversity of available products. This has led to a rich data environment that requires more sophisticated and accurate modelling approaches for electricity spot prices. This paper deals with an analysis of continuous-time stochastic volatility jump-diffusion processes in the context of pricing of futures contracts written on electricity spots. We formulate a variety of nested models which aim to capture the most prominent characteristics and stylised facts of the electricity spot market including mean reversion, seasonality, extreme volatility, and spikes. The proposed modelling framework extends the existing models by incorporating mean reversion, stochastic volatility, and jumps in both the underlying spot price process and its volatility. The modelling parameters are estimated using the Markov Chain Monte Carlo (MCMC) technique for the Australian electricity market. We find that incorporating stochastic volatility and jumps in both the underlying electricity spot price and its volatility is absolutely essential to accurately fit the observed electricity spot prices. We derive futures prices in a semi-closed form and confirm flexibility of the proposed models by their ability to fit the observed spot and futures prices in the Australian electricity market.
Bayesian Factor Models for Probabilistic Cause of Death Assessment with Verbal Autopsies

Tsuyoshi Kunihama
Kwansei Gakuin University

The distribution of deaths by cause provides crucial information for public health planning, response, and evaluation. About 60% of deaths globally are not registered or given a cause, limiting our ability to understand disease epidemiology. Verbal autopsy (VA) surveys are increasingly used in such settings to collect information on the signs, symptoms, and medical history of people who have recently died. This article develops a novel Bayesian method for estimation of population distributions of deaths by cause using verbal autopsy data. The proposed approach is based on a multivariate probit model where associations among items in questionnaires are flexibly induced by latent factors. Using the Population Health Metrics Research Consortium labeled data that include both VA and medically certified causes of death, we assess performance of the proposed method. Further, we estimate important questionnaire items that are highly associated with causes of death. This framework provides insights that will simplify future data collection.
The Challenges of Analyzing Drug Safety Data with Competing Risk Events and a Bayesian Mixture Model

Aileen Zhu
China Novartis Institutes for Biomedical Research Co., Ltd.

Drug sponsors are often requested to do an investigation of serious safety events, such as cardiovascular events, in clinical trials. However, such an analysis is often hampered by the presence of competing risk events, e.g., non-event related death, that preclude the observation of the event types of interest. The competing risk events are especially of concern when the event rates are not balanced between active and control arms. We were recently requested by a health authority to address such a concern, with the suggested subdistribution proportional hazard model. However, this model only considers patients experiencing competing risk events to remain in the risk set, without investigating how likely these patients could have experienced the events of interest if they had not had the competing risk events. The yielded results being liberal for the arm with more competing risk events, often do not address health authority’s concerns. In this presentation, an alternative approach is proposed. In particular, a mixture model implemented in the Bayesian framework, by considering patients with an event of interest prior to study completion or not in two separate distributions. A gamma distribution is assumed for patients with an event of interest before study completion, and a logistic regression model is fitted to indicate whether patients had an event of interest or not. The estimation is later on used to impute the data for patients who had a competing risk event and who dropped out due to other reasons. The final estimation is then based on the multiply imputed data from the mixture model, which accounts for the uncertainty of whether an event of interest could have occurred for those who dropped out early (due to competing risk events or not).

Bayesian Modeling of Graph Laplacians

Leo Duan
University of Florida

It has become very common to observe multiple graphs at the same time. They can either come from repeated measurement of a cohort (e.g. brain networks related to a certain disease), or from distinct data sources involving the same nodes (e.g. different activities for people on a social network). While there is a rich literature of generative models describing how each edge is formed, the variation across graphs is often astonishingly large, leading to serious concerns about model misspecification. Comparatively, the graph Laplacian has been rigorously studied in graph theory and shows mathematical link to community structure. Motivated to exploit those properties for improved robustness, while allowing uncertainty quantification, we propose a new generative model called the ‘spiked Laplacian graph’. At its core, it models the normalized Laplacian as a noisy transform of a low rank matrix. This parameterization characterizes the strength of connectivity between
and within sub-graphs, without putting restrictive assumption on the edge locations. To infer community membership for each node, one can simply use the signs in the eigenvectors, eliminating the need for any clustering algorithm. For multiple graphs, a Bayesian treatment is taken to allow a hierarchical structure for borrowing information across graphs. With the posterior distribution available in closed-form, one can rapidly quantify the uncertainty using Markov chain Monte Carlo. Theory establishes the trade-off between model resolution and estimation accuracy. A neuroscience application demonstrates the strength of this model and the flexibility to handle a set of graphs from heterogeneous sources.
Robust Bayesian Regression with Shrinkage Priors
Shonosuke Sugasawa
The University of Tokyo

When the number of covariates is large compared with a sample size, shrinkage estimation of coefficients in regression models is known to be useful. However, shrinkage methods under parametric assumptions would be influenced by outliers. While several robust methods have been proposed based on frequentist approaches, Bayesian methods would be more preferable in terms of easiness of uncertainty quantification of the shrinkage estimation as well as reasonable choice of tuning parameters by assigning prior distributions. We develop a robust Bayesian method based on quasi-posterior distribution with shrinkage priors for regression coefficients, and provide an efficient method for posterior computation. The proposed method is demonstrated via numerical studies based on synthetic and real datasets.

The Berry-Esseen Type Bound for the Bernstein-von Mises Theorem in Moderately High Dimensions
Keisuke Yano
The University of Tokyo

We present the Berry-Esseen bound on the Bernstein-von Mises theorem in approximately liner regression models with moderately high dimensions. Specifically, we work with quasi-Bayesian approaches to handle possibly non-Gaussian error terms. We apply our bound to the finite sample evaluation of the frequentist coverage errors of Bayesian credible sets. In particular, our bound implies that Castillo-Nickl credible bands in Gaussian white noise models and linear inverse problems have coverage errors (for the true function) decaying polynomially fast in the sample size. This result shows advantages of Bayesian credible bands over confidence bands based on extreme value theorem. This talk is based on the joint work with Kengo Kato (Cornell University).
Singular Value Shrinkage Priors for Bayesian Prediction
Takeru Matsuda
The University of Tokyo

We develop singular value shrinkage priors for the mean matrix parameters in the matrix-variate normal model with known covariance matrices. Our priors are super-harmonic and put more weight on matrices with smaller singular values. They are a natural generalization of the Stein prior. Bayes estimators and Bayesian predictive densities based on our priors are minimax and dominate those based on the uniform prior in finite samples. In particular, our priors work well when the true value of the parameter has low rank.

