Bayesian Estimation of a Meta-analysis model using Gibbs sampler

Junaidi 
*University of Newcastle*

Darfiana Nur 
*University of Newcastle*

Elizabeth Stojanovski 
*University of Newcastle*

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

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.

This conference paper is available at Research Online: http://ro.uow.edu.au/asearc/5
Bayesian Estimation of a Meta-analysis model using Gibbs sampler

Junaidi$^{1,2}$, D. Nur$^1$, E. Stojanovski$^1$

$^1$The University of Newcastle, Callaghan, NSW, 2308, AUSTRALIA  
$^2$The 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

\[ Y_i \sim N(\theta_i, \sigma^2_i), \quad i = 1, 2, ..., n \]  
\[ \theta_i \sim N(\mu, \sigma^2), \]  
\[ \mu \sim N(0, D \to \infty). \]

Level 1 (1) of the model represents study data, where $Y_i$ denotes the observed study statistics with variance-covariance matrix, $\sigma^2_i$ and $n (i = 1, 2, ..., n)$ denotes the number of studies. Level 2 (2) represents the study-specific parameters, denoted $\theta$ and associated variance-covariance matrix $\sigma^2$. The overall mean is denoted by $\mu$. The prior distribution for $\mu$ is assumed to follow the normal distribution with mean, zero, and variance $D \to \infty$ 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:

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

The prior distributions for parameters \( \theta_i \) and \( \mu \) follow the normal distribution while the prior for \( \sigma_A^2 \) follows the inverse gamma:

\[ P(\sigma_A^2) = \frac{1}{\Gamma(\alpha)} (\frac{\sigma_A^2}{\beta})^{\alpha - 1} e^{-\beta/2\sigma_A^2} \]

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

\[ P(\theta_{1:n}, \mu, \sigma_A^2 | Y_{1:n}) = \frac{P(Y_{1:n} | \theta_{1:n}, \mu, \sigma_A^2)P(\theta_{1:n}, \mu, \sigma_A^2)}{P(Y_{1:n})} \propto P(\theta_{1:n}, \mu, \sigma_A^2 | Y_{1:n}) \]

The joint posterior becomes

\[ P(\theta_{1:n}, \mu, \sigma_A^2 | Y_{1:n}) = P(\mu)P(\sigma_A^2)P(\theta_{1:n} | \mu, \sigma_A^2) \]

(assuming independence on \( \mu \) and \( \sigma_A^2 \))

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

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

(7)

To apply the Gibbs sampler we need to obtain the conditional posterior for each parameter \( \theta_i, \mu, \sigma_A^2 \).

Conditional posterior of \( \theta_i \)

Using equation (7), we calculate the conditional posterior of \( \theta_i \) by considering \( \theta_i \) as a random variable and \( \mu, \sigma_A^2 \) as constants.

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

(7)

Completing square to obtain the mean for the normal distribution is formed by summing and subtracting with \( \hat{\mu}_i \) and corresponding variance-covariance matrix, \( \hat{\sigma}_i^2 \) where

\[ \hat{\mu}_i = \frac{\mu \sigma_A^2_i + \sigma_A^2 \theta_i}{\sigma_A^2_i + \sigma_A^2} \quad \text{and} \quad \hat{\sigma}_i^2 = \frac{\sigma_A^2 \sigma_A^2_i}{\sigma_A^2_i + \sigma_A^2} \]

Conditional posterior of \( \mu \)

Using equation (7), the conditional posterior of \( \mu \) is calculated by considering \( \mu \) as a random variable and \( \theta_i, \sigma_A^2 \) as constants. Hence:

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

(7)
Multiplying and dividing by \( \frac{1}{\sigma_\theta^2} \) and letting

\[
\hat{\mu}_\mu = \frac{\sum_{i=1}^n \theta_i + \frac{1}{\sigma_\theta^2}}{\frac{n}{\sigma_\theta^2} + \frac{1}{\sigma_\theta^2}}
\]

produces

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

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

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

Therefore, \( \mu | \theta_i, \sigma_\theta^2, Y_{1:n} \) follows the normal distribution with mean \( \hat{\mu}_\mu \) and corresponding variance-covariance matrix, \( \hat{\sigma}_\mu^2 \)

