

# **Fusion Learning & BFF (Bayesian, frequentist and fiducial) Inferences and Statistical Foundations**

**Department of Statistics & DIMACS  
Rutgers University**

*April 11-13, 2016*

*CoRE Auditorium, Busch Campus, Rutgers University*

## **Abstracts**

1. Title: Synthesizing Information and Making Local Conclusions: Multivariate Inference, Multiple Tests, and Not so Many Assumptions  
Speaker: Arne Bathke

Abstract: Recent developments in multivariate statistics allow for the possibility to draw valid inference from data that are not multivariate normally distributed, and, more importantly, don't need to have equal covariance matrices for different groups. Different robust non- and semiparametric approaches are presented, along with procedures maintaining strong error control in selecting key variables and factor levels.

2. Title: The Use of Rejection Odds and Rejection Ratios in Testing Hypotheses  
Speaker: Jim Berger

Abstract: Much of science is (rightly or wrongly) driven by hypothesis testing. Even in situations where the hypothesis testing paradigm is correct, the common practice of basing inferences solely on p-values has been under intense criticism for over 50 years. We discuss, as an alternative, the use of the odds of a correct rejection of the null hypothesis to incorrect rejection. Both pre-experimental versions (involving the power and Type I error) and post-experimental versions (depending on the actual data) are considered. Implementations are discussed that range from depending only on the p-value to consideration of full Bayesian analysis. A surprise is that all implementations -- even the full Bayesian analysis -- have complete frequentist justification. Versions of these techniques can be implemented that require only minor modifications to existing practices, yet overcome some of their most severe shortcomings.

3. Title: A new test for functional one-way ANOVA with application to ischemic heart screening  
Speaker: Ming-Yen Cheng

Abstract: Motivated by an ischemic heart screening problem, we study a new global test for one-way ANOVA in functional data analysis, an increasingly important area in the new era of big data. The test statistic is taken as the maximum of the pointwise  $F$ -test statistic over the interval the functional responses are observed. Nonparametric bootstrap is employed to approximate the null distribution and to obtain an approximate critical value. The asymptotic distribution of the test statistic is derived and the asymptotic power of the proposed test is studied. In particular, under mild conditions, asymptotically our test has the correct level and is root- $n$  consistent in detecting local alternatives. Parallel results are also given when only discretized functional data are available, which is the common case in applications. Simulation studies show that, in terms of both level accuracy and power, the proposed test outperforms several existing competitors when the within-subject correlation is high or moderate, and it is comparable with the latter otherwise. Application to an ischemic heart dataset suggests that resting electrocardiogram signals alone could be used as an effective tool in clinical ischemic heart screening, without the need of further stress tests as required in the current standard procedure.

4. Data-based distributions for unknown parameters: always, sometimes, never?  
Speaker: Sir David R. Cox

Abstract: A brief review is made of the role of formal theory in the analysis and interpretation of data. The possibility of formal probabilistic-like statements about an unknown parameter without the introduction of a prior distribution seems to have been suggested first by R.A. Fisher (1930). Neyman's (1935) formal theory of confidence intervals had quite a different conceptual base. The advantages and limitations of such probabilistic-like statements are discussed with illustrations.

5. Title: Edge exchangeability: a new foundation for modeling network data  
Speaker: Harry Crane

Abstract: Exchangeable models for vertex labeled graphs cannot replicate the large sample behaviors of sparsity and power law degree distributions observed in many network datasets. Out of this mathematical impossibility emerges the question of how network data can be modeled in a way that reflects known empirical behaviors and respects basic statistical principles. We address this question with the new invariance principle of edge exchangeability, which unlike its vertex exchangeable counterpart admits models for networks with sparse and/or power law structure. With this, we settle a longstanding question in statistical network modeling. (Joint work with Walter Dempsey)

6. Title: Bayesian Inference Using Bregman Divergence Measures  
Speaker: Dipak K. Dey

Abstract: Many applications of the Bregman divergence have played a key role in recent advances in machine learning. We turn the spotlight on the Bregman divergence and its applications in Bayesian inference; Bayesian point estimation, Bayesian model selection,

and simultaneous Bayesian estimation and variable selection for sparse high-dimensional data. Since the Bregman divergence includes many well-known loss functions as special cases, our proposed methodologies unify and generalize many existing methodologies in Bayesian decision theoretical framework. Our simulation studies demonstrate the validity and the reliability of the proposed methods. Its generality and superiority is observed in many practical examples with real data and in many parametric statistical inference problems. (Joint work with Gyuhyeong Goh)