The Laplace Approximation to a High Dimensional Model
Yoichi Miyata
Takasaki City University of Economics

For recent years, much attention has been paid on high dimensional models, in which the dimension of a parameter vector grows with the sample size, in the both fields of Bayesian and frequentist statistics. To evaluate each of those models, its marginal likelihood could be one of useful tools. In this talk, we consider a high dimensional linear model in which the number of covariates increase with the sample size, and a prior density with the Laplace distribution having a nondifferentiable point. Then, the Laplace approximation using the LASSO (Least Absolute Shrinkage and Selection Operator) estimator is derived for the marginal likelihood under some suitable conditions. Furthermore, we use this approximation to present a way to implement an importance sampling method effectively.
On Comparing Asset Pricing Models
Lingxiao Zhao
Washington University in St. Louis

We revisit the framework of Barillas and Shanken (2018) (BS henceforth) to point out that the Bayesian marginal likelihood based model comparison method in that paper is unsound. We show that in this asset pricing comparison, the priors on the nuisance parameters across models must satisfy a certain change of variable property for densities, summarized in Proposition 2, that is violated by the off-the-shelf Jeffreys priors used in the BS method. Hence, the BS marginal likelihoods are non-comparable across models and cannot be used to locate the risk factors. We conduct extensive simulation exercises in two designs: one with 8 potential pricing factors and a second with 12 factors, in each case matching the factors to real world factors that arise in this setting. The empirical performance of the BS method is shown to be dismal, even when epic (and practically unattainable) sample sizes of .12 and 1.2 million are used to conduct the model comparisons. In a notable advance with many practical ramifications, we derive a new class of improper priors on the nuisance parameters that satisfy our Proposition 2, leading to valid marginal likelihoods, and valid model comparisons. The empirical performance of our new marginal likelihoods is stunningly different. Our method open doors to a new exciting wave of reliable Bayesian work on the comparison of asset pricing models.
A New Monte Carlo Method for Estimating Marginal Likelihoods

Lynn Kuo
University of Connecticut

Evaluating the marginal likelihood in Bayesian analysis is essential for model selection. There are existing estimators based on a single Markov chain Monte Carlo sample from the posterior distribution, including the harmonic mean estimator and the inflated density ratio estimator. We propose a new class of Monte Carlo estimators based on this single Markov chain Monte Carlo sample. This class can be thought of as a generalization of the harmonic mean and inflated density ratio estimators using a partition weighted kernel (likelihood times prior). We also show that our estimator is consistent and has better theoretical properties than the harmonic mean and inflated density ratio estimators. In addition, we provide guidelines on choosing the optimal weights. A simulation study is conducted to examine the empirical performance of the proposed estimator. We further demonstrate the desirable features of the proposed estimator with two real data sets: one is from a prostate cancer study using an ordinal probit regression model with latent variables; the other is for the power prior construction from two Eastern Cooperative Oncology Group phase III clinical trials using the cure rate survival model with similar objectives.

Inflated Density Ratio and Its Variation and Generalization for Computing Marginal Likelihoods

Yu-Bo Wang
Clemson University

In the Bayesian framework, the marginal likelihood plays an important role in variable selection and model comparison. The marginal likelihood is the marginal density of the data after integrating out the parameters over the parameter space. However, this quantity is often analytically intractable due to the complexity of the model. In this paper, we first examine the properties of the inflated density ratio (IDR) method, which is a Monte Carlo method for computing the marginal likelihood using a single MC or Markov chain Monte Carlo (MCMC) sample. We then develop a variation of the IDR estimator, called the dimension reduced inflated density ratio (Dr.IDR) estimator. We further propose a more general identity and then obtain a general dimension reduced (GDr) estimator. Simulation studies are conducted to examine empirical performance of the IDR estimator as well as the Dr.IDR and GDr estimators. We further demonstrate the usefulness of the GDr estimator for computing the normalizing constants in a case study on the inequality-constrained analysis of variance.
Monte Carlo Methods for Computing Marginal Likelihoods with Applications to Item Response Theory Models

Ming-Hui Chen
University of Connecticut

Nowadays, Bayesian methods are routinely used for estimating parameters of item response theory (IRT) models. However, the marginal likelihoods are still rarely used for comparing IRT models due to their complexity and a relatively high dimension of the model parameters. In this paper, we review Monte Carlo (MC) methods developed in the literature in recent years and provide a detailed development of on how these methods are applied to the IRT models. In particular, we focus on the “best possible” implementation of these MC methods for the IRT models. These MC methods are used to compute the marginal likelihoods under the one-parameter IRT model with the logistic link (1PL model) and the two-parameter logistic IRT model (2PL model) for a real English Examination dataset. We further use the widely applicable information criterion (WAIC) and deviance information criterion (DIC) to compare the 1PL model and the 2PL model. The 2PL model is favored by all of these three Bayesian model comparison criteria for the English Examination data.
Learning Semiparametric Regression with Missing Covariates Using Gaussian Process Models

Dipak K. Dey
University of Connecticut

Missing data often appear as a practical problem while applying classical models in the statistical analysis. Here, we consider a semiparametric regression model in the presence of missing covariates for nonparametric components under a Bayesian framework. As it is known that Gaussian processes are a popular tool in nonparametric regression because of their flexibility and the fact that much of the ensuing computation is parametric Gaussian computation. However, in the absence of covariates, the most frequently used covariance functions of a Gaussian process will not be well defined. We propose an imputation method to solve this issue and perform our analysis using Bayesian inference, where we specify the objective priors on the parameters of Gaussian process models. Several simulations are conducted to illustrate effectiveness of our proposed method and further, our method is exemplified via two real datasets, one through Langmuir equation, commonly used in pharmacokinetic models, and another on Auto-mpg data.
Measuring Cross-country Interconnectedness with a Panel Unobservable Component Model

Gianni Amisano
Federal Reserve Board

We use a panel unobserved component model for output and unemployment to study the interconnectedness among several countries, along different dimensions. In particular, our model allows for commonalities in cyclical and long-term movements. Using a time varying parameter version of the model, we are also able to analyze the evolution of cross-country linkages. A further extension of the model allows us to examine whether commonalities can be modelled with correlation across idiosyncratic components. The model and its extensions are given a network interpretation.

Disjunct Support Prior for Practically Significant Variable Selection in Regression

Daniel Andrade *
SOKENDAI (and NEC)

Sparseness of the regression coefficient vector is often a desirable property, since it helps to improve interpretability. However, in practice, we may have to trade in a small reduction in prediction accuracy for an increase in sparseness. For that purpose, we consider here approaches that ignore negligible small, but non-zero regression coefficients. Spike-and-slab priors, as proposed by Chipman et al., 2001, can potentially handle such a trade-off between prediction accuracy and sparseness. Given a threshold parameter delta on the coefficients, they propose to adjust the variance of the spike and slab prior such that coefficients with absolute value smaller than delta have more probability mass on the spike prior, and vice versa. However, for large enough sample sizes, the Bayes Factors of any two models converges to 1. This is clearly an undesirable property for Bayesian hypotheses testing. We propose to resolve this inconsistency issue by using disjunct support priors on the regression coefficients. We prove the consistency of our proposed method, and experimentally confirm our results on several data sets.
Quantile Nonlinear Effects of Return and Abnormal Trading Volume for the Realized Kernel

Dong Manh Cuong *
Feng Chia University

Realized volatility is a new rising concept in modern finance, with one popular approach for measuring daily volatility being the realized kernel (RK) estimator of Barndorff-Nielsen et al. (2008), which is robust to microstructure noise and jumps. Our research examines the nonlinear responses of stock markets’ RK to its potential factors across different quantile levels. Using the threshold quantile regression model with GARCH specification, we investigate the relationship between RK and its lagged one autoregressive effect of the RK, open-to-closed returns, and abnormal trading volume. Applying a Bayesian adaptive sampling scheme via the Markov chain Monte Carlo method, our results confirm the existences of stock market volatility clustering, a negative and asymmetric leverage effect, and negative and asymmetric impacts of trading volume to market volatility. These results provide an in-depth understanding about how stock volatility reacts to its determinants and how stock markets operate in different market conditions.

