BFF4: Fourth Bayesian, Fiducial, and Frequentist Workshop Program

Hosted by Harvard University

Monday, May 1 to Wednesday, May 3, 2017
Event Hall, Student Organization Center at Hilles (SOCH)
59 Shepard Street, Cambridge, MA 02138, USA

Table of Contents

Schedule

All events are located in Hilles Event Hall unless otherwise noted.

Monday, May 1

8:00 am to 8:45 am	Conference registration and coffee
8:45 am to 9:00 am	Opening Remarks **Xiao-Li Meng**, Harvard University
9:00 am to 10:15 am	Featured Discussion: *What Bayes did, and (more to my point) what Bayes did not do* Speaker: **Arthur Dempster**, Harvard University Discussant: **Glenn Shafer**, Rutgers Business School
10:15 am to 10:30 am	Coffee break
10:30 am to 12:00 noon	Invited Session **Ryan Martin**, North Carolina State University, *“Confidence, probability, and plausibility”* **Jan Hannig**, University of North Carolina Chapel Hill, *“Generalized Fiducial Inference: Current Challenges”* **Nanny Wermuth**, Chalmers University of Technology/Gutenberg-University, *“Characterising model classes by prime graphs and by statistical properties”*
12:00 noon to 1:30 pm	Poster session with lunch provided Hilles, 1st Floor
1:30 pm to 2:45 pm	Featured Discussion: *Using rates of incoherence to refresh some old “foundational” debates* Speaker: **Teddy Seidenfeld**, Carnegie Mellon University Discussant: **Christian Robert**, University of Warwick/Paris-Dauphine
2:45 pm to 3:00 pm	Coffee break
3:00 pm to 4:00 pm	Invited Session **Alfred Hero**, University of Michigan, “*Continuum limits of shortest paths*” **Daniel Roy**, University of Toronto, “*On Extended Admissible Procedures and their Nonstandard Bayes Risk*”
4:00 pm to 5:30 pm	Panel: *Views from the Rising Stars* Panelists: **Ruobin Gong**, Harvard University; **Jan Hannig**, University of

	North Carolina Chapel Hill; **Keli Liu**, Stanford University; **Ryan Martin**, North Carolina State University; **Tyler VanderWeele**, Harvard TH Chan School of Public Health Moderator: **Pierre Jacob**, Harvard University
7:00 pm to 9:00 pm	Banquet (registration required): *Risky Business* **Stephen Stigler**, University of Chicago *Hilles Performance Hall, Penthouse Level*

Tuesday, May 2

8:30 am to 9:00 am	Coffee
9:00 am to 10:15 am	Featured Discussion: *The Secret Life of I.J. Good* Speaker: **Sandy Zabell**, Northwestern University Discussant: **Cynthia Dwork**, Harvard University
10:15 am to 10:30 am	Coffee break
10:30 am to 12:00 noon	Invited Session **Vladimir Vovk**, University of London, *“Nonparametric predictive distributions”* **Don Fraser**, University of Toronto, *“Distributions for theta: Validity and Risks”* **Antonietta Mira**, Università della Svizzera Italiana, *“Deriving Bayesian and frequentist estimators from time-invariance estimating equations: a unifying approach”*
12:00 noon to 1:30 pm	Break for lunch
1:30 pm to 2:45 pm	Featured Discussion: *BFF Four--Are We Converging?* Speaker: **Nancy Reid**, University of Toronto Discussant: **Deborah Mayo**, Virginia Tech
2:45 pm to 3:00 pm	Coffee break
3:00 pm to 4:00 pm	Invited Session **James M. Robins**, Harvard TH Chan School of Public Health, “*Counterexamples to Bayesian, Pure-Likelihoodist, and Conditional Inference in Biased-Coin Randomized Experiments and Observational Studies: Implications for Foundations and for Practice*” **Larry Brown**, Wharton School of the University of Pennsylvania, “*Empirical Bayes Prediction under Check Loss*”

4:00 pm to 5:30 pm	Panel: *Perspectives of the Pioneers* **Jim Berger**, Duke University; **Larry Brown**, Wharton School of the University of Pennsylvania; **David Cox**, Oxford University (via remote participation); **Don Fraser**, University of Toronto; **Nancy Reid**, University of Toronto Moderator: **Vijay Nair**, University of Michigan

