The Bayesian Estimation of Teaching Satisfaction Based on MCMC Method

Qi Zhao

Abstract: In this paper, using MCMC method, parameter estimation problem of the Bayesian model of teaching satisfaction is discussed. Firstly, we introduced the Bayesian model of teaching satisfaction and MCMC method. Afterwards, example calculation is given to check the feasibility of MCMC method to solve such problems.

Keywords: teaching satisfaction; Bayesian model; MCMC method; OpenBUGS software

1 Introduction

With the development of the higher education in China, the evaluation of teaching satisfaction has become an important evaluation work of the university. Teaching satisfaction is a comprehensive evaluation of teacher's basic quality, teaching level, teaching content, teaching method and so on. Teaching satisfaction is not only an important measure of teachers' level and teaching effect, also an important basis for schools to strengthen the construction of teaching staff, improve teaching quality and improve teaching management. At home and abroad, the research on teaching satisfaction has attracted the attention of many scholars and institutions. At present, the research on teaching satisfaction can be divided into two categories: (1) Study on the determinants and the adjustment variables of teaching satisfaction, for example, literature[1-2]; (2) Construction of evaluation model of teaching satisfaction, for example: Structural equation model[3], orderly Logistic regression model[4], Fuzzy set model[5], etc. Zhu xin ling[6] proposed the Bayesian model of teaching satisfaction. Based on Markov chain monte carlo method, this paper solves parameter estimation problem of the Bayesian model.
2 The Bayesian model of teaching satisfaction

2.1 The basic idea of Bayesian statistics

The Bayesian statistical method was first proposed by British scholar Thomas Bayes. It developed rapidly in the second half of the twentieth century. It is now a very influential statistical school Bayesian school. The original significance of Bayesian statistics is to put forward the prior distribution or subjective probability. Suppose \( \theta \) is an unknown quantity that we're interested (The set of all its possible values is called the parameter space, Which is denoted as \( \Theta \)), According to point of view of Bayesian, \( \theta \) is a random variable, so it can be described by a probability distribution. This distribution is called a prior distribution. Bayesian formula combines the prior and sample information to obtain posterior distribution, the posterior distribution is denoted as \( \pi(\theta | x) \). All decisions and inferences are based on posterior distribution. The Bayesian formula is given below\[7\].

Suppose \( \pi(\theta) \) is prior distribution of \( \theta \), the sample \( X \) depend on \( \theta \), and its conditional probability density function is \( p(x | \theta) \). The joint density function of \( X \) and \( \theta \) is

\[
h(x, \theta) = p(x | \theta)\pi(\theta)
\]

The marginal probability density function of \( X \) as follows

\[
m(x) = \int_{\Theta} h(x, \theta)d\theta = \int_{\Theta} p(x | \theta)\pi(\theta)d\theta
\]

So

\[
\pi(\theta | x) = \frac{h(x, \theta)}{m(x)} = \frac{p(x | \theta)\pi(\theta)}{\int_{\Theta} p(x | \theta)\pi(\theta)d\theta}
\]

This equation is called Bayes’ rule of density function form. The posterior distribution combines prior \( (\theta) \) and sample information \( (X) \), so the posterior distribution is closer to the actual situation than the prior distribution. Generally, a digital feature of the posterior distribution is chosen as the estimate of \( \theta \), for example mathematical expectation or quantile. This paper adopts mathematical expectation \( \theta = E[\pi(\theta | x)] \). Bayesian statistics is very rich in content, The Bayesian model is based on Bayesian statistical theory, The basic architecture of the Bayesian model as follows:

Prior distribution \( \oplus \) Sample information \( \rightarrow \) The model of posterior distribution \( \rightarrow \) Parameter estimation of the model \( \rightarrow \) Model test

There in “\( \oplus \)” should be understood as the effect of Bayesian formula. Parameter estimation of the posterior distribution is the key element in Bayesian inference.

