Fifth Bayesian, Fiducial, and Frequentist (BFF5) Conference: Foundations of Data Science

Hosted by University of Michigan

Sunday, May 6 to Wednesday, May 9, 2018

School of Public Health I (May 6)
Rackham Graduate School Building (May 7-9)

Table of Contents

Conference Information

WiFi: MGuest

Organizing Committee:

Mathieu Bray, PhD Candidate, Department of Biostatistics
Yang Chen, Assistant Professor, Department of Statistics
Emily Hector, PhD Candidate, Department of Biostatistics
Peter X-K. Song, Professor, Department of Biostatistics
Lu Tang, PhD Candidate, Department of Biostatistics
Lu Wang, Associate Professor, Department of Biostatistics

We thank the following sponsors for their support of BFF5:

Departments of Biostatistics & Statistics, University of Michigan
Michigan Institute for Data Science (MIDAS)
University of Michigan Office of Research

Schedule

Sunday, May 6

Room 1690, School of Public Health I

If you have trouble accessing the School of Public Health, please call one of the organizing committee members below for assistance: Mathieu Bray: 734-747-0939; Emily Hector: 734-239-3148; Lu Tang: 434-466-6323

9:00 am - 12:00 pm Short Course on Fiducial Statistics
Jan Hannig, University of North Carolina
(Coffee Break: 10:30 am - 10:45 am)

12:00 pm - 2:00 pm Lunch Break

2:00 pm - 5:00 pm Short Course on Confidence Distributions
Min-ge Xie, Rutgers University
(Coffee Break: 3:30 pm - 3:45 pm)

Monday, May 7

Rackham Graduate School Building, 4th Floor

Conference talks and panel sessions will take place in the Rackham Amphitheatre; Poster sessions will take place in the Assembly Hall; Coffee breaks will take place in the Assembly Hall on Monday, May 7 and Wednesday, May 9, and in the East Conference Room on Tuesday, May 8.

9:00 am - 9:15 am Opening Remarks
Peter Song, University of Michigan
Jack Hu, University of Michigan, **Xiao-Li Meng**, Harvard University

9:15 am - 10:00 am Tutorial on Fiducial Statistics/Confidence Distributions
Jan Hannig & Min-ge Xie
Chair: Yang Chen, University of Michigan

10:00 am - 10:15 am Coffee Break

10:15 am - 11:45 am Invited Session: Revisiting Fiducial Statistics
Thomas Lee, University of California Davis, "*Fiducial Made Sexy*"
Ryan Martin, North Carolina State University, "*Probability Dilution, False Confidence, and Non-additive Beliefs*"
Paul Edlefsen, Fred Hutchinson Cancer Research Center, "*Personalism and Dempster-Shafer Analysis for the 21st Century*"
Chair: Jonathan Terhorst, University of Michigan

11:45 am - 12:30 pm Poster Session

12:30 pm - 2:00 pm Lunch Break

2:00 pm - 3:00 pm Invited Keynote
Roderick Little, University of Michigan, "*On Statistics, Study Design and Data Science: the Prediction and Modeling Cultures*"
Chair: Lu Wang, University of Michigan

3:00 pm - 4:30 pm Invited Session: Foundations of Data Science
Harry Crane, Rutgers University, "*The Fundamental Principle of Data Science*"
H.V. Jagadish, University of Michigan, "*Responsible Data Science*"
Fabio Cuzzolin, Oxford Brookes University, "*Random Sets at the Interface of Statistics and AI*"
Chair: Long Nguyen, University of Michigan

4:30 pm - 4:45 pm Coffee Break

4:45 pm - 5:30 pm Panel Session
Senior Panelists: **Fabio Cuzzolin, Roderick Little, Xiao-Li Meng,**
Junior Panelists: **Philip Boonstra, Yang Chen, Lu Tang**, University of Michigan
Moderator: Georges Monette, York University

5:30 pm - 6:30 pm Poster Session

Banquet

Ballroom, The Graduate Hotel

6:45 pm - 10:00 pm Banquet (Registration Required)
Glenn Shafer, Rutgers Business School, "*So Much Data. Have We Been Here Before?*"
Roderick Little & Bhramar Mukherjee, University of Michigan, *"Skit – To Bayes or Not To Bayes?"*
Chair: Peter Song, University of Michigan

Tuesday, May 8

9:00 am - 10:00 am Invited Keynote
Jim Berger, Duke University, "*BFF Challenges in Dealing with Multiplicities*"
Chair: Seung Yeoun Lee, Sejong University