Wednesday, May 3

8:30 am to 9:00 am	Coffee
9:00 am to 10:15 am	Featured Discussion: *Randomisation isn't perfect but doing better is harder than you think* Speaker: **Stephen Senn**, Luxembourg Institute of Health Discussant: **Ned Hall**, Harvard University
10:15 am to 10:30 am	Coffee break
10:30 am to 12:00 noon	Invited Session **Jim Berger**, Duke University, "*An Objective Prior for Hyperparameters in Normal Hierarchical Models*" **Harry Crane**, Rutgers University, "*Probabilities as Shapes*" **Peter Song**, University of Michigan, "*Confidence Distributions with Estimating Functions: Efficiency and Computing on Spark Platform*"
12:00 noon to 1:30 pm	Break for lunch
1:30 pm to 2:45 pm	Featured Discussion: *Modeling Imprecise Degrees of Belief* Speaker: **Susanna Rinard**, Harvard University Discussant: **Andrew Gelman**, Columbia University
2:45 pm to 3:00 pm	Coffee break
3:00 pm to 4:00 pm	Invited Session **Nils Lid Hjort**, University of Oslo, "*Data Fusion with Confidence Distributions: The II-CC-FF Paradigm*" **Gunnar Taraldsen**, Norwegian University of Science and Technology, "*Improper priors and fiducial inference*"
4:00 pm to 5:30 pm	Panel: *The Scientific Impact of Foundational Thinking* Panelists: **Emery Brown**, Massachusetts Institute of Technology and Massachusetts General Hospital; **Paul Edlefsen**, Fred Hutchinson Cancer Research Center; **Andrew Gelman**, Columbia University; **Regina Liu**,

Rutgers University; **Donald B. Rubin**, Harvard University
Moderator: **Min-ge Xie**, Rutgers University

Abstracts

Featured Discussions, in order of appearance

Monday, May 1

What Bayes Did and What Bayes Did Not Do
Speaker: **Arthur Dempster**
Discussant: **Glenn Shafer**

The basic probe model used for illustration in Thomas Bayes's 1763 posthumous paper can be weakened by removing the controversial prior distribution. The example then becomes a useful illustration of basic features of the Dempster-Shafer (DS) theory of statistical inference. DS provides modeling and reasoning tools for advising users about whether assertions of fact are true or false. The weakening leads directly to relaxing the strict Bayesian requirement of "probabilities of everything", while inferential probabilities p and q that quantify "surely true" and "surely false" are complemented by a new species of inferential probability r that quantifies "neither" option, signaling inadequacy of underlying evidence. Whereas the familiar or ordinary calculus of probability (OCP) has $p + q = 1$, DS theory has $p + q + r = 1$ creating an extended calculus of probability (ECP).

Scientific applications depend on basic choices among competing assumed models. Since these choices are proposed, discussed, and accepted, by analysts, they are personal in nature, so lead to personal probability inferences. Familiar inferential outputs such as significance tests and interval estimates are restated in DS terms. While mathematical and computational aspects of DS often require heavy lifting, much of the abstract framework of the ECP is essentially the same as that of the OCP, the difference being that the former distributes probability components called masses over subsets of a state space, while the latter distributes probabilities directly over singleton elements of the state space. Although unfamiliar to most statisticians, ECP mathematics is more streamlined than OCP mathematics, with only two basic operators, namely, projection and combination. Modern MCMC techniques developed for the OCP can be adapted for use with the ECP.

Using rates of incoherence to refresh some old "foundational" debates
Speaker: **Teddy Seidenfeld** (in collaborated work with Jay Kadane and Mark Schervish -- CMU)
Discussant: **Christian Robert**

This talk summarizes our approach for extending de Finetti's dichotomous criterion of coherent/incoherent (precise) previsions by introducing a family of indices for identifying degrees of incoherence, which apply also in the imprecise case. That is, we argue against a unique index.

We apply this approach to some old debates about Bayesian vs Frequentist statistical practices. I discuss details of one such application, regarding the familiar issue about Fixed-level Testing.

