Statistical Science
2019, Vol. 34, No. 2, 234–235
https://doi.org/10.1214/19-STS674REJ
Main article: https://doi.org/10.1214/18-STS674

Rejoinder: Bayes, Oracle Bayes, and Empirical Bayes

Bradley Efron

This paper was originally a talk at the 2017 JSM, presenting a personal point of view on the current state of empirical Bayes inference. It is not surprising that the discussants, all of whom have written important papers on the subject, should have different points of view—from each other's and from mine. I don't have serious disagreements with any of these, but rejoinders have the nice property of being unchallengeable, at least in the short run, so I won't pass up my chance to get in a few free shots.

The paper and the commentaries touch on a range of related dichotomies, some of which have dogged discussions of empirical Bayes since its earliest days:

1. Hierarchical Bayes or frequentist empirical Bayes?
2. Omnibus loss functions or individual parameter inferences?
3. g-modeling or f-modeling?
4. Smooth parametric priors or the NPMLE?
5. Relevance considerations or inferences from the full data?
6. Random parameters θ or the compound decision model?
7. Finite sample performance or asymptotics?

Robbins' original work began at the apogee of statistical frequentism, with Bayesian thinking playing a decidedly minor role. Professors Greenshtein and Ritov's comments are fully frequentist (which attracts them to the compound decision framework) while Professor van der Vaart follows the hierarchical Bayes route. My paper tries to have it both ways. I never use hierarchical priors but, in Section 6, I employ Laird and Louis' Type III bootstrapping as a poor man's substitute. Van der Vaart is correct, the Dirichlet prior approach *is* pretty, but that doesn't mean it is right. Uninformative priors aren't guaranteed to produce accurate inferences—see Figure 13.7 of [3]—though here, in expert hands, it is probably fine. (Notice that in Professors Kroenker and Gu's careful calculations Dirichlet priors took more than an hour of computer time, compared to a few seconds for the bootstrap methods; van der Vaart is more optimistic about DP computation.)

Early empirical Bayes work focused on omnibus loss functions, for example, ASE in equation (5) of the paper, equation (1) in Professor Jiang's comments, and (1) in Greenshtein and Ritov. As emphasized in Section 2, omnibus loss favors the frequentist side of empirical Bayes; Section 6 redresses the balance in its "finite Bayes" calculations, where Bayesian ideas are dominant. Both van der Vaart and Greenshtein and Ritov consider empirical Bayes confidence intervals, more individual-parameter than omnibus in nature (though Jiang's confidence interval setup is omnibus). Sections 6 and 7, where individual confidence intervals are examined, were my own favorite parts of the paper.

The results are in for g-modeling vs f-modeling: none of the discussants had much good to say for f-modeling. As Professor Laird points out, f-modeling requires large numbers of perfectly parallel situations for its work, as well as a specialized problem set; g-modeling requires large numbers too, but not necessarily parallel ones. General g-modeling is discussed in [2], including an example incorporating covariate information (answering Greenshtein and Ritov's "plain vanilla" critique). Professor Louis' Section 2 nicely sums up the prosecution's case against f-modeling. And yet, many of the well-known empirical Bayes applications, from the butterflies and Robbins' formula and the baseball players up to false discovery rates, have depended on f-modeling, so it would be premature to banish it from the empirical Bayes toolkit. (I may be over-defensive on this point: my 2011 book used only f-modeling.)

Laird's 1978 paper [8] provided the key NPMLE theorem, while Koenker and Mizera's 2014 paper [7] translated the theory into a practical applied tool. Overall, the discussants' preferences tipped toward nonparametric maximum likelihood estimation (NPMLE) rather than the smooth parametric models in the paper. Koenker and Gu's commentary, a model for the

Bradley Efron is Professor, Department of Statistics, Stanford University, Sequoia Hall, 390 Serra Mall, Stanford, California 94305-4020, USA (e-mail: brad@stat.stanford.edu).