On Bayesian Analysis of Generalized Linear Models Using the Jacobian Technique

Sourish Das and Dipak K. Dey

In this article, we obtain an estimator of the regression parameters for generalized linear models, using the Jacobian technique. We restrict ourselves to the natural exponential family for the response variable and choose the conjugate prior for the natural parameter. Using the Jacobian of transformation, we obtain the posterior distribution for the canonical link function and thereby obtain the posterior mode for the link. Under the full rank assumption for the covariate matrix, we then find an estimator for the regression parameters for the natural exponential family. Then the proposed estimator is specially derived for the Poisson model with log link function, and the binomial response model with the logit link function. We also discuss extensions to the binomial response model when covariates are all positive. Finally, an illustrative real-life example is given for the Poisson response model when covariates are all positive. Furthermore, an illustrative real-life example is given for the Poisson response model when covariates are all positive.

KEY WORDS: Bernstein-von Mises theorem; Canonical Link; Conjugate prior; Log link; Natural exponential family; Posterior mode; Regression parameter.

1. INTRODUCTION

Bayesian analysis of generalized linear models (GLIM) requires specification of a prior for the regression parameters used in the model. Uniform priors are very commonly used as conventional noninformative priors. However, Ibrahim and Laud (1991) showed that the uniform prior for the regression parameters in a GLIM could lead to improper posterior distributions, thus making the uniform prior undesirable. They further proved two theorems that support the use of Jeffreys’s priors for GLIMs with intrinsically fixed or known scale parameters. The posterior calculations in such types of problem are very complicated and require a sophisticated algorithm to estimate the regression parameters. These algorithms often face the problem of lack of convergence. If someone is interested in developing the maximum likelihood estimates, then Fisher’s scoring method can be used, which also often suffers from lack of convergence. Our motivation in this article is to find a simple technique for estimating the regression parameters from a Bayesian viewpoint.

In the next section, we develop a methodology for finding posterior mode for a link function, using the Jacobian technique in a GLIM. In this methodology, the distribution for the response belongs to the natural exponential family. We restrict ourselves to a canonical link and choose the conjugate prior for the natural parameter.

2. JACOBIAN TECHNIQUE

Suppose \( y_1, y_2, \ldots, y_n \) are independent observations, where \( y_i \) has the density from the natural exponential family

\[
 f(y_i | \theta_i) = \exp \{ \theta_i y_i - \psi(\theta_i) + c(y_i) \},
\]

where \( i = 1, 2, \ldots, n \). The density in (1) is parameterized by the canonical parameter \( \theta_i \). The \( \psi(\cdot) \), and \( c(\cdot) \) are known functions and the \( \theta_i \)'s are related to the regression coefficients by the link function

\[
 \theta_i = g(\eta_i),
\]

where \( i = 1, 2, \ldots, n \), and

\[
 \eta_i = x_i' \beta
\]

is the systematic component of GLIM. In (3) \( x_i = (x_{i1}, x_{i2}, \ldots, x_{ip}) \) is a \( 1 \times p \) vector denoting the \( i \)th row of \( n \times p \) matrix of covariates \( X, \beta = (\beta_1, \ldots, \beta_p)' \) is a \( p \) vector of regression coefficient, and \( g(\cdot) \) is a monotonic differentiable function. The model given by (1), (2), and (3) is called the generalized linear model (GLIM). The Normal, logistic, and Poisson regression models are special cases of the GLIM; see McCullagh and Nelder (1989) for more details.

It follows from Diaconis and Ylvisker (1979) that the conjugate prior distribution for \( \theta_i \) is

\[
 \pi(\theta_i) = K \exp \{ m \mu_0 \theta_i - m \psi(\theta_i) \},
\]

where the normalizing constant \( K = K(m, \mu_0) \) is chosen such that \( \pi(\theta_i) \) is a proper density function. Here \( \pi(\theta_i) \) is a two-parameter natural exponential family of densities for \( \theta_i \) with the natural parameters \( m \) and \( \mu_0 \).
An important issue in any Bayesian analysis is the specification of a prior distribution. A commonly used prior in Bayesian analysis is the conjugate prior. It follows immediately from (4) that the posterior distribution of $\theta_i$ is

$$\pi(\theta_i \mid y_i) = K \left( \mu_0 + 1, \frac{y_i + m\mu_0}{1 + m} \right) \times \exp \{ (y_i + m\mu_0)\theta_i - (1 + m)\psi(\theta_i) \}. \quad (5)$$

Now the link function in (2) can be considered as a simple monotonic transformation of $\theta_i$. Then it follows clearly that the Jacobian of transformation from $\theta_i$ to $\eta_i$ is

$$J \left( \frac{\theta_i}{\eta_i} \right) = g'(\eta_i).$$

Hence we have the posterior density for $\eta_i$ is

$$\pi_1(\eta_i \mid y_i) = K \left( \mu_0 + 1, \frac{y_i + m\mu_0}{1 + m} \right) \times \exp \{ (y_i + m\mu_0)\eta_i - (1 + m)\psi(\eta_i) \} g'(\eta_i), \quad (6)$$

where $K(k, a, b) = (1 + b)^{-a} \Gamma(k)$. The following theorem provides the posterior mode for $\eta_{i}, i = 1, 2, \ldots, n$.