Bayesian Inferences of Multiple Structural Change GARCH Model with Skew Student t Errors

Bonny Y.F. Lee *
Feng Chia University

This research considers a piecewise autoregressive GARCH model with exogenous variables and skew Student t errors, which we call this model a segmented ARX-GARCH model and use it to make inferences about all unknown parameters and to identify the location of structural breaks. It fills the gap in existing literature to cover skew Student t errors. Compared with other distributions, the fat-tailed skew Student t distribution performs well at describing financial time-series datasets in financial markets. We employ the segmented ARX-GARCH model with skew Student t errors, and estimate the model parameters via Bayesian inference, in order to show the validity and reliability of the Bayesian methods. We then utilize the adaptive Metropolis-Hastings (MH) MCMC algorithm, which combines with the random walk MH algorithm and the independent kernel MH algorithm to accelerate convergence. Our goal is to know how many breakpoints and the location of the breakpoints. We first assume the number of breakpoints is prefixed and employ deviance information criterion to decide the optimal number of breakpoints. We also extend the segmented GARCH model to an asymmetric GARCH one. As an illustration, we provide a simulation study to examine the credibility of our MCMC sampling scheme. For real data analysis, we examine the impact of daily crude oil returns and gold returns on stock S&P 500 returns during 2007 to 2018.
Bayesian Spatiotemporal Varying Coefficients Models for Ischemic Stroke Study
Kuo-Jung Lee
National Cheng-Kung University

A Bayesian spatiotemporal generalized linear regression with varying-coefficient model is proposed to examine geographic variation of the medical prescription use for ischemic stroke and to have a clearer understanding of association of the relevant risk factors such as comorbidities, medication and environmental society with mortality rate in ischemic stroke. The spatial heterogeneity of coefficients for important factors that may affect the mortality in ischemic stroke across 349 townships in Taiwan. By applying spatial-temporal models, we can understand the spatial variation in risk to ischemic stroke. It then turns out that, based on the findings, we can properly arrange medical resources and reduce the life-threatening damage caused by the uneven distribution of medical resources. The data was collected from a retrospective cohort study using 2004-2012 National health insurance research database.

Objective Priors in Robust Quasi-Bayesian Inference Using the Divergences
Tomoyuki Nakagawa
Tokyo University of Science

In Bayesian analysis, it is well known that ordinary Bayesian estimator is not robust against outliers. Ghosh and Basu (2016) proposed the robust Bayesian estimation against outliers by using the density power divergence. They characterized the robustness in terms of the influence function. However, this estimator does not work well for the estimation of the scale parameter, and unstable when the contamination ratio is not small. On the other hand, in frequentist viewpoint, it was shown that an estimator using the gamma-divergence can make a stable estimation even when the contamination ratio is not small. Therefore, in this presentation, we propose a robust estimation using gamma-divergence. We also construct moment matching priors which is a kind of objective priors for the proposed models, and illustrate their performances in some simulation studies.
Bayesian Analysis of Intraday Stochastic Volatility Models with Skew Heavy-tailed Error and Smoothing Spline Seasonality

Makoto Nakakita *
Keio University

The aim is to extend the stochastic volatility (SV) model for application with intraday high frequency data of asset returns. It is well-known that intraday high frequency data of asset returns exhibit not only stylized characteristics (e.g., volatility clustering, heavy-tailed distribution) but also cyclical fluctuation in return volatility, which is called intraday seasonality. In a typical trading day, the volatility tends to be higher immediately after the opening or near the closing, but it tends to be lower in the middle of the trading hours. Our modeling strategy is two-fold. First, we model the intraday seasonality of return volatility with a B-spline polynomial and estimate it along with the stochastic volatility simultaneously. Second, we incorporate a possibly skew and heavy-tailed error distribution into the SV model by assuming that the error distribution belongs to a family of generalized hyperbolic (GH) distribution such as Laplace, variance-gamma and Students t. We develop an efficient Markov chain Monte Carlo (MCMC) sampling algorithm for Bayesian inference of the proposed model and apply it to intraday data of Japanese stocks.

Bayesian and Decomposition Analysis for Health Inequality in Japan

Haruhisa Nishino
Hiroshima University

Japan is now confronting a severely ageing population. Health conditions are attracting more concern, particularly among aged people. Then health inequality in Japan becomes an important topic as well as income equality. Self-reported health data in Japan enable us to analyse health inequality. The data are ordinal categorical data, that is, qualitative data. We calculate the inequality index in Japan based on ordinal categorical health data and decompose the health inequality by age groups and sexes using the Bayesian method. This poster presentation shows the changes and decompositions in the health inequality in Japan.
Information Geometry of Complex Autoregressive Models and Its Positive Superharmonic Priors

Hidemasa Oda *
University of Tokyo

Complex-valued stochastic processes are useful models for parameterizing complex or bivariate signals. We expanded the theory of $\alpha$-geometry of real time series, introduced by Amari in 1987, for complex time series. For most of the parts, generalized results of real time series for complex time series are obtained. We also investigated the difference of the estimations between real and complex time series. It is natural to expect the statistical manifold to be a complex manifold since it has a richer structure. In particular, we are interested in the case when it is a Kähler manifold. Choi and Mullhaupt showed that the information geometry of complex time series is Kähler in 2015. We further investigated the structure of the complex autoregressive models and its positive superharmonic priors. We expect that Ricci-free $\alpha$-Kähler structure plays an important role in the theory of complex information geometry. We will discuss the application of $\alpha$-Kähler geometry for general complex linear systems.

---

Novel Conjugate Analysis in Unknown Dimensional Multinomial Probabilities

Toru Ogura
Mie University

The Dirichlet-multinomial distribution is known as finite discrete multivariate probability distribution. Ferguson (1973) extended it to the infinite discrete multivariate probability distribution using Bayesian method. By using flexible prior density, Bayesian methods enable various extensions. We discuss an unknown dimensional Dirichlet-multinomial distribution. This means although the dimension is finite, but the number of dimension is unknown. We estimate the probability of each discrete from the obtained samples. In the study of such estimation, the maximum likelihood method is often used, but when the number of discrete samples is zero, the probability of discrete is estimated to be zero. To address this problem, we propose an estimation of the probability using the posterior mean of the canonical parameter. Ogura and Yanagimoto (2018) showed that it is better to use the posterior mean of the canonical parameter than the posterior mean of the mean parameter for simultaneous estimation in the Poisson distribution. The reason is that there was a large difference in the estimation when the observed value was zero. Since it is considered that there are some observed as zero in the Dirichlet-multinomial distribution, we used the posterior mean of the canonical parameter. We made numerical comparisons in various settings to confirm the effectiveness.
A More Stable and Scalable Posterior Computation for Bayesian Graphical Models

Sakae Oya *
Keio University

Wang (2012) proposed a novel approach to posterior computation for a Bayesian graphical LASSO model. In this paper, we propose modifications to Wang (2012)’s approach so that the precision matrix would never fail to be positive definite. The positive definiteness is assured by sampling off-diagonal elements of the precision matrix from the domain where the precision matrix is positive definite. Furthermore, using matrix partition slightly different from Wang (2012)’s approach, our proposed algorithm generates each of the off-diagonal elements exactly once in a Gibbs sampling cycle and reduces the cost of random number generation. With these modifications, our proposed algorithm is more stable in the sense that the sampling procedure will not halt due to numerical exceptions related to lack of positive definiteness. This numerical stability contributes to the algorithms scalability. In our experience, it can be applicable to high-dimensional graphical models with thousands of nodes. Furthermore, it is straightforward to extend our proposed algorithm to a graphical model with the horseshoe shrinkage prior. In the rest of the paper, we demonstrate the efficacy of our proposed algorithm in simulation experiments with various graphical structures.