Here are a three related publications, linked to my homepage.
Measures of Incoherence: How not to gamble if you must. Bayesian Statistics 2002 Valencia.
A Rate of Incoherence Applied to Fixed-Level Testing. Phil. Sci. (2002) 69: S248-S264.
What Kind of Uncertainty is that? J.Phil (2012) 516-533.

Tuesday, May 2

The Secret Life of I. J. Good
Speaker: **Sandy Zabell**
Discussant: **Cynthia Dwork**

I. J. ("Jack") Good was an important Bayesian statistician for more than half a century after World War II, and played an important role in the (eventual) post-war Bayesian revival. But his graduate training had been in mathematical analysis (one of his advisors had been G. H. Hardy); what was responsible for this metamorphosis from pure mathematician to statistician?

As Good only revealed in 1976, during the war he had initially served as an assistant to Alan Turing at Bletchley Park, working on the cryptanalysis of the German Naval Enigma, and it was from Turing that he acquired his life-long Bayesian philosophy. Declassified documents now permit us to understand in some detail how this came about, and indeed how many of the ideas Good discussed and papers he wrote in the initial decades after the war in fact presented in sanitized form results that had had their origins in his wartime work. In this talk, drawing on these newly available sources, I will discuss the daily and very real use of Bayesian methods that Turing and Good employed, and how this was very gradually revealed by Good over the course of his life (including revealing his return to classified work in the 1950s).

BFF Four — Are We Converging?
Speaker: **Nancy Reid**
Discussant: **Deborah Mayo**

This talk will give an overview of some of the developments in the BFF series of meetings and related research literature, including confidence distributions, generalized fiducial inference, and inferential models, with particular emphasis on the nature of their probability statements, their calibration properties, and their common elements. I will try to situate the discussion in the context of the current explosion of interest in big data and data science.

Wednesday, May 3

Randomisation isn't perfect but doing better is harder than you think
Speaker: **Stephen Senn**
Discussant: **Ned Hall**

Some criticisms of randomisation can be shown to be both unfair and irrelevant because critics who have voiced them have failed to walk the talk. For example, the criticism that indefinitely many confounders invalidate claims that randomisation balances for covariates would be realized to be false by any critic who went through the exercise of trying to write a simulation to prove it true and would be seen to be irrelevant by considering the probabilistic statements that statistical analyses provide. Claims that allowing patients to choose their own treatment are as valid from the Bayesian point of view as randomisation do not survive the associated calculation if one bothers to do it.

I shall provide theoretical arguments and one disturbing practical example as evidence that although randomisation does not solve all problems of confounding, in trying to do better you can easily do worse.

Modeling Imprecise Degrees of Belief
Speaker: **Susanna Rinard**
Discussant: **Andrew Gelman**

Human agents do not have precise, real-valued degrees of belief. Thus, their belief states are better modeled by a set of probability functions, rather than a single function. However, some have argued that these models lead to decision theoretic incoherence. I defend the set of functions model against this charge. Additionally, I argue against the widespread view that maximal uncertainty is best represented by the maximally wide interval [0, 1].

Invited Talks, in order of appearance

Monday, May 1

Confidence, probability, and plausibility
Ryan Martin

Confidence is a fundamental concept in statistical inference, and there is a tendency to try to relate this notion to probability, despite the warnings given in introductory statistics texts and in the recent literature. In this talk, I will argue that plausibility provides a more appropriate description of confidence than probability, both intuitively and mathematically. Having made this connection, it is natural to ask what insights can be gleaned. For example, what properties of the plausibility function correspond to the desired coverage probability condition? How can a plausibility function with these properties be constructed? The recent work on inferential models (IMs) provides an answer to both of these questions, but suggests a follow-up question: are there good confidence regions/plausibility functions that cannot be reached via IMs? To answer this question, I will present a new complete-class theorem for IMs. The take-way message is that, not only can the IM approach be used, e.g., to construct suitable confidence regions, any other suitable confidence region—Bayes, fiducial, etc.—can be obtained from an IM approach.

