

University of Wollongong

Research Online

Applied Statistics Education and Research
Collaboration (ASEARC) - Conference Papers

Faculty of Engineering and Information
Sciences

2012

Bayesian Estimation of a Meta-analysis model using Gibbs sampler

Junaidi

University of Newcastle

Darfiana Nur

University of Newcastle

Elizabeth Stojanovski

University of Newcastle

Follow this and additional works at: <https://ro.uow.edu.au/asearc>

Recommended Citation

Junaidi; Nur, Darfiana; and Stojanovski, Elizabeth, "Bayesian Estimation of a Meta-analysis model using Gibbs sampler" (2012). *Applied Statistics Education and Research Collaboration (ASEARC) - Conference Papers*. 5.

<https://ro.uow.edu.au/asearc/5>

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au

Bayesian Estimation of a Meta-analysis model using Gibbs sampler

Abstract

A hierarchical Bayesian model is investigated. This model can accommodate study heterogeneity in meta-analyses. The joint posterior distribution is derived by multiplying the likelihood and priors on this model. The conditional posterior distribution of all parameters is obtained for Gibbs sampler algorithm. A simulation study is then performed to demonstrate the validity of the Gibbs sampler in terms of parameter estimation.

Keywords

Meta-analysis, Bayes estimates, Hierarchical model, Gibbs sampler

Publication Details

Junaidi; Nur, Darfiana; and Stojanovski, Elizabeth, Bayesian Estimation of a Meta-analysis model using Gibbs sampler, Proceedings of the Fifth Annual ASEARC Conference - Looking to the future - Programme and Proceedings, 2 - 3 February 2012, University of Wollongong.

Bayesian Estimation of a Meta-analysis model using Gibbs sampler

Junaidi^{1,2}, D. Nur¹, E. Stojanovski¹

¹The University of Newcastle, Callaghan, NSW, 2308, AUSTRALIA

²The University of Tadulako, Kampus Bumi Tadulako Tondo, Palu, 94118, INDONESIA

Junaidi@uon.edu.au, Darfiana.Nur@newcastle.edu.au, Elizabeth.Stojanovski@newcastle.edu.au

Abstract

A hierarchical Bayesian model is investigated. This model can accommodate study heterogeneity in meta-analyses. The joint posterior distribution is derived by multiplying the likelihood and priors on this model. The conditional posterior distribution of all parameters is obtained for Gibbs sampler algorithm. A simulation study is then performed to demonstrate the validity of the Gibbs sampler in terms of parameter estimation.

Keywords: Meta-analysis, Bayes estimates, Hierarchical model, Gibbs sampler.

1. Introduction

Meta-analysis can be presented either in the Frequentist or Bayesian framework. Within the Frequentist framework, statistical inference for model parameters is based only on data from studies while the Bayesian framework considers parameters as random variables. Heterogeneity is commonly encountered and arises when effect sizes are different or studies differ on some aspect. Heterogeneity can arise from differences in statistical sources (such as populations), methodologies (such as design) and clinical aspects (such as treatment durations) [5].

A Hierarchical Bayesian model (HBM) is used in a study [2], as an example, to assess heterogeneity within sarcoma subtypes. The same model was implemented in [8] to combine data due to differences between studies to assess the relationship between chlorinated by-product exposure and adverse reproductive health effect.

In the present paper, a Bayesian estimation of a hierarchical model that was developed by Dumouchel [6] is investigated further using Gibbs sampler. The model is presented in Section 2; followed by the analytical derivation of all conditional posterior distributions, and its application into the simulation study, respectively, in Section 3.

2. Methods

Hierarchical Bayesian Model

The standard HBM proposed by Dumouchel provides the following distributional assumptions

$$Y_i \sim N(\theta_i, \sigma_{Y_i}^2) \quad i = 1, 2, \dots, n \quad (1)$$

$$\theta_i \sim N(\mu, \sigma_{\theta}^2) \quad (2)$$

$$\mu \sim N(0, D \rightarrow \infty).$$

Level 1 (1) of the model represents study data, where Y_i denotes the observed study statistics with variance-covariance matrix, $\sigma_{Y_i}^2$ and n ($i = 1, 2, \dots, n$) denotes the number of studies. Level 2 (2) represents the study-specific parameters, denoted θ and associated variance-covariance matrix σ_{θ}^2 . The overall mean is denoted by μ . The prior distribution for μ is assumed to follow the normal distribution with mean, zero, and variance $D \rightarrow \infty$ indicating elements of D as tending to infinity.