The Hedonic Regression with Bayesian Hierarchical Model for Japanese Rice Wine. What Affects Price?

Wakuo Saito *
Keio University

In recent days, the amount of exported sake have been increasing, and sake comes to get prestige as Japanese Rice Wine. Thus, people expect the development of sake with various commodities. There are no econometric approach, however, for the sake pricing. We developed a hedonic price function for Japanese Sake with hierarchical Bayesian model. From the model estimation, we found that quality indices, such as polished rice ratio, sake degree, and acidity, have high effects on sake price. In addition, there are some differences in material and regional indices.
Hierarchical Factorial Structure of WISC-4 Data

Kazuo Shigematsu
Keio University

It has been a big issue in psychology of intelligence whether there exists g-factor. To address this problem, the Bayesian approach was employed to analyze the data gathered for standardization of Japanese version of WISC-4. The g-factor can be interpreted to be a second order factor or one of the factors which directly influences all observed variables. This research compared a number of possible intelligence models (e.g. CHC model) in terms of WAIC and LOO-CV. Also, the theoretical problem associated with the model evaluation of the latent variables models.

Modeling Financial Durations with Limit Order Book Information

Tomoki Toyabe *
Keio University

It is a stylized fact that durations between executions in financial markets have intraday seasonality and autocorrelation. The Autoregressive Conditional Duration (ACD) model has been widely used to capture these characteristics. However, durations are also supposed to be affected by liquidity in the market. We propose a new ACD model that utilizes the limit order book information for reflecting the liquidity. In our empirical analysis, we applied the proposed ACD model to high-frequency stock price data in the Tokyo Stock Exchange and estimated it with an efficient Markov chain Monte Carlo method. We also conducted model comparison among different specifications of the proposed model.

Modeling Temporal Data by Using Mixture for Marked Poisson Processes

Toshiya Yoshida *
Keio University

This paper reviews the performance of the method of modeling temporal data proposed by Taddy and Kottas (2012). This method utilizes nonparametric Dirichlet mixture modeling for the simultaneous estimation of the density function and the intensity function. We are then able to estimate the conditional expectation for the marks in a direct and unequivocal manner. In addition to modeling the temporal data, we will assess the goodness of fit by creating Q-Q plots. Among the most remarkable is the coherence of the posterior simulation, the inference about the functionals of marked Poisson processes, and the model checking. A demonstration of this method using real data sets will also be described.
Bayesian Learning in High-dimensional State-space Models

Hedibert Freitas Lopes
INSPER Institute of Education and Research

Applied Bayesian Statistics has benefited greatly from the avalanche of Monte Carlo-based tools for approximate posterior inference in highly complex and structured scientific models. From environmental and health studies to financial applications, virtually all areas of science where evidence-based scrutiny is mandatory to validate scientific hypotheses have benefitted from such technological explosion.

My research also exemplifies these trends. I will discuss work on high-dimensional state-space models with particular attention to time-varying covariance learning. In one direction, my co-authors and I deal with the curse of dimensionality via parameter reduction, such as those found in the factor modelling literature. In another direction, we heavily regularize the estimation of parameters. I will review the challenges we faced when dealing with such high-dimensional state-space models. Two or three motivating examples from my recent research will be used throughout the talk. I finish with a few directions of my future research in these and related areas.
A Semiparametric Mixture Method for Local False Discovery Rate Estimation from Multiple Studies

Woncheol Jang
Seoul National University

We propose a two-component semiparametric mixture model to estimate local false discovery rates in multiple testing problems. The two pillars of the proposed approach are Efron’s empirical null principle and log-concave density estimation for the alternative distribution. Our method outperforms other existing methods, in particular when the proportion of null is not that high. It is robust against the misspecification of alternative distribution. A unique feature of our method is that it can be extended to compute the local false discovery rates by combining multiple lists of \( p \)-values. We demonstrate the strengths of the proposed method by simulation and several case studies.

Amortized Variational Inference Methods for Bayesian Model Criticism

Seongil Jo
Chonbuk National University

Implementing common methods for Bayesian model criticism can be computationally demanding in some applications. Here we consider applying recently developed amortized variational inference methods to this task in the context of certain latent variable models, such as random effects models. In amortized variational inference, local variational parameters determining posterior approximations for the latent variables are parametrized more parsimoniously as a function of global variational parameters and local data. The reduction of the number of variational parameters that amortization brings leads to very fast algorithms for fitting computationally expensive models. The power of these algorithms is yet to be exploited in applications to Bayesian model criticism. As an application, we consider the detection of outlying clusters in generalized linear mixed models using an approach where repeated computations of posterior distributions for random effects is needed.
Analysis of Longitudinal Binary and Survival Time Data Using Joint Models with General Random Effects Covariance Matrix

Keumbaik Lee
Sungkyunkwan University

In clinical trials, longitudinal data are collected repeated over time. However, these data have a tendency to show many missing values because of censoring problems, and these missing values are not negligible. In order to address this problem, this paper deals with joint models of longitudinal data and survival data. In the process, generalized linear mixed model (GLMM) is used as longitudinal sub-model and survival sub-model is Cox proportional hazard model. Moreover, random effects are used to account for association between the longitudinal and survival models. However, estimation of the random effects covariance matrix is not easy because the matrix is high-dimensional, and the matrix should be positive definite. In this regard, modified Cholesky decomposition (MCD) is utilized to overcome the limitation of the random effects covariance matrix estimation. In real data study, Markov Chain Monte Carlo algorithm is used to verify parameters. we employ random-walk Metropolis Hastings algorithm to estimate parameters and random effects. CD4 cell count data, which is related to HIV, is analyzed.