10:00 am - 10:15 am Coffee Break

10:15 am - 12:15 pm	Invited Session: Advances in Bayesian Statistics **Subhadeep Mukhopadhyay**, Temple University, "*Bayesian Modeling via Goodness-of-fit*" **Radu Craiu**, University of Toronto, "*Adaptive Ideas in ABC-MCMC*" **Dongchu Sun**, University of Missouri, "*Bayesian Analysis of Covariance Matrix of Multivariate Normal Distribution with a New Class of Priors*" **Gunnar Taraldsen**, Norwegian University of Science and Technology, "*Bayesian Machine Learning with Improper Posteriors*" Chair: Lili Zhao, University of Michigan
12:15 pm - 2:00 pm	Lunch Break
2:00 pm - 3:00 pm	Invited Keynote **Nancy Reid**, University of Toronto, "*Some Challenges for Inference*" Chair: Elizaveta Levina, University of Michigan
3:00 pm - 4:30 pm	Invited Session: Developments in Inference **Todd Kuffner**, Washington University in St. Louis, "*Philosophy of Science, Principled Statistical Inference, and Data Science*" **Peter Grunwald**, CWI and Leiden University, "*Safe Probability*" **Jun Zhang**, University of Michigan, "*Geometry of Maximum Entropy Inference*" Chair: Hui Jiang, University of Michigan
4:30 pm - 4:45 pm	Coffee Break
4:45 pm - 5:30 pm	Panel Session: Senior Panelists: **Jim Berger, Radu Craiu, Nancy Reid** Junior Panelists: **Emily Hector**, University of Michigan, **Keli Liu**, Stanford University, **Zhenke Wu**, University of Michigan Moderator: Michael Levine, Purdue University

Wednesday, May 9

9:00 am - 10:00 am	Invited Keynote **Alfred Hero**, University of Michigan, "*Meta-learning: Predicting Performance Limits from Data*" Chair: Michael Elliot, University of Michigan
10:00am - 10:15 am	Coffee Break

10:15 am - 12:15 pm	Invited Session: Additional BFF Perspectives **Donald Fraser**, University of Toronto, "*What Can I Get From Likelihood?*" **Ambuj Tewari**, University of Michigan, "*Low Assumptions Learning Theory*" **Ruobin Gong**, Harvard University, "*Judicious Judgment Meets Unsettling Updating*" **Ling Zhou**, University of Michigan, "*Scalable and Efficient Statistical Inference with Estimating Functions in the MapReduce Paradigm for Big Data*" Chair: Mathieu Bray, University of Michigan
12:15pm - 12:30pm	Closing Remarks **Peter Song, Xiao-Li Meng**

Abstracts

Sunday, May 6 - Short Courses

Short Course on Fiducial Statistics

Jan Hannig, University of North Carolina Chapel Hill

R. A. Fisher, the father of modern statistics, developed the idea of fiducial inference during the first half of the 20th century. While his proposal led to interesting methods for quantifying uncertainty, other prominent statisticians of the time did not accept Fisher's approach as it became apparent that some of Fisher's bold claims about the properties of fiducial distributions did not hold up for multi-parameter problems. Beginning around the year 2000, the authors and collaborators started to re-investigate the idea of fiducial inference and discovered that Fisher's approach, when properly generalized, would open doors to solve many important and difficult inference problems. They termed their generalization of Fisher's idea as generalized fiducial inference (GFI). The main idea of GFI is to carefully transfer randomness from the data to the parameter space using an inverse of a data generating equation without the use of Bayes theorem. The resulting generalized fiducial distribution (GFD) can then be used for inference. After more than a decade of investigations, the authors and collaborators have developed a unifying theory for GFI, and provided GFI solutions to many challenging practical problems in different fields of science and industry. Overall, they have demonstrated that GFI is a valid, useful, and promising approach for conducting statistical inference. The goal of this tutorial is to explain the definition and basic workings of GFI, survey some theoretical results in relation to frequentist inference, and list some related open research problems.

Short Course on Confidence Distributions

Min-ge Xie, Rutgers University

In this half-day short course, we introduce the concept of the confidence distribution, describe its role in statistical foundation, link it to existing inferential approaches (including bootstrap, Bayesian and fiducial methods), and show how it can be applied broadly to solve a wide range of problems in inference, prediction and computing. The confidence distribution (CD) is a sample-dependent distribution function on the parameter space that can represent confidence intervals (regions) of all levels for a parameter of interest. It provides "simple and interpretable summaries of what can reasonably be learned from data (and an assumed model)" (Cox, 2013), and, in turn, meaningful answers to all questions related to statistical inference. In this short course, we review the development of CD, and introduce how to how derive CDs and make inference and predictions using CDs. An emphasis is to demonstrate that CD can yield useful statistical inference tools for many statistical problems where methods with desirable properties have been lacking or not be easily available.

Invited Talks, in order of appearance

Monday, May 7

Fiducial Made Sexy

Thomas Lee, University of California Davis

In this talk we report our on-going work on applying Fisher's fiducial idea to some statistical and machine learning problems, including scalable methods for high-dimensional variable selection and matrix completion. Joint work with Qi Gao, Jan Hannig, Randy Lai and Chunzhe Zhang

Probability Dilution, False Confidence, and Non-additive Beliefs

Ryan Martin, North Carolina State University

In the context of statistical inference, data is used to construct degrees of belief about the quantity of interest. If the beliefs assigned to certain hypotheses tend to be large, not because the data provides supporting evidence, but because of some other structural deficiency, then inferences drawn would be questionable. Motivated by the paradoxical probability dilution phenomenon arising in satellite collision analysis, I will introduce a notion of false confidence and show that all additive belief functions have the aforementioned structural deficiency. Therefore, in order to avoid false confidence, a certain class of non-additive belief functions are required, and I will describe these functions and how they are constructed.