Generalized fiducial Inference: Current Challenges
Jan Hannig

R. A. Fisher, the father of modern statistics, proposed the idea of fiducial inference in the 1930's. While his proposal led to some interesting methods for quantifying uncertainty, other prominent statisticians of the time did not accept Fisher's approach because it went against the ideas of statistical inference of the time. Beginning around the year 2000, the presenter 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). After more than a decade of investigations, the presenter 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. In this talk we provide short introduction to GFI. Most of the talk will be devoted to the biggest challenge we see right now: How to incorporate additional vague information such as smoothness and sparseness into the GFI paradigm. It is our hope that the contributions to GFI will stimulate the growth and usage of this exciting approach for statistical inference.

Characterising model classes by prime graphs and by statistical properties
Nanny Wermuth

Graphical criteria to characterise a statistical model class are attractive since they can lead to computationally fast recognitions, especially for simplifying structures of conditional independences. A prominent example are chordal graphs which have single-node elimination schemes so that estimation can be in terms of sequences of single response regressions. Their main disadvantage for inference is that they do not permit to model many types of interventions, those which affect several connected responses at the same time. This is however an important and frequent feature in causal investigations.

With an important old result from graph theory for the so-called prime graphs, another subclass of undirected graphs results, which may contain chordless cycles. These have traditionally be thought of as cumbersome to fit, especially for data with many variables. This subclass of undirected graphs, named hollow trees, has some astonishing features for symmetric binary variables. They have node-set elimination schemes which permit to model joint responses and lead to the equivalence of sequences of logit regressions which are equivalent to linear least-squares regressions.

These results are explained in more detail. They are expected to lead to extensions and to further insights.

Continuum limits of shortest paths
Alfred Hero

Many statistical applications involve computing minimal paths over a graph relative to a measure of pairwise dissimilarity between the nodes of the graph. These include shortest paths in exploratory data mining and non-dominated anti-chains in Markowitz portfolio selection. When the node attributes are random vectors and the dissimilarity is locally monotone in Euclidean distance these minimal paths can have continuum limits as the number of nodes approaches infinity. The continuum limit perspective connects geometry of data to geometry of probability. This perspective can also lead to low complexity variational approximations to the solution of combinatorial minimum path problems.

On Extended Admissible Procedures and their Nonstandard Bayes Risk
Daniel Roy

For finite parameter spaces under finite loss, every Bayes procedure derived from a prior with full support is admissible, and every admissible procedure is Bayes. This relationship already breaks down once we move to finite-dimensional Euclidean parameter spaces. Compactness and strong regularity conditions suffice to repair the relationship, but without these conditions, admissible procedures need not be Bayes. For parametric models under strong regularity conditions, admissible procedures can be shown to be the limits of Bayes procedures. Under even stricter conditions, they are generalized Bayes, i.e., they minimize the Bayes risk with respect to an improper prior. In both these cases, one must venture beyond the strict confines of Bayesian analysis.

Using methods from mathematical logic and nonstandard analysis, we introduce the class of nonstandard Bayes decision procedures---namely, those whose Bayes risk with respect to some prior is within an infinitesimal of the optimal Bayes risk. Among procedures with finite risk functions, we show that a decision procedure is extended admissible if and only if its nonstandard extension is nonstandard Bayes. This result assumes no regularity conditions and makes no restrictions on the loss or model. In particular, it holds in nonparametric models.

For problems with continuous risk functions defined on metric parameter spaces, we derive a nonstandard analogue of Blyth's method that can be used to establish the admissibility of a procedure. We also apply the nonstandard theory to derive a purely standard theorem: when risk functions are continuous on a compact Hausdorff parameter space, a procedure is extended admissible if and only if it is Bayes.

Joint work with Haosui (Kevin) Duanmu. Preprint available at https://arxiv.org/abs/1612.09305

Tuesday, May 2

Nonparametric predictive distributions
Vladimir Vovk

In this talk I will introduce and discuss predictive distributions for the labels of future observations in nonparametric regression that satisfy a natural property of validity in terms of guaranteed coverage for IID observations. Such predictive distributions are well known in Bayesian and parametric statistics but not, to the best of our knowledge, in nonparametric statistics. It turns out that it is sufficient to assume that the observations are produced independently from the same probability measure, without making any further assumptions. One of our specific algorithms, which we call the Least Squares Prediction Machine (LSPM), is motivated by the method of Least Squares in linear regression and generalizes a classical procedure due to Fisher, Dempster, and Hill. If the standard parametric assumptions for Gaussian linear regression hold, the LSPM is as efficient asymptotically as the Fisher-Dempster-Hill procedure; and if those parametric assumptions fail, the LSPM is still valid, provided the observations are IID. If time permits, I will also briefly discuss more flexible and computationally efficient alternatives to the LSPM, satisfying the same property of validity under the general IID assumption. (Based on joint work with Jieli Shen, Valery Manokhin, and Min-ge Xie)

