



Discussion of 'Prior-based Bayesian information criterion (PBIC)'

Jan Hannig

To cite this article: Jan Hannig (2019) Discussion of 'Prior-based Bayesian information criterion (PBIC)', *Statistical Theory and Related Fields*, 3:1, 30-31, DOI: [10.1080/24754269.2019.1611144](https://doi.org/10.1080/24754269.2019.1611144)

To link to this article: <https://doi.org/10.1080/24754269.2019.1611144>



Published online: 30 Apr 2019.



Submit your article to this journal [↗](#)



Article views: 17




View related articles [↗](#)



View Crossmark data [↗](#)



Discussion of 'Prior-based Bayesian information criterion (PBIC)'

Jan Hannig 

Department of Statistics and OR, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

ARTICLE HISTORY Received 6 April 2019; Accepted 22 April 2019

KEYWORDS Generalised fiducial inference; non-local prior; likelihood principle

The authors should be congratulated on a stimulating piece of work. There is definitely a great need to understand information-based criteria such as BIC better. I especially enjoyed the way they simplified the problem through LaPlace approximation allowing for closed form calculations. This made me remember fondly some old papers of R. A. Fisher (1922). However, as any good academician, I have a few comments:

First, I do not believe that the name of the method is very descriptive. When I first read the title of the paper, I thought that it is going to be about introducing a user selected prior information into the BIC. Instead, the proposed PBIC criterion is based on a very particular prior. This prior is flat for parameters that are common to all the models, and Cauchy-like (Equation (9)) for orthogonal transformation of parameters under selection. I view this both as the biggest strength and biggest weakness of this paper. On one side, the fact that classical BIC does not depend on the prior selected can be viewed as its big advantage. There is no room for the user to doubt the compatibility of a particular prior with the problem considered. Consequently, it also makes it easier for BIC to be accepted in the frequentist community. On the other, the particular prior selected for PBIC, allows for more detailed calculations. This leads to a more detailed formula, as compared to BIC, accounting for different precision in different parameters. In any case, I believe the name should reflect this specific choice of prior explicitly. What about *Cauchy prior-based Bayesian Information Criteria* (CBIC)?

Second, PBIC use of a particular prior opens it to criticism of the prior selected. In particular, Johnson and Rossell (2010) argue that for model selection it is better to use non-local priors, i.e. priors that do not put mass near the origin. One could implement this by

replacing Equation (9) with

$$\pi_i^R(\xi) = \int_0^1 \frac{\xi^2}{(d_i + b_i)/(2\lambda_i) - d_i} N \times (\xi \mid 0, (d_i + b_i)/(2\lambda_i) - d_i) \frac{1}{2\sqrt{\lambda_i}} d\lambda_i.$$

Similar calculation as those made in the paper lead to

$$m(\mathbf{x}) = e^{l(\hat{\theta})} \frac{|\hat{I}|^{-1/2}}{|D|^{1/2}} \prod_{i=1}^p \frac{\sqrt{2}((1 + n_i^e)v_i - 3 + e^{-v_i})}{(3 - (-2 + n_i^e)v_i - (-1 + n_i^e)v_i^2)} \times \frac{1}{(n_i^e + 1)^{3/2}v_i^2} \times (1 + o(1)). \quad (1)$$

While (1) is somewhat longer than PBIC, it is just as intuitive. Would it make sense to consider this or some other similar formula?

Third, as discussed in the paper, the selection of n_i^e is the weakest point. Unfortunately, I could not come any more clever, less ad-hoc way to do this than the authors.

Finally, I would like to ask the authors if they believe there is a room for a fiducial information criterion.

Disclosure statement

No potential conflict of interest was reported by the author.

Funding

Jan Hannig's research was supported in part by the National Science Foundation [grant number 1512945 and 1633074].

ORCID

Jan Hannig  <http://orcid.org/0000-0002-4164-0173>

References

- Fisher, R. A. (1922). On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 222, 309–368.
- Johnson, V. E., & Rossell, D. (2010). On the use of non-local prior densities in Bayesian hypothesis tests. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 72, 143–170.