Gibbs sampler

The Gibbs sampler is an algorithm to generate a sequence of samples from the joint distribution from two or more random variables [3]. It calculates the full conditional distribution in a sampling scheme by sampling one parameter at a time given the data and other parameters.

A study applied the Gibbs sampler to estimate heritability and genetic correlations between clinical mastitis and retained placenta for the first lactations of

Holstein dairy cows [7]. The posterior means helped identify the importance of health traits for dairy cows. In another study, [12] implemented Gibbs sampling to estimate genetic parameters for growth traits of cultured clam. Four size traits at different growth stages were observed. The growth traits showed high-magnitude heritabilities, from the credible interval, indicating these traits as important in genetic improvement programmes.

The objective of this paper is to compute the joint posterior distribution of Dumouchel's model to identify parameter stability by applying the Gibbs sampler algorithm.

3. Results

Posterior Derivation

Within the Bayesian framework, combining the prior probability distribution with the likelihood distribution produces the posterior probability distribution. The joint likelihood, which represents study data, follows the normal distribution and for the current model is presented as follows:

$$P(Y_{1:n} | \theta_{1:n}, \mu, \sigma_{\theta}^2) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma_{Y_i}^2}} \exp\left[-\frac{(Y_i - \theta_i)^2}{2\sigma_{Y_i}^2}\right], \quad (3)$$

The prior distributions for parameters θ_i and μ follow the normal distribution while the prior for σ_{θ}^2 follows the inverse gamma:

$$P(\theta_{1:n} | \mu, \sigma_{\theta}^2) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma_{\theta}^2}} \exp\left[-\frac{(\theta_i - \mu)^2}{2\sigma_{\theta}^2}\right], \quad (4)$$

$$P(\mu) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left[-\frac{(\mu - \theta_0)^2}{2\sigma_0^2}\right], \quad (5)$$

$$P(\sigma_{\theta}^2) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} (\sigma_{\theta}^2)^{-\alpha-1} e^{-\beta/\sigma_{\theta}^2}. \quad (6)$$

The joint posterior distribution is defined as follows, where $P(Y_i)$ is a normalising constant.

$$P(\theta_{1:n}, \mu, \sigma_{\theta}^2 | Y_{1:n}) = \frac{P(\theta_{1:n}, \mu, \sigma_{\theta}^2, Y_{1:n})}{P(Y_{1:n})} \propto P(\theta_{1:n}, \mu, \sigma_{\theta}^2, Y_{1:n})$$

The joint posterior becomes

$$= P(\theta_{1:n} | \mu, \sigma_{\theta}^2) P(\mu) P(\sigma_{\theta}^2) P(Y_{1:n} | \theta_{1:n}, \mu, \sigma_{\theta}^2)$$

(assuming independence on μ and σ_{θ}^2)

Using (4), (5), (6) and (3) respectively, we have the full joint posterior as follows:

$$P(\theta_{1:n} | \mu, \sigma_{\theta}^2, Y_{1:n}) = \frac{1}{\sqrt{2\pi\sigma_{\theta}^2}} \exp\left[-\frac{(\theta_i - \mu)^2}{2\sigma_{\theta}^2}\right] \times$$

$$\frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left[-\frac{(\mu - \theta_0)^2}{2\sigma_0^2}\right] \times \frac{\beta^{\alpha}}{\Gamma(\alpha)} (\sigma_{\theta}^2)^{-\alpha-1} e^{-\beta/\sigma_{\theta}^2} \times$$

$$\prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma_{Y_i}^2}} \exp\left[-\frac{(Y_i - \theta_i)^2}{2\sigma_{Y_i}^2}\right] \quad (7)$$

To apply the Gibbs sampler we need to obtain the conditional posterior for each parameter ($\theta_i, \mu, \sigma_{\theta}^2$).

