samsi
NSF·Duke·NCSU·UNC

Causal Inference Program Opening Workshop
December 9-11, 2019

SPEAKER TITLES/ABSTRACTS

Edoardo Airoldi
Temple University

"Designing Experiments on Social and Information Networks"

Classical approaches to causal inference largely rely on the assumption of "lack of interference", according to which the outcome of an individual does not depend on the treatment assigned to others, as well as on many other simplifying assumptions, including the absence of strategic behavior. In many applications, however, such as evaluating the effectiveness of health-related interventions that leverage social structure, assessing the impact of product innovations and ad campaigns on social media platforms, or experimentation at scale in large IT companies, several common simplifying assumptions are untenable. Moreover, the effect of interference itself is often an inferential target of interest, rather than a nuisance. In this talk, we will formalize issues that arise in estimating causal effects when interference can be attributed to a network among the units of analysis, within the potential outcomes framework. We will introduce and discuss several strategies for experimental design in this context centered around a useful role for statistical models. In particular, we wish for certain finite-sample properties of the estimator to hold even if the model catastrophically fails, while we would like to gain efficiency if certain aspects of the model are correct. We will then contrast design-based, model-based and model-assisted approaches to experimental design from a decision theoretic perspective.

Guillaume Basse
Stanford University

"Estimating Displacement Effects in a Hot Spot Policing Intervention in Medellin, Colombia"

In hot spot policing, resources are targeted at specific locations predicted to be at high risk of crime; so-called ``hot spots." Rather than reduce overall crime, however, there is a concern that these interventions simply displace crime from the targeted locations to nearby non-hot spots. We address this question in the context of a large-scale randomized experiment in Medellin, Colombia, in which police were randomly assigned to increase patrols at a subset of possible hotspots. Estimating the displacement effects on control locations is difficult because the probability that a nearby hot spot is treated is a complex function of the underlying geography. While existing methods developed for this ``general interference" setting, especially Horvitz-Thompson (HT) estimators, have attractive theoretical properties, they can perform poorly in practice, due in part to extreme weights. We make several contributions in this paper. First, we characterize the performance of HT estimators under different interference settings and compare that to modified estimators, like the H{\`a}jek estimator, that trade off small amounts of bias for large reductions in variance. Second, we suggest a simulation-based approach that uses the structure of the application of interest --- in our case, the Medellin street network --- for assessing the trade-offs of different randomization-based estimators. Finally, we explore some implications for experimental design in this setting. We use these insights to re-analyze displacement effects in the hot spot policing experiment.

Alexander D'Amour
Google Research

"Latent Variable Models, Causal Inference, and Sensitivity Analysis"

Recently, the machine learning community has expressed strong interest in applying latent variable modeling strategies to causal inference problems with unobserved confounding. Here, I discuss one of the big debates that occurred over the past year, and how we can move forward. I will focus specifically on the failure of point identification in this setting, and discuss how this can be used to design flexible sensitivity analyses that cleanly separate identified and unidentified components of the causal model.

Tirthankar Dasgupta
Rutgers University

"Fisher Randomization Test: A Confidence Distribution Perspective and Applications to Large Experiments"

Fisher randomization tests (FRT) are flexible tools because they are a model free, permit assessment of causal effects of interventions on ANY type of response for ANY assignment mechanism using ANY test statistic. The tremendous development of computing resources has recently sparked a huge interest in using FRT to test complex causal hypotheses that can arise from massive studies. In spite of its wide applicability and recent surge of interest, several aspects of the theoretical properties of randomization tests still remain unclear, somewhat limiting its applicability. This research provides a theoretical inferential framework for FRT by combining two fundamental ideas: potential outcomes and confidence distributions. It also demonstrates how such a connection can be exploited to combine causal inference from multiple experiments with different structures and complexities and also to â€œdivide and conquerâ€□ randomization-based inference arising from large experiments. (Based on joint work with Minge Xie, Xiaokang Luo, and Regina Liu.)