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

Conditional posterior of \( \sigma_\theta^2 \)

Using equation (7), we calculate the conditional distribution of \( \sigma_\theta^2 \) by considering \( \sigma_\theta^2 \) as a random variable and \( \mu, \theta_i \) as constants. Hence:

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

\[
\frac{\beta^2}{\sqrt{\alpha} (\sigma_\theta^2)^{\alpha-1}} e^{-\beta/\sigma_\theta^2}
\]

\[
= \frac{\beta^2}{\sqrt{\alpha} (\sigma_\theta^2)^{\alpha-1}} e^{-\beta/\sigma_\theta^2} \exp \left[ -\frac{\sigma_\theta^2 \sum_{i=1}^n (\theta_i - \mu)^2}{\alpha \sigma_\theta^2} \right]
\]

\[
= \frac{\beta^2}{\sqrt{\alpha} (\sigma_\theta^2)^{\alpha-1}} e^{-\beta/\sigma_\theta^2} \exp \left[ -\frac{\sum_{i=1}^n (\theta_i - \mu)^2}{\alpha \sigma_\theta^2} \right]
\]

Therefore

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

The Gibbs sampler algorithm

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

1. Initialising \( \mu \) and \( \sigma_\theta^2 \) from distribution

\[
\mu^{(0)} \sim N(\hat{\mu}_0, \sigma_\theta^2)
\]

\[
\sigma_\theta^{2(0)} \sim IG(\alpha, \beta)
\]

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

\[
\theta_i^{(l)} | \mu^{(l)}, \sigma_\theta^{2(l)} \sim N(\mu^{(l)}, \sigma_\theta^{2(l)})
\]

where \( \mu_0 = \frac{\mu^{(l)} \sigma_\theta^{2(l)} + 1 \sigma_\theta^{2(l)}}{\sigma_\theta^{2(l)} + \sigma_\theta^{2(l)}} \) and \( \sigma_\theta^{2(l)} = \frac{\sigma_\theta^{2(l)} \sigma_\theta^{2(l)}}{\sigma_\theta^{2(l)} + \sigma_\theta^{2(l)}} \)

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

\[
\mu^{(l)} | \theta_i^{(l)}, \sigma_\theta^{2(l)} \sim N(\hat{\mu}_\mu, \hat{\sigma}_\mu^2)
\]

where \( \hat{\mu}_\mu = \frac{\sigma_\theta^{2(l)} \sum_{i=1}^n \theta_i + \frac{1}{\sigma_\theta^{2(l)}}}{\frac{n \sigma_\theta^{2(l)}}{\sigma_\theta^{2(l)}} + \frac{1}{\sigma_\theta^{2(l)}}} \) and \( \hat{\sigma}_\mu^2 = \frac{\sigma_\theta^{2(l)} \sigma_\theta^{2(l)}}{\frac{n \sigma_\theta^{2(l)}}{\sigma_\theta^{2(l)}} + \frac{1}{\sigma_\theta^{2(l)}}} \)

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

\[
\sigma_\theta^{2(l)} | \theta_i^{(l)}, \mu^{(l)} \sim IG(\alpha + \frac{n}{2}, \beta + \frac{\sum_{i=1}^n (\theta_i^{(l)} - \mu^{(l)})^2}{2})
\]

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 \( \mu \), and the variance-covariance matrices \( \sigma_{\theta_i}^2 \) and \( \sigma_{\mu}^2 \).

**Step 2.** Generate \( \theta_i \) based on \( \mu \) (where \( n \) is the number of studies, \( i = 1, 2, \ldots, n \)).

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

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 matrices \( (\sigma_{\theta_i}^2) \) as data and changing constants (on the diagonal) for each study, respectively.

**Step 2.** Generate all parameters \( (\theta_i, \mu, \sigma_\theta^2) \) 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 \( (\mu) \) is inside the credible interval consistently and close to the
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
