Bayesian Analysis (2015) **10**, Number 2, pp. 501–504

Comment on Article by Dawid and Musio*

Matthias Katzfuss[†] and Anirban Bhattacharya[‡]

The authors consider the interesting and important issue of Bayesian inference based on objective functions other than the likelihood. They focus on model selection in the low-dimensional setting using prequential local proper scoring rules.

1 General non-likelihood-based inference

There is a large and disparate literature on inference based on objective functions other than the likelihood. We will briefly mention some examples here, but we believe that a more thorough review and comparison would be a worthy endeavor.

Numerous objective functions have been proposed to replace the (log-)likelihood in pursuit of various inference goals. Proper scoring rules are a natural choice for serving as such objective functions, due to their property of being minimized (in expectation) under the true model. Depending on the goal of the analysis, certain well-known proper scoring rules can achieve robustness (e.g., continuous ranked probability score, or CRPS), have simple closed-form expressions (e.g., Dawid–Sebastiani score), or do not require densities (e.g., CRPS) or normalizing constants (e.g., Hyvärinen score, as in the present paper). See Gneiting and Katzfuss (2014) for a recent review of these and other scoring rules.

In a frequentist context, examples of approaches falling into this category of scoring-rule-based inference are minimum contrast estimation (e.g., Pfanzagl, 1969; Birgé and Massart, 1993), composite likelihood (e.g., Lindsay, 1988), and M-estimation (e.g., Huber and Ronchetti, 2009). Some further review is given in Dawid et al. (2014).

There have also been related approaches in the Bayesian framework. Shaby (2014) provides a nice review of Bayesian inference using general objective functions and, based on results of Chernozhukov and Hong (2003), he proposes an "open-faced sandwich adjustment" to obtain pseudo-posteriors with properly calibrated frequentist properties. Further, the "Gibbs posterior" (Jiang and Tanner, 2008; Li et al., 2013) has received considerable interest, where the negative log-likelihood is replaced by some "empirical risk" R_n (usually targeting the specific parameter to be estimated) to construct a pseudo-posterior of the form

$$Q(\theta) \propto \exp\{-\lambda R_n(\theta)\}\pi(\theta), \tag{1}$$

where λ is a positive scaling constant (often called "temperature"). Sampling from the pseudo-posterior Q can be performed via standard MCMC algorithms.

*Main article DOI: 10.1214/15-BA942.
[†]Department of Statistics, Texas A&M University, katzfuss@tamu.edu
[‡]Department of Statistics, Texas A&M University, anirbanb@stat.tamu.edu

 DOI: 10.1214/15-BA942A