Peng Ding
University of California, Berkeley

"A Bracketing Relationship between Difference-in-Differences and Lagged-Dependent-Variable Adjustment"

Difference-in-differences is a widely used evaluation strategy that draws causal inference from observational panel data. Its causal identification relies on the assumption of parallel trends, which is scale-dependent and may be questionable in some applications. A common alternative is a regression model that adjusts for the lagged dependent variable, which rests on the assumption of ignorability conditional on past outcomes. In the context of linear models, Angrist and Pischke (2009) show that the difference-in-differences and lagged-dependent-variable regression estimates have a bracketing relationship. Namely, for a true positive effect, if ignorability is correct, then mistakenly assuming parallel trends will overestimate the effect; in contrast, if the parallel trends assumption is correct, then mistakenly assuming ignorability will underestimate the effect. We show that the same bracketing relationship holds in general nonparametric (model-free) settings. We also extend the result to semiparametric estimation based on inverse probability weighting.

Frederick Eberhardt
Caltech University

"Causal Discovery in Neuroimaging Data"

Many different measurement techniques are used to record neural activity in the brains of different organisms, including fMRI, EEG, MEG, lightsheet microscopy and direct recordings with electrodes. Each of these measurement modes have their advantages and disadvantages concerning the resolution of the data in space and time, the directness of measurement of the neural activity and which organisms they can be applied to. For some of these modes and for some organisms, significant amounts of data are now available in large standardized open-source datasets. I will report on our efforts to apply causal discovery algorithms to, among others, fMRI data from the Human Connectome Project, and to lightsheet microscopy data from zebrafish larvae. In particular, I will focus on the challenges we have faced both in terms of the nature of the data and the computational features of the discovery algorithms, as well as the modeling of experimental interventions.

Avi Feller
University of California, Berkeley

"Synthetic Control and Weighted Event Study Models with Staggered Adoption"

Estimating the impact of policy changes across states and other jurisdictions remains a challenging methodological problem. In our motivating example, 33 states granted collective bargaining rights to teachers at different times between 1964 and 1987; the question is whether this increased teacher salaries and school spending. Regression-based methods, like event study models, rely on a "parallel trends" assumption that is implausible in this application. Alternative estimators, like the synthetic control method (SCM), were designed for a single treated unit rather than for multiple units adopting over time. In this paper, we generalize SCM to this staggered adoption setting. Fitting separate SCM weights for each treated unit and then averaging does not necessarily achieve good balance for the average of the treated units. We propose a new pooled SCM estimator for the average treatment effect as well as a hybrid estimator that balances both the average treated unit and

each treated unit separately. We draw on recent results to show that our hybrid estimator has an interpretation as a generalized propensity score model with partially pooled coefficients. Finally, we combine our staggered adoption SCM approach with event study modeling to obtain an augmented estimator that improves over either SCM weighting or event study modeling alone. We assess the performance of the proposed method via extensive simulations and apply our results to the teacher collective bargaining example. We implement the proposed method in the augsynth R package. Co-authors: Eli Ben-Michael and Jesse Rothstein

Colin Fogarty
MIT

"Testing Weak Nulls in Matched Observational Studies"

We develop sensitivity analyses for weak nulls in matched observational studies while allowing unit-level treatment effects to vary. In contrast to randomized experiments and paired observational studies, we show for general matched designs that over a large class of test statistics, any valid sensitivity analysis for the weak null must be unnecessarily conservative if Fisher's sharp null of no treatment effect for any individual also holds. We present a sensitivity analysis valid for the weak null, and illustrate why it is conservative if the sharp null holds through connections to inverse probability weighted estimators. An alternative procedure is presented that is asymptotically sharp if treatment effects are constant, and is valid for the weak null under additional assumptions which may be deemed reasonable by practitioners. The methods may be applied to matched observational studies constructed using any optimal without-replacement matching algorithm, allowing practitioners to assess robustness to hidden bias while allowing for treatment effect heterogeneity.

Laura Forastiere
Yale University

"Â Heterogeneous Causal Effects under Network Interference"

Spillovers are a crucial component in understanding the full impact of interventions at the population-level. Information about spillovers of health interventions would support decisions about how best to deliver interventions and can be used to guide public funds allocation. In fact, policy makers can gain from understanding the heterogeneity of spillover effects to identify the most contagious or influential individuals andÂ those who are more susceptible. Social network targeting shows great promise in behavioral change interventions and policy makers are in need of guidance on how best to design their programs so as to use social networks to maximize adoption of healthy behaviors for improving community health.

