Journal of Econometrics 189 (2015) 54–69

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

A Bayesian chi-squared test for hypothesis testing☆

Yong Li[a,*], Xiao-Bin Liu[b], Jun Yu[b,c]

[a] Hanqing Advanced Institute of Economics and Finance, Renmin University of China, Beijing, 100872, PR China
[b] School of Economics, Singapore Management University, 90 Stamford Road, Singapore 178903, Singapore
[c] Lee Kong Chian School of Business, Singapore Management University, 90 Stamford Road, Singapore 178903, Singapore

ARTICLE INFO

Article history:
Received 2 June 2014
Received in revised form
24 January 2015
Accepted 21 June 2015
Available online 7 July 2015

JEL classification:
C11
C12

Keywords:
Bayes factor
Decision theory
EM algorithm
Lagrange multiplier
Markov chain Monte Carlo
Latent variable models

ABSTRACT

A new Bayesian test statistic is proposed to test a point null hypothesis based on a quadratic loss. The proposed test statistic may be regarded as the Bayesian version of the Lagrange multiplier test. Its asymptotic distribution is obtained based on a set of regular conditions and follows a chi-squared distribution when the null hypothesis is correct. The new statistic has several important advantages that make it appealing in practical applications. First, it is well-defined under improper prior distributions. Second, it avoids Jeffrey–Lindley's paradox. Third, it always takes a non-negative value and is relatively easy to compute, even for models with latent variables. Fourth, its numerical standard error is relatively easy to obtain. Finally, it is asymptotically pivotal and its threshold values can be obtained from the chi-squared distribution. The method is illustrated using some real examples in economics and finance.

1. Introduction

This paper is concerned with statistical testing of a point null hypothesis after a Bayesian Markov chain Monte Carlo (MCMC) method has been used to estimate the models. Testing for a point null hypothesis is prevalent in economics although its importance is debatable. In the meantime, Bayesian MCMC methods have found more and more applications in economics because they make it possible to fit increasingly complex models, including latent variable models (Shephard, 2005), dynamic discrete choice models (Imai et al., 2009) and dynamic general equilibrium models (DSGE) (An and Schorfheide, 2007).

In the Bayesian paradigm, the Bayes factor (BF) is the gold standard for Bayesian model comparison and Bayesian hypothesis testing (Kass and Raftery, 1995; Geweke, 2007). Unfortunately, the BF is not problem-free. First, the BF is sensitive to the prior and subject to Jeffreys–Lindley's paradox; see for example, Kass and Raftery (1995), Poirier (1995) and Robert (1993, 2001). Second, the calculation of the BF for hypothesis testing generally requires the evaluation of marginal likelihood which is a marginalization over the unknown quantities. In many cases, the evaluation of marginal likelihood is difficult. Not surprisingly, alternative strategies have been proposed to test a point null hypothesis in the Bayesian literature. These methods can be classified into two classes.

In the first class, refinements are made to the BF to overcome the theoretical and computational difficulties. For example, to reduce the influence of the prior on the BF, one may split the data into two parts, a training sample and a sample for statistical analysis. The training sample is used to update the non-informative prior and to obtain a new proper informative prior, as in the fractional BF (O'Hagan, 1995). In practice, however, this strategy is not always satisfactory because it relies on an arbitrary division of the data. To alleviate this difficulty, Berger and Perrichi (1996) proposed the so-called intrinsic BF which is based on the minimal training

☆ Li gratefully acknowledges the financial support of the Chinese Natural Science Fund (No. 71271221), Program for New Century Excellent Talents in University, Beijing Higher Education Young Elite Teacher Project and Beijing Social Sciences Federation Young Teacher Project (No. 2014SKL020). Yu would like to acknowledge the financial support from Singapore Ministry of Education Academic Research Fund Tier 2 under the grant number MOE2011-T2-2-096 and Tier 3 under the grant number MOE2013-T3-1-017. We would like to acknowledge helpful comments from the co-editor, an associate editor, two referees, Tao Zeng and the seminar and conference participants at Peking University, Princeton University, the inaugural China Meeting of the Econometric Society.

