September 15

Speaker: Suvra Pal
Title: Support Vector Machine-Based Cure Rate Models
Abstract: In this talk, I will present a new promotion time cure model (PCM) that uses the support vector machine (SVM) to model the probability of cure. The proposed model inherits the features of the SVM and provides flexibility in capturing non-linearity in the data. Furthermore, the new model can incorporate potentially large number of covariates. For the estimation of model parameters, I will discuss the steps of an expectation maximization algorithm where I will make use of the sequential minimal optimization technique together with the Platt scaling method. Next, I will present the results of a detailed simulation study and show that the proposed model outperforms the existing logistic regression-based and spline regression-based PCM models, specifically when the true classification boundary is non-linear. I will also show that the proposed model's ability to capture complex classification boundaries can improve the estimation results related to the survival distribution of the uncured. Finally, I will analyze a data from leukemia cancer study and show that the proposed model results in improved predictive accuracy.

September 22

Speaker: Mohamad Kazem Shirani Faradonbeh
Title: Learning from Multiple Multivariate Time-Series
Abstract: Autoregressive models are among the most popular ones for analyzing temporally dependent data. However, little is known about learning from multiple time-series trajectories, especially when it comes to multi-dimensional and non-stationary data. We study this problem and propose a joint estimation method for learning transition matrices of multiple vector autoregressive models that are unknown linear combinations of some unknown basis matrices. The setting is technically challenging due to high dimensionality of the parameter space as well as the compound nature of the uncertainty. Still, our theoretical analysis shows that the proposed joint estimator has an optimal sample-complexity and excels individual learning methods. Furthermore, applications to data-driven stabilization of dynamical systems through exogenously designed input experiments will be discussed.

September 29

Speaker: Noirrit Chandra
Title: Bayesian Nonparametric Common Atoms Regression for Generating Synthetic Controls in Clinical Trials
Abstract: The availability of electronic health records (EHR) has opened opportunities to supplement increasingly expensive and difficult to carry out randomized controlled trials (RCT) with evidence from readily available real world data. In this paper, we use EHR data to construct synthetic control arms for treatment-only single arm trials. We propose a novel nonparametric Bayesian common atoms mixture model that allows us to find equivalent population strata in the EHR and the treatment arm and then resample the EHR data to create equivalent patient populations under both the single arm trial and the resampled EHR. Resampling is implemented via a density-free importance sampling scheme. Using the synthetic control arm, inference for the treatment effect can then be carried out using any method available for RCTs. Alternatively the proposed nonparametric Bayesian model allows straightforward model-based inference. In simulation experiments, the proposed method exhibits higher power than alternative methods in detecting treatment effects for complicated response functions. We apply the method to supplement single arm treatment-only glioblastoma studies with a synthetic control arm based on historical trials.
October 13

**Speaker:** Jungsik Noh  
**Title:** Granger-causality Inference Framework to Study the Actin Cytoskeleton Regulatory System from Live Cell Fluorescence Imaging Data  
**Abstract:** Many cell regulatory systems implicate nonlinearity and redundancy among components. The regulatory network governing actin cytoskeleton structures at the cell edge is prototypical of such a system, containing tens of actin-nucleating and -modulating molecules with functional overlap and feedback loops. Due to instantaneous and long-term compensation, phenotyping the system response to perturbation provides limited information on the targeted component’s roles in the unperturbed system. Accordingly, how individual actin regulators contribute to actin cytoskeleton dynamics remains ambiguous. Here, we present a perturbation-free reconstruction of cause-effect relations among actin regulators by applying Granger-causal inference to constitutive image fluctuations that indicate regulator recruitment as a proxy for activity. Using multivariate time series representing the spatiotemporal molecular and cell morphological dynamics of the system, we identify spatially confined causal relationships between the local activities of actin and its regulators and the corresponding local cell edge motion.

October 20 – Special event: ASA Travel Course

**Speaker:** Babette Brumback  
**Title:** Fundamentals of Causal Inference: With R  
**Abstract:** One of the primary motivations for clinical trials and observational studies of humans is to infer cause and effect. Disentangling causation from confounding is of utmost importance. Fundamentals of Causal Inference: With R explains and relates different methods of confounding adjustment in terms of potential outcomes and graphical models, including standardization, doubly robust estimation, difference-in-differences estimation, and instrumental variables estimation. Several real data examples, simulation studies, and analyses using R motivate the methods throughout. The course assumes familiarity with basic statistics and probability, regression, and R. The course will be taught with a blend of lecture and worked examples. More details here: https://www.amstat-nt.org/events/short-courses