Bayesian Analysis (2020) **00**, Number 0, pp. 1

Contributed Discussion

Ian H. Jermyn[*] and Karthik Bharath[†]

The authors propose a nice procedure for prior construction using maximum entropy under a constraint on the local 'non-uniformity' of the density (and potentially further constraints such as concavity or monotonicity). Use of the resulting priors in mixture models, model selection criteria, and their (possible) propriety, make them attractive.

We are concerned, however, by the lack of 'invariance' (more accurately, equivariance under diffeomorphisms of parameter space) of the approach. This condition is essential for a well-defined method, and hence for one that claims to be 'objective'. One can see this from several points of view.

1. Mathematically, the lack of equivariance implies that the method is not well-defined as it stands, since different parameterizations will lead to different probability measures on the parameter space, and therefore different posteriors and inferences. Only if a procedure for selecting a distinguished parameterization (or, more precisely, as discussed below, an underlying measure) is specified does the method become well-defined. Such a procedure is not specified in the paper.

2. Example: If the prior density only depends on the parameter space, then the prior density on $\mathbb{R}_+ \times \mathbb{R}_+$ will have the same functional form whether we parameterize a Gamma likelihood with 'shape' and 'scale' or 'shape' and 'rate'. This one functional form corresponds to two different prior probability measures, leading to different posteriors. Which should we use? In a similar but physically motivated example: if two researchers choose to parameterize a model using temperature T or inverse temperature $\beta = 1/T$, both common choices, the parameter space will be $\mathbb{R}_+$ in both cases leading to densities of the same functional form, and therefore different prior probability measures.

3. The lack of equivariance is equivalent to the lack of a well-defined underlying measure against which to define a density. It is well known (and obvious) that the expression for the entropy, I_E, is not well-defined unless the logarithm contains a ratio of p to another reference density m, both of which are defined with respect to an (arbitrary) measure $\mathrm{d}x$, meaning that $m(x)\,\mathrm{d}x$ is the underlying measure with respect to which the density is defined. Similarly, I_F is not well-defined unless p is replaced by p/m.

4. The paper attempts to avoid this issue by invoking Lebesgue measure in Definition 1. We note, however, that Lebesgue measure is itself *not* a well-defined quantity until the additive algebraic (as opposed to topological) structure of $\mathbb{R}$ (or, alternatively, the action of the translation group on $\mathbb{R}$) is defined. This structure is rarely present *a priori*, yet without it, the underlying measure is arbitrary.

[*]Department of Mathematical Sciences, Durham University, UK
[†]School of Mathematical Sciences, University of Nottingham, UK

 DOI: 0000

5. Finally, in a *reductio ad absurdum*, we note that lack of equivariance is the only argument against using the uniform density in a particular parameterization to represent ignorance. If equivariance is no longer a requirement, then this choice is once again on the table. Indeed, the method predicts this: it is easy to see that the *global* minimizer of $I(p)$ (minimizing over boundary conditions as well as densities) is simply the uniform density.

The proposed method, maximum entropy under a constraint on the local 'non-uniformity' of the density, together with further constraints, in the form of boundary conditions or otherwise, might be described as 'objective', in Jaynes' sense (when discussing maximum entropy more generally) of depending on objectively-defined constraints only, were it not for the arbitrariness introduced by the lack of a well-defined underlying measure (equivalently, lack of equivariance to diffeomorphisms of the parameter space). This makes the proposed method dependent on arbitrary choices, which is the opposite of 'objective'. However, procedures for defining such measures do already exist: in particular, group invariance and the use of the likelihood, as in Jeffreys' prior, provide solutions.

The use of the likelihood for this purpose should not be scorned. If we know nothing about a parameter *a priori*, from whence does its connection to reality, its meaning, arise? This can only come from the likelihood connecting the parameter to current data; this is all that remains to define, for example, the difference between temperature and inverse temperature. In this situation, it is not only unsurprising, it is inevitable, that any prior will depend on the model.

The introduction of such a model-dependent measure as the underlying measure for the definition of the entropy and non-uniformity terms in the proposed method would result in a well-defined method with great utility.