Distributions for theta: Validity and Risks
Don Fraser

We give a brief overview of distributions for y and distributions for theta. Then discuss the general methods available for theta and how they relate to the given statistical model.

Exponential models provide a widely general family for investigation and linear and curved parameters are examined for amenability to the use of parameter distributions; the failure for curved parameters provides a serious risk for such distributions.

In addition likelihood analysis determines where information concerning scalar parameters is located and finds that in some generality the usual Bases procedure can be ineffective in eliciting this information. This raises serious questions concerning the validity of parameter distributions for inference.

Deriving Bayesian and frequentist estimators from time-invariance estimating equations: a unifying approach
Antonietta Mira

Time-invariance estimating equations (Baddeley, 2000) are a recipe for constructing estimators of the parameter of any stochastic model, using properties of auxiliary Markov chains associated with the model. In this paper we extend the time-invariance framework to a Bayesian context. We recover some well known Bayesian estimators and construct new ones. We find some interesting relationships between Bayesian and frequentist estimators and set up a unifying approach. (Collaboration with Adrian Baddeley, Curtin University)

Counterexamples to Bayesian, Pure-Likelihoodist, and Conditional Inference in Biased-Coin Randomized Experiments and Observational Studies: Implications for Foundations and for Practice
James M. Robins

I review the failure of Bayesian, pure-likelihoodist, and conditional Inference in biased-coin randomized experiments and observational studies. I show how unpacking the reasons for their failure leads to inferential approaches that increase efficiency by taking as much as possible from these three paradigms without sacrificing the vastly greater robustness of simple unconditional frequentist methods such as Horvitz-Thompson estimation.

Empirical Bayes Prediction under Check Loss
Larry Brown

Observe *n* independent normal samples, each with their own mean value. The objective is to predict new values for each of the samples under a check-loss error measure. [“Check-loss” is linear in either under- or over-estimation, but with possibly different multiplicative constants for each of the two possible types of error.] This is equivalent to a desire to estimate a pre-specified quantile of the predictive distribution. Our interest is on a big data regime with n large. We develop and use an Empirical Bayes methodology that minimizes a new, uniformly efficient asymptotic risk estimate.

In common with many other problems we find that empirical Bayes shrinkage provides better performance than simple coordinate-wise rules. However, the problem here differs in fundamental respects from estimation or prediction under the quadratic losses considered in most of the previous literature. This necessitates different strategies for creation of effective empirical Bayes predictors. The hyper-parameter estimator we develop involves an appropriate use of Hermite polynomial expansions for the relevant stochastic functions. The results here extend an interesting feature also found in Xie, Kou and Brown (2012, 2016) for estimation under quadratic loss – the strategy to use an (asymptotic) risk estimate has an asymptotic optimality property not shared by E-B strategies that use maximum likelihood or method-of-moments to estimate the hyperparameters.

This is joint work with Gourab Mukherjee and Paat Rusmevichientong
(both at University of Southern California).

Wednesday, May 3

An Objective Prior for Hyperparameters in Normal Hierarchical Models
Jim Berger

Hierarchical models are the workhorse of much of Bayesian analysis, yet there is uncertainty as to which objective priors to use for hyperparameters (parameters at higher levels of the hierarchical model). Formal approaches to objective Bayesian analysis, such as the Jeffreys-rule approach or reference prior approach, are only implementable in simple hierarchical settings (such as the one-way model). Thus it is common to use less formal approaches, such as utilizing formal priors from non-hierarchical models in hierarchical settings. This can be fraught with danger, however. For instance, non-hierarchical Jeffreys-rule priors for variances or covariance matrices result in improper posterior distributions if they are used at higher levels of a hierarchical model. Thus, such less formal approaches must be carefully evaluated, and not just from the perspective of posterior propriety.