2.2 Parameter estimation problem of the Bayesian model of teaching satisfaction

Essentially, the Bayesian estimate of teaching satisfaction is a multi-grade scored model. The Bayesian estimate of multi-grade scored model has been proved in detail by Wu\[8\]. This article cites its main conclusions. Under the primary indicator system, each indicator may be have grades \( 0,1,\ldots,K \). Suppose the indicator system consists of \( n \) indicators, there are \( x_o \)
indicators for evaluation of 0, ..., there are $x_k$ indicators for evaluation of $K$. For each reviewer, his score is recorded as: $x = (x_0, x_1, ..., x_K)^\top$, then $\sum_{i=0}^{K} x_i = n$. Suppose that the score of a certain reviewer is $T = (T_0, T_1, ..., T_K)^\top$, mark $\theta_0 = \frac{T_0}{n}, \theta_k = \frac{T_k}{n}, \theta = (\theta_0, ..., \theta_K)^\top$, then $\theta$ can be used as a probability of scoring. When $\theta$ is known, $X$ obeys multinomial distribution:

$$p(x \mid \theta) = n! \prod_{i=0}^{K} \frac{\theta^x_i}{x_i!}$$ (4)

Among them $x_i = 0, 1, ..., n(i = 0, 1, ..., K)$, $\sum_{i=0}^{K} x_i = n, \sum_{i=0}^{K} \theta_i = 1$; When the survey was conducted on m evaluator, denote sample as $X = (x^{(1)}, ..., x^{(m)}), x^{(j)} = (x_{0j}, ..., x_{Kj})^\top, j = 1, ..., m$. Where $x^{(j)}$ represents the score of the jth evaluator, So there is always: $\sum_{i=0}^{K} x_{ij} = n, j = 1, ..., m$. Suppose $\pi(\theta)$ is prior distribution of $\theta$, when $\theta$ is known, $X$ obeys joint distribution as follows:

$$f(x \mid \theta) = (n!)^m \prod_{i=0}^{K} \frac{\theta^x_i}{\prod_{j=1}^{m} x_{ij}!}$$ (5)

We will study the Bayes estimation of $\theta$ based on MCMC method.

3 MCMC method and OpenBUGS software

Bayesian model is simple and the probability form is elegant, but, obtaining the posterior distribution often requires the integration of high-dimensional functions. This can be computationally very difficult, so it is necessary for us to discuss some new methods of calculation. We focus here on Markov Chain Monte Carlo (MCMC) methods. MCMC methods have their roots in the Metropolis algorithm (Metropolis et al. 1953), Since then, many scholars have developed this method and apply it to various fields. Loosely speaking, we will simple throughout $\pi(x)$, one way of doing this is through a Markov Chain having $\pi(x)$ as its stationary distribution. In other words, Markov chain monte carlo method is essentially a monte carlo comprehensive program. The generation of random sample is associated with a markov chain. Gibbs sampling is the most commonly used MCMC method, its potential of markov chains is established by decomposition a series of conditional distribution, here's a brief introduction:

Let $\pi(x_1, ..., x_m)$ denotes m-dimensional joint distribution, constructor the transition nuclear in Gibbs sampling as follows:

$$P_{x,y} \triangleq P(x,y) = \prod_{k=1}^{n} \pi(y_k \mid y_1, ..., y_{k-1}, x_{k+1}, ..., x_m)$$ (6)
among them \( x = (x_1, \ldots, x_m) \), \( y = (y_1, \ldots, y_m) \), \( x_i \in D \), \( y_i \in D \) (\( D \) denotes a \( m \)-dimensional region), 
\( \pi(y_1 \mid y_2, \ldots, y_m, x_1, \ldots, x_m) \) is the conditional distribution. Gibbs sampling specific steps are as follows, By Markov chain \( X_t(\omega) \) sample, a sample of \( X_{t+1}(\omega) \) can be obtained according to the following procedure:

(1) first, \( y_1 \) is obtained by a random variable \( X_{n+1}(\omega) \), \( X_{n+1}(\omega) \) obey the distribution of \( \{\pi(y_1 \mid x_1, \ldots, x_m), y_1 \in D\} \) \( (x_2, \ldots, x_m \) from \( X_n(\omega) \))