* Corresponding author.
E-mail addresses: gibbsli@ruc.edu.cn (Y. Li), yujun@smu.ed.sg (J. Yu).
URL: http://www.mysmu.edu/faculty/yujun/ (J. Yu).

http://dx.doi.org/10.1016/j.jeconom.2015.06.021
0304-4076/

sample that results in proper posteriors. In general, the minimal training sample is not unique. Hence, the intrinsic BF is obtained by averaging the partial BFs calculated from all possible minimal training samples. Unfortunately, the intrinsic BF is computationally demanding, especially for latent variable models. O'Hagan (1995) discussed properties of the fractional and the intrinsic BFs.

In the second class, instead of refining the BF methodology, several interesting Bayesian approaches have been proposed for hypothesis testing based on the decision theory. For example, Bernardo and Rueda (2002, BR hereafter) showed that the BF for the Bayesian hypothesis testing can be regarded as a decision problem with a simple zero–one discrete loss function. However, the zero–one discrete function requires the use of non-regular (not absolutely continuous) prior and this is why the BF leads to Jeffreys–Lindley's paradox. BR further suggested using a continuous loss function, based on the well-known continuous Kullback–Leibler (KL) divergence function. As a result, it was shown in BR that their Bayesian test statistic does not depend on any arbitrary constant in the prior. However, BR's approach has some disadvantages. First, the analytical expression of the KL loss function required by BR is not always available, especially for latent variable models. Second, the test statistic is not a pivotal quantity. Consequently, BR had to use subjective threshold values to test the hypothesis.

To deal with the computational problem in BR in latent variable models, Li and Yu (2012, LY hereafter) proposed a new test statistic based on the $\mathcal{Q}$ function in the Expectation–Maximization (EM) algorithm of Dempster et al. (1977). LY showed that the new statistic is well-defined under improper priors and easy to compute for latent variable models. Following the idea of McCulloch (1989), LY proposed to choose the threshold values based on the Bernoulli distribution. However, like the test statistic proposed by BR, the test statistic proposed by LY is not pivotal. Moreover, it is not clear if the test statistic of LY can resolve Jeffreys–Lindley's paradox.

Based on the difference between the deviances, Li et al. (2014, LZY hereafter), developed another Bayesian test statistic for hypothesis testing. This test statistic is well-defined under improper priors, free of Jeffreys–Lindley's paradox, and not difficult to compute. Moreover, its asymptotic distribution can be derived and one may obtain the threshold values from the asymptotic distribution. Unfortunately, in general the asymptotic distribution depends on some unknown population parameters and hence the test is not pivotal.

In the present paper, we propose an asymptotically pivotal Bayesian test statistic, based on a quadratic loss function, to test a point null hypothesis within the decision-theoretic framework. The new statistic has several nice properties that makes it appealing in practice after the models are estimated by Bayesian MCMC methods. First, it is well-defined under improper prior distributions. Second, it is immune to Jeffreys–Lindley's paradox. Third, it is easy to compute. The main computational effort is to get the first derivative of the likelihood function with respect to the parameters. For latent variable models, the first derivative can be easily evaluated from the MCMC output with the help of the EM algorithm. Fourth, its numerical standard error (NSE) can be relatively easy to obtain. Finally, the asymptotic distribution of the test statistic follows the chi-squared distribution and hence the test is asymptotically pivotal.

Under a set of regularity conditions, we show that if the null hypothesis is correct our test statistic is asymptotically equivalent to the Lagrange multiplier (LM) statistic, a very popular test statistic in the frequentist's paradigm for testing a point null hypothesis. However, our proposed test has several important advantages over the LM test. First, it can incorporate the prior information to improve statistical inference. Second, the implementation of the LM test requires maximum likelihood (ML) estimation of the model under the null hypothesis. For some models, such as latent variable models and DSGE models, it is generally hard to do ML and, hence, to compute the LM statistic. Bayesian MCMC has been used to fit models with increasing complexity. The proposed test is the by-product of the Bayesian posterior output and hence easier to implement than the LM test. Third, unlike the LM test that can take a negative value in finite sample, our test always takes a nonnegative value. Finally, unlike the LM test, the new test does not need to invert any matrix. This advantage is useful when the dimension of the parameter space is high.