7. Rigorizing and Extending the Cox–Jaynes Derivation of Probability: Implications for Statistical Practice

Speaker: David Draper

Abstract: There have been three attempts to date to establish foundations for the discipline of probability, namely the efforts of Kolmogorov (who rigorized the frequentist approach), de Finetti (who gave Bayesian notions of belief and betting odds a formal treatment) and RT Cox/ET Jaynes (who developed a Bayesian theory of probability based on reasonable expectation (Cox) and the optimal processing of information (Jaynes)). The original “proof” of the validity of the Cox–Jaynes approach has been shown to be incomplete, and attempts to date to remedy this situation are themselves not entirely satisfactory. Here we offer a new axiomatization that both rigorizes the Cox–Jaynes derivation of probability and extends it --- from apparent dependence on finite additivity to (1) countable additivity and (2) the ability to simultaneously make uncountably many probability assertions in a logically-internally-consistent manner --- and we discuss the implications of this work for statistical methodology and applications. This topic is sharply relevant for statistical practice, because continuous expression of uncertainty --- for example, taking the set  $\Theta$  of possible values of an unknown  $\theta$  to be  $(0, 1)$ , or  $\mathbb{R}$ , or the space of all cumulative distribution functions on  $\mathbb{R}$  --- is ubiquitous, but has not previously been rigorously supported under at least one popular Bayesian axiomatization of probability. The most important area of statistical methodology that our work has now justified from a Cox–Jaynes perspective is Bayesian non-parametric (BNP) inference, a topic of fundamental importance in applied statistics. We present two interesting foundational findings: (1) Kolmogorov’s probability function  $P_K(A)$  of the single argument  $A$  is isomorphic to a version of the Cox–Jaynes two-argument probability map  $P_{CJ}(A|B)$  in which Kolmogorov’s  $\Omega$  has been hard-wired to coincide with his sample space  $\Omega$ , and (2) most or all previous BNP work has actually been foundationally supported by a hybrid frequentist-Bayesian version of Kolmogorov’s probability function in which parameters are treated as random variables (an unacceptable move from the frequentist perspective); this previous BNP work is methodologically sound but is based on an awkward blend of frequentist and Bayesian ideas (whereas our Cox–Jaynes BNP is entirely Bayesian, which has interpretational advantages). (Join work with Alexander Terenin)

8. Title: The general Univariate Dempster-Shafer Model and its survival analysis counterpart for evaluating HIV-1 vaccine efficacy when censorship is not random.

Speaker: Paul Edlefsen

Abstract: In clinical trials such as those for evaluating the efficacy of vaccines to prevent acquisition of HIV-1, censorship mechanisms include loss-to-followup that is plausibly well-characterized by a mechanism independent of treatment assignment. However in two (of the only six) vaccine efficacy trials for HIV, participants were unblinded mid-study on

recommendation of the monitoring board due to efficacy futility. Recent analyses of unblinded followup of these studies is complicated by potential violations of the typical assumptions that missingness is random and uninformative. In this talk I present the beginnings of an approach to addressing missingness in survival analysis using the "don't know" concept from Dempster-Shafer Analysis. I provide a brief background of the logic and some foundations of Dempster-Shafer inference, introduce the general Univariate DSM for evaluation of iid samples from an unknown univariate distribution, show that this can be transformed to an analysis of survival, and present a proof-of-concept analysis of the vaccine efficacy in the RV144 trial, which is the only HIV vaccine efficacy trial to show even modest evidence for protection.

9. Confidence densities, uninformative priors, and the bootstrap

Speaker: Bradley Efron

Abstract: A multiparameter family of densities  $f(X|\mu)$  yields observation  $x$ , from which the statistician wishes to infer the value of a real-valued parameter  $\theta = t(\mu)$ . For example  $X$  might be multi-variate normal,  $X \sim N(m, V)$ ,  $\mu = (m, V)$ , and  $\theta$  the trace of  $V$ . A statistical holy grail task is to find a convincing posterior density of  $\theta$  given  $x$ , when there is no prior information on the distribution of  $\mu$ . A suite of more or less related answers have been proposed: uninformative priors, matching priors, fiducial methods, and confidence densities (the last being derivatives of confidence distributions.) This talk reviews the various theories, connecting them to bootstrap methods for their implementation.