Brown (1971) approached the question of choice of hyperpriors in normal hierarchical models by looking at the frequentist notion of admissibility of resulting estimators. The motivation was that hyperpriors that are too diffuse result in inadmissible estimators, while hyperpriors that are concentrated enough result in admissible estimators. Hyperpriors that are `on the boundary of admissibility' are sensible choices for objective priors, being as diffuse as possible without resulting in inadmissible procedures.

In this talk, we propose a particular objective prior for use in all normal hierarchical models, based on considerations of admissibility, ease of implementation (including computational considerations), and performance.

Probabilities as Shapes
Harry Crane

Probabilities are conventionally treated as numbers. But why? What are the essential features of probability? How do numbers capture these features? And in what way are numbers essential for doing so? I explore these questions from the standpoint of interpreting ‘probability’ in terms of ‘evidence,’ so that an assertion is ‘probable’ just in case there is ‘evidence’ for it. I first argue that numbers are not essential for expressing the concept of probability as it arises in much of everyday reasoning. I then propose an alternative view in which probabilities are regarded as abstract shapes. In the interest of exposition and time, I suppress the formalism of abstract shapes and instead focus on the main implications of this approach. As a consequence of trying to recover classical 'numeric' probabilities from the 'probabilities as shapes' framework, I arrive at the Dempster-Shafer axiomatization of belief functions instead of the usual probability axioms. This conclusion can perhaps be seen as further support for the Dempster-Shafer theory as the appropriate formalism for measuring evidence.

Confidence Distributions with Estimating Functions: Efficiency and Computing on Spark Platform
Peter Song

This talk focuses on the generalization of confidence distributions to the setting of estimating functions where the Fisher's likelihood is unavailable. In this case, we consider two types of formulations of confidence distributions: one is based on the asymptotic normality of estimators and the other is based on the asymptotic normality of estimating functions. The former is termed as Wald-type confidence distribution and the latter is termed as Rao-type confidence distribution. The primary goal is to compare the efficiency between these two types of confidence distribution approaches in both finite-sample and large-sample scenarios. The implementation of these two types of inference approaches in the Big Data Spark platform is also discussed along with extensive numerical illustrations.

Data Fusion With Confidence Distributions: The II-CC-FF Paradigm
Nils Lid Hjort

I introduce and develop a general paradigm for combining information across diverse data sources. In broad terms, suppose ϕ is a parameter of interest, built up via components $\psi_1,\ldots,\psi_k$ from data sources $1,\ldots,k$. The proposed scheme has three steps. First, the Independent Inspection (II) step amounts to investigating each separate data source, translating statistical information to a confidence distribution $C_j(\psi_j)$ for the relevant focus parameter ψ_j associated with data source j. Second, Confidence Conversion (CC) techniques are used to translate the confidence distributions to confidence log-likelihood functions, say $\ell_{c,j}(\psi_j)$. Finally, the Focused Fusion (FF) step uses relevant and context-driven techniques to construct a confidence distribution for the primary focus parameter $\phi=\phi(\psi_1,\ldots,\psi_k)$, acting on the combined confidence log-likelihood.