Personalism and Dempster-Shafer Analysis for the 21st Century

Paul T Edlefsen, Fred Hutchison Cancer Center

Personalism, like its more familiar cousins objectivism and subjectivism, is a perspective on the role of the statistician (or more generally, the scientist: "you") in the conduct of science. The differences among these perspectives usually impact little on the daily conduct of statisticians, to our great relief. We are already sufficiently aware of our expected role on a paper or grant to ensure that the science is reproducible. However we are also not unaware of the several difficult issues that we face in statistical science including our contribution to the too-often disappointing performance of phase 3 trials, which results in a high cost for the biomedical enterprise and possibly in lost opportunity for quality human living. There are a host of thorny issues we must face, including how to maximize information yield from research investments while accounting for multiple testing and post-hoc inference. Furthermore, "p>>n" problems (high dimensional covariates with low numbers of observations) arise increasingly often in clinical trials. Can "Fisher's Greatest Blunder" (which is how Brad Efron described Fisher's "Fiducial" methodology) possibly offer an insight into a solution to all of these problems? I will begin by introducing the personalist perspective on the role that "you" play in the scientific process and I will briefly describe Dempster-Shafer (DS) methodology for statistical inference and prediction. I will argue that DS potentially offers a path toward a truly cohesive statistical inference framework for the 21st century.

On Statistics, Study Design and Data Science: The Prediction and Modeling Cultures
Roderick Little, University of Michigan

Leo Breiman's provocative Statistical Science paper in 2001, "Statistical Modeling: The Two Cultures", contrasts machine learning, with its focus on prediction, with more classical parametric modeling approaches to statistics. I am more in the parametric modeling camp, but I appreciate the prediction perspective as yielding a simple and unified approach to problems in statistics – the overarching objective being to predict the things you don't know, with appropriate measures of uncertainty. Philosophically I try to follow the "calibrated Bayes" perspective of Box and Rubin. I discuss this viewpoint, tying it to Rubin's seminal papers on design, and some recent applications to missing data and causal inference.

The Fundamental Principle of Data Science
Harry Crane, Rutgers University

I define the Fundamental Principle of Data Science: "If the data provides sufficient evidence in favor of a proposition A, then inference in A is justified." A "data science" is any systematic protocol for implementing this principle. For example, the standard approach to hypothesis testing, whereby a small P-value is interpreted as sufficient evidence that the null hypothesis is false, is a straightforward illustration of this principle. I will discuss some formal aspects of the Fundamental Principle as well as the implications of this principle for the foundations of data science, with some discussion about the extent to which conventional statistical frameworks (e.g., Bayesian, frequentist, inferential models) live up to this principle.

Responsible Data Science
H. V. Jagadish, University of Michigan

Data Science is having a huge impact on society. As this impact grows, it becomes increasingly important to practice Data science responsibly, to minimize harms that can result while maximizing benefits. It is no longer enough to consider individual steps in the Data Science pipeline: we must take responsibility end-to-end. Concerns include bias in algorithmic decisions, lack of diversity in results, and transparency of the entire process. This talk will describe some of the technical problems in addressing these concerns and practicing responsible data science.

Random Sets at the Interface of Statistics and AI
Fabio Cuzzolin, Oxford Brookes University

Random set theory, originally born within the remit of mathematical statistics, lies nowadays at the interface of statistics and AI. Arguably more mathematically complex than standard probability, the field is now facing open issues such as the formulation of generalised laws of probability, the generalisation of the notion of random variable to random set spaces, the extension of the notion of random process, and so on. Frequentist inference with random sets can be envisaged to better describe common situations such as lack of data and set-valued observations. To this aim, parameterised families of random sets (and Gaussian random sets in particular) are a crucial area of investigation. In particular, we will present some recent work on the generalisation of the notion of likelihood, as the basis for a generalised logistic regression framework capable to better estimate rare events; a random set-version of maximum-entropy classifiers; and a recent generalisation of the law of total probability to belief functions. In a longer-term perspective, random set theory can be instrumental to new robust foundations for statistical machine learning allowing the formulation of models and algorithms able to deal with mission-critical applications 'in the wild', in a mutual beneficial exchange between statistics and artificial intelligence.

So Much Data. Have We Been Here Before?
Glenn Shafer, Rutgers Business School

Statisticians have been dealing with the paradoxes of ever more data since the French Revolution. The bigger the data, the more we know, and the more we think we know. Can we still go to mathematical statistics to understand what we don't know?

Tuesday, May 8

BFF Challenges in Dealing with Multiplicities
Jim Berger, Duke University

In the frequentist approach to handling of multiplicities, one develops a procedure that, e.g., controls for the false positive rate. In the Bayesian approach, one ignores certain multiplicities and deals with others only through the assignment of prior probabilities. These apparently conflict with frequentist analysis, so the BFF challenge is to find common ground. We begin by reviewing two situations in which common ground is thought to be not possible: sequential (or interim) analysis and sequential endpoint testing, where the multiplicities arise through multiple looks at the data and through having multiple decision endpoints, respectively. While the first is a resolvable BFF problem, the second does not appear to be so. We then turn to two multiple testing scenarios where agreement is possible, testing mutually exclusive hypotheses and simultaneous multiple testing. As time permits, we consider additional topics such as variable selection, model averaging, and subgroup analysis, where it is unclear how a BFF consensus could be reached.

