Statistical Science
2019, Vol. 34, No. 2, 206–208
https://doi.org/10.1214/19-STS705
Main article: https://doi.org/10.1214/18-STS674

Comment: Bayes, Oracle Bayes, and Empirical Bayes

Nan Laird

Efron has provided us with an interesting overview of several newer analytical developments for Empirical Bayes (EB) applications. He begins by telling us that empirical Bayes is new, but then immediately acknowledges that it is not so new. This paper makes several points that illustrate this dichotomy. First, there are new statistical methods to improve/sharpen inferences in the empirical Bayes setting. Second, both Bayesians and frequentists can benefit from using these approaches, and finally, the big-data era offers many new possibilities for their application. All of these points mean that we should see a lot more of empirical Bayes in practice. I agree with Efron than this should be true, although I do not feel as sanguine as Efron does. The ability to convert a complex data set into the simple EB framework as described by models 1 and 2 in Efron requires a lot of clever insight (for example, casting FDR as an empirical Bayes approach to multiple hypothesis testing) and we do not have good recipes for that part of the job. In addition, whether or not these techniques are widely accepted still suffers partly from the lack of a clear frequentist or Bayesian identity, partly on having reliable and readily available software, but also on our being able to convince potential users of their advantages, especially if the methods require complex computations and are not easy to explain. Efron's paper makes a lot of progress on all of these fronts.

Fred Mosteller introduced me to empirical Bayes ideas when I was a graduate student. Mosteller is not usually mentioned in the context of empirical Bayes, although his famous work with Wallace on determining the authorship of the disputed federalist papers had a decidedly empirical Bayes flavor (Mosteller and Wallace, 1964). Their method was widely characterized as Bayes, but they used data from papers of known authorship to estimate the "prior odds" for the two authors under consideration for the disputed papers. Their work is another good example of what I would characterize as "clever insight."

The first part of Efron's paper concerning Oracle Bayes has a decidedly frequentist bent and uses the ASE as an optimality criterion. I admit to being a frequentist, because it is generally the most practical, but I cannot get excited about the ASE (Average Squared Error). I can see it is possibly attractive in some settings, but with death rates, cure rates, hospital performance measures, or even gene expression, I find we are more interested in features of the ensemble, such as the extremes, thresholding, or in ordering the θ's. Thus my remarks will focus more on estimation of the mixing or prior density g, and on interval estimates for the θ's, such as those discussed in Efron (1996).

Estimating the prior, or mixing, distribution clearly arises in the EB setting, but also has broader application. Many of the real-life applications I have been involved with are more concerned with estimating g rather than the individual θ's (DerSimonian and Laird (1983)). For example, Mosteller and his collogues, Gilbert and McPeek (Gilbert, McPeek and Mosteller (1977)) were interested in how to quantitatively characterize progress in surgery and anesthesia. They sampled the surgical literature and obtained 13 randomized clinical trials (RCTs), each producing an estimate of the improvement in cure rate of an innovation over a standard therapy. Their objective was to use the estimates from these 13 RTCs to characterize the level of improvement.

As Efron notes, estimating g plays a central role in Bayes empirical Bayes inference, but using NonParametric Maximum Likelihood Estimation (NPMLE) for g is not attractive for this application because of its sparseness. I mention the progress in surgery example because it illustrates that sometimes sparseness is precisely what we want. It also illustrates that in real

Nan Laird is the Harvey V. Fineberg Research Professor of Biostatistics at Harvard School of Public Health, 677 Huntington Ave, Boston Massachusetts 02115, USA (e-mail: laird@biostat.hsph.edu).