10. TITLE: Combining One-Sample Confidence Interval Procedures for Valid Non-Asymptotic Inference in the Two-Sample Case

Speaker: Michael P. Fay

Abstract: In this talk, I will present a simple general method for combining two one-sample confidence interval procedures to obtain inferences in the two-sample problem. If we have a monotonic function of the one sample parameters such as the difference, and if the one-sample confidence interval procedures both guarantee coverage and are nested, then our method automatically creates a confidence interval on the function of the two parameters (e.g., the difference) that appears to guarantee coverage. For example, for the two sample binomial problem, two exact Clopper-Pearson confidence intervals for single binomial parameters may be combined by this method to get a confidence interval on the difference in those parameters, and the resulting confidence interval matches inferences using a one-sided Fisher's exact test. The method may be applied to other applications, including a new confidence interval for a difference in medians and a new interval for difference in right censored survival distributions at a fixed point in time. I relate the method to confidence distributions, which are a frequentist framing of some fiducial ideas. (Joint work with Michael Proschan and Erica Brittain.)

11. Title: What can we expect from distributions for parameters?

Speaker: Don Fraser

Abstract: Are all distributions for parameters of the same substance?  
Are some more informative than others?  
Are they accurate summaries of available information?

What can they reliably be used for?

We give a brief history and overview on the preceding questions

12. TITLE: Taking Bayesian Inference Seriously

Speaker: Andrew Gelman

Abstract: Over the years I have been moving toward the use of informative priors in more and more of my applications. I will discuss several examples from theory, application, and computing where traditional noninformative priors lead to disaster, but a little bit of prior information can make everything work out. Informative priors also can resolve some of the questions of replication and multiple comparisons that have recently shook the world of science. It's funny for me to say this, after having practiced Bayesian statistics for nearly thirty years, but I'm only now realizing the true value of the prior distribution.

13. Title: Fast Bayesian Factor Analysis via Automatic Rotations to Sparsity

Speaker: Ed George

Abstract: Rotational transformations have traditionally played a key role in enhancing the interpretability of factor analysis via post-hoc modifications of the factor model orientation. Regularization methods also serve to achieve this goal by prioritizing sparse loading matrices. In this work, we cross-fertilize these two paradigms within a unifying Bayesian framework. Our approach deploys intermediate factor rotations throughout the learning process, greatly enhancing the effectiveness of sparsity inducing priors. These automatic rotations to sparsity are embedded within a PXL-EM algorithm, a Bayesian variant of parameter-expanded EM for posterior mode detection. By iterating between soft-thresholding of small factor loadings and transformations of the factor basis, we obtain (a) dramatic accelerations, (b) robustness against poor initializations and (c) better oriented sparse solutions. For accurate recovery of factor loadings, we deploy a two-component refinement of the Laplace prior, the spike-and-slab LASSO prior. This prior is coupled with the Indian Buffet Process (IBP) prior to avoid the pre-specification of the factor cardinality. The ambient dimensionality is learned from the posterior, which is shown to reward sparse matrices. Our deployment of PXL-EM performs a dynamic posterior exploration, outputting a solution path indexed by a sequence of spike-and-slab priors. A companion criterion, motivated as an integral lower bound, is provided to effectively select the best recovery. The potential of the proposed procedure is demonstrated on both simulated and real high-dimensional data, which would render posterior simulation impractical. (Joint work with Veronika Rockova).

14. New Challenges in Generalized Fiducial Inference

Speaker: Jan Hannig

Abstract: Generalized fiducial inference is an outgrowth of Fisher's fiducial inference and Weerahandi's generalized inference. Over the last 10 years it has been shown to have many attractive properties especially in parametric problems. In this talk we will discuss some recent results and challenges stemming from application of generalized fiducial inference to some classical non-parametric problems.