Bayesian Modeling via Goodness-of-fit
Subhadeep (Deep) Mukhopadhyay, Fox School of Business - Temple University

The two key issues of modern Bayesian statistics are: (i) establishing principled approach for distilling statistical prior that is consistent with the given data from an initial believable scientific prior; and (ii) development of a consolidated Bayes-frequentist data analysis workflow that is more effective than either of the two separately. In this paper, we propose the idea of "Bayes via goodness-of-fit" as a framework for exploring these fundamental questions, in a way that is general enough to embrace almost all of the familiar probability models. Several examples, spanning application areas such as clinical trials, metrology, insurance, medicine, and ecology show the unique benefit of this new point of view as a practical data science tool.

Adaptive ideas in ABC-MCMC
Radu Craiu, University of Toronto

Approximate Bayesian Computation (ABC) can be a useful computational tool when the data distribution cannot be evaluated, but can be simulated from. An important challenge is to design MCMC algorithms that can make ABC more generally applicable. We will discuss some of the existent ideas along with recent work aimed towards an adaptive MCMC-ABC sampler. Joint work with Evgeny Levi.

Bayesian Analysis of Covariance Matrix of Multivariate Normal Distribution with a New Class of Priors
Dongchu Sun, University of Missouri

Objective Bayesian analysis for unknown covariance matrix of multivariate normal has received a lot of attention in the last two decades. Discovering an unconstrained and statistically interpretable covariance matrix is still an open problem in statistics. In this paper, we proposed a new class of the priors including both inverse Wishart and reference priors as special cases. It is showed that for a modified reference prior, the unknown precision matrix can be estimated even when there are only three observations. A new MCMC algorithm is implemented for the class of priors and is shown to be more efficient comparing with the existing hit-and-run and Metropolis-Hasting algorithms.

Bayesian Machine Learning with Improper Posteriors
Gunnar Taraldsen, Norwegian University of Science and Technology

Improper priors have been criticized based on many different grounds: (i) They are not probabilities (ii) The posterior may be improper (iii) Many undesirable features may appear: marginalization paradox, Jeffreys-Lindley paradox, etc. Improper priors are nonetheless an integral part of current statistical practice and theory: (I) Construction of "non-informative" priors (II) Necessary to obtain a complete class of admissible decision rules.(III) Limits of proper priors. Bayesian machine learning can be considered as machine learning on a higher level. In the simplest case this is exemplified by replacing a prediction by a probability distribution for the prediction to quantify the uncertainty in the prediction. The general Bayesian algorithm is given by combining a model with the prior to obtain the posterior. A key problem is to obtain a prior in cases with many unknown parameters, and in particular so when the parameters have no obvious interpretation as in a deep neural network. We present a mathematical model for statistics based on replacing the axioms of Kolmogorov with the concept of a Renyi space. The main mathematical result is a well-defined conditional expectation in this more general model. The resulting theory includes improper priors and also improper posteriors.

Some Challenges for Inference
Nancy Reid, University of Toronto

At BFF4, David Cox submitted a series of simple examples that focus attention on particular challenges for theories of inference. In this talk I will use these examples to highlight similarities and differences in modes of inference, and try to connect this discussion to more realistic applied problems. Joint work with: Don Fraser

Philosophy of Science, Principled Statistical Inference, and Data Science

Todd Kuffner, Washington University in St. Louis

Statistical reasoning and statistical inference have strong historical connections with philosophy of science. In this talk, the new paradigm of data-driven science is examined through comparison with principled statistical approaches. I will review the merits and shortcomings of principled statistical inference. I will offer some comments on post-selection inference, inference for black box algorithms, and a survey of future challenges.

Safe Probability

Peter Grunwald, CWI Amsterdam and Leiden University

We formalize the idea of probability distributions that lead to reliable predictions about some, but not all aspects of a domain. We obtain a hierarchy of 'safety' notions. On top are valid distributions that can safely be used for all types of predictions. Next come calibrated distributions (a distribution might be wrong yet still calibrated), fiducial (strictly weaker than calibrated) and marginal distributions, and at the bottom level are distributions that only yield 'safe' predictions relative to a single specific loss function. We can formally express notions like 'fiducial distributions are safe for confidence sets with a given stopping rule but not if the stopping rule is unspecified' and start thinking about, for example, extensions of the fiducial idea that can deal with optional stopping. Re-thinking probability distributions in this way allows us to get much of the sting out of many a foundational discussion concerning inference under misspecification, objective Bayes, and the benefits and problems of working with sets of prior distributions (Robust Bayes, 'imprecise' probability).

Geometry of Maximum Entropy Inference

Jun Zhang, University of Michigan

We revisit classic framework of maximum entropy inference. It is well-known that the MaxEnt solution is the exponential family (log-linear model), which can be characterized by the dually-flat Hessian geometry. Here, we provide a generalization to the classic formulation by using a general form of entropy function, which leads to the deformed-exponential family (i.e. generalized linear model) as its solution. The resulting geometry may still be Hessian, but there is an extra degree of freedom in specifying the underlying geometry. Our framework can cover various generalized entropy functions, such as Tsallis entropy, Renyi entropy, phi-entropy, and cross-entropy functions widely used in machine learning and information sciences. It is an elementary application of concepts from Information Geometry.

Wednesday, May 9

Meta-learning: Predicting Performance Limits from Data

Alfred Hero, University of Michigan

The objective of meta-learning to use a random data sample to learn about fundamental limits on performance of a classifier or other statistical inference procedure. The challenge is to do this in the non-parametric setting without having to approximate and implement an optimal classifier algorithm. Meta-learning is important in data science since empirical performance prediction is crucial to optimizing the data life cycle. Examples where meta-learning is applied include sequential design of experiments, reinforcement learning and sensor management in the fields of statistics, machine learning and systems engineering, respectively. We will review the meta-learning problem and highlight recent progress in the context of theory and application.

