Abstracts

Mobile health technology in health applications

Anna Choi[a] and Tze Leung Lai[b]

[a]*Stanford University, USA and The Chinese University of Hong Kong, Shenzhen, China*
[b]*Stanford University, USA*

ABSTRACT

With the explosion of digital data, high performance computing, data processing, development and management of databases, data warehousing, mathematical representations, statistical modeling and analysis, and visualization are crucial in extracting information from the data collected for domain-specific applications. Mobile health (mHealth) technologies are ubiquitous in the delivery of healthcare services via mobile communication devices which include mobile phones, patient monitoring devices, and wearable devices such as smart watches. These devices can be used to collect personal health data to improve wellness and to monitor and track environmental exposures that may affect disease and epidemic outbreak. In this tutorial, we will discuss the use of mHealth technologies in two health applications: (1) use of geographic information system (GIS) and spatial modeling to ascertain the relationship between residential proximity and the site with high levels of an environmental contaminant; and (2) online delivered treatment on the improvement of mental wellness among subjects with depression.

Contextual multi-armed bandits in precision medicine and health

Anna Choi[a] and Tze Leung Lai[b]

[a]*Stanford University, USA and The Chinese University of Hong Kong, Shenzhen, China*
[b]*Stanford University, USA*

ABSTRACT

Multi-armed bandit with side information (covariate/contextual bandits) arise in many fields of application. The development of personalized strategies and recommender systems requires both exploration and exploitation. Personalized marketing uses websites to track a customer's interests and purchasing records and thereby to market products individualized for the customer, whereas recommender systems select items such as movies and news for users based on the features (covariates) of the users and items. An example of a trial that used biomarker-integrated approach of targeted therapy for lung cancer elimination will be discussed. New definitive theory of multi-armed bandits with covariates can provide theoretical support for previous experimental studies in personalized strategies and recommender systems, and can be used to develop new "learn-as-we-go" strategies. By incorporating statistical and machine learning approaches, the covariate bandit theory can advance "Big Data" analytics for many applications. Implementation of the theory involves modern developments in statistics, reinforcement learning and computer science.

Dynamic treatment regimes and “SMART” design

Bibhas Chakraborty

Duke-NUS Medical School, National University of Singapore, Singapore

ABSTRACT

Effective treatment of chronic conditions typically requires ongoing interventions where clinicians sequentially make therapeutic decisions, adapting the type, dosage and timing of treatment according to evolving patient characteristics. The framework of dynamic treatment regimes (DTRs) formalizes this sequential decision-making process prevalent in clinical practice. Constructing data-driven DTRs from either observational data or sequentially randomized trials comprise a cutting-edge area of research in modern biostatistics. This area brings together concepts from dynamic programming, reinforcement (machine) learning, causal inference, design of clinical trials, and non-regular asymptotic theory, thus offering ample opportunities for statisticians and data scientists.

This tutorial lecture will provide a comprehensive overview of the above topic, beginning with a discussion of relevant data sources for constructing DTRs; in particular, we will focus on a cutting-edge study design called sequential multiple-assignment randomized trial (SMART). Examples of SMART involving mobile-based interventions will be discussed. We will then turn our attention to estimation of optimal DTRs using a method called *Q-learning* originally developed in computer science. Finally, we will briefly discuss non-regular inferential challenges and present a feasible solution using the m-out-of-n bootstrap in this context.

Introduction to causal inference in complex longitudinal settings

Jonathan Huang

*Singapore Institute for Clinical Sciences (SICS), A*STAR, Singapore*

ABSTRACT

Researchers are often interested in the effect of treatments or exposures over time. In simple settings of point treatments or where exposures are not expected to vary over time (*i.e.* static), traditional statistical approaches are often sufficient for desired estimands. However, often treatments may affect key mediating covariates which, in turn may affect future treatment. In such settings of time-varying confounding, alternative approaches and estimators are required, particularly in analyses of observational data. This tutorial will introduce g-computation via the parametric g-formula, first introduced by Robins (1986) [1] but more recently made computationally accessible (*e.g.* Daniel, *et al.* 2011 [2]). Worked examples in R using simulated and publically-available data from chronic infectious disease (HIV) and cardiovascular health will be provided. Some support in STATA is also possible. The tutorial will also touch on loss to follow-up, repeated outcome measures, and time-to-event studies, as well as alternative estimators such as inverse-probability weighted marginal structural models and implementation of cross-validated ensemble machine learning (as time allows). A basic knowledge of the counterfactual approach to causal inference (*i.e.* Rubin causal model) is assumed, however a quick refresher will be provided as a matter of background.

Note: Participants wishing to follow along with the code should bring a laptop with wifi / internet capabilities. No pre-installed software will be required, as the project will be shared via RStudio Cloud.

References

[1] Robins JM. A new approach to causal inference in mortality studies with sustained exposure periods - Application to control of the healthy worker survivor effect. *Mathematical Modelling*. 1985. 7:1393-1512.

[2] Daniel RM, De Stavola BL, Cousens SN. gformula: Estimating causal effects in the presence of time-varying confounding or mediation using the g-computation formula. *The Stata Journal*. 2011. 11(4):479-517.

Cognitive interviewing in e- & mHealth research

Andre Matthias Müller

Saw Swee Hock School of Public Health, National University of Singapore, Singapore

ABSTRACT

There is an increased interest in using digital technologies in health-related research. Unfortunately, most researchers do not explore how potential participants interact with and cognitively process the digital content developed for their studies. Commonly, content, features and website- or app-flows are developed based on what researchers think is 'right' and 'easy' to understand. This might not be in line with what actual participants perceive. In this workshop I will introduce the think aloud technique, a cognitive interviewing approach that provides in-depth insights into participants' perceptions and experiences when using e- & mHealth approaches. I will first provide an overview of the technique and explain its general principles. Secondly, participants will learn how to prepare, execute and analyse think-aloud interviews that are aimed at improving comprehension of digital content. Third, we will have break-out sessions during which participants practice the think-aloud interview technique (scenario will be provided). Participants will then share their results and experiences. Finally, I will provide tips and hints to data management and analysis, before concluding the workshop with a discussion on pros and cons of think-aloud interviews.

Note: Please bring a laptop to the workshop for the hands-on activities.

Introduction to causal inference

BaoLuo Sun

Department of Statistics and Applied Probability
National University of Singapore, Singapore

ABSTRACT

This tutorial provides a basic introduction to causal inference under the potential outcomes framework [1, 2, 3]. A fundamental problem of causal inference is that we can observe only one potential outcome for each person [4]; estimation of causal effects can therefore be considered as an attempt to infer properties of the missing potential outcomes. We will introduce some research designs and statistical methods for causal inference in both randomized experiments and observational studies under a single-stage exposure setting. Examples drawn from various fields in the health and social sciences will be used as illustrations.

References

[1] Splawa-Neyman, J. *On the application of probability theory to agricultural experiments. Essay on principles. Section 9.*, Statistical Science 5, pp. 465–472.

[2] Rubin, Donald B. *Estimating causal effects of treatments in randomized and nonrandomized studies.*, Journal of educational Psychology 66(5), p. 688.

[3] Robins, J.M. and Greenland, S. *Causal inference without counterfactuals: comment.*, Journal of the American Statistical Association 90, pp. 122-129.

[4] Holland, Paul W. *Statistics and causal inference*, Journal of the American Statistical Association 81, pp. 945-960.