(2) secondly, \( y_2 \) is obtained by a random variable \( X_{n+1,2}(\omega) \), \( X_{n+1,2}(\omega) \) obey the distribution of \( \{\pi(y_2 \mid y_1, x_2, \ldots, x_m), y_2 \in D_2\} \), in this down, \( y_k (k = 1, \ldots, m-1) \) is obtained by a random variable \( X_{n+1,k}(\omega) \), \( X_{n+1,k}(\omega) \) obey the distribution of \( \{\pi(y_k \mid y_1, \ldots, y_{k-1}, x_k, \ldots, x_m), y_k \in D_k\} \)

(3) In the end, \( y_m \) is obtained by a random variable \( X_{n+1,m}(\omega) \), \( X_{n+1,m}(\omega) \) obey the distribution of \( \{\pi(y_m \mid y_1, \ldots, y_{m-1}), y_m \in D\} \), then \( (y_1, \ldots, y_m) \) is a sample of \( X_{n+1}(\omega) \)

Now, Take an initial value \( X_0(\omega) = y^{(0)} \), \( y^{(l)} \) is obtained by the random variable \( X_{l}(\omega) \) according to the above methods. To get the sample \( y^{(1)}, \ldots, y^{(n)} \) from \( X_{1}(\omega), \ldots, X_{n}(\omega) \), when \( n \) is large enough, the distribution of \( X_{n}(\omega) \) approximation of \( \pi(x_1, \ldots, x_m) \), can approximate thought, \( y^{(n)} \) is a sample obey the distribution of \( \pi(x_1, \ldots, x_m) \).

Many software and applications for the MCMC approach have been developed, for example, OpenBUGS\cite{11-12} is the proprietary software that applies the MCMC method to the Bayes model. Using OpenBUGS makes it easy to sample a number of commonly used models and distributions, The user does not need to know the parameter's prior density or the exact expression of the likelihood function When using OpenBUGS. The user simply sets the prior distribution of the variables and describes the model in general, then Bayes analysis of the model can be easily realized. No complex programming is required.

The basic operations of OpenBUGS include commands such as File, Edit, Model, Doodle, Attributes, Inference, and Info.

File: Used for new files, open files, save files, etc

Edit: Use the Edit command to copy files, cut files, paste files, delete files, and so on.

Model: The Specification command in the Model is used to validate the model, includes model checking, data loading, compilation, and generation of initial values. Set the iteration times and step lengths through the Update command.

Doodle: Used to build the Doodle model diagram, The Doodle model diagram appears as nodes, arrows, and tablets.

Attributes: This command is used to change the font, format, font size, and color of the selected text, etc.

Inference: Output for iterative calculation results.

Info: This command is used to obtain information about nodes and data.
In addition, OpenBUGS also has Tools, maps, Windows, Examples, Manuals, and Help commands, we won't go into details here.

We can use a directed graph model (the Doodle model) in OpenBUGS, or write code directly by hand.

4 Examples of application

We take the example of literature [6] as an example to illustrate the MCMC method and the application process of OpenBUGS software.

This example is a teaching evaluation survey of a university for the second semester of the 2004-2005 academic year. The investigators collected information on the teaching evaluation of a class of 70 students. The scoring rules are as follows: 1-great; 2-good; 3-better; 4-normal; 5-poor.

The survey statistics form is shown in document [6], we won't repeat it here.

We assume that the prior distribution of \( \theta \) is Dirichlet distribution, The OpenBUGS program is compiled as follows:

```r
model
  for(i in 1:70){
    x[i,1:5]~dmulti(theta[,20])
  }
  theta[1:5]~ddirich(prior[]);
}
```

To ensure the convergence of the parameters, we performed 1000 pre-iterations in the process of running the model. The following is the partial information of \( \theta \).