Under a causal inference framework, we develop machine learning methods to assess the heterogeneity of treatment and spillover effects in a two-stage randomized experiment with clustered networks.

Laura Hatfield
Harvard Medical School

"Difference-in-differences: more than meets the eye"

The world of health care is full of policy interventions: a state expands eligibility rules for its Medicaid program, a medical society changes its recommendations for screening frequency, a hospital implements a new care coordination program. After a policy change, we often want to know, "Did it work?" This is a causal question; we want to know whether the policy CAUSED outcomes to change. One popular way of estimating causal effects of policy interventions is a difference-in-differences study. In this controlled pre-post design, we measure the change in outcomes of people who are exposed to the new policy, comparing average outcomes before and after the policy is implemented. We contrast that change to the change over the same time period in people who were not exposed to the new policy. The differential change in the treated group's outcomes, compared to the change in the comparison group's outcomes, may be interpreted as the causal effect of the policy. To do so, we must assume that the comparison group's outcome change is a good proxy for the treated group's (counterfactual) outcome change in the absence of the policy. This conceptual simplicity and wide applicability in policy settings makes difference-in-differences an appealing study design. However, the apparent simplicity belies a thicket of conceptual, causal, and statistical complexity. In this talk, I will introduce the fundamentals of difference-in-differences studies and discuss recent innovations including key assumptions and ways to assess their plausibility, estimation, inference, and robustness checks.

Luke Keele
University of Pennsylvania

"Bracketing Bounds for Differences-in-Differences Methods"

The method of differences-in-differences (DID) is widely used to estimate causal effects. The primary advantage of DID is that it can account for time-invariant bias from unobserved confounders. However, the standard DID estimator will be biased if there is an interaction between history in the after period and the groups. That is, bias will be present if an event besides the treatment occurs at the same time and affects the treated group in a differential fashion. We present a method of bounds based on DID that accounts for an unmeasured confounder that has a differential effect in the post-treatment time period. These DID bracketing bounds are simple to implement and only require partitioning the controls into two separate groups. We also develop two key extensions for DID bracketing bounds. First, we develop a new falsification test to probe the key assumption that is necessary for the bounds estimator to provide consistent estimates of the treatment effect. Next, we develop a method of sensitivity analysis that adjusts the bounds for possible bias based on differences between the treated and control units from the pretreatment period. We apply these DID bracketing bounds and the new methods we develop to an application on the effect of voter identification laws on turnout. Specifically, we focus estimating whether the enactment of voter identification laws in Georgia and Indiana had an effect on voter turnout.

Edward Kennedy
Carnegie Mellon University

“Optimal Causal Inference with High-Dimensional Discrete Data”

In this work we consider estimation of causal effects in a setting where the covariates required for confounding adjustment are discrete but arbitrarily high-dimensional. Contrary to classical results that consider dimension d fixed and sample size n tending to infinity, we give non-asymptotic risk bounds for arbitrary n and d, and consider non-classical regimes where d can grow with or exceed n. We study the plug-in estimator of the average treatment effect, showing that this estimator is only consistent in the regime where the dimension grows slower than the sample size. We also consider a propensity-weighted effect estimator, which has some surprising ability to adapt to both exact and approximate effect homogeneity. Then we go on to characterize the minimax lower bounds for the average effect under heterogeneity and homogeneity, indicating the fundamental limits of causal inference in high dimensions as well as when an estimator can no longer be improved upon.

Michael Kosorok
University of North Carolina

“Some Applications of Reinforcement Learning in Precision Medicine”

We provide an overview of some recent developments in machine learning tools for dynamic treatment regime discovery in precision medicine. The first development is a new off-policy reinforcement learning tool for continual learning in mobile health to enable patients with type 1 diabetes to exercize safely. The second development is a new inverse reinforcement learning tools which enables use of observational data to learn how clinicians balance competing priorities for treating depression and mania in patients with bipolar disorder. Both practical and technical challenges are discussed.