15. Title: Confidence distributions for change-points and regime shifts

Speaker: Nils Hjort

Abstract: Suppose observations  $y_1, \dots, y_n$  stem from a parametric model  $f(y, \theta)$  with the parameter taking one value  $\theta_L$  for  $y_1, \dots, y_{\tau}$  and another value  $\theta_R$  for  $y_{\tau+1}, \dots, y_n$ . I will provide and examine two different general strategies for not merely estimating the break point  $\tau$  but also to complement such an estimate with full confidence distributions, both for the change-point and for associated measures of differences between the two levels of  $\theta$ . The first idea involves testing homogeneity for the two segments to the left and the right of a candidate change-point value at a fine-tuned level of significance. Carrying out such a scheme requires having a goodness-of-fit test for constancy of the parameter over a segment of indices, and I also develop classes of such tests. The second general method uses the log-likelihood function, profiled over the other parameters, and I show how this may lead to confidence inference for  $\tau$ . The methods are illustrated for a couple of real data stories, with these meeting different types of challenges. (Joint work with Celine Cunen and Gudmund Hermansen)

16. TITLE: Statistical inference theories, multiple uses of the same data, and past-realized data  
Speaker: Benjamin Holcblat and Steffen Grønneberg

Abstract: In many fields, such as nonexperimental fields, test statistics and confidence regions are typically probabilistically dependent on previously examined data. The Bayesian and Neyman-Pearson inference theories are known to be inadequate for such situations. We formalize these inadequacies, and also show that the Neyman-Pearson justification for usual tests and confidence regions is inadequate for past-realized data. Finally, we show that GFI (Generalized Fiducial Inference), which has recently emerged in the statistical literature (e.g., Cisewski and Hannig, 2012; Taraldsen and Lindqvist, 2013), can yield an inference approach that is adequate in these situations. This finding is good news, which supports the famous Efron's conjecture regarding Fisher's fiducial theory of inference "Maybe Fisher's biggest blunder will become a big hit in the 21st century!" (Efron, 1998): GFI is inspired by Fisher's fiducial theory of inference (Hannig, 2009), although there are important differences between GFI and the common interpretations of Fisher's fiducial theory. The presentation is based on the working paper "Statistical Theories, Multiple Uses of the Same Data, and Past-Realized Data," which is available at <http://ssrn.com/abstract=2663452>

17. Title: A Sequential Split-Conquer-Combine Approach for Gaussian Process Modeling using Confidence Distributions  
Speaker: Ying Hung

Abstract: Gaussian process (GP) models are widely used in the analysis of spatial data, computer experiments, and machine learning. However the computational issue that hinders GP from broader application is generally recognized, especially for massive data observed on irregular grids. In this talk, we introduce a sequential split-conquer-combine (SSCC) approach to tackle this problem. This SSCC approach can substantially reduce the computation at the same time provides an estimation result that is asymptotically equivalent to the one obtained from the entire data. Furthermore, the uncertainty of the proposed GP predictor is quantified using confidence distributions. We illustrate the proposed method by a data center example based on tens of thousands of computer experiments generated from a computational fluid dynamic simulator. (Joint work with Chengrui Li and Min-ge Xie)

18. Title: A Problem in Forensic Science? Whose Prior, Whose Bayes Factor, and Who Are You Kidding?  
Speakers: Hari Iyer and Steve Lund

Abstract: Lindley (1979, *Biometrika*, pp 207-213) laid out a rational and coherent statistical framework to evaluate the weight of evidence in forensic science based on a subjective Bayesian modeling of the problem. The approach exploits the odds form of the Bayes rule which says

$$\text{Posterior Odds} = \text{Prior Odds} \times \text{Bayes Factor}$$

or, using the prevalent terminology among forensic scientists,

$$\text{Posterior Odds} = \text{Prior Odds} \times \text{Likelihood Ratio}.$$

The idea is that the forensic expert will limit himself/herself to the calculation of the likelihood ratio (LR) and communicate the result to the court and the triers of fact (TOF) can modify their respective prior odds as to the guilt/innocence of the defendant by applying Bayes rule and arrive at their posterior odds. Thus, it appears that the assessment of the value of evidence by the forensic expert can be separated from subjective impressions regarding the guilt/innocence of the defendant. The approach is currently being evaluated as a candidate framework for adoption in the United States.

Whereas the implementation of the Lindley framework appears to be straightforward, in practice it is fraught with challenges as it involves many subjective choices including choice of priors and of models. While any individual need only be satisfied with their own choices, the act of advising others as to how they may interpret available information carries a greater burden in conveying, at each point, what choice was made and its effect relative to what other choices might have been made instead.

We provide examples that illustrate how complex this process can be even in simple scenarios. In spite of this complexity, it is our view that the evidence interpreter has the responsibility to expend the necessary effort and tell the entire story behind any numerical summary such as the LR and, if justice is to be served, communicate as accurately as one can, the level of subjectivity and uncertainty accompanying such summaries.