| parameter | mean   | SD    | MC error | 2.5%   | Median | 97.5%  | begin | sample |
|-----------|--------|-------|----------|--------|--------|--------|-------|--------|
| \( \theta_1 \) | 0.2444 | 0.01137 | 1.068E-4 | 0.2225 | 0.2442 | 0.2669 | 1001  | 10000  |
| \( \theta_2 \) | 0.2671 | 0.01177 | 1.095E-4 | 0.2444 | 0.2669 | 0.2906 | 1001  | 10000  |
| \( \theta_3 \) | 0.2548 | 0.0115 | 1.184E-4 | 0.2325 | 0.2546 | 0.2775 | 1001  | 10000  |
| \( \theta_4 \) | 0.1564 | 0.009644 | 8.876E-5 | 0.1379 | 0.1564 | 0.1757 | 1001  | 10000  |
| \( \theta_5 \) | 0.07729 | 0.007118 | 6.698E-5 | 0.0639 | 0.07706 | 0.09152 | 1001  | 10000  |

We can see it in the table, 24.44 percent of the class said the teacher had a "very good" teaching effect, 26.71% of the class said “good”, 25.48% of the class said “better”, 15.64% of the class said “normal”, 7.729% of the class said “poor”. The above results show that the class students are generally satisfied with the teaching effect of the course teachers. The overall unsatisfactory ratio (that is, the ratio of "poor" teaching effect) is less than 8%. This is consistent with the results of literature [6].

The OpenBUGS software also gives information about the standard deviation and median of the posterior distribution of the parameter. In addition, OpenBUGS software can also give a series of information such as the kernel density estimation graph of the posterior distribution, Gibbs sampling dynamic graph of parameters, iterative history and convergence statistical diagnosis graph. This information makes the sampling result more intuitive and reliable.
5 Concluding remarks

The birth of MCMC has made Bayesian statistical have a great development, this genetic algorithm with markov property is almost applicable to all Bayes models. The application of OpenBUGS software based on MCMC method makes Bayes estimation of model parameters avoid tedious high-dimensional integral calculation, the application of Bayes method is programmed. The MCMC method is applied to the estimation model of teaching satisfaction, which shows the wide applicability of this method. The research of this paper provides enlightenment for the application of this method to more Bayes models.

Reference:
[1] H.J. Xiong, D.L. Ma. (2013), The Analysis of Undergraduate Teaching Satisfaction from Simultaneous Analysis of Several Groups. Education Science. 29: (5)24-32.
[2] R.L. Sa, X.F. Yang, Z.Y. Yang. (2010), Factor Analysis of University Students’ Satisfaction with Teaching. Journal of Research on Education for Ethnic Minorities. 6 (21): 79-81.
[3] X.H. Ma. (2012), Structural equation model of teaching satisfaction in colleges and universities. Statistics and Decision Making. 20: 66-68.
[4] S.X. Wan. (2012), Case teaching satisfaction study based on orderly Logistic regression model. Journal of Changchun University of Science and Technology (Social Sciences Edition). 5: 171-173.
[5] Q. Xu. (2006), Research on fuzzy comprehensive model of teaching satisfaction assessment. Journal of Chongqing Industrial and Commercial University (Natural Science Edition). 23 (6): 596-599.
[6] X.L. Zhu, P. Li. (2005), The Bayes estimation of teaching satisfaction. Statistics and decision making. 24: 149-150.
[7] Sh. S. Mao, J.L. Wang, X.L. Pu. (1998), Advanced Mathematical Statistics, Higher Education Press, Beijing, 1998.
[8] D.W. Wu. (2002), The Bayesian estimate of multi-grade scored model. Journal of Huaiyin Teachers College (Natural Science Edition).
[9] Metropolis N., Rosenbluth A.W., Rosenbluth M.N., Teller A.H., Teller E. (1953), Equations of state calculations by fast computing machines [J]. J Chemical Phys. 21: 1087-1091.
[10] Geman S., Geman D. (1984), Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence. 6: 721-741.
[11] J.W. Zhang, W.L. Gao, etc. (2017), Open BUGS software introduction and application. Health statistics in China. 34 (1): 170-172.
[12] W.W. Yang, P. Liu, etc. (2016), A powerful tool for Bayesian statistical analysis Open BUGS software. Health statistics in China. 33 (3); 510-513.