What Can I Get From Likelihood?

Don Fraser, University of Toronto

There's 'always' been that enigma: How can you put an arbitrary input into the conditional probability lemma and then claim the output is probability? The other input of course is likelihood, so the effective question then is: How can you process likelihood into probability? We consider regular statistical models and show that scalar parameters have a prior that converts likelihood into posterior probabilities and that these are verifiable by routine simulations; and we give an explicit expression for the prior. We also show that vector parameters cannot in general be so converted to probabilities, although you might be lucky, With the widespread use of Bayes this raises professional credibility issues that reflect on the discipline of statistics itself, and the issue shouldn't be swept away by citing cases where there are genuine priors that should be part of the model building. And even then good arguments can be given for separately recording inference results from the separate sources, which might have entirely different background support. Joint work with M Bédard.

Low Assumptions Learning Theory

Ambuj Tewari, University of Michigan

Statistical learning theory assumes both stationarity and independence of the data generating process. How far can we go in developing a formal theory of learning without such assumptions? Can we define interesting notions of learnability even if the data is generated according to an arbitrary stochastic process? In this talk, I will describe some of our recent attempts to define general notions of learnability and to characterize classes of functions that are learnable under these notions. It turns out that uniform martingale laws of large numbers have a key role to play in this area. Joint work with Philip Dawid.

Judicious Judgment Meets Unsettling Updating

Ruobin Gong, Harvard University

Statistical learning using imprecise probabilities is gaining more attention because it presents an alternative strategy for reducing irreplicable findings by freeing the user from the task of making up unwarranted high-resolution assumptions. However, model updating as a mathematical operation is inherently exact, hence updating imprecise models requires the user's judgment in choosing among competing updating rules. These rules often lead to incompatible inferences, and can exhibit unsettling phenomena like dilation, contraction and sure loss, which cannot occur with the Bayes rule and precise probabilities. We explore behavioral discrepancies between the generalized Bayes rule, Dempster's rule and the Geometric rule as alternative posterior updating rules for Choquet capacities of order 2. We show that both the generalized Bayes rule and Geometric rule are incapable of updating without prior information regardless of how strong the information in our data is, and that Dempster's rule and the Geometric rule can mathematically contradict each other with respect to dilation and contraction. Our findings show that unsettling updates reflect a collision between the rules' assumptions and the inexactness allowed by the model itself, highlighting the invaluable role of judicious judgment in handling low-resolution information, and the care we must take when applying learning rules to update imprecise probabilities.

Scalable and Efficient Statistical Inference with Estimating Functions in the MapReduce Paradigm for Big Data

Ling Zhou, University of Michigan

The theory of statistical inference along with the strategy of divide-and-conquer for large- scale data analysis has recently attracted considerable interest due to great popularity of the MapReduce programming paradigm in the Apache Hadoop software framework. The central analytic task in the development of statistical inference in the MapReduce paradigm pertains to the method of combining results yielded from separately mapped data batches. One seminal solution based on the confidence distribution has recently been established in the setting of maximum likelihood estimation in the literature. This paper concerns a more general inferential methodology based on estimating functions, termed as the Rao-type confidence distribution, of which the maximum likelihood is a special case. This generalization provides a unified framework of statistical inference that allows regression analyses of massive data sets of important types in a parallel and scalable fashion via a distributed file system, including longitudinal data analysis, survival data analysis, and quantile regression, which cannot be handled using the maximum likelihood method. This paper investigates four important properties of the proposed method: computational scalability, statistical optimality, methodological generality, and operational robustness. In particular, the proposed method is shown to be closely connected to Hansen's generalized method of moments (GMM) and Crowder's optimality. An interesting theoretical finding is that the asymptotic efficiency of the proposed Rao-type confidence distribution estimator is always greater or equal to the estimator obtained by processing the full data once. All these properties of the proposed method are illustrated via numerical examples in both simulation studies and real-world data analyses.

Poster Abstracts

On a Nonstandard Minimax Theorem

Haosui Duanmu, UC Berkeley

For statistical decision problems with finite parameter space, the minimax value equals to the maxmin value. One usually need to relax the notion of priors to regain such equivalence for infinite parameter spaces. Various extension of this classical result has been established, but they are subject technical conditions. By using nonstandard analysis, we show that, for every statistical decision problem, the standard minimax value equals to the maxmin value over all nonstandard priors. Thus, we obtain a nonstandard minimax theorem under complete generality. Using our nonstandard minimax theorem, we are able to derive a standard minimax theorem on compact parameter space, a finitely additive minimax theorem with bounded
risk functions and a standard minimax theorem on totally bounded parameter space using sequences of priors. Joint work with Robert Anderson (UC Berkeley), Daniel Roy (University of Toronto) and Qiang Sun (University of Toronto).

