Review of Hybrid Bayesian Inference and Its Applications

Ao Yuan¹, Xi Zhang² and Gang Han³✉

¹Department of Bioinformatics and Biostatistics, Georgetown University, USA
²Department of Industrial Engineering and Management, Peking University, China
³Department of Biostatistics, Yale University School of Public Health, USA

Abstract

Incorporating both Bayesian and frequentist methods in the same model the hybrid Bayesian statistical inference possesses some advantages over either of the methods used individually. It has been increasingly applied to biomedical research including genomics, viral load modeling, biomedical engineering, and survey. In this article we briefly describe the theoretical framework of hybrid Bayesian inference and review its applications. We also discuss the potential future research topics in this area.

INTRODUCTION

The Bayesian and frequentist methods each can be more favorable in practice, and the user normally needs to choose one or the other method. In some situations, however, only some, not all, of the parameters have prior information. Such examples arise in genomics [1-3], correlated data modeling [4,5], and the design and analysis of computer experiments [21]. Motivated by the need to incorporate Bayesian and frequentist parameters in one statistical model and simultaneously make inference about all model parameters, [6] developed a theoretical framework for the Bayesian-frequentist hybrid inference, which is also referred to as hybrid Bayesian inference in this article. Thereafter a growing number of research papers have been published to extend the theory or to apply the inference methodology in different research fields. Within the analysis of various types of data, this approach demonstrates the benefits such as assimilating all prior information and obtaining less small sample bias compared with the alternative hierarchical Bayesian analysis.

In this article we give a short summary of the methodology and its recent applications. We also discuss the future research potentials of the Bayesian hybrid inference. In Sections 2 and 3 we review the hybrid Bayesian theoretical developments and applications in various research fields, respectively. Discussion about some conceptual and computational issues and concluding remarks are given in Section 4.

THE HYBRID BAYESIAN THEORY

Motivated by scenarios where in the same model some of the parameters are favored by frequentist statistical inference while others by Bayesian inference, [6] proposed a simultaneous inference approach to combine these two types of parameters. Unlike the profile likelihood approaches [7,8] and other approaches compromising the Bayesian and frequentist estimation [9,10], the hybrid Bayesian inference searches the joint optimum, i.e., making inference about the Bayesian parameters based on generalized Bayes rule and making inference about the frequentist parameters by maximizing the likelihood [6], derived the first order and second order properties of the hybrid Bayesian inference and proved that the computation of the hybrid inference is similar to that of the maximum likelihood estimate (MLE) under the 0-1 loss. Although the solution to the hybrid Bayesian estimation usually has no closed form and the computation is time consuming (comparable to the full Bayesian estimation), the theoretical results in [6] can greatly simplify the computational algorithm under the proper choice of the loss function.

As an extension to the framework in [6,11] studied the hybrid model under the 0-1 loss, and proved that under general conditions, the asymptotic variance of the Bayesian parameters under the 0-1 loss in the hybrid inference is the same as that of the MLE, and both estimations are efficient, which makes it similar to the hybrid inference under the general convex loss. The work in [12] indicates a close relationship to the hybrid Bayesian framework in [6], for testing statistical hypotheses in that: (1) allowing partial knowledge of the prior for a parameter; and (2) constructing a blended posterior to form the Bayes factor in consideration of the Bayesian and frequentist aspects.

APPLICATIONS

We discuss the extensive applications of the hybrid Bayesian inference by research areas. In genomics, [1] applied the hybrid Bayesian inference to analyze the gene copy number changes, where a mixture model is used to infer three status of each gene. In [1], decent prior information about the gene effects was available, whereas there was not such prior information for the mixing proportions and model variances. Thus the gene
effects were modeled using Bayesian approach, and the mixing proportions and model variances modeled using the frequentist approach. [13] constructed empirical Bayes confidence limits as an adjusted local false discovery (IFDR) rate when the number of hypotheses is relatively small compared to normal microarray study settings for applying the IFDR [14]. [2] developed a hybrid multiple testing procedure based on the Bayesian hybrid inference, where the posterior interval is negatively influenced by the unreliable prior and the fixed-parameter confidence intervals lead to excessive IFDRs when the proportion of true null hypothesis is high. [3] employed the hybrid model when studying the genetic liability, in which the prior knowledge about the disease prevalence odds is available, but not available for the other parameters. Also, the prior information in [3] grows along with the current sample size so that the setting of [15] was applied, and the prior information is asymptotically retained to achieve a higher efficiency.

The hybrid Bayesian inference was also applied to mixed effects models and correlated data analysis. One area is to model the HIV viral load based on the CD4 and CD8 cell counts, where the non-linear mixed effects model involves a number of parameters. Some of the parameters have prior knowledge based on previous studies but others are not [4] applied the hybrid Bayesian inference framework and a Monte Carlo EM algorithm [17,18] to model the dynamic viral load using CD4 cell counts at baseline and the follow-up visits in 6 months. [19] applied the hybrid Bayesian method to the random effects model with longitudinal data, where the prior information on some of the regression parameters was available, but not for all the parameters. The inference results were compared to those of the classical MLE, and the proposed inference showed improvements under scenarios of small sample sizes.

Another area of application is the missing data analysis for test data or survey. [5] applied the hybrid inference method to test data (GPA or TOFEL) with missing values in blocks. In their linear regression model, coefficients of some attributes required frequentist estimation while coefficients of other attributes were found to prefer Bayesian methods with estimated priors. Under the condition that some of the covariates were missing in their data, they used the non-missing covariates and the corresponding responses to construct the prior distribution for part of the regression parameters and applied the 0-1 loss function to compute the hybrid estimator. Similar to other applications [1,4,5] used 0-1 loss function and derived the asymptotic properties of the final estimates.