Mark van der Laan
University of California, Berkeley

“Targeted Learning for Causal Inference Based on Real World Data”

We discuss a general roadmap for generating causal inference based on observational studies used to general real world evidence. We review targeted minimum loss estimation (TMLE), which provides a general template for the construction of asymptotically efficient plug-in estimators of a target estimand for realistic (i.e, infinite dimensional) statistical models. TMLE is a two stage procedure that first involves using ensemble machine learning termed super-learning to estimate the relevant stochastic relations between the treatment, censoring, covariates and outcome of interest. The super-learner allows one to fully utilize all the advances in machine learning (in addition to more conventional parametric model based estimators) to build a single most powerful ensemble machine learning algorithm. We present Highly Adaptive Lasso as an important machine learning algorithm to include.

In the second step, the TMLE involves maximizing a parametric likelihood along a so-called least favorable parametric model through the super-learner fit of the relevant stochastic relations in the observed data. This second step bridges the state of the art in machine learning to estimators of target estimands for which statistical inference is available (i.e, confidence intervals, p-values etc). We also review recent advances in collaborative TMLE in which the fit of the treatment and censoring mechanism is tailored w.r.t. performance of TMLE. We also discuss asymptotically valid bootstrap based inference. Simulations and data analyses are provided as demonstrations.

Lihua Lei
University of California, Berkeley

"Model-Free Assessment of Population Overlap in Observational Studies"

The reliability of causal inference in observational studies crucially relies on the overlap of baseline covariates (a.k.a. positivity or common support) between different treated groups. Without sufficient overlap, the internal validity is hampered unless the causal estimand is redefined. The current empirical assessment of overlap is typically based on estimated propensity scores. This approach is meaningful only when the model specification for propensity scores is correct and it typically has no formal statistical guarantee due to the lack of uncertainty quantification. In this work we treated the overlap condition as a composite null hypothesis (e.g. that propensity scores lie in $[\eta 0, 1 - \eta 0]$ for all units in study population with a given $\eta 0$) and developed a family of tests with Type-I error control in finite samples without any assumption on the data generating process, if only the observations are independent and identically distributed. The test provides a provably reliable alarming system that helps practitioners detect the lack of overlap if a desirable overlap condition is rejected. Our test exploits the duality between overlap and performance of the optimal classifier that distinguishes treatment groups using baseline covariates [D'Amour et al., 2017], as well as technical tools from information theory and empirical process theory.

Michael Leung
University of Southern California

"Causal Inference under Approximate Neighborhood Interference"

This paper studies causal inference in randomized experiments with network interference. To reduce the dimensionality of the problem, most of the existing literature assumes that the treatment assigned to alters only affect the ego's response through a low-dimensional exposure mapping. We instead study models satisfying a substantially weaker approximate neighborhood interference assumption in which the dependence of the ego's response on treatments assigned to alters outside of the ego's K-neighborhood is increasingly negligible as K grows. We show that this assumption is satisfied for well-known models of social interactions, in contrast to the exposure mapping approach. We then prove that standard Horovitz-Thompson estimators used in the literature consistently estimate causal effects and are asymptotically normal under restrictions on the network topology. For practical inference, we show that an estimator based on a wild bootstrap conservatively estimates the variance. Finally, we illustrate our results in a simulation study and empirical application.

Jared Murray
University of Texas, Austin

"Bayesian Nonparametric Models for Treatment Effect Heterogeneity: model parameterization, prior Choice, and posterior summarization"

We describe different approaches for specifying models and prior distributions for estimating heterogeneous treatment effects using Bayesian nonparametric models. We make an affirmative case for direct, informative (or partially informative) prior distributions on heterogeneous treatment effects, especially when treatment effect size and treatment effect variation is small relative to other sources of variability. We also consider how to provide scientifically meaningful summaries of complicated, high-dimensional posterior distributions over heterogeneous treatment effects with appropriate measures of uncertainty.

Elizabeth Ogburn
Johns Hopkins University

"Social Network Dependence, the Replication Crisis, and (In)Valid Inference"

In the first part of this talk, we show that social network dependence can result in *confounding by network structure*, akin to confounding by population structure in GWAS studies, potentially contributing to replication crises across the health and social sciences. Researchers in these fields frequently sample subjects from one or a small number of communities, schools, hospitals, etc., and while many of the limitations of such convenience samples are well-known, the issue of statistical dependence due to social network ties has not previously been addressed. A paradigmatic example of this is the Framingham Heart Study (FHS). Using a statistic that we adapted to measure network dependence, we test for network dependence and for possible confounding by network structure in several of the thousands of influential papers published using FHS data. Results suggest that some of the many decades of research on coronary heart disease, other health outcomes, and peer influence using FHS data may suffer from spurious estimates of association and anticonservative uncertainty quantification due to unacknowledged network structure.