Conditional posterior of θ_i

Using equation (7), we calculate the conditional posterior of θ_i by considering θ_i as a random variable and μ, σ_{θ}^2 as constants.

$$P(\theta_i | \mu, \sigma_{\theta}^2, Y_{1:n}) = \frac{1}{\sqrt{2\pi\sigma_{\theta}^2}} \exp\left[-\frac{(\theta_i - \mu)^2}{2\sigma_{\theta}^2}\right] \times \frac{1}{\sqrt{2\pi\sigma_{Y_i}^2}} \exp\left[-\frac{(Y_i - \theta_i)^2}{2\sigma_{Y_i}^2}\right]$$

$$= \exp\left[-\frac{1}{2\sigma_{\theta}^2}(\theta_i - \mu)^2 - \frac{1}{2\sigma_{Y_i}^2}(\theta_i - Y_i)^2\right]$$

$$= \exp\left[-\frac{1}{2}\left[\theta_i^2\left(\frac{1}{\sigma_{\theta}^2} + \frac{1}{\sigma_{Y_i}^2}\right) - 2\theta_i\left(\frac{\mu}{\sigma_{\theta}^2} + \frac{Y_i}{\sigma_{Y_i}^2}\right)\right]\right]$$

Multiplying and dividing by $\left(\frac{1}{\sigma_{\theta}^2} + \frac{1}{\sigma_{Y_i}^2}\right)^{-1}$ and letting

$$\hat{\mu}_{\theta} = \frac{\left(\frac{\mu}{\sigma_{\theta}^2} + \frac{Y_i}{\sigma_{Y_i}^2}\right)}{\left(\frac{1}{\sigma_{\theta}^2} + \frac{1}{\sigma_{Y_i}^2}\right)}$$

produces

$$P(\theta_i | \mu, \sigma_{\theta}^2, Y_{1:n}) = \exp\left[-\frac{1}{2}\left(\frac{1}{\sigma_{\theta}^2} + \frac{1}{\sigma_{Y_i}^2}\right)[\theta_i^2 - 2\theta_i \hat{\mu}_{\theta}]\right]$$

Completing square to obtain the mean for the normal distribution is formed by summing and subtracting with $\hat{\mu}_{\theta}^2$

$$= \exp\left[-\frac{1}{2}\left(\frac{1}{\sigma_{\theta}^2} + \frac{1}{\sigma_{Y_i}^2}\right)[\theta_i^2 - 2\theta_i \hat{\mu}_{\theta} + \hat{\mu}_{\theta}^2]\right]$$

$$= \exp\left[-\frac{1}{2}\left(\frac{1}{\sigma_{\theta}^2} + \frac{1}{\sigma_{Y_i}^2}\right)[\theta_i - \hat{\mu}_{\theta}]^2\right]$$

Therefore, $\theta_i | \mu, \sigma_{\theta}^2, Y_i$ follows the normal distribution with mean $\hat{\mu}_{\theta}$ and corresponding variance-covariance matrix, $\hat{\sigma}_{\theta}^2$ where

$$\hat{\mu}_{\theta} = \frac{\mu\sigma_{Y_i}^2 + Y_i\sigma_{\theta}^2}{\sigma_{Y_i}^2 + \sigma_{\theta}^2} \quad \text{and} \quad \hat{\sigma}_{\theta}^2 = \frac{\sigma_{Y_i}^2\sigma_{\theta}^2}{\sigma_{Y_i}^2 + \sigma_{\theta}^2}$$

Conditional posterior of μ

Using equation (7), the conditional posterior of μ is calculated by considering μ as a random variable and $\theta_i, \sigma_{\theta}^2$ as constants. Hence:

$$P(\mu | \theta_{1:n}, \sigma_{\theta}^2, Y_{1:n}) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma_{\theta}^2}} \exp\left[-\frac{(\theta_i - \mu)^2}{2\sigma_{\theta}^2}\right] \times$$

$$\frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left[-\frac{(\mu - \theta_0)^2}{2\sigma_0^2}\right]$$

$$= \exp\left[-\frac{1}{2\sigma_{\theta}^2} \sum_{i=1}^n (\theta_i - \mu)^2 - \frac{1}{2\sigma_0^2} (\mu - \theta_0)^2\right]$$

$$= \exp\left[-\frac{1}{2}\left[\mu^2\left(\frac{n}{\sigma_{\theta}^2} + \frac{1}{\sigma_0^2}\right) - 2\mu\left(\frac{\sum_{i=1}^n \theta_i}{\sigma_{\theta}^2} + \frac{\theta_0}{\sigma_0^2}\right)\right]\right]$$

