

Average Method of Mixed Sampling Distribution Bayesian Model

Ying Zhang

School of Mathematics, University of Jinan, Shandong Jinan 250022, China

Abstract. Based on the study of Bayesian model average (BMA), this paper proposes to mix the prior distribution and sampling distribution to obtain the average method of the mixed sampling distribution Bayesian model overcoming the problem that traditional econometric modeling method does not explicitly consider the uncertainty of the model. If all the alternative models have the same parametric form, then the new Bayesian estimation will degenerate into the BMA estimator. The empirical results show that the method is better than Bayesian model average.

1 Introduction

It is the mainstream analytical paradigm in the field of empirical macroeconomics to depict, describe and simulate the economic reality by econometric method. However, due to the complexity of the economic system and the openness of economic theory, the modelers often face the problem of uncertain model in actual modeling.

Model uncertainty is a potential universal problem in econometric modeling. In the uncertain environment of the model, the traditional modeling method seriously threatens the scientificity and robustness of econometric modeling. Therefore, established measurement model can produce serious errors in the analysis of prediction, policy evaluation, etc., and even the wrong conclusions can be drawn. Therefore, how to study and explore new modeling methods is the real problem and main challenge in the development of econometrics in the environment of model uncertainty.

Leamer (1983) first pointed out the problem of model uncertainty in the process of econometric modeling, and proposed the method of Extreme-Bounds Analysis to solve the problem of model uncertainty [1]. Levine & Renelt (1992) improved and adjusted the extremum boundary analysis method presented by Leamer (1983), and proposed the adjustment of extremum boundary analysis method [2]. However, the above analysis method is not ideal in solving the problem of uncertain model. Draper (1995) proposed the famous Bayesian model of average modeling thought [3], which provided a statistical theoretical basis for solving the problem of model uncertainty in econometric process. Subsequently, the Bayesian model averaging method has been rapidly developed and widely used in econometric modeling in the uncertain environment.

Bayesian model averaging method is a comprehensive method of handling model uncertainty. It is based on

Bayesian theorem to comprehensively consider the prior information of the model and parameters and the information provided by the samples, thus eliminating the uncertainty of the model as much as possible.

2 Model building

Firstly, the basic idea of Bayesian model average is briefly reviewed, then the mixed average Bayesian model is presented.

2.1 Bayesian model average

In the Bayesian analysis framework, suppose we need to estimate some interest variables Δ from the sample data. The general distribution of these samples is unknown, and these distributions need to be set. There is usually a wide selection of theoretical distribution models, and in general, different models have different estimates. The basic idea of Bayesian average is to take a weighted average of all Δ estimated from each model and use the final mean as an estimate of Δ , instead of just using the estimate of a model as an estimate of Δ .

Consider the data sample D and a collection of models that contain K models $M = (M_1, M_2, \dots, M_k)$. It contains a real model M_i . If Δ is the interest variable dependent on D , then Δ 's posterior distribution is given as follows.

$$P(\Delta | D) = \sum_{k=1}^K P(\Delta | D, M_k) P(M_k | D) \quad (1)$$

According to formula (1), we can think of $P(\Delta | D)$ as the weighted average of the posterior distribution $P(\Delta | D, M_k)$ ($k=1, \dots, K$), where $P(M_k | D)$ is the k

weight and the posterior probability under the hypothesis of M_k as the real model.

The posterior model probability is derived.

$$P(M_k | D) = \frac{P(D | M_k)P(M_k)}{\sum_{j=1}^K P(D | M_j)P(M_j)} \quad (2)$$

$P(M_k)$ represents the prior probability distribution of M_k as the real model. Given the data D , integrate the parameters β_k ,

$$P(D | M_j) = \int P(D | \beta_k, M_k)P(\beta_k | M_k)d\beta_k \quad (3)$$

The $P(D | \beta_k, M_k)$ in the integral formula (3) is the sampling distribution of the conditions in which M_k is the true distribution, where β_k represents the parameter vector. $P(\beta_k | M_k)$ represents the prior distribution of parameter β_k with real distribution M_k as the condition. If $P(\beta_k | M_k)$ is a discrete distribution, then the integral in (3) can be replaced by the sum.

In addition, if M_k is a true model, then the posterior distribution of Δ is

$$P(\Delta | D, M_k) = \int P(\Delta | \beta_k, D, M_k)P(\beta_k | D, M_k)d\beta_k \quad (4)$$

After the above posterior distribution of Δ , the posterior mean and the posteriori difference are obtained.

$$\begin{aligned} E(\Delta | D) &= \sum_{k=1}^K E(\Delta | D, M_k)P(M_k | D) \\ &= \int_{-\infty}^{+\infty} \Delta \left[\sum_{k=1}^K P(\Delta | D, M_k)P(M_k | D) \right] d\Delta \\ &= \sum_{k=1}^K P(M_k | D) \int_{-\infty}^{+\infty} \Delta P(\Delta | D, M_k) d\Delta \\ &= \sum_{k=1}^K \hat{\Delta}_i P(M_k | D) \end{aligned}$$

$$\begin{aligned} Var(\Delta | D) &= \sum_{k=1}^K P(M_k | D) \left[\frac{(E(\Delta | D, M_k) - E(\Delta | D))^2}{+Var(\Delta | D, M_k)} \right] \end{aligned}$$

Among them, $\hat{\Delta}_i = E(\Delta | D, M_k)$.