Post-processed Posteriors for Band Structured Covariances

Jaeyong Lee
Seoul National University

We consider two classes of band structured covariances, classes of banded and bandable covariances. Due to the difficulty of constructing priors with computational efficiency and theoretical optimality, Bayesian inference for band structured covariances remains elusive. In this paper, we propose post-processed posteriors for the banded and bandable covariances. The post-processed posterior is obtained by post-processing the conjugate inverse-Wishart posterior for the covariance without any structural restriction. The structural restriction of the posterior is satisfied by the post-processing. We show that the proposed post-processed posteriors have optimal minimax rate for the bandable covariances and nearly optimal minimax rate for banded covariances. A simulation study and a real data analysis are given.
Use of Two Non-informative Priors in an Empirical Bayes Estimator of Multiple Poisson Means

Takemi Yanagimoto and Toru Ogura

1 The Institute of Statistical Mathematics; 2 Mie University

The simultaneous estimation problem of the multiple Poisson means arises often in various application fields including epidemiological studies. It looks obvious that the empirical Bayesian approach is promising. However, there exist many to be done to construct an explicit form of this estimator. Let $x$ be a sample of size 1 from the Poisson density $\text{Po}(\lambda)$. Our preliminary studies indicate that the posterior mean of the canonical parameter $\theta = \log \lambda$ under the reference prior performs better than the maximum likelihood estimator (MLE) and also than the traditional Bayesian estimator. In order to develop this observation further, we attempt to construct the empirical Bayes method for estimating $K$ ($\geq 2$) means in independent Poisson densities $\text{Po}(\lambda_k)$’s. A family of conjugate prior densities of the form

\[ \mathcal{P} = \{ \pi(\lambda; m, \delta) = \exp\{-\delta D(m, \lambda)\}b(\lambda)k(m, \delta) | \delta \geq 0 \} \tag{1} \]

is assumed, where $D(\cdot, \cdot)$ denotes the Kullback-Leibler divergence. Our primary interest is in the choice of the supporting density of the form

\[ b(\lambda) = \lambda^{a-1}. \tag{2} \]

Specifically, the cases of $a = 0$ and 0.5 in (2) attract our attention. When $\delta = 0$, they are familiar non-informative prior densities for $\theta$ as the uniform and the reference prior densities, respectively. As in the usual empirical Bayes method, the hyperparameters $m$ and $\delta$ are determined by maximizing the marginal likelihood, and the mean parameters $\lambda_k$’s are estimated through the posterior means of $\theta_k$’s under the given determined hyperparameters. We conduct numerical and simulation comparison studies of selected empirical Bayes estimators, including the use of two different priors for estimating the hyperparameters and the multiple Poisson means. The studies indicate that the empirical Bayes estimator under the uniform prior for
determining the hyperparameters and then under the reference prior for estimating the multiple means shows good performance. This finding is very interesting, since the use of two priors in an empirical Bayes method can be promising. It is our understanding that the empirical Bayes method is a different discipline from the original Bayesian theory. Recall that the introduction of an unknown hyperparameter can be subject to criticism. Thus the use of two priors in an empirical Bayes method is not to be avoided. On the other hand, it looks tough to justify the use of two priors. In this concern we discuss here two points regarding the possible justification. One pertains to the dual structure observed in the proposed empirical Bayes method. In fact, the marginal likelihood is obtained by the $m$-mixture of the sampling densities under a prior density, while the posterior mean of the canonical parameter $\theta$ relates with the $e$-mixture of the sampling densities under the posterior density. This difference gives an interpretation on our use of two priors. Note that the DIC by Spiegelhalter et al. (2002, 2014) is based on the $e$-mixture of sampling densities, but it does not look that the direct application it to the present problem is successful. The other concerns the relation of Bayesian estimators with the MLE. When $\delta = 0$, it is widely known that the posterior mean of the mean parameter under the uniform prior density coincides with the MLE. On the other hand, the posterior mean of the canonical parameter under the reference prior density is shown to be asymptotically equivalent with the MLE up to $O(n^{-2})$.

---

**Bayesian Methods for Accelerated Materials Discovery**

Ryo Yoshida

The Institute of Statistical Mathematics

The ability of machine learning (ML) models, which are trained on massive amounts of data, has reached or even outperformed humans in intellectually demanding tasks across various fields. As such, ML has received much attention as a key driver to the next frontier of materials science, which can contribute to substantial savings, in terms of both time and costs that are involved in the development of new materials. In this talk, I will describe some key drivers of ML technologies to achieve this goal.

The first topic is the ML-assisted materials design. In general, the material spaces are considerably high-dimensional. For instance, the chemical space of small organic molecules is known to contain as many as $10^{60}$ candidates. The problem entails a considerably complicated combinatorial optimization where it is impractical to fully explore the vast landscape of structure-property relationships. The aim of this study is to create a novel material design method by the integration of machine learning and quantum chemistry calculation. The method begins by obtaining a set of machine learning models to forwardly predict properties of input material structures for multiple design objectives. These models are inverted to the backward model through Bayes law. Then we have a posterior probability distribution which is conditioned by desired properties. Exploring high probability regions of the posterior, it is expected to identify new materials possessing the desired target properties. The emergence of such ML algorithms to exhaustively search in such a
huge space is expected to accelerate the pace of expanding the frontier in the vast universe of materials. Under industry-academia partnerships, we are putting into practice this Bayesian material design method.

The second topic is a subject of data scarcity. There are growing needs for the use of ML to derive fast-to-evaluate surrogate models for materials properties. In recent years, a broad array of materials property databases has begun to emerge towards digital transformation of materials science. However, the volume and diversity of materials data remain far from fully enjoying technological advances recently made in ML. A ML framework called transfer learning has the great potential to break the barrier of limited amounts of materials data in which various kinds of properties are physically interrelated. For a given target property to be predicted from a limited supply of training data, models on related proxy properties are pre-trained using an enough amount of data, which capture common features relevant to the target task. Re-purposing such machine-acquired features on the target task brings an outstanding achievement in the prediction performance even with exceedingly small data as if highly experienced human experts can perform rational inferences even on considerably less experienced tasks. To facilitate the widespread use of transfer learning, we have developed a pre-trained model library, called XenonPy.MDL. In this first release, the library constitutes more than 100,000 pre-trained models for various properties of small molecules, polymers, and inorganic solid-state materials. Along with this library, I will demonstrate some outstanding successful applications of transfer learning.

References
3. iQSPR: https://github.com/yoshida-lab/iqspr

High-dimensional Motion Segmentation with Semi-Markov Latent Gaussian Processes

Daichi Mochihashi
The Institute of Statistical Mathematics

Humans perceive continuous high-dimensional time series by dividing it into meaningful segments such as words and units of motion. We believe that such unsupervised segmentation is also important for robots to learn topics such as language and motion. To this end, we previously proposed a hierarchical Dirichlet process-Gaussian process hidden semi-Markov model (HDP-GP-HSMM). However, an important drawback to this model is that it cannot divide high-dimensional time-series data and low-dimensional features must be extracted in advance. Segmentation largely depends on the design of features, and it is difficult to design effective features, especially in the case of high-dimensional data. To overcome this problem, we propose a hierarchical Dirichlet process variational autoencoder Gaussian process hidden semi-Markov model (HVGH). The parameters of the proposed HVGH
are estimated through a mutual learning loop of the variational autoencoder and our previously proposed HDP-GP-HSMM. Hence, HVGH can extract features from high-dimensional time-series data, while simultaneously dividing it into segments in an unsupervised fashion. In an experiment, we used various motion-capture data to show that our proposed model estimates the correct number of classes and more accurate segments than baseline methods. Moreover, we show that the proposed method can learn latent compressed space suitable for segmentation. This is a joint work with Masatoshi Nagano, Tomoaki Nakamura (The University of Electro-Communications), Takayuki Nagai, Wataru Takano (Osaka University), and Ichiro Kobayashi (Ochanomizu University).
Quantile Forecasting Based on a Multivariate Hysteretic Autoregressive Model with GARCH Errors and Time-varying Correlations

