PhD and MS ideas

Ken Rice, kenrice@uw.edu

December 11, 2020

My interests lie mainly in:

- Use of large observational datasets, usually for epidemiology
- Statistical genetics, in particular use of Whole Genome Sequencing data to better understand complex diseases
- Connections between Bayesian and frequentist ideas, particularly better understanding shrinkage
- Meta-analysis

Some recently-completed student work:

- Tyler Bonnett (MS, 2018). Tyler's MS thesis was on *A decision theoretic approach to statistical testing: Connections between frequentist and Bayesian measures of evidence.* He received the 2018 Biostatistics Senior Student Award and the 2018 SPH Outstanding MS Student Award. His thesis is the basis of a recent RSS discussion paper.
- Zora Yang (MS, 2019). Zora's MS thesis was on *Evaluating the Accuracy of Approximate Power and Sample Size Calculations for Logistic Regression.* She is now working in industry but her Shiny App and R package are available, and have considerable potential for applications.
- Ljubomir Miljacic (MS, 2019). Ljubomir's MS thesis was on *Performance of the NCI method of dietary intakes in small sample sizes.* This work, evaluating a complex hiearchical model used in dietary intake studies, will be submitted for publication shortly. He is now awaiting visa clearance to work in the US as a statistical consultant.

Currently I am working with:

- Kendrick Li (PhD candidate). Kendrick has published new methods to improve the small-sample performance of meta-analysis of 2×2 tables. He is working on a heteroskedasticity-robust version of the LASSO, and a Bayesian analog of quasi-likelihood and robust covariance estiamtes, motivated via decision theory.

- Chloe Krakauer (PhD candidate). Following on from Tyler's MS work (above) where Chloe researched connections between Bayesian decision theory and severity, she has performed a substantial review paper on criticism of significance tests – i.e. methods that address whether a familiar "$p < \alpha$?" decision is risky, or not. (This area is remarkably controversial.) She is also addressing the role of decision theory in post-test inference, in particular trying to derive Bayes rules that behave like Zhong and Prentice's elegant "Winner's Curse" correction.

- Spencer Hansen is reading with me on meta-analysis of proportions – in particular, on methods for exact confidence intervals (or close to exact) for the average of several Binomial observations, where the 'success probability' p differs between them. This form of average is very widely-used in practice, but an interval that works well in small samples is not available. We will also look at a proposed method of meta-analyzing confidence intervals, focusing on its efficiency and possibly proposing alternative methods. In some early results, it appears the classical binomial test may be valid for inference on the overall proportion, even under heterogeneity.

Some possible projects:

- Causal inference as a decision problem. Causal inference is generally approached from either a graph-based approach (due to Pearl, e.g. Pearl (2016)) or a potential outcomes (due to Rubin, e.g. Imbens and Rubin (2015)). However, a third theory of causal inference has been developed by Phil Dawid and colleagues (see e.g. Dawid (2000), with discussion, Dawid (2015), Dawid (2020)), in which causal effects are identified via decision-theoretic arguments. The first task in this project is to understand this framework and to use it in some preliminary examples, comparing it to use of DAGs and potential outcomes. The second task is to incorporate recent decision-theoretic advances (e.g. the choice of making No Decision, or the extension of decision losses to also perform loss estimation) into Dawid's framework. **Note:** Niki Petrakos recently started reading with me on this topic.

- Decision theory for sequential trials. In Rice 2010 I show how Wald tests – comparing Z statistics to reference values – have a simple Bayesian analog, that can be motivated using a slightly-modified version of quadratic loss. An extended version of the Wald test is used in sequential trials, where at pre-assigned points in the course of a trial, interim analyses are performed. Based on the Z statistic we may choose to continue the trial, or stop early for efficacy or futility. Beyond carefully controlling the Type I error rate over the duration of the study, the *boundaries* for the Z statistics are typically motivated in an *ad hoc* way, and several families of boundary are available. In this project we will devise loss functions that motivate specific boundaries, and so show by what criteria the various existing boundaries are better, or worse.

- Intervals for Stein estimates. The James-Stein estimate is a landmark result in statistics: its construction showed that intuitively-sensible and optimal estimates of single parameters could not be 'pasted together' to give optimal estimates in higher dimensions. It also shows up the ill-posed nature of frequentist optimality criteria. Constructing uncertainty intervals around it (or indeed around essentially all shrinkage estimates except ridge regression) remains an active area. In this project recent derivations of a Bayesian analog of the Stein positive part estimator as a Bayes rule will be used to construct such

intervals: specifically, it will evaluate the set of points in Θ which contain 95% support for the truth but which, as decisions, are minimally worse than the Bayes rule. As well as direct evaluation of this novel interval, its operating characteristics will be evaluated.

- Admissibility. Recent work by Chloe Krakauer and Kendrick Li uses a loss function that produces point estimate and corresponding variance estimates. It would be of interest here to know whether the famous James-Stein inadmissibility result (see previous topic) still holds. Specifically, is the sample mean and scaled covariance admissible, or not? If it is admissible, this overturns Stein's major result, and would be a major finding. If not, does domination happen if we shrink both the estimate and corresponding covariance, or just one?

- Testing interval null hypotheses. Testing is important in practice, so understanding it better is important to improve statistical practice. While foundational work on testing often revolves around differences between Bayesian and Frequentist methods, this is something of a red herring – for one-sided tests it is well-known that the two paradigms are compatible. For two-sided tests of null hypotheses, the work by Tyler and Chloe (above) also illustrates Bayes/frequentist compatibility, though it requires unusually careful statements of what is being assessed. However, for interval null hypotheses – for example whether $\theta \in [-0.1, 0.1]$, that the parameter lies in some range of values that for practical purposes might as well be zero – the situation is much less clear. In a short-but-classic 1996 article, Mark Schervish shows how (recommended) UMPU frequentist tests give p-values that do not behave coherently, and simply fail as measures of evidence. Bayesian methods are automatically coherent, so should not fail in this way. In this project we will explore the extension of Tyler/Chloe's work to interval nulls. It could be that the UMPU test is insufficiently like *any* Bayesian procedure to ensure coherence. Or that, with careful specification, we can find some Bayesian analog of it does exist, that differs just enough to avoid incoherence. Either way is interesting, and would make a nice follow-up paper to Schervish 1996. Applications to clinical trial-based tests of non-inferiority, and high-throughput genetic screening.

Most of these projects would be appropriate for either an MS thesis or as a chapter in a PhD dissertation. Students considering them should be aware that many of them require creativity. Also – as is usual in this sort of research – there is some risk the result the project looks for is not actually obtainable. If this happens, another project could be required, which of course takes time. Accepting the risk of that may be easier for PhD students than MS students.

The projects are described in terms of statistical methods, but all results would have practical applications, and implications. However, the projects may not be a great fit for students who want their research to be more strongly application-led.

If these projects or anything similar appeals to you, please get in touch with me.