In the practical application of Bayesian model average, several kinds of problems are usually solved:

The first is the setting of the prior distribution.

For the setting of prior distribution, there are two different prior information. One is the prior distribution $P(M_k)$ of the model M_k , and the other is the conditional prior distribution $P(\beta_k | M_k)$ of the parameter vectors β_k .

For the problem of the prior distribution of the model, Raftery (1997) proposes the uniform prior of the model, which is to specify the same prior probability for all models: $P(M_k) = \frac{1}{K}$ ($k = 1, 2, \dots, K$), where K is the

total number of alternative models in the model space [4].

For the problem of setting the prior distribution of parameters, Zellner (1986) proposed the setting of g prior distribution for the regression model, which is called Zellner's g priori [5].

The second is the setting of the sampling distribution, which is the basic model that needs to be integrated in the model average. For example, it could be a linear regression model or a generalized linear model, or a survival analysis model.

The third is to solve the marginal likelihood problem. Newton and Raftery (1994) proposed using MCMC method to estimate the marginal likelihood [6].

Fourth, search model space. By using the methods of approximate or simulation, we can get the marginal likelihood of a single model. However, when the number of alternative models is large, it is impossible to calculate the marginal likelihood and even posterior probability of all models. In this case, the Markov Chain Monte Carlo Model Composition (MC3) model search strategy can be adopted. Searching for an important model constitutes a subset of the model space, and then on the basis of this subset, the Bayesian model average is used. This is an exhaustive search method to find the global optimal model.

2.2 Mixed Bayesian model average

Based on the above classic Bayesian model average theory, we propose the following mixed Bayesian model average.

Consider the data sample D and a collection of models that contain K models $M = (M_1, M_2, \dots, M_k)$, then, the posterior distribution of the interest variable Δ is

$$P(\Delta | D) = \frac{P(D | \Delta)P(\Delta)}{\int_{\Gamma} P(D | \Delta)P(\Delta)d\Delta} \quad (5)$$

The likelihood probability $P(D | \Delta)$ is defined as follows:

$$P(D | \Delta) = \sum_{k=1}^K P(D | \Delta, M_k)P(M_k) \quad (6)$$

In (6), $P(D | \Delta, M_k)$ is the sampling distribution of model M_k . $P(M_k)$ represents the prior probability distribution of M_k as the real model.

The prior distribution of interest variables Δ is defined as follows:

$$P(\Delta) = \sum_{k=1}^K P(\Delta | M_k)P(M_k) \quad (7)$$

Where $P(\Delta | M_k)$ is the conditional prior distribution of interest variables Δ under known M_k conditions.

Under (6) and (7), suppose that for all K and

$j, k \neq j,$

$$p_{kj}(D) = \int_{\Gamma} P(D|\Delta, M_k)P(\Delta|M_j)d\Delta < \infty \text{ holds,}$$

then the posterior distribution of the interest variable Δ in (8) is

$$P(\Delta|D) = \frac{\sum_{k=1}^K P(D|M_k)P(\Delta|D, M_k)P^2(M_k) + \sum_{k=1}^K \sum_{j=1; j \neq k}^K P(D|\Delta, M_k)P(\Delta|M_j)P(M_k)P(M_j)}{\sum_{k=1}^K P(D|M_k)P^2(M_k) + \sum_{k=1}^K \sum_{j=1; j \neq k}^K P(D)P(M_k)P(M_j)} \quad (8)$$

A detailed derivation of equation (8) is given below.
 In fact,

$$\begin{aligned} & P(D|\Delta)P(\Delta) \\ &= \left(\sum_{k=1}^K P(D|\Delta, M_k)P(M_k)\right)\left(\sum_{k=1}^K P(\Delta|M_k)P(M_k)\right) \\ &= \sum_{k=1}^K P^2(M_k)P(\Delta|M_k)P(D|\Delta, M_k) \\ &+ \sum_{k=1}^K \sum_{j=1; j \neq k}^K P(M_k)P(M_j)P(D|\Delta, M_k)P(\Delta|M_j) \end{aligned}$$

By Bayesian criterion,

$$\begin{aligned} & P(\Delta|M_k)P(D|\Delta, M_k) \\ &= P(D|M_k)P(\Delta|D, M_k) \end{aligned}$$

Substitute it in

$$\begin{aligned} & P(D|\Delta)P(\Delta) \\ &= \sum_{k=1}^K P^2(M_k)P(D|M_k)P(\Delta|D, M_k) \\ &+ \sum_{k=1}^K \sum_{j=1; j \neq k}^K P(M_k)P(M_j)P(D|\Delta, M_k)P(\Delta|M_j) \end{aligned}$$

And because,

$$\begin{aligned} P(D) &= \int P(D|\Delta)P(\Delta)d\Delta \\ &= \sum_{k=1}^K P^2(M_k)P(D|M_k) + \\ &\sum_{k=1}^K \sum_{j=1; j \neq k}^K p_{kj}(D)P(M_k)P(M_j) \end{aligned}$$

So, combining the previous formula (5), you can get equation (8).