Safe Testing

Peter Grunwald, CWI and Leiden University

A large fraction (some claim > 1/2) of published research in top journals in applied sciences such as medicine and psychology is irreproduceable. In light of this 'replicability crisis', standard p-value based hypothesis testing has come under intense scrutiny. One of its many problems is the following: if our test result is promising but nonconclusive (say, p = 0.07) we cannot simply decide to gather a few more data points. While this practice is ubiquitous in science, it invalidates p-values and error guarantees. Here we propose an alternative hypothesis testing methodology based on supermartingales - it has both a gambling and a data compression interpretation. This method allows us to consider additional data and freely combine results from different tests by multiplication (which would be a mortal sin for p-values!), and avoids many other pitfalls of traditional testing as well. If the null hypothesis is simple (a singleton), it also has a Bayesian interpretation, and essentially coincides with a proposal by Vovk (1993), and is similar to a proposal by Berger, Brown and Wolpert (1994). We work out the case of composite null hypotheses, which allows us to formulate safe, nonasymptotic versions of the most popular tests such as the t-test and the chi square tests. Safe tests for composite H0 are not always Bayesian, but rather based on the 'reverse information projection', an elegant concept with roots in information theory rather than statistics. Joint work with Wouter Koolen and Rianne de Heide.

Using Synthetic Data to Incorporate External Information into Regression Model Estimation

Tian Gu, University of Michigan

We consider the situation where there is a well-established regression model [Y|X], using a set of commonly available risk predictors X to predict an important outcome Y. A modest sized dataset of size n containing Y, X, and B is available, where B is a new variable that is thought to be important and would enhance the prediction of Y. The challenge is to build a good model for [Y|X,B] that uses both the available dataset and the known model for [Y|X]. One popular proposal in the literature to achieve this is the constrained maximum likelihood (CML) approach, by maximizing the likelihood for [Y|X,B] subject to the constraints on the parameters from [Y|X]. We propose a synthetic data approach, which consists of creating m additional synthetic data observations, and then analyzing the combined dataset of size n+m to estimate the parameters of the model [Y|X,B]. In two special cases we show that the synthetic data approach with large m gives identical asymptotic variance for the parameters of the [Y|X,B] model as the CML approach. This provides some theoretical justification for the synthetic data approach, and given its broad applicability makes the approach very appealing. Joint work with Jeremy M. G. Taylor, Bhramar Mukherjee and Wenting Cheng.

Bayesian Low Rank Matrix Learning with Application to Image Analysis
Yijun Li, University of Michigan

Singular value decomposition (SVD) has been widely adopted for modeling the matrix-variate data with a low rank approximation. Many frequentist methods have been proposed to estimate the "true" rank of the matrix parameter, however there is lack of statistical inference for it. In this work, we resort to a Bayesian SVD (Hoff, 2009) method to make probabilistic inference on the matrix rank to which we assign a discrete prior on a set of pre-specified integer values. We develop an efficient posterior computation algorithm based on importance sampling and Gibbs sampling. The proposed algorithm can be implemented through embarrassingly parallel computing. Simulation studies show that the proposed method achieve a high accuracy in computing posterior probabilities on the matrix rank. We illustrate the proposed method with image data analysis. Joint work with Jian Kang

High Dimensional Inference via Variational Autoencoders.
Jiapeng Liu, Purdue University

We propose a new approach to high-dimensional inference for linear regression model, motivated by the variational autoencoders (VAEs). A two-stage implementation is developed to obtain the estimators of the regression coefficients. First, the high-dimensional linear regression is transformed into a Many-Normal-Means problem, in which we define a vector of auxiliary variables and derive its posterior distribution via VAEs. Next, the distribution of the regression coefficients is obtained by applying an l_0-penalized least square to the posterior distribution of the auxiliary variables in the first stage. An extensive set of experiments are performed, in which we show that our approach outperforms other current methods in both point estimation and inference. Joint work with Yixuan Qiu and Xiao Wang.

Development of a Predictive Model for Deceased Donor Organ Yield
Wesley J. Marrero, University of Michigan

Organ transplantation is a lifesaving intervention for patients with organ failure. However, there is a gap between the supply and demand of organs in the U.S. and worldwide. Optimizing organ yield (number of organs transplanted per donor) is a potentially modifiable way to increase the number of organs available for transplant. Historically, models to predict donor organ yield have been developed based ordinary least squares regression and ordinal logistic regression, however an alternative modeling methodology may be superior to conventional approaches. We aimed to compare different linear and non-linear statistical models to predict deceased donor organ yield. We used data from the OPTN database from 2000-2014 to parameterize our models. We used 80% of the data as a derivation cohort for a cross-validation procedure and the rest of the data as an external validation set. The cross-validation procedure was replicated 50 times and the random hold-outs consisted of 20% of the derivation cohort. Among the models considered, a Bayesian additive regression trees (BART) model resulted in the lowest error on predicting the number of organs transplanted per deceased donor. Two-sample t-tests showed that the BART had significantly lower mean absolute error when predicting deceased donor organ yield (all $p < 0.001$). The BART would improve prediction from at least 63 organs per 1000 donors (compared to an ordinary least squares regression) to at most 120 organs per 1000 donors (compared to an ordinal logistic regression). A model to accurately predict deceased donor organ yield can serve as an aid to assess organ procurement performance and forecast future organ availability. Joint work with Mariel S. Lavieri, Seth D. Guikema, David W. Hutton, and Neehar D. Parikh.