Cathy W.S. Chen
Feng Chia University

To understand and predict chronological dependence in the second-order moments of asset returns, we consider a multivariate hysteretic autoregressive (HAR) model with GARCH specification and time-varying correlations, by providing a new way to describe a nonlinear dynamic structure of the target time series. The hysteresis variable governs the nonlinear dynamics of the proposed model in which the regime switch can be delayed if the hysteresis variable lies in a hysteresis zone. The proposed model combines on three useful model components for modeling economic or financial data: (1) the multivariate HAR model, (2) the asymmetric multivariate volatility models, and (3) a dynamic conditional correlation structure. We incorporate an adapted multivariate Student-t innovation based on a scale mixture normal presentation in the HAR model to tolerate for dependence and different shaped innovation components. We carry out multivariate volatilities, Value-at-Risk, and marginal expected shortfall based on a Bayesian sampling scheme through adaptive Markov chain Monte Carlo (MCMC) methods, which allow us to statistically estimate all unknown model parameters and forecasts simultaneously. We illustrate the proposed methods herein by using both simulated and real examples and measure for industry downside tail risk jointly.
Bayesian Analysis of Realized Matrix-Exponential GARCH Models
Manabu Asai
Soka University

The paper develops a new realized matrix-exponential GARCH (MEGARCH) model, which uses the information of returns and realized measure of co-volatility matrix simultaneously. The paper also considers an alternative multivariate asymmetric function to develop news impact curves. We consider Bayesian MCMC estimation to allow non-normal posterior distributions, and illustrate the usefulness of the algorithm with numerical simulations for two assets. For three US financial assets, we compare the realized MEGARCH models with existing multivariate GARCH class models. The empirical results indicate that the realized MEGARCH models outperform the other models regarding in-sample and out-of-sample performance. The news impact curves based on the posterior densities provide reasonable results.

A Generalized Threshold Stochastic Volatility Model Incorporating with Realized Measures
Feng-Chi Liu
Feng Chia University

A generalized threshold stochastic volatility (THSV) model jointly modeled with realized measures (RMs) is considered in this study. Advancing from trading technology, high frequency data is used to construct RMs for accurate volatility forecasting. This study proposes a threshold-type RM equation to jointly model with a generalized THSV model. Based on the Bayesian approach, parameters of the proposed model are estimated by the designed Markov chain Monte Carlo method. In the real data analysis, we employ the risk measures of value-at-risk (VaR) and expected shortfall (ES) to evaluate the performance of volatility forecasting. The results show that the proposed model can produce more accurate volatility forecasting than the model with a simple RM equation.
Nonparametric Bayesian Two-part Random Effects Model for Longitudinal Semi-continuous Data Analysis

Yeongseung Chung
Korea Advanced Institute of Science and Technology

Longitudinal semi-continuous data are frequently encountered in many fields of research. Two-part random effects model (TP-REM) has been widely used to analyze the semi-continuous data and covariates. The TP-REM separates the semi-continuous data into the binary and continuous parts, and fits random effects regression to each part separately. The existing TP-REM typically assumes that the random effects follow a normal distribution, which is limited to flexibly characterize the between-subject heterogeneity and identify the underlying subpopulation structure. Moreover, the existing model incorporates covariates in parametric manners, and time-varying covariates in a limited way that their effects do not change over time. Motivated by these limitations, we propose a novel nonparametric Bayesian two-part random effects model for analyzing longitudinal semi-continuous data and various types of covariates. In the proposed model, we depict the longitudinal trajectories of the binary and continuous parts separately using splines with subject-specific basis coefficients (i.e., random effects). Then, the random effects from both parts are jointly modeled using a Dirichlet process mixture of normals to account for the between-part correlations and identify the subpopulation structure. Furthermore, the random effects are associated with different types of covariates in a nonparametric manner, so that complex relationships between longitudinal semi-continuous data and time-fixed or time-varying covariates are modeled. The method is illustrated through a simulation study and is applied to the social insurance expenditure data obtained from the KWEPS study.

Coauthor: Jinsu Park, Korea Advanced Institute of Science and Technology
Use of Model Reparametrization to Improve Variational Bayes

Linda S. L. Tan
National University of Singapore

We propose using model reparametrization to improve variational Bayes inference for a class of models whose variables can be classified as global (common across observations) or local (observation specific). Posterior dependency between local and global variables is reduced by applying an invertible affine transformation on the local variables. The functional form of this transformation is deduced by approximating the posterior distribution of each local variable conditional on the global variables by a Gaussian distribution via a second order Taylor expansion. Variational Bayes inference for the reparametrized model is then obtained using stochastic approximation. Our approach can be readily extended to large datasets via a divide and recombine strategy. Using generalized linear mixed models, we demonstrate that reparametrized variational Bayes provides improvements in both accuracy and convergence rate compared to state of the art Gaussian variational approximation methods.

Nonparametric Bayesian Modeling in Government Statistics: Recent Developments in Imputation, Editing, and Data Protection

Hang J. Kim
University of Cincinnati

Federal statistical agencies are demanded to disseminate accurate data to the public while information provided by survey respondents often suffer from a large amount of incorrect and missing responses. Increasing concern about disclosing identities of individuals and their sensitive data is a relatively new assignment imposed on the statistical agencies. In the last five years, nonparametric Bayesian modeling has been suggested as a solution to address the multiple, competing goals in data dissemination. In this talk, we will review recent developments in nonparametric Bayesian modeling for government statistics, which is able to (a) capture a flexible joint distribution for the underlying true values of the input data, (b) account for uncertainty introduced during the data dissemination process, and (c) generate synthetic data for privacy protection. The talk will also introduce several potential applications and future research topics including differential privacy and the impact of sampling design information on data protection.
Estimation and Prediction for Spatial Generalized Linear Mixed Models
Vivekananda Roy
Iowa State University

Spatial generalized linear mixed models (SGLMMs) are popular for analyzing non-Gaussian spatial data. Some popular link functions, such as the Box-Cox, used in SGLMMs are inappropriate because they are inconsistent with the Gaussian assumption of the spatial field. We present sensible choices of parametric link functions which possess desirable properties. It is important to estimate the parameters of the link function and spatial covariance, rather than assume known values. To that end, we present efficient generalized importance sampling estimators based on multiple Markov chains and suitable reparameterizations for empirical Bayes analysis of SGLMMs. We develop a methodology for selecting models with appropriate link function family, which extends to choosing a spatial correlation function as well. We present an ensemble prediction of the mean response by appropriately weighting the estimates from different models. The proposed methodology is illustrated using simulated and real data examples.
A Bayesian Approach for Spatial Cluster Detection of Regression Coefficients

Huiyan Sang
Texas A&M University

In this work, we propose a new Bayesian spatially clustered coefficient (BSCC) regression model, to detect spatial clustering patterns in the associations between response variables and covariates. In BSCC, regression coefficients are assumed to be constants within each spatially contiguous cluster. To model the clustering patterns, we develop a novel and flexible space partitioning prior based on Euclidean spanning trees, which is capable of capturing irregularly shaped clusters. An efficient Reversible Jump Markov chain Monte Carlo (MCMC) algorithm is designed to estimate the clustered coefficient values and their uncertainty measures. Finally, we illustrate the performance of the model with simulation studies and a real data analysis of temperature-salinity relationship in the Atlantic Ocean.