But data with network dependence abounds, and in many settings researchers are explicitly interested in learning about social network dynamics. Therefore, there is high demand for methods for causal and statistical inference with social network data. The second part of the talk describes recent work on causal inference for observational data from a single social network, focusing on (1) new types of causal estimands that are of interest in social network settings, and (2) conditions under which central limit theorems hold and inference based on approximate normality is licensed.

Georgie Papadogeorgou
Duke University

"Mitigating Unobserved Spatial Confounding Bias with Mixed Models"

Confounding by unmeasured spatial variables has received some attention in the spatial statistics and causal inference literatures, but concepts and approaches have remained largely separated. We bridge these distinct strands of statistics by considering unmeasured spatial confounding within a formal causal inference framework, and estimating effects using modifications of outcome regression tools popular within the spatial literature. We show that using spatially correlated random effects in the outcome model, an approach common among spatial statisticians, does not mitigate bias due to spatial confounding. Motivated by the bias term of commonly-used estimators, we propose an affine estimator which addresses this deficiency. We discuss how unbiased estimation of causal parameters in the presence of unmeasured spatial confounding can only be achieved under an untestable set of assumptions which will often be application-specific. We provide one set of assumptions that is sufficient for identification of the causal effect based on the observed data. These assumptions describe how the exposure and outcome of interest relate to the unmeasured variables. Estimation of the model components necessary for unbiased estimation of the causal effect proceeds using tools common in the spatial statistics literature, and specifically via a regularized restricted maximum likelihood approach employing weakly informative priors to avoid degenerate estimates. This work is motivated by and used to estimate the causal effect of county-level (a) exposure to emissions from coal-powered electricity generating units, and (b) relative humidity on particulate matter across the New England area in the United States, and to investigate the potential threat from unmeasured spatial confounders in this context.

Thomas Richardson
University of Washington

"On Testing Marginal versus Conditional Independence"

Motivated by the problem of causal discovery, we consider testing marginal independence versus conditional independence in the Gaussian setting. The two models are non-nested and their intersection is a union of two marginal independences. To control the error even when the two models are indistinguishable, rather than insist on a dichotomous choice, we propose a procedure that will choose either or both models. We present a procedure with uniform error guarantees, including under local alternatives.

Michael Rosenblum
Johns Hopkins University

"Adaptive Design in Surveys and Clinical Trials: similarities, differences and opportunities for cross-fertilization"

Adaptive designs involve preplanned rules for modifying an on-going study based on accruing data. We compare the goals and methods of adaptation for trials and surveys, identify similarities and differences, and make recommendations for what types of adaptive approaches from one domain have high potential to be useful in the other. For example, clinical trials could benefit from recently developed survey methods for monitoring which groups have low response rates and intervening to fix this. Clinical trials may also benefit from more formal identification of the target population, and from using paradata (contextual information collected before or during the collection of actual outcomes) to predict participant compliance and retention and then to intervene to improve these. Surveys could benefit from stopping rules based on information monitoring, applying techniques from sequential multiple-assignment randomized trial designs to improve response rates, prespecifying a formal adaptation protocol and including a data monitoring committee.We conclude with a discussion of the additional information, infrastructure and statistical analysis methods that are needed when conducting adaptive designs, as well as benefits and risks of adaptation.
Joint work with Peter Miller, Benjamin Reist, Elizabeth A. Stuart, Michael Thieme, and Thomas A. Louis. Paper: https://doi.org/10.1111/rssa.12438

Michael E. Sobel
Columbia University

"Estimating Causal Effects in Studies of Human Brain Function: new models, methods and estimates"