Predicting Mood Using Multivariate Mobile Sensor Data Streams for Medical Interns

Timothy NeCamp, University of Michigan

There is a critical need to understand the temporal dynamics of depression using real-time objective measures. The Intern Health Study (IHS) seeks to identify predictors of depressive symptomsby following a large cohort of medical in- terns and using mobile technology to measure self-reported mood scores (1-10), minute level activity data, nightly sleep time and duration, andheart rate. To improve understanding of depression and inform future interventions, we aim to discover which past variables are most predictive of an intern's mood. We develop a variable selection methodthat is able to determine which variables at which time lags are most predictive of current mood. Our method is advantageous because it: (1) Performs a truncation in order to eliminate irrelevant time lags. (2) Takes advantage of the temporal structure of our data set. Joint work with Srijan Sen, Edward Ionides, and Zhenke Wu.

Exact and Efficient Inference for Partial Bayes Problems

Yixuan Qiu, Purdue University

Bayesian methods are useful for statistical inference. However, real-world problems can be challenging using Bayesian methods when the data analyst has only limited prior knowledge. In this paper we consider a class of problems, called Partial Bayes problems, in which the prior information is only partially available. Taking the recently proposed Inferential Model approach, we develop a general inference framework for Partial Bayes problems, and derive both exact and efficient solutions. In addition to the theoretical investigation, numerical results and real applications are used to demonstrate the superior performance of the proposed method. Joint work with Lingsong Zhang and Chuanhai Liu.

Existence of Matching Priors on Compact Spaces

Daniel Roy, University of Toronto

An exact matching prior at level $1-\alpha$ is a prior such that an associated $1-\alpha$ credible set is also a $1-\alpha$ confidence set. We study the existence of matching priors for general families of credible regions. Our main result gives sufficient conditions for exact matching priors to exist on compact metric spaces, extending recent results for finite spaces. Using our results, we design two families of credible regions that admit matching priors that approximate posterior credible balls and highest-posterior-density regions, respectively, and show that they share some of their most important properties. Our proof of the main theorem uses tools from nonstandard analysis and establishes new results about the nonstandard extension of the Wasserstein metric which may be of independent interest. Joint work with Haosui Duanmu and Aaron Smith.

Non-penalized Variable Selection via Generalized Fiducial Inference

Jonathan P. Williams, UNC Chapel Hill

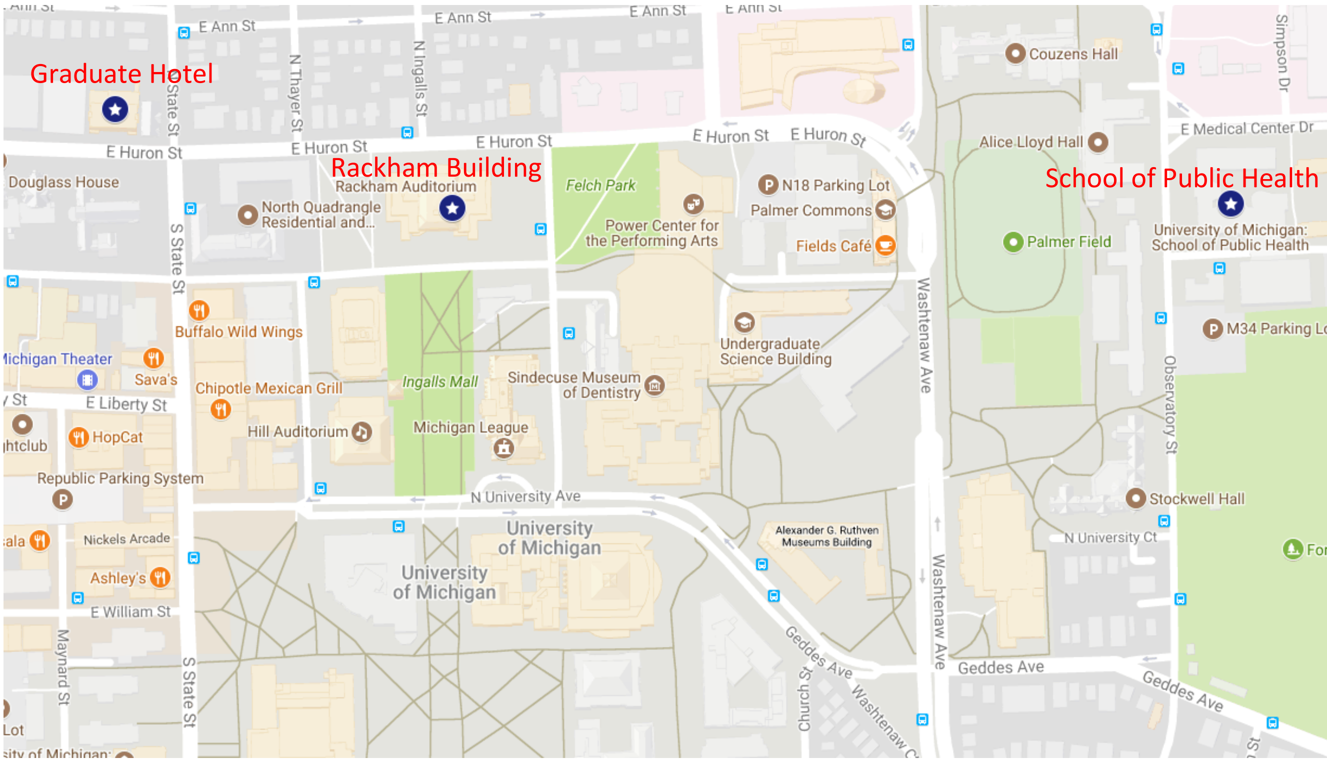
Standard penalized methods of variable selection and parameter estimation rely on the magnitude of coefficient estimates to decide which variables to include in the final model. However, coefficient estimates are unreliable when the design matrix is collinear. To overcome this challenge an entirely new perspective on variable selection is presented within a generalized fiducial inference framework. This new procedure is able to effectively account for linear dependencies among subsets of covariates in a high-dimensional setting where p can grow almost exponentially in n, as well as in the classical setting where $p \le n$. It is shown that the procedure very naturally assigns small probabilities to subsets of covariates which include redundancies by way of explicit L_{0} minimization. Furthermore, with a typical sparsity assumption, it is shown that the proposed method is consistent in the sense that the probability of the true sparse subset of covariates converges in probability to 1 as $n \to \infty$, or as $n \to \infty$ and $p \to \infty$. Very reasonable conditions are needed, and little restriction is placed on the class of possible subsets of covariates to achieve this consistency result. Joint work with Jan Hannig.