The posterior distribution obtained in equation (8) is known as the mixed Bayesian model average (MBMA).

In fact, the mixed Bayesian model average proposed in this paper is the generalization of the classical BMA model, or the BMA model is an exception to the mixed Bayesian model average proposed in this paper. In fact, if we assume that all models have the same sampling distribution.

That is $P(D|\Delta, M_k) = P(D|\Delta, M_j)$, then MBMA degenerates into BMA.

In fact, in the numerator of the equation (8)

$$\begin{aligned} & \sum_{j=1; j \neq k}^K P(D|\Delta, M_k)P(\Delta|M_j)P(M_k)P(M_j) \\ &= \sum_{k=1}^K P(\Delta|M_k)P(D|\Delta, M_k)P(M_k) \sum_{j=1; j \neq k}^K P(M_j) \\ &= \sum_{k=1}^K P(\Delta|M_k)P(D|\Delta, M_k)P(M_k)(1-P(M_k)) \end{aligned}$$

So, the numerator of the equation (8)

$$\begin{aligned} &= \sum_{k=1}^K P(D|M_k)P(\Delta|D, M_k)P(M_k) \\ &= \int_{\Gamma} P(D|\Delta, M_k)P(\Delta|M_k)d\Delta = P(D|M_k). \end{aligned}$$

With the same treatment as above, for the denominator of the equation (8)

Therefore,

$$P(D) = \sum_{k=1}^K P(D|M_k)P^2(M_k) +$$

$$\sum_{k=1}^K \sum_{j=1; j \neq k}^K p_{kj}(D)P(M_k)P(M_j)$$

$$= \sum_{k=1}^K P(D|M_k)P(M_k)$$

$$P(\Delta|D) = \frac{\sum_{k=1}^K P(D|M_k)P(\Delta|D, M_k)P(M_k)}{\sum_{k=1}^K P(D|M_k)P(M_k)}$$

$$= \sum_{k=1}^K \frac{P(D|M_k)P(M_k)}{\sum_{k=1}^K P(D|M_k)P(M_k)} P(\Delta|D, M_k)$$

$$= \sum_{k=1}^K P(\Delta|D, M_k)P(M_k|D)$$

$$= P(\Delta|D)$$

3 Empirical test of the model

The estimate of the mixed Bayesian model average is verified by using the daily data of RMB against the us dollar.

The data used in this paper are the daily settlement rate of RMB against us dollar, from January 1, 2012 to December 31, 2016 (the target is their earnings valuation). Laplace (μ, σ^2) distribution and normal distribution $N(\mu, \sigma^2)$ are chosen for the sampling distribution. Let's say the mean is known, the interest variable is $\Delta = \sigma^2$. The prior distribution of σ^2 of the two models is set to Gamma (α, β) .

Set the same prior distribution for M_1 and M_2 , i.e. 0.5. After observing the data, compared with M_2 (0.23), M_1 (0.77) is more likely to be a real model. Compared

with BMA method, the conditional variance obtained by MBMA method is smaller (0.04).

4 Conclusion

Based on the average theory of Bayesian model, an average method of mixed Bayesian model is proposed, and the detailed derivation process and empirical test are given. The empirical analysis shows that the average method of mixed Bayesian model can reduce the variance of estimation and improve the accuracy of estimation.

References

1. Leamer E. Let's Take the Con out of Econometrics [J]. *American Economic Review*, 1983, 73 (1): 31-43.
2. Levine R, Renelt D. A Sensitivity Analysis of Cross-country Growth Regressions [J]. *American Economic Review*, 1992, 82 (3): 942-963.
3. Draper D. Assessment and Propagation of Model Uncertainty [J]. *Journal of the Royal Statistical Society*, 1995, 57: 45-70.
4. Raftery A.E., Madigan D., Hoeting J. Bayesian Model Averaging for Linear Regression Models [J]. *Journal of the American Statistical Association*, 1997, 92: 179-191.
5. Zellner. On Assessing Prior Distributions and Bayesian Regression Analysis with G-prior Distributions [J]. *Bayesian Inference & Decision Techniques*, 1986, 6: 233-243.
6. Newton M.A., Raftery A.E. Approximate Bayesian Inference with the Weighed Likelihood Bootstrap [J]. *Journal of the Royal Statistical Society. Series B (Methodological)*, 1994 (1).