Multiplying and dividing by $\left(\frac{n}{\sigma_\theta^2} + \frac{1}{\sigma_0^2}\right)^{-1}$ and letting

$$\hat{\mu}_\mu = \frac{\left(\frac{\sum_{i=1}^n \theta_i + \theta_0}{\sigma_\theta^2} + \frac{\theta_0}{\sigma_0^2}\right)}{\left(\frac{n}{\sigma_\theta^2} + \frac{1}{\sigma_0^2}\right)} \text{ produces}$$

$$P(\mu | \theta_{1:n}, \sigma_\theta^2, Y_{1:n}) = \exp \left[-\frac{1}{2} \left(\frac{n}{\sigma_\theta^2} + \frac{1}{\sigma_0^2} \right) [\mu^2 - 2\mu\hat{\mu}_\mu] \right]$$

Completing square to obtain the mean for the normal distribution is formed by summing and subtracting with $\hat{\mu}_\mu^2$

$$\begin{aligned} &= \exp \left[-\frac{1}{2} \left(\frac{n}{\sigma_\theta^2} + \frac{1}{\sigma_0^2} \right) [\mu^2 - 2\mu\hat{\mu}_\mu + \hat{\mu}_\mu^2] \right] \\ &= \exp \left[-\frac{1}{2} \left(\frac{n}{\sigma_\theta^2} + \frac{1}{\sigma_0^2} \right) [\mu - \hat{\mu}_\mu]^2 \right] \end{aligned}$$

Therefore, $\mu | \theta_i, \sigma_\theta^2, Y_i$ follows the normal distribution with mean $\hat{\mu}_\mu$ and corresponding variance-covariance matrix, $\hat{\sigma}_\mu^2$ where

$$\hat{\mu}_\mu = \frac{\sigma_0^2 \sum_{i=1}^n \theta_i + \theta_0 \sigma_\theta^2}{n\sigma_0^2 + \sigma_\theta^2} \quad \text{and} \quad \hat{\sigma}_\mu^2 = \frac{\sigma_0^2 \sigma_\theta^2}{n\sigma_0^2 + \sigma_\theta^2}$$

Conditional posterior of σ_θ^2

Using equation (7), we calculate the conditional distribution of σ_θ^2 by considering σ_θ^2 as a random variable and μ, θ_i as constants. Hence:

$$P(\sigma_\theta^2 | \theta_i, \mu, Y_{1:n}) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma_\theta^2}} \exp \left[-\frac{(\theta_i - \mu)^2}{2\sigma_\theta^2} \right] x$$

$$\begin{aligned} &\frac{\beta^\alpha}{\Gamma(\alpha)} (\sigma_\theta^2)^{-\alpha-1} e^{-\beta/\sigma_\theta^2} \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} (\sigma_\theta^2)^{-(\alpha+\frac{n}{2})-1} \exp \left[-\frac{\beta + \frac{\sum_{i=1}^n (\theta_i - \mu)^2}{2}}{\sigma_\theta^2} \right] \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} (\sigma_\theta^2)^{-(\alpha+\frac{1}{2})-1} \exp \left[-\frac{\beta + \frac{\sum_{i=1}^n (\theta_i - \mu)^2}{2}}{\sigma_\theta^2} \right] \end{aligned}$$

Therefore

$$\sigma_\theta^2 | \theta_{1:n}, \mu, Y_{1:n} \sim \text{IG} \left(\alpha + \frac{n}{2}, \beta + \frac{\sum_{i=1}^n (\theta_i - \mu)^2}{2} \right)$$

The Gibbs sampler algorithm

The parameters of interest for the model are θ_i, μ , and σ_θ^2 . The aim of the Gibbs sampler is to simulate the joint posterior, $P(\theta_{1:n}, \mu, \sigma_\theta^2 | Y_{1:n})$ based on all the conditional posterior distributions. The Gibbs sampler algorithm for the present model comprises:

1. Initialising μ and σ_θ^2 from distribution

$$\mu^{(0)} \sim N(\theta_0, \sigma_0^2)$$

$$\sigma_\theta^{2(0)} \sim \text{IG}(\alpha, \beta)$$

2. Generate $\theta_i^{(j)}$ based on $\mu^{(j)}$ and $\sigma_\theta^{2(0)}$

$$\theta_i^{(1)} | \mu^{(0)}, \sigma_\theta^{2(0)} \sim N(\hat{\mu}_\theta, \hat{\sigma}_\theta^2)$$

where $\hat{\mu}_\theta = \frac{\mu^{(0)} \sigma_{Y_i}^2 + Y_i \sigma_\theta^{2(0)}}{\sigma_{Y_i}^2 + \sigma_\theta^{2(0)}}$ and $\hat{\sigma}_\theta^2 = \frac{\sigma_{Y_i}^2 \sigma_\theta^{2(0)}}{\sigma_{Y_i}^2 + \sigma_\theta^{2(0)}}$

3. Given $\theta_i^{(j)}$ and $\sigma_\theta^{2(0)}$, then generate $\mu^{(j)}$

$$\mu^{(j)} | \theta_i^{(j)}, \sigma_\theta^{2(0)} \sim N(\hat{\mu}_\mu, \hat{\sigma}_\mu^2)$$

where $\hat{\mu}_\mu = \frac{\sigma_0^2 \sum_{i=1}^n \theta_i^{(j)} + \theta_0 \sigma_\theta^{2(0)}}{n\sigma_0^2 + \sigma_\theta^{2(0)}}$ and $\hat{\sigma}_\mu^2 = \frac{\sigma_0^2 \sigma_\theta^{2(0)}}{n\sigma_0^2 + \sigma_\theta^{2(0)}}$

4. Given $\theta_i^{(j)}$ and $\mu^{(j)}$, then generate $\sigma_\theta^{2(j)}$

$$\sigma_\theta^{2(j)} | \theta_i^{(j)}, \mu^{(j)} \sim \text{IG} \left(\alpha + \frac{n}{2}, \beta + \frac{\sum_{i=1}^n (\theta_i^{(j)} - \mu^{(j)})^2}{2} \right)$$

Simulation study

A simulation study for the model was presented in [9]. Generated were 1,000 random samples from 30 studies using R, from the multivariate normal distribution. By fixing the overall mean and variance-covariance matrix, estimates of the true mean effect were generated. These estimates will be compared to the true effect to assess bias. Steps for the simulation study are as follows

Step 1. Fix the value of the overall mean (μ), and the variance-covariance matrices $\sigma_{Y_i}^2$ and σ_θ^2 .

Step 2. Generate θ_i based on μ (where n is the number of studies, $i = 1, 2, \dots, n$).

Step 3. Values of the observed statistics (Y_i) will be obtained based on θ_i

Estimation of Parameters

Parameters were estimated applying the Gibbs sampler algorithm using R. Steps conducted to estimate parameters are as follows

Step 1. Read the observed statistics (Y_i) and the variance-covariance matrix ($\sigma_{Y_i}^2$) as data and changing constants (on the diagonal) for each study, respectively.

Step 2. Generate all parameters ($\theta_i, \mu, \sigma_\theta$) using the Gibbs sampler algorithm in the above by fixing the initial values for the priors.

Step 3. The estimated overall mean is obtained by calculating the mean in step 2.

The length of 10,000 iterations were generated with a burn-in of 9,000 iterations conducted in order to achieve convergence of parameters. The results show the value of the overall mean ($\mu = 2.54$) as close to the true value (2.56). The credible interval and standard deviation are 2.11–2.98 and 0.23, respectively. Trace and density plots (Figure A and B, respectively) show convergence. The estimated overall mean (μ) is inside the credible interval consistently and close to the

target (true value). Geweke's convergence diagnostic test is based on the equality of the means of the first (by default the first 10%) and last part (by default the last 50%). Heidelberger and Welch's convergence diagnostic test is based on stationarity testing of the series [4]. The tests pointed out that the overall estimated mean to converge, indicating parameter stability.