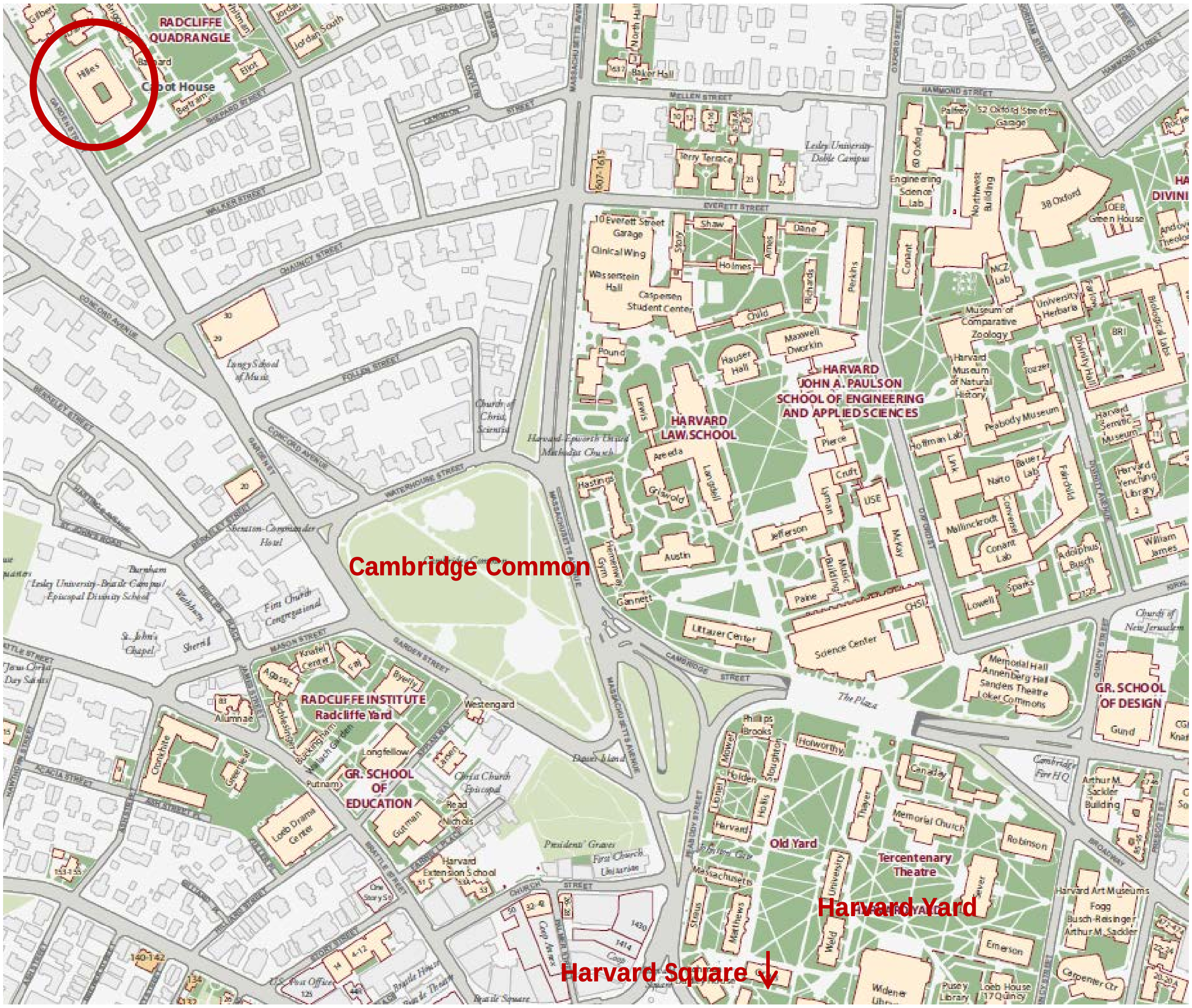
In simpler setups, the II-CC-FF strategy amounts to versions of meta-analysis, but its potential lies in applications to harder problems. It can accommodate both nonparametric and Bayesian components. Illustrations are presented, related to actual applications.

Improper priors and fiducial inference
Gunnar Taraldsen

The use of improper priors flourish in applications and is as such a central part of contemporary statistics. Unfortunately, this is most often presented without a theoretical basis: "Improper priors are just limits of proper priors ... "

We present ingredients in a mathematical theory for statistics which generalize the axioms of Kolmogorov so that improper priors are included. A particular by-product is an elimination of the famous marginalization paradoxes in Bayesian and structural inference. Secondly, we demonstrate that structural and fiducial inference can be formulated naturally in this theory of conditional probability spaces. A particular by-product is then a proof of conditions which ensure coincidence between a Bayesian posterior and the fiducial distribution. The concept of a conditional fiducial model is introduced, and the interpretation of the fiducial distribution is discussed. It is in particular explained that the information given by the prior distribution in Bayesian analysis is replaced by the information given by the fiducial relation in fiducial inference.

Map



Workshop Center: Student Organization Center at Hilles (SOCH) Event Hall (1st Floor) and Performance Hall (Penthouse Level); 59 Shepard Street, Cambridge, MA 02138

Notes