19. Title: Sparse Simultaneous Signal Detection and Its Applications in Genomics  
Speaker: Hongzhe Li; University of Pennsylvania

Abstract: The increasing availability of large-scale genomic data has made possible an integrative approach to studying disease. Such research seeks to uncover disease mechanisms by combining multiple types of genomic information, which may be collected on multiple sets of patients. I will focus on a study that integrates GWAS and eQTL data collected from two different sets of subjects to find transcripts potentially functionally relevant to human heart failure. I will first formalize a model that defines important transcripts as those whose expression levels are associated with SNPs that are simultaneously associated with disease. I will then propose a new procedure to test for the existence of these simultaneous signals, show that the test statistic is asymptotically optimal under certain conditions and provide a procedure to obtain finite-sample p-values. I will

apply the proposed test to a heart failure study at Penn (MAGNet) to identify potentially important transcripts that are mechanistically associated with human heart failure. Finally, I will also briefly present a related problem of optimal detection of weak positive dependence between two mixture distributions and show its application in gene set enrichment analysis.

20. Title: Fusion Learning: combining of inferences from diverse sources using data depth and confidence distribution

Speaker: Dungang Liu

Abstract: For the purpose of combining inferences from several nonparametric studies for a common hypothesis, we develop a new methodology using the concepts of data depth and confidence distribution. A confidence distribution (CD) is a sample-dependent distribution function that can be used to estimate parameters of interest. It is a purely frequentist concept yet can be viewed as a “distribution estimator” of the parameter of interest. In this project, we use the concept of CD, coupled with data depth, to develop a new approach for combining several nonparametric studies for a common multivariate parameter. This approach has several advantages. First, it allows us to resample directly from the empirical distribution, rather than from the estimated population distribution satisfying the null constraints. Second, it enables us to obtain test results directly without having to construct an explicit test statistic and then establish or approximate its sampling distribution. The proposed method provides a valid inference approach for a broad class of testing problems involving multiple studies where the parameters of interest can be either finite or infinite dimensional. The method will be illustrated using simulations and flight data from the Federal Aviation Administration (FAA). (Joint work with Regina Liu and Min-ge Xie)

21. Title: Can Big Data Help Us Better Understand Statistical Foundations?

Speaker: Keli Liu

Abstract: It is frequently asked what statistics can do for Big Data but can Big Data return the favor and do something for statistical foundations, specifically, for objective inference? How to define objectivity has long remained a conundrum. Nonetheless statisticians have developed a gauntlet of tests that any contender for the title of objective inference must pass. The challenges brought about by Big Data give us ripe opportunity to add to this gauntlet. For example, the informative nature of “flat priors” becomes startlingly obvious in high dimensional problems. In this talk, we consider the problem of parallelizable inference: the data is so big that analyses can only be run on pieces of the data. We use this setting as a prism to reveal insights into the similarities and differences between Fiducial and objective Bayesian inference. (Joint work with Xiao-Li Meng)

22. Title: On beliefs, validity, and the foundations of statistics.

Speaker: Ryan Martin

Abstract: The foundations of statistics are surely important, but there seems to be no consensus in the literature on what are the goals of these foundations and how they should affect statistical practice. I claim that the primary goal of the foundations of statistics is to provide practitioners with a set of guiding principles which, if followed, will ensure that their analysis is valid. In this talk, I will present some recent progress towards this goal, based on a relatively new approach to inference, centered around the construction of valid

belief functions. Examples will be presented throughout to give the discussion context.  
(Joint work with Chuanhai Liu)

23. Title: Let's Believe Belief Functions: A Paradigm for Multi-resolution Probabilistic Inference  
Speaker: Xiao-Li Meng

Abstract: Since its introduction in 1960s, the Dempster-Shafer (D-S) theory of belief function has inspired many researchers in domains such as signal processing and artificial intelligence, yet significantly fewer in statistics. We believe that this polarized interest is likely due to the theory's apparent reliance on rules beyond ordinary probability calculus, which can be uplifting or upsetting depending on whether one views probability as the fundamental grammar of inference under uncertainty. In this paper we argue that statisticians should not find the D-S theory less acceptable than Heitjan and Rubin's (H-R; 1991, *Annals of Statistics*) modeling of coarse data, because the only essential difference between the two is that the latter operates on a data space with resolution lower than intended, and the former on a parameter space with resolution lower than desired, including at zero, i.e., the state of complete ignorance. Since H-R theory only requires ordinary probability rules, so does D-S theory. In particular the belief of a set  $A$ , as defined in D-S, is the joint probability of the event that both the parameter and its coarsening set are contained in  $A$ . Through recasting belief functions on the joint space using purely ordinary probabilities, we hope to vitalize its statistical interpretability and help facilitate the following:

- Promote the D-S theory in the statistical community by articulating its logic, assumptions and rules in ordinary probabilistic language;
- Propose extensions to existing D-S rules using insights gained from the joint-space formulation;
- Raise the awareness of the versatility of belief functions and its potential contributions to modes of inference, especially in (the ubiquitous) settings calling for partial a-priori knowledge;
- Synergize D-S theory with Bayesian, Fiducial and Frequentist (BFF) paradigms via the multi-resolution framework.

(Joint work with Robin Gong)

24. Title: The Backward Reliability Curve and its practical usefulness  
Speaker: Mounir Mesbah

Abstract: Statistical Validation of Quality of Life Measurements is mainly done through the validation of some specific measurement models relating the observed outcomes to the unobserved theoretical latent construct (the HrQoL variable that scientist aim to assess). Validation of such models, based on goodness of fit (GOF) tests, is not straight forward, mainly because the set of variables involved in the models is partly unobserved. Goodness of fit tests in the latent context still remains an issue.

A second, more controversial, issue in that context is the respective role of models and data. Traditionally in statistical applications, the choice of a model is left to the actual data. A model is validated if the real data support it! A model is validated if it is chosen among the other rival models by the real data!

If we adopt this traditional approach without precautions, we could validate any HrQoL questionnaire, to the delight, of course, of the person who developed it. So, in our context, the right approach is to use a good model, previously selected for its underlying theoretical measurement properties, in order to evaluate the questionnaire with the help of the observations recorded. We have to keep in mind that the goal is to evaluate the instrument (the questionnaire), not to evaluate the measurement model.

For that purpose, the Backward Reliability Curve is very useful. We will show in this talk, how and why it can be used to detect graphically non unidimensional instrument, and other departures from underlying theoretical measurement properties. Some mathematical results are presented. Intensive simulations in various contexts are shown. Application to real data sets from Quality of Life, Education and Genetics are presented.

25. Title: Credibility of Confidence Sets in Nonstandard Econometric Problems

Speaker: Ulrich Müller

Abstract: Confidence intervals are commonly used to describe parameter uncertainty. In nonstandard problems, however, their frequentist coverage property does not guarantee that they do so in a reasonable fashion. For instance, confidence intervals may be empty or extremely short with positive probability, even if they are based on inverting powerful tests. We apply a betting framework and a notion of bet-proofness to formalize the “reasonableness” of confidence intervals as descriptions of parameter uncertainty, and use it for two purposes. First, we quantify the violations of bet-proofness for previously suggested confidence intervals in nonstandard problems. Second, we derive alternative confidence sets that are bet-proof by construction. We apply our framework to several nonstandard problems involving weak instruments, near unit roots, and moment inequalities. We find that most previously suggested confidence intervals are not bet-proof, and numerically determine alternative bet-proof confidence sets.

26. Title: The Spike-and-Slab LASSO

Speaker: Veronika Rockova

Abstract: Despite the wide adoption of spike-and-slab methodology for Bayesian variable selection, its potential for penalized likelihood estimation has largely been overlooked. The first part of the talk bridges this gap by cross-fertilizing these two paradigms with the *Spike-and-Slab LASSO* procedure for variable selection and parameter estimation in linear regression. A new class of self-adaptive penalty functions is introduced that arise from a fully Bayes spike-and-slab formulation, ultimately moving beyond the separable penalty framework. A virtue of these non-separable penalties is their ability to borrow strength across coordinates, to adapt to ensemble sparsity information and to exert multiplicity adjustment. The *Spike-and-Slab LASSO* procedure harvests efficient Bayesian EM and coordinate-wise implementations with a path-following scheme for dynamic posterior exploration. As will be shown on simulated data, the fully Bayes penalty mimics oracle performance, providing a viable alternative to cross-validation. Asymptotic theory for the separable and non-separable variants of the penalty is developed, showing rate-optimality of the global mode as well as optimal posterior concentration when  $p > n$ . Finally, I will introduce the *Particle EM algorithm*, a new ensemble optimization approach for learning multi-modal spike-and-slab posteriors. Particle EM performs efficient posterior mode detection through a set of self-avoiding particles. These particles are geared towards

uncharted areas of the posterior, providing a more comprehensive summary of posterior topography than independently initialized mode hunting algorithms.