A Bayesian Approach to Assess Intervention Effects on Opiate Overdose Incidents in Space and Time

Richard Zehang Li
Yale University

Opioid use and overdose have become an important public health issues in the United States. However, understanding the spatial and temporal dynamics of opioid overdose incidents and effects of public health interventions and policy changes can be challenging. Effects may be heterogeneous across space and time, and may exhibit spillover into regions in which the intervention did not take place. Using a publicly available dataset consisting of the time/date, location, and nature of heroin-related emergency calls in the city of Cincinnati, Ohio, we propose a Bayesian hierarchical model to characterize and predict the risks of overdose incidents in small areas over time, incorporating geographic, social, and demographic covariates. We characterize the predictive performance of this model, and outline a framework for estimating causal impacts of public health interventions in spatial-temporal settings. We discuss assumptions and interpretations for the causal effect estimated using the proposed mixed-effect model approach.
Bayesian Generalized Regression Models with Gaussian Process Priors

Xia Wang
University of Cincinnati

In many scientific fields, it is a common practice to collect a sequence of categorical responses, binary or ordinal, across time, space, or along with a collection of covariates. Researchers are interested in finding out how the expected outcome is related to covariates, and aim at better prediction in the future categorical outcomes. It is critical to appropriately modeling three important components in the generalized regression model, including the link function, the systematic component and the random effects. We propose a flexible generalized regression model with families of parametric link functions and a Gaussian process prior on the latent structure. Commonly adopted link functions such as probit or logit links have fixed skewness and lack the flexibility to allow the data to determine the degree of the skewness. Gaussian processes have been widely used to model nonlinear systems or dependence structure. Extensive simulation studies and real data applications show that the combination of the parametric link functions and the non-parametric Gaussian process leads to a family of very flexible generalized regression models. Bayesian computation is employed in model estimation. Posterior consistency of the resulting posterior distribution is demonstrated in the binary response case.
[Session 5c] Recent Advances in Bayesian Predictive Inference

Organizer
Yuzo Maruyama, University of Tokyo

Session Chair
Yuzo Maruyama, University of Tokyo

Bayesian predictive distributions for Poisson and negative binomial models when the parameter spaces are restricted

Yasuyuki Hamura
University of Tokyo

Predictive probability estimation for Poisson and negative binomial distributions is addressed when the parameter spaces are restricted. Bayesian predictive probabilities against priors on the restricted spaces are compared with the corresponding non-restricted Bayesian predictive probabilities. It is shown that the former dominate the latter under some conditions when the predictive probabilities are evaluated by the risk function relative to the Kullback-Leibler divergence. This result is proved by first showing the corresponding dominance results for estimating the restricted parameters and then translating them into the framework of predictive probability estimation. In particular, for the negative binomial case, an identity that relates Bayesian predictive probability estimation to Bayesian point estimation is derived and utilized. Such identities are known in the cases of normal and Poisson distributions and the identity given here is an extension of the result to the negative binomial case.
Asymptotic Properties of Bayes Estimators Based on Shrinkage Priors for Curved Exponential Families

Michiko Okudo
University of Tokyo

We investigate asymptotic properties of plug-in distributions with Bayes estimators for curved exponential families. We consider orthogonal shift of plug-in distributions in direction orthogonal to the curved exponential family and parallel to the full exponential family. It is shown that plug-in distributions with asymptotically efficient estimators can be improved in terms of Kullback-Leibler risk by shifting them in those directions. It is also shown that the optimal shift coincides with the shift from the plug-in distribution with Bayes estimator for the curved exponential family to the plug-in distribution with Bayes estimator for the full exponential family. We also consider priors for curved exponential families and show the condition where shrinkage priors dominate Jeffreys prior.

Dual Roles of Maximizing Likelihood and Shannon Entropy under Alpha-divergence Loss

Toshio Ohnishi
Kyushu University

The maximization of the likelihood and that of the Shannon entropy are the most famous principles in statistical inference. This paper reveals notable duality of these two important notions under alpha-divergence loss function in the Bayesian prediction problem. The best prediction is obtained by maximizing the likelihood and the Shannon entropy respectively for alpha in [-1, 1) and for alpha = 1. The minimum prediction in the sense that it derives the worst member of a class of predictors is derived by maximizing the Shannon entropy and the likelihood respectively for alpha = -1 and for alpha in (-1, 1].
[Session 6a] Monte Carlo Methods for Complicated Target Distributions

Organizer
Kengo Kamatani, Osaka University

Session Chair
Kengo Kamatani, Osaka University

Analysis of Markov Chain Monte Carlo Method with Heavy-tailed Target Distributions
Kengo Kamatani
Osaka University

In this talk, we will discuss Markov chain Monte Carlo (MCMC) methods with heavy-tailed invariant probability distributions. When the invariant distribution is heavy-tailed the algorithm has difficulty reaching the tail area. We will study the effect by using the high-dimension scaling limit. We also study ergodic properties of some MCMC methods with heavy-tailed invariant distributions.

Bayesian Inference for Intractable Likelihood Models
Krzysztof atuszyski
Warwick University

Constructing MCMC algorithms for Bayesian inference in intractable likelihood models is problematic since difficulties in evaluating the likelihood make the application of the standard Metropolis-Hastings acceptance formula impossible. This enforces either (1) approximate methods that introduce bias of unknown magnitude, or (2) the pseudomarginal approach that is exact, but slows down MCMC convergence, sometimes dramatically. In this talk I will present a new approach (3) based on unbiased estimators of the likelihood and the Barkers acceptance ratio. The approach is exact and retains the per iteration convergence rate comparable to that of the standard Metropolis-Hastings. I will illustrate the approach with examples of exact inference for stochastic differential equations.