Geweke test

Fraction in 1st window = 0.1
 Fraction in 2nd window = 0.5

var1
 1.32

Heidelberger-Welch test

Stationarity test	start iteration	p-value
passed	1	0.596

Halfwidth test	Mean	Halfwidth
passed	2.54	0.0268

Figure A. Trace plot of μ

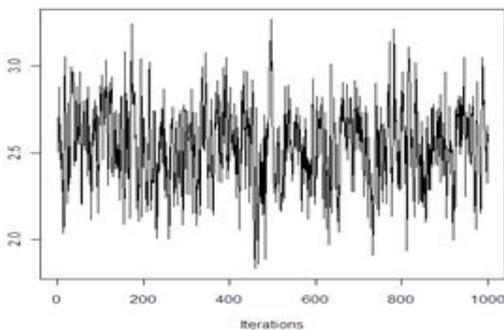
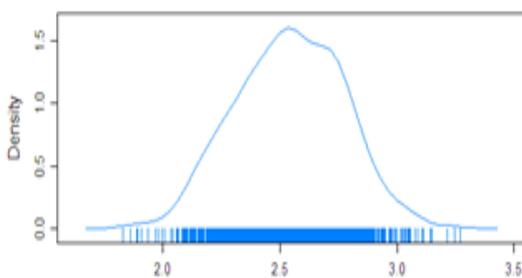


Figure B. Density plot of μ



4. Conclusion

We have presented the analytical form of a posterior distribution for a Bayesian model in meta-analysis. The conditional distributions of all parameters were obtained using the Gibbs sampling approach. Results from the simulation study showed the overall

estimated mean as close to the true value, indicating the estimator as consistent and stable.

5. Future Work

Although derivation of the conditional distributions can be relatively easy by applying Gibbs sampler, it is not always possible to find an efficient way to sample from these distributions [11]. Work in progress investigates another approach [10] to obtaining the posterior distribution.

Acknowledgments

The first author in this research was supported by a Directorate General of Higher Education Indonesia (DIKTI) scholarship.

References

- [1] Chen, Y., and Pei, J, "An assessment of a TNF polymorphic marker for the risk of HCV infection: Meta-analysis and a new clinical study design." *Infection, Genetics and Evolution*, **9**: 1356-1363, 2009.
- [2] Chugh, R., Wathen, J.K., Maki, R.G. et al., "Phase II Multicenter Trial of Imatinib in 10 Histologic Subtypes of Sarcoma Using a Bayesian Hierarchical Statistical Model." *Journal of Clinical Oncology*, **27**: 3148-3153, 2009.
- [3] Congdon, P., *Bayesian Statistical Modelling*, John Wiley & Sons, 2006.
- [4] Cowles, M, K., Carlin, B. P., "Markov Chain Monte Carlo Convergence Diagnostics: A comparative review", *Journal of the American Statistical Association*, **91**: 883-904, 1996..
- [5] Dohoo, I., Stryhn, H. et al., "Evaluation of underlying risk as a source of heterogeneity in meta-analyses: A simulation study of Bayesian and frequentist implementations of three models." *Preventive Veterinary Medicine*, **81**(1-3): 38-55, 2007.
- [6] Dumouchel, W. "Bayesian meta-analysis", *Statistical Methodology in the Pharmaceutical Sciences*, 509-529, 1990.
- [7] Hossein-Zadeh, N. G., "Bayesian estimate of genetic parameters for metritis, retained placenta, milk fever, and clinical mastitis in Holstein dairy cows via Gibbs sampling." *Research in Veterinary Science*, **90**:146-149, 2011.
- [8] Jones, D. R. et al., "Interspecies extrapolation in environmental exposure standard setting: A Bayesian synthesis approach." *Regulatory Toxicology and Pharmacology*, **53**(3): 217-225, 2009.
- [9] Junaidi, Stojanovski, E. Nur, D., "Prior Sensitivity Analysis for a Hierarchical Model", *Proceeding of the Fourth Annual ASEARCH Conference*, 64-67,2011.
- [10] Lindley, D. V., Smith, A. F. M., "Bayes Estimates for the Linear Model." *Journal of the Royal Statistical Society*, **34**(1): 1-41, 2011.
- [11] Tanner, M. A., *Tools for statistical inference: methods for the exploration of posterior distribution and likelihood function*, Springer, New York, 1996.
- [12] Wang, H., Chai, X., Liu, B., "Estimation of genetic parameters for growth traits in cultured clam Meretrix meretrix (Bivalvia: Veneridae) using the Bayesian method on Gibbs sampling." *Aquaculture Research*, **42**: 240 – 247, 2011.