Another application of the hybrid Bayesian approach is the simultaneous determination of calibration and tuning parameter for computer experiments [20,21] proposed a model for complex computer experiments where calibration parameters (unknown or unmeasured quantities in the physical experiments) were estimated based on a hierarchical Bayesian approach and the tuning parameters (values for setting the computer simulator) were determined by minimizing the L2 discrepancy between the computer output and the unknown true physical response. This is essentially the hybrid Bayesian inference if minimizing the L2 discrepancy is equivalent to maximizing the likelihood. [21] applied this approach to the optimal design of knee prosthesis and showed that this hybrid approach could lead to more stable and accurate estimates than the hierarchical Bayesian approach.

We finalize this section by describing two applications of the Hybrid Bayesian inference in practice:

**Example 1: Gene copy number change analysis.** In this analysis, the comparative genomic hybridization (CGH) data is collected. In practice, the true copy number changes cannot be observed directly, but rather, with a CGH data, in which the fluorescence ratios between two samples, case and control, are measured across a genomic region. For loci with copy number deletion/amplification, the corresponding log-ratio measurement tends to be lower/higher. We are to answer the question of what is the probability that a given gene or region has increased or decreased copy number changes. In CGH data analysis, often a three-state mixture model is used: deletion state, normal state and amplification state, and we arbitrarily label them as state 1, 2 and 3. Genes with copy number deletion tend to have smaller log-ratio measurements, those with normal status tend to have median measurements, and those with amplification tend to have larger measurements. The goal is to classify each observed data into one of the 3 states. Without knowing these state memberships, the density for the data can be specified as a 3-component mixture model, in which there are three sets of parameters: the means of the 3 states sub-model, the corresponding variances, and the mixing proportions. Commonly for this problem, we have or can find good prior knowledge on the means from existing studies, as summarized by a prior density. But often we don't have enough experience for the parameters of mixing proportions and variances. One may put a non-informative prior on these parameters to formulate a full Bayesian model. But even a non-informative prior can have misleading effects for small sample size, and it is known that if the prior is too “thin” around true parameter, the Bayes estimate can be inconsistent [22]. Also we will see below that by using the hybrid model, the parameter estimation can be formulated as a EM algorithm, which typically simplifies the computation considerably for mixture model as compared to a full Bayesian model. Thus we use a hybrid model with Bayesian estimate for the means and the MLE for the mixing proportions and variances. To simplify computation, we use the 0-1 loss on the means, the corresponding Bayesian solution is the posterior mode, thus the hybrid estimator is formulated as the maximizer of the augmented likelihood, with computation parallels that of the full MLE, and much simpler than the classical full Bayesian estimator, in which, the estimator is often computed via the Markov chain Monte Carlo, a computation intensive method, and often subject to additional sampling error. However, it is known that parameters in mixture model are not easy to compute directly, no matter what estimation method used. But if the state status of each observation is known, then the corresponding “complete data” model is easy to fit. For this reason, we treat the status as missing data, and use the EM algorithm to compute the hybrid estimator in the model. For details of the model and computation, see [1].

**Example 2: HIV viral dynamic modeling:** HIV viral dynamic modeling is critical in understanding the interaction between CD4 cell count and viral load over time to evaluate the HIV treatment effect. [4] implemented hybrid Bayesian inference to
make inference about the parameters in the longitudinal mixed effects model. We use this example to illustrate the computation procedure of the hybrid Bayesian inference for mixed effects. Under the 0-1 loss, their computation is based on a Monte Carlo EM algorithm with the following steps:

[Step 1.] Set the initial values of all model parameters and set the number of iteration.

[Step 2.] For the current parameter values, perform the expectation step of the EM algorithm incorporating the prior information using the Monte Carlo samples.

[Step 3.] Maximization: search the parameter values to maximize the expectation of the conditional marginal density [4], Equation 11.

[Step 4.] Repeat Steps 2 and 3 at each iteration until the estimates converge.

**DISCUSSION AND CONCLUSION**

The hybrid model bridges the two major statistical methodologies, the frequentist and the Bayesian. A few concepts may be questioned in practice and thus the clarification is necessary: Some may argue that by placing an informative prior on the part of the parameters and a non-informative prior on the rest parameters, one can perform a full Bayesian inference, instead of the hybrid inference. However, even a non-informative prior may have misleading effects under small sample size scenarios [6], and it is known that if the prior is too “thin” around true parameter, the Bayes estimate might be inconsistent [22]. In our view a hybrid model is better in this regard.

Another argument is that without an objective prior, placing a uniform prior on the part of parameters and using a 0-1 loss will lead to a full Bayesian model, which is the same as the hybrid model, and consequently the hybrid model is essentially Bayesian. We first point out that the classical frequentist MLE is a Bayesian procedure in this view: it is the Bayes estimator under some other losses, its computation complexity is approximately the same as that of the Bayes estimator. However, if we use the 0-1 loss, the computation is parallel to that of the MLE, which is known to be much simpler than the Bayes estimator under some other losses, and the corresponding likelihood has the interpretation of an augmented likelihood by the prior density. The same discussion applies when the hybrid concept is extended to other settings, such as the M-estimation and estimation equations.

**ACKNOWLEDGEMENTS**

We give thanks to Mr. Changyue Song for editorial assistance and comments. We would like to thank Dr. Wenjiang Fu, Editor-in-Chief, and the reviewer for valuable suggestions that improved the quality of this article.

**REFERENCES**


2. Bickel DR. Empirical Bayes Interval Estimates that are Conditionally Equal to Unadjusted Confidence Intervals or to Default Prior Credibility Intervals. Statistical Applications in Genetics and Molecular Biology. 2012a; 11.