Neuroscientists often use functional magnetic resonance imaging (fMRI) to infer effects of treatments on neural activity in brain regions. In a typical fMRI experiment, each subject is observed at several hundred time points. At each point, the blood oxygenation level dependent (BOLD) response is measured at 100,000 or more locations (voxels). Typically, these responses are modeled treating each voxel separately, and no rationale for interpreting associations as effects is given. Building on Sobel and Lindquist (2014), who used potential outcomes to define unit and average effects at each voxel and time point, we define and estimate both \point" and \cumulated" effects for brain regions. Second, we construct a multi-subject multi-voxel multi-run whole brain causal model with explicit parameters for regions. We justify estimation using BOLD responses averaged over voxels within regions, making feasible estimation for all regions simultaneously, thereby also facilitating inferences about association between effects in different regions. We apply the model to a study of pain, finding effects in standard pain regions. We also observe more cerebellar activity than observed in previous studies using prevailing methods.

Co-author: Martin A. Lindquist, Johns Hopkins University

Peter Spirtes
Carnegie Mellon University

"Simplicity Concepts for Causal Inference"

Consistency proofs for causal inference algorithms generally require some assumptions that make simple causal theories preferable to more complex causal theories. There are a number of different simplicity assumptions that have been proposed, including a variety of faithfulness assumptions (that all conditional independence relations that hold in the population are entailed by the true causal graph) of various strength, parametric assumptions, and sparsity assumptions. I will discuss the strengths and weaknesses of the various simplicity assumptions, and whether there is some unifying concept underlying them.

Jennifer Starling
University of Texas

"Smooth Extensions to BART for Heterogeneous Treatment Effect Estimation, with Applications to Women's Healthcare Practice and Policy"

Bayesian Additive Regression Trees (BART) has been shown to be an effective framework for modeling nonlinear regression functions, with strong predictive performance in a variety of contexts. The BART prior over a regression function is defined by independent prior distributions on tree structure and leaf or end-node parameters. In observational data settings, Bayesian Causal Forests (BCF) has successfully adapted BART for estimating heterogeneous treatment effects, particularly in cases where standard methods yield biased estimates due to strong confounding.

We introduce BART with Targeted Smoothing, an extension which induces smoothness over a single covariate by replacing independent Gaussian leaf priors with smooth functions. We then introduce a new version of the Bayesian Causal Forest prior, which incorporates targeted smoothing for modeling heterogeneous treatment effects which vary smoothly over a target covariate. We demonstrate the utility of this approach by applying our model to a timely women's health and policy problem: comparing two dosing regimens for an early medical abortion protocol, where the outcome of interest is the probability of a successful early medical abortion procedure at varying gestational ages, conditional on patient covariates. We discuss the benefits of this approach in other women's health and obstetrics modeling problems where gestational age is a typical covariate.

Eric Tchetgen Tchetgen
University of Pennsylvania

"A Semiparametric Instrumental Variable Approach to Optimal Treatment Regimes under Endogeneity"

There is a fast-growing literature on estimating optimal treatment regimes based on randomized trials or observational studies under a key identifying condition of no unmeasured confounding. Because confounding by unmeasured factors cannot generally be ruled out with certainty in observational studies or randomized trials subject to non-compliance, we propose a general instrumental variable approach to learning optimal treatment regimes under endogeneity.
Speci fically, we provide sufficcient conditions for the identificcation of both value function for a given regime and of the optimal treatment regime with the aid of a binary instrumental variable,

when no unmeasured confounding fails to hold. We establish consistency of the proposed weighted estimators. We also extend the proposed method to identify and estimate the optimal treatment regime among those who would comply to the assigned treatment under monotonicity. In this latter case, we establish the somewhat surprising result that the complier optimal regime can be consistently estimated without directly collecting compliance information. Furthermore, we propose novel semiparametric locally efficient and multiply robust estimators. Our approach is illustrated via extensive simulation studies and a data application on the e ffect of child rearing on labor participation.

This is joint work with Yifan Cui

Caroline Uhler
MIT

"From Causal Inference to Gene Regulation"

A recent break-through in genomics makes it possible to perform perturbation experiments at a very large scale. The availability of such data motivates the development of a causal inference framework that is based on observational and interventional data. We first characterize the causal relationships that are identifiable from interventional data. In particular, we show that imperfect interventions, which only modify (i.e., without necessarily eliminating) the dependencies between targeted variables and their causes, provide the same causal information as perfect interventions, despite being less invasive. Second, we present the first provably consistent algorithm for learning a causal network from a mix of observational and interventional data. This requires us to develop new results in geometric combinatorics. In particular, we introduce DAG associahedra, a family of polytopes that extend the prominent graph associahedra to the directed setting. We end by discussing applications of this causal inference framework to the estimation of gene regulatory networks.

