Review of “On Bayesian nonparametric modeling of two correlated distributions”

Executive Summary:

This paper is clearly written, publication ready, and could be accepted as a review paper with no modifications. In this reviewer’s humble opinion, this paper is not particularly novel nor is it likely to be particularly impactful – many of the ideas in the paper are very closely related to prior art and the claims are small. That said, smaller increments have been published and the authors commendably write in a simple and straightforward way that is likely to garner the attention of a larger number of readers than the closely related prior art has or will. The largest doubts in this reviewer’s mind revolve around novelty and impact.

Summary of the paper:

This paper considers the problem of modeling two related distributions using a hierarchical Bayesian nonparametric prior. The paper consists of a review of related work, the definition of the simple model and some of its properties, an introduction of MCMC methods for inference (split-merge), and experiments on synthetic and real data.

With observables in two populations or treatments given likelihood Y\_1i ~ F(t\_1i) whose parameter is given a prior t\_1i ~ F\_1’ (replace 1 with 2 for the second population). The hierarchical prior considered is a simplification of the prior art in Muller, Quintana, and Rosner (cited) to two populations (rather than J populations), namely:

F\_1’ = e\_1F\_0 + (1-e\_1)F\_1

F\_2’ = e\_2F\_0 + (1-e\_2)F\_2

with F\_0, F\_1, F\_2 given a shared DP prior with base measure H. Note that atoms will be shared between F\_1 and F\_2 and that, with e\_1,e\_2 > 0, the weights given to those atoms will be correlated. Among other things, this paper characterizes that correlation.

Gripes and misgivings:

What follows here is perhaps my single biggest misgiving with the paper: sections one and two describe a novel normalization model (from Gamma processes) for constructing F\_1’ and F\_2’ which emphasizes the fact that two parameters e\_1, and e\_2 (condensed into a single parameter for most of the paper e) are used to control the mixtures of F\_0 and the population specific distributions F\_1 and F\_2 (so as to differentiate this model from that of Muller et al which given a superficial reading most readers would identify as the same). Then, surprisingly, the paper discards the two parameter “normalization model” (sans the particulars of the method of constructing the model, the apparent significant novelty of this paper) for the single e model (Muller et al’s model) until page 14 when an MCMC sampler for the novel model is given. Even more surprising is the apparent skipping of the “novel” model in the experiments section except for lines 43-46 of page 18 in which a textual description of the normalization model’s sampler performance is given and one paragraph (and a figure) starting on page 22 on the novel normalization model in an applied setting where its value and interpretation is at best sketchy.

Page 9, lines 52 onward is the description of the difference between this and the prior art: “the construction method used here is a more systematic one, and induces some nice properties.” What? This reviewer respectfully suggests that the distinction between the model presented in this paper and models appearing in prior art should be made more clear and that, if the novelty of the model (rather than simply a nicer construction) is to remain a claim in the paper, substantial experimental evidence of the importance of the extra parameter should be provided. If the distinction is merely in the method of constructing the model and all other differences are relatively trivial, this reviewer respectfully suggests that this should be straightforwardly admitted. It is the opinion of this reviewer that the paper should be accepted on the construction and write-up alone – its simplicity and clarity is refreshing. This reviewer spent the bulk of his time reviewing the manuscript trying to figure out what was novel and what was not. The only conclusion drawn about this was that there seems to be little novel about the model (or even inference), but reasonable novelty in terms of construction methodology, etc.

Suggestions for improvement (written in the imperative, but to be taken as suggestions):

Adding graphical aides (graphical models for instance, like those in Muller et al) would help readers substantially.

A mention of non-conjugate models should be included.

The additional split/merge step in the normalization model (page 15, lines 42-end) should be given the same kind of extended explanation as was given the earlier split merge steps.

An actual explanation of why the posterior mass for the weight is assigned to the minimum of weights (page 6, line 55) should be given. The referenced prior art does a poor job of explaining this.

Many of the titles, fonts, and so forth on the included figures need to be fixed.

The hospital efficiency data study could be omitted without loss in this reviewer’s opinion. Ideally it would be replaced with a dataset whose effects were clearer and for which the e\_1, e\_2 extra parameterization would be clearly of greater value.