This is joint work with Flavio Goncalves, Gareth Roberts and Dootika Vats.
Randomized Hamiltonian Monte Carlo as Scaling Limit of the Bouncy Particle Sampler

Daniel Paulin
Oxford University

The Bouncy Particle Sampler is a Markov chain Monte Carlo method based on a nonreversible piecewise deterministic Markov process. In this scheme, a particle explores the state space of interest by evolving according to a linear dynamics which is altered by bouncing on the hyperplane tangent to the gradient of the negative log-target density at the arrival times of an inhomogeneous Poisson Process (PP) and by randomly perturbing its velocity at the arrival times of an homogeneous PP. Under regularity conditions, we show here that the process corresponding to the first component of the particle and its corresponding velocity converges weakly towards a Randomized Hamiltonian Monte Carlo (RHMC) process as the dimension of the ambient space goes to infinity. RHMC is another piecewise deterministic non-reversible Markov process where a Hamiltonian dynamics is altered at the arrival times of a homogeneous PP by randomly perturbing the momentum component. We then establish dimension-free convergence rates for RHMC for strongly log-concave targets with bounded Hessians using coupling ideas and hypocoercivity techniques. Joint work with George Deligiannidis, Arnaud Doucet and Alexandre Bouchard-Ct.
Shrinkage in Bayesian Shape Constrained Inference
Debdeep Pati
Texas A&M University

We show that any lower-dimensional marginal density obtained from truncating multivariate normal distributions to the positive orthant exhibits a mass-shifting phenomenon. Despite the truncated multivariate normal having a mode at the origin, the marginal density assigns increasingly small mass near the origin as the dimension increases. The phenomenon is accentuated as the correlation between the random variables increases; in particular we show that the univariate marginal assigns vanishingly small mass near zero as the dimension increases provided the correlation between any two variables is greater than 0.8. En-route, we develop precise comparison inequalities to estimate the probability near the origin under the marginal distribution of the truncated multivariate normal. This surprising behavior has serious repercussions in the context of Bayesian shape constrained estimation and inference, where the prior, in addition to having a full support, is required to assign a substantial probability near the origin to capture flat parts of the true function of interest. Without further modifications, we show that commonly used priors are not suitable for modeling flat regions and propose a novel alternative strategy based on shrinking the coordinates using a multiplicative scale parameter. The proposed shrinkage prior guards against the mass shifting phenomenon while retaining computational efficiency.
Bayesian Nonparametric Clustering Analysis for Single Cell RNA Sequencing Data

Mengjie Chen
University of Chicago

Single cell RNA sequencing (scRNAseq) technique has emerged as a powerful tool in genomics. scRNAseq is capable of providing gene expression measurements at single cell level and at genome-wide scale, thus allowing in-depth characterization of a potentially heterogeneous cell population. A key analytic task in scRNAseq involves classifying cells into sub-populations, which requires the development of statistical methods that can perform effective and accurate unsupervised clustering. However, standard clustering methods are not directly applicable to scRNAseq data, as these methods often fail to account for the high measurement noise and an abundance of drop-out events encountered in scRNAseq data, which are results of the low capture efficiency and low amount of input material in scRNAseq. Here, we present a Bayesian nonparametric method that is tailored for clustering analysis in scRNAseq. Our method accounts for both high measurement noise and drop-out events, and is capable of automatically inferring the number of cell sub-populations from the data at hand. Our method incorporates the sparse factor model and the Dirichlet process normal mixture clustering model into a same joint framework, and effectively performs clustering on a low dimensional informative manifold inferred from the noisy data thus enabling accurate clustering performance. To make our method scalable, we also develop a computationally efficient variational algorithm for model inference. With extensive simulations and a real data application, we show that our method outperforms several competing methods in clustering scRNAseq data.

Bayesian Generative Training is Robust

Chao Gao
University of Chicago

In this talk, I will discuss a Bayesian training procedure that is inspired by ABC (approximate Bayesian computation) and GAN (generative adversarial nets). I will discuss the robustness property of the posterior distribution compared with classical Bayesian methods that use explicit likelihood functions.
Bayesian Estimation of Population Moments and Parameters in Biased Sampling

Yuya Shimizu
Keio University

We propose an estimation method of population moments or parameters in “biased sampling data” in which for some units of data, not only the variable of interest but also the covariates, have missing observations and the proportion of “missingness” is unknown. We use auxiliary information such as the distribution of covariates or their moments in random sampling data in order to correct the bias. Moreover, with additional assumptions, we can correct the bias even if we have only the moment information of covariates. The main contribution of this research is the development of a Bayesian estimator for biased sampling data. For example, the customer data of a company can be considered as this biased sampling data when the researcher cannot observe the unit-level data of other companies, and only obtains the customer data of a company; however, the researcher can often obtain census data as auxiliary information, and can correct the bias by using the proposed method. Both the simulation and empirical application for the marketing data demonstrate that the proposed estimation method can correct the bias.

This is joint work with Takahiro Hoshino (Keio University).
Modeling Heterogeneous Impacts of Mental Accounting and Household Stock to Consumers Inter-shopping Duration

Kazuhiro Miyatsu
Nielsen Company Japan

Consumers do not always behave as rationally as to maximize their utilities, but they often take actions in a seemingly irrational way. Behavioral economics describes such consumer anomalies with the concept of Mental Accounting. In this notion, consumers have different criteria values for goods depending on purpose of its use as well as circumstances at purchase. Our modeling of inter-shopping duration accommodates mental condition changes captured by two latent variables, i.e. mental accounting and household stock. Mental accounting is modeled as cumulative purchase amount from consumer payday until the next payday comes attempting to seize consumers pressure in spending at each occasion. Our research aims to comprehend consumer inter-shopping behaviors in four-regime dimensions where they could be seemingly irrational from traditional economics viewpoints. The model is derived from threshold-based modeling framework that incorporates consumer heterogeneity in a hierarchical Bayesian manner, and modeling parameters are estimated using Markov chain Monte Carlo (MCMC) methods. Empirical studies have been exploited with scanner panel data of a retailer shop, and results indicate that our model outperforms by having consumer mental conditions changes into consideration at time of purchase.

This is joint work with Tadahiko Sato (University of Tsukuba).

Dynamic Two-Stage Modeling for Category-Level and Brand-Level Purchases with a Bayes Inference

Kei Miyazaki
Kansai University

We propose an econometric two-stage model for category-level and brand-level purchases that allows for simultaneous brand purchases when analyzing scanner panel data. The proposed approach is consistent with the stage-wise decision process of the Engel Blackwell Miniard model, which is well established in consumer behavior theory. According to the EBM model, in the case of low-involvement products, consumers first determine whether they will make product-category purchases based on problem recognition, such as a lack of household inventory. This decision is independent of brand loyalty, utility, or preferences. Our proposed model is consistent with this idea. Such modeling approaches have not been used in econometric models. We conduct a Bayesian estimation using the Markov Chain Monte Carlo algorithm for our proposed model. In simulation studies, we showed that parameters can be recovered with good accuracy. The hit rate for the proposed model is superior to the considered alternative models. The results in real-data analysis also indicate that the proposed method outperforms all of the considered alternatives. Whereas the latter models are unable to examine switching behaviors.
between the main brands and the baseline brand, the proposed model is able to do so. The proposed method can be used for other purposes beyond examining brand-switching behaviors. For example, it is possible to analyze store choice and multiple-product-category purchase data.

This is joint work with Takahiro Hoshino (Keio University) and Ulf Bockenholt (Northwestern University).
Within the framework of partial exchangeability, we review models based on dependent completely random measures (or suitable transformations thereof). Some of their marginal and conditional distributional properties are presented with focus on additive, hierarchical and nested constructions. These distributional results provide insight into the inferential implications of the considered models and allow to derive effective sampling schemes. Popular nonparametric models are obtained as special cases. Illustrations related to species sampling problems and survival analysis are provided.