Bayesian Variable Selection and Frequentist Post-Selection Inference

Qiyiwen Zhang, Washington University in St. Louis

Frequentist post-selection inference is an increasingly important problem due to the ever-growing popularity of high-dimensional linear regression models and estimators. By contrast, Bayesian procedures which combine variable selection and inference typically do not acknowledge the selection uncertainty in the final inference. Despite this, frequentist repeated-sampling and large-sample properties of Bayesian variable selection, parameter estimation and inference have been deemed crucial for gaining acceptance, even by Bayesian statisticians. We study two popular Bayesian variable selection procedures and their frequentist consistency properties, as well as the relevance of explicit incorporation of selection uncertainty for the final Bayesian inference. Joint work with Todd Kuffner.

Map, Locations and Parking



Short Courses: School of Public Health Building I
1415 Washington Heights, Ann Arbor, MI 48109
Free parking available on Sundays at M28 lot, 1503 Washington Heights, Ann Arbor, MI 48109.

Conference: Rackham Graduate School Building
915 E Washington St, Ann Arbor, MI 48109
On-campus visitor parking available at Palmer Structure, Palmer Dr, Ann Arbor, MI 48109; Additional off-campus parking available at Liberty Square Structure, 510 E Washington St, Ann Arbor, MI 48104, and Republic Parking, 324 Maynard St, Ann Arbor, MI 48104.

Banquet: The Graduate Ann Arbor Hotel
615 E Huron St, Ann Arbor, MI 48104
Hotel provides valet parking; Self-parking available at Liberty Square Structure, 510 E Washington St, Ann Arbor, MI 48104.

Some Recommended Nearby Restaurants:

Sava's (216 S State St)
Mani Osteria and Bar (341 E Liberty St)
Tomukun Noodle Bar (505 E Liberty St)
Tomukun Korean BBQ (505 E Liberty St)
Slurping Turtle (608 E Liberty St)
Zingerman's Delicatessen (422 Detroit St)

Schedule at a Glance

Sunday, May 6

School of Public Health I

If you have trouble accessing the School of Public Health, please call one of the organizing committee members below for assistance: Mathieu Bray: 734-747-0939; Emily Hector: 734-239-3148; Lu Tang: 434-466-6323

9:00 am - 12:00 pm	Short Course on Fiducial Statistics: Jan Hannig
12:00 pm - 2:00 pm	Lunch Break
2:00 pm - 5:00 pm	Short Course on Confidence Distributions: Min-ge Xie

Monday, May 7

9:00 am - 9:15 am	Opening Remarks	Rackham Bldg., 4th Floor
9:15 am - 10:00 am	Tutorial on Fiducial Statistics and Confidence Distributions	
10:00 am - 10:15 am	Coffee Break/Poster Session	
10:15 am - 11:45 pm	Invited Session: Revisiting Fiducial Statistics	
11:45 pm - 12:30 pm	Poster Session	
12:30 pm - 2:00 pm	Lunch Break	
2:00 pm - 3:00 pm	Invited Keynote: Roderick Little	
3:00 pm - 4:30 pm	Invited Session: Foundations of Data Science	
4:30 pm - 4:45 pm	Coffee Break	
4:45 pm - 5:30 pm	Panel Session	
5:30 pm - 6:30 pm	Poster Session	
6:45 pm - 10:00 pm	Banquet	The Graduate Hotel

Tuesday, May 8

Rackham Bldg., 4th Floor

9:00 am - 10:00 am	Invited Keynote: Jim Berger
10:00 am - 10:15 am	Coffee Break
10:15 am - 12:15 pm	Invited Session: Advances in Bayesian Statistics
12:15 pm - 2:00 pm	Lunch Break
2:00 pm - 3:00 pm	Invited Keynote: Nancy Reid
3:00 pm - 4:30 pm	Invited Session: Developments in Inference
4:30 pm - 4:45 pm	Coffee Break
4:45 pm - 5:30 pm	Panel Session

Wednesday, May 9

Rackham Bldg., 4th Floor

9:00 am - 10:00 am	Invited Keynote: Alfred Hero
10:00 am - 10:15 am	Coffee Break
10:15 am - 12:15 pm	Invited Session: Additional BFF Perspectives
12:15 pm - 12:30 pm	Closing Remarks