Stefan Wager
Stanford University

"Experimenting in Equilibrium"

We study experimental design in large-scale stochastic systems with substantial uncertainty and structured cross-unit interference. We consider the problem of a platform that seeks to optimize supply-side payments p in a centralized marketplace where different suppliers interact via their effects on the overall supply-demand equilibrium, and propose a class of local experimentation schemes that can be used to optimize these payments without perturbing the overall market equilibrium. We show that, as the system size grows, our scheme can estimate the gradient of the platformâ€™s utility with respect to p while perturbing the overall market equilibrium by only a vanishingly small amount. We can then use these gradient estimates to optimize p via any stochastic first-order optimization method. These results stem from the insight that, while the system involves a large number of interacting units, any interference can only be channeled through a small number of key statistics, and this structure allows us to accurately predict feedback effects that arise from global system changes using only information collected while remaining in equilibrium.

Jingshen Wang
University of California, Berkeley

"Inference on Treatment Effects after Model Selection with Application to Subgroup Analysis"

Inferring cause-effect relationships between variables is of primary importance in many sciences. In this talk, we start by discussing two approaches for making valid inferences on treatment effects when a large number of covariates are present. The first approach is to perform model selection and then to deliver inference based on the selected model. If the inference is made ignoring the randomness of the model selection process, then there could be severe biases in estimating the parameters of interest. While the estimation bias in an under-fitted model is well understood, a lesser-known bias that arises from an over-fitted model will be addressed. The over-fitting bias can be eliminated through data splitting at the cost of statistical efficiency, and we propose a repeated data splitting approach to mitigate the efficiency loss. The second approach concerns the existing methods for debiased inference. We show that the debiasing approach is an extension of OLS to high dimensions. A comparison between these two approaches provides insights into their intrinsic bias-variance trade-off, and the debiasing approach may lose efficiency in observational studies. For the second part of the talk, we discuss a generalization on how to estimate treatment effects for groups of individuals that share similar features in observational studies. More importantly, even if the estimated treatment effects suggest a promising subgroup (i.e. the group with the maximal treatment effect), we address the question of how good the subgroup really is.

Lu Wang
University of Michigan

"New Statistical Learning Methods for Estimating the Optimal Dynamic Treatment Regime"

We present recent advances and statistical developments for evaluating Dynamic Treatment Regimes (DTR), which allow the treatment to be dynamically tailored according to evolving subject-level data. Identification of an optimal DTR is a key component for precision medicine and personalized health care. Specific topics covered in this talk include several recent projects with robust and flexible methods developed for the above research area. We will first introduce a dynamic statistical learning method, adaptive contrast weighted learning (ACWL), which combines doubly robust semiparametric regression estimators with flexible machine learning methods. We will further develop a tree-based reinforcement learning (T-RL) method, which builds an unsupervised decision tree that maintains the nature of batch-mode reinforcement learning. Unlike ACWL, T-RL handles the optimization problem with multiple treatment comparisons directly through a purity measure constructed with augmented inverse probability weighted estimators. T-RL is robust, efficient and easy to interpret for the identification of optimal DTRs. However, ACWL seems more robust against tree-type misspecification than T-RL when the true optimal DTR is non-tree-type. At the end of this talk, we will also present a new Stochastic-Tree Search method called ST-RL for evaluating optimal DTRs.

Yanxun Xu
Johns Hopkins University

“When and How to Treat Patients?”

Patients that undergo kidney transplantation are at risk for a number of complications and graft rejection after surgery, which could lead to death. In order to prevent graft rejection, immuno-suppressive therapy such as tacrolimus is administered to patients post-surgery. The patients are monitored over time with repeated follow-up records (e.g., tacrolimus blood levels, creatinine levels, BMI) after transplantation and the dosage levels of the immunosuppresive drugs at each visitation can be adjusted by the clinician. Based on patients’ baseline information and the followup data, We develop a statistical joint modeling framework to construct an optimal longitudinal treatment strategy for each individual patient by combining a longitudinal model for patients’ creatinine levels, a survival model with the endpoint being patient's death or graft failure, and a marked point process for clinical decisions (how often the patient is instructed to followup, and drug dosage adjustments). Our method shows promising performance on a real kidney transplantation dataset, and outperforms alternatives on synthetic datasets.