The paper is organized as follows. Section 2 reviews the Bayesian literature on testing a point null hypothesis from the viewpoint of the decision theory. Section 3 develops the new Bayesian test statistic, establishes its asymptotic properties, discusses how to compute it and its NSE from the MCMC outputs. Section 4 illustrates the new method by using three real examples in economics and finance. Section 5 concludes the paper. Appendix collects the proof of all the theoretical results and the derivation of the test statistic in the examples.

2. Bayesian hypothesis testing under decision theory

2.1. Testing a point null hypothesis

Let the observable data, $\mathbf{y} = (y_1, y_2, \ldots, y_n)' \in \mathbf{Y}$. A probability model $M \equiv \{p(\mathbf{y}|\boldsymbol{\theta}, \boldsymbol{\psi})\}$ is used to fit the data. We are concerned with a point null hypothesis testing problem which may arise from the prediction of a particular theory. Let $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ denote a vector of p-dimensional parameters of interest and $\boldsymbol{\psi} \in \boldsymbol{\Psi}$ a vector of q-dimensional nuisance parameters. The problem of testing a point null hypothesis is given by

$$\begin{cases} H_0: & \boldsymbol{\theta} = \boldsymbol{\theta}_0 \\ H_1: & \boldsymbol{\theta} \neq \boldsymbol{\theta}_0 \end{cases}. \tag{1}$$

The hypothesis testing may be formulated as a decision problem. It is obvious that the decision space has two statistical decisions, to accept H_0 (name it d_0) or to reject H_0 (name it d_1). Let $\{\mathcal{L}[d_i, (\boldsymbol{\theta}, \boldsymbol{\psi})], i = 0, 1\}$ be the loss function of statistical decision. Hence, a natural statistical decision to reject H_0 can be made when the expected posterior loss of accepting H_0 is sufficiently larger than the expected posterior loss of rejecting H_0, i.e., when

$$\mathbf{T}(\mathbf{y}, \boldsymbol{\theta}_0) = \int_{\Theta}\int_{\Psi} \{\mathcal{L}[d_0, (\boldsymbol{\theta}, \boldsymbol{\psi})] - \mathcal{L}[d_1, (\boldsymbol{\theta}, \boldsymbol{\psi})]\}\, p(\boldsymbol{\theta}, \boldsymbol{\psi}|\mathbf{y}) \mathrm{d}\boldsymbol{\theta}\mathrm{d}\boldsymbol{\psi} > c \geq 0,$$

where $\mathbf{T}(\mathbf{y}, \boldsymbol{\theta}_0)$ is a Bayesian test statistic; $p(\boldsymbol{\theta}, \boldsymbol{\psi}|\mathbf{y})$ is the posterior distribution with some given prior $p(\boldsymbol{\theta}, \boldsymbol{\psi})$; c is a threshold value. Let $\triangle\mathcal{L}[H_0, (\boldsymbol{\theta}, \boldsymbol{\psi})] = \mathcal{L}[d_0, (\boldsymbol{\theta}, \boldsymbol{\psi})] - \mathcal{L}[d_1, (\boldsymbol{\theta}, \boldsymbol{\psi})]$ be the net loss difference function which can generally be used to measure the evidence against H_0 as a function of $(\boldsymbol{\theta}, \boldsymbol{\psi})$. Hence, the Bayesian test statistic can be rewritten as

$$\mathbf{T}(\mathbf{y}, \boldsymbol{\theta}_0) = E_{\boldsymbol{\vartheta}|\mathbf{y}}\left(\triangle\mathcal{L}[H_0, (\boldsymbol{\theta}, \boldsymbol{\psi})]\right).$$

2.2. A literature review

The BF is defined as the ratio of the two marginal likelihood functions, namely,

$$BF_{01} = \frac{p(\mathbf{y}|M_0)}{p(\mathbf{y}|M_1)},$$