27. Title: What does “frequentist” mean?

Speaker: Glenn Shafer

Abstract: The term “frequentist” was apparently introduced by an American philosopher, Ernest Nagel, in 1936 and first used by a British statistician, M. G. Kendall, in 1949. By the 1960s, however, Jerzy Neyman and other non-Bayesians embraced the label. As a result of this history, the word does not have exactly the same meaning for philosophers and statisticians.

28. Title: On the Missing F in BFF

Speaker: Nozer Singpurwalla

Abstract: Lurking in the minds of many in the large communities of computer scientists and natural language processors, lies another F. Recently, a few statisticians and have started to pay attention to it, albeit, using a terminology that appears to be not as provocative as the unexpurgated F. The purpose of this *expository* talk, is to raise awareness to this missing F, and to open the floor for a discussion about its need and merits.

A motivation for the appearance of the said F is the claim that classical probability theory in the framework of Laplace, Bayes, and Kolmogorov, upon which BFF hinges, is unable to address all forms of uncertainty and that alternatives are needed. In this talk we attempt to make the case that the missing F can be made to work in concert with classical probability to enhance the latter's effectiveness.

29. Confidence Estimating Functions

Speaker: Peter Song

Abstract: The method of confidence distribution is a primary method of choice to combine parameter estimates obtained from multiple sources. In the literature of likelihood estimation and inference, many researchers have advocated the use of score functions to make statistical inference (e.g. Rao's score test), which has been shown to be advantageous to methods based directly on the likelihood estimator (e.g. Wald test) in many situations. In a similar spirit, we develop a new approach to constructing the confidence distribution based on the pivotal distributions of unbiased estimating functions. This is different from the currently available method of confidence distribution construction using the asymptotic distributions of estimators. We show that when combining independent data sources by the proposed method of confidence estimating functions, the resulting estimation is equivalent to the estimation given by Hansen's generalized method of moments. Using this analytic connection, we show that the estimator given by the method of confidence estimating functions is optimal in the sense of minimum variance. In contrast, the estimator directly obtained from the combined data cannot achieve the same estimation efficiency, except when the estimating function is the score function. Some numerical examples will be provided to illustrate the proposed method. (Joint work with Ling Zhou)

30. Title: Still Researching on Asymptotic Methods? Try Generalized Inference!

Speaker: Sam Weerahandi

Abstract: At this juncture when superior generalized inferences have been developed for variety statistical problems, developments of asymptotic methods have become redundant for problems involving continuous distributions. If you continue to do so in challenging problems, another researcher will develop generalized tests and intervals that are superior in terms of Type I error, Power, coverage, and so on.

To encourage the proposed transition, an overview of Generalized Tests and Confidence intervals will be given with application to a variety applications ranging from non-normal distributions to inference on Mixed Models. Advantages of the generalized inference over Bayesian inference in the absence of informative priors will also be discussed. Some areas requiring further research on statistical tests based on exact probability statements will be mentioned.

31. Title: We R “BFF” (Best Friends Forever) on road to BFF (Bayesian, Frequentist, Fiducial) inferences  
Speaker: Min-ge Xie

Abstract: Once upon a time, there is a group who traveled to conferences to present their research from different statistical paradigms (Bayesian, Frequentist, Fiducial, plus more). They have grown to like each other’s work (even though they occasionally still secretly like more of their own). They called themselves the traveling circus of BFF (best friends forever). And, thus the journey of this BFF circus began, and the BFF (Bayesian, Frequentist, Fiducial) workshop series was born. The first two were held in Shanghai, now the third in Rutgers, and many more yet to come. It is the dear wish of the BFF circus group to grow its group as well as the BFF workshop in years to come.

In the second and also the technical part, I will introduce the concept of confidence distribution (CD) and articulate the logic behind the CD developments. Furthermore, I will show how CD can potentially bridge probabilistic inferences in Bayesian, frequentist and fiducial (BFF) schools, in almost all aspects including estimation, testing and prediction. If times allows, I will also highlight a few sample projects carried out by Rutgers CD group, to showcase the several added values of the CD development in solving a range of problems, including fusion learning from big data.