Kun Zhang
Carnegie Mellon University

“Learning Hidden Causal Variables and Relations”

Over the past several decades, there has been much progress in discovering causal relations from observational data. Most of the work has been focusing on causal relations between observed variables, while in reality we often have causal interactions between hidden variables underlying the observed ones. In this talk I will present some methods for discovering such hidden variables and estimating their causal relations.

Qingyuan Zhao
University of Cambridge

“Bootstrapping Sensitivity Analysis for Inverse Probability Weighting Estimators”

To identify the estimand in missing data problems and observational studies, it is common to base the statistical estimation on the â€œmissing at randomâ€□ and â€œno unmeasured confounderâ€□ assumptions. However, these assumptions are unverifiable using empirical data and pose serious threats to the validity of the qualitative conclusions of the statistical inference. A sensitivity analysis asks how the conclusions may change if the unverifiable assumptions are violated to a certain degree. In this paper we consider a marginal sensitivity model which is a natural extension of Rosenbaumâ€™s sensitivity model that is widely used for matched observational studies. We aim to construct confidence intervals based on inverse probability weighting estimators, such that asymptotically the intervals have at least nominal coverage of the estimand whenever the data generating distribution is in the collection of marginal sensitivity models. We use a percentile bootstrap and a generalized minimax/maximin inequality to transform this intractable problem to a linear fractional programming problem, which can be solved very efficiently. We illustrate our method using a real dataset to estimate the causal effect of fish consumption on blood mercury level.

Corwin Zigler
University of Texas

"Bipartite Causal Inference with Interference for Evaluating Air Pollution Regulations"

A fundamental feature of evaluating causal health effects of air quality regulations is that air pollution moves through space, rendering health outcomes at a particular population location dependent upon regulatory actions taken at multiple, possibly distant, pollution sources. Motivated by studies of the public-health impacts of power plant regulations in the U.S., this talk introduces the novel setting of bipartite causal inference with interference, which arises when 1) treatments are defined on observational units that are distinct from those at which outcomes are measured and 2) there is interference between units in the sense that outcomes for some units depend on the treatments assigned to many other units. Interference in this setting arises due to complex exposure patterns dictated by physical-chemical atmospheric processes of pollution transport, with intervention effects framed as propagating across a bipartite network of power plants and residential zip codes. New causal estimands are introduced for the bipartite setting, along with an estimation approach based on generalized propensity scores for treatments on a network. The new methods are deployed to estimate how emission-reduction technologies implemented at coal-fired power plants causally affect health outcomes among Medicare beneficiaries in the U.S.

Jose Zubizarreta
Harvard University

"Complex Discontinuity Designs Using Covariates"

Regression discontinuity designs are extensively used for causal inference in observational studies. However, they are usually confined to settings with simple treatment rules, determined by a single running variable, with a single cutoff. In this paper, we propose a new framework and methods for general discontinuity designs that encompasses complex treatment rules. These rules may be determined by multiple running variables, each with many cutoffs, and that possibly lead to the same treatment. Moreover, the running variables may be discrete and the treatments do not need to be binary. In this framework, the observed covariates play a central role for identification, estimation, and generalization of causal effects. Identification essentially relies on a local unconfoundedness assumption. Estimation proceeds as in any observational study under the strong ignorability assumption, yet in a neighborhood of the cutoffs of the running variables. We discuss estimation approaches based on matching and weighting, including additional regression adjustments in doubly robust estimators. We present assumptions for generalization; that is, for identification and estimation of average treatment effects for target populations beyond the study sample that resides in a neighborhood of the cutoffs. We also examine a new approach to select the neighborhood for the analyses and assess the plausibility of the assumptions. We argue that, in a sense, traditional continuity and local randomization frameworks for regression discontinuity designs are particular cases of our proposed framework. We motivate and illustrate this framework with an example of the impact of grade retention on juvenile crime.