**Theorem 1.** Suppose the posterior distribution of $\eta_i$ is given by (6) and $g(\cdot)$ is monotonic twice differentiable function and $g'(\eta_i) \neq 0$. Then the solution of the following equation

$$\psi'(g(\eta_i)) = \frac{y_i + m\mu_0}{1 + m} + \frac{g''(\eta_i)}{g'(\eta_i)^2} \frac{1}{(1 + m)} \quad (7)$$

gives the posterior mode of $\eta_i$.

**Proof:** Taking logarithm on both sides of the posterior distribution of $\eta_i$ as given in (6), it follows that

$$\log(\pi_1(\eta_i \mid y_i)) = C + \{(y_i + m\mu_0)\psi(\eta_i) \} + \log(g'(\eta_i)), \quad \text{where } C \text{ is a constant.}$$

Thus, setting

$$\frac{\partial \log \pi_1(\eta_i \mid y_i)}{\partial \eta_i} = 0$$

implies that

$$(y_i + m\mu_0) + g'(\eta_i) - (1 + m)\psi'(g(\eta_i))g'(\eta_i) + \frac{g''(\eta_i)}{g'(\eta_i)^2} = 0,$$

which in turn is equivalent to \{(y_i + m\mu_0) - (1 + m)\psi'(g(\eta_i)) + \frac{g''(\eta_i)}{g'(\eta_i)^2} \} = 0, since $g'(\eta_i) \neq 0$. Hence the posterior mode for $\eta_i$ is obtained through (7).

Once we have the posterior mode of the $\eta_i$, we obtain the estimate of the regression parameters under the full rank assumption of the covariate matrix. In the next section, we develop the methodology for the Bayesian estimation of the regression parameters for a GLIM.

### 3. ESTIMATION OF REGRESSION PARAMETERS

Let $\hat{\eta}_i$ be the posterior mode of $\eta_i$, obtained through (7). It follows from (3) that

$$\hat{\eta}_i = x_i'\hat{\beta},$$

where $i = 1, \ldots, n$ and using vector notation we can express it as

$$\hat{\eta} = X\hat{\beta}, \quad (8)$$

where $X = (x_1', x_2', \ldots, x_n')'$ and $\beta = (\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_k)'$.

Thus, $X'\hat{\eta} = X'X\hat{\beta}$ and under the full rank assumption we have

$$\hat{\beta} = (X'X)^{-1}X'\hat{\eta}. \quad (9)$$

This newly proposed estimator looks like the least squares estimator of $\beta$ under a linear model. In the next section, we derive $\hat{\beta}$ for some standard problems. We consider the binomial response model with logit link function, Poisson model, with log link function, and some other instances of the GLIM.

### 4. POSTERIOR ESTIMATOR OF THE REGRESSION PARAMETERS FOR SOME STANDARD INSTANCES OF GLIM

In this section, we find the posterior estimator for the regression parameters for the Binomial response model with logit link function, the Poisson model with log link function, the Negative binomial response with corresponding canonical link function and also for one extension of the standard GLIM. In the following example we consider the Poisson response model with sample size $n$ in each of $k$ groups with log links.

#### 4.1 Poisson Response Model with Sample of Size $n$ in Each of $k$ Groups with Log Links

Suppose $Y_{ij} \sim \text{Poisson}(\lambda_i), i = 1, \ldots, n, j = 1, \ldots, k$. Then $T_i = \sum_{j=1}^{k} Y_{ij} \sim \text{Poisson}(k\lambda_i)$, which is also sufficient for $\lambda_i$. Therefore the likelihood of $T_i$ is

$$f(t_i \mid \theta_i) = \frac{\exp \{ t_i \log k\lambda_i - k\lambda_i \}}{t_i!} = \frac{\exp \{ t_i \theta_i - e^\theta_i \}}{t_i!},$$

where $\theta_i = \log(k\lambda_i)$ is the natural parameter of the natural exponential family. Hence the required likelihood function is $l(t_i \mid \theta_i) \propto \exp \{ t_i \theta_i - \psi(\theta_i) \}$. Consider the log-link as

$$\log(k\lambda_i) = \eta_i = \theta_i; i = 1, \ldots, n \text{ with systematic component as } \eta_i = x_i'\beta, i = 1, \ldots, n.$$ 

Now we choose the conjugate prior for $\lambda_i$ as Gamma(a, b) which leads to the posterior distribution of $\lambda_i$ as Gamma($t_i + a, (1 + \frac{1}{k})^{-1}$). By Theorem 1, the posterior mode is

$$\hat{\eta}_i = \log \left( \frac{t_i + a}{1 + \frac{1}{k}} \right), \quad i = 1, \ldots, n.$$ 

The matrix of covariates is reduced to $X = (\{x_{il}\})$, where

$$\pi_{il} = \frac{1}{k} \sum_{j=1}^{k} x_{ilj}; \quad i = 1, \ldots, n; \quad l = 1, \ldots, p(\leq k).$$

Hence the required estimator of $\beta$ is

$$\hat{\beta} = (X'X)^{-1}X'\hat{\eta} = (X'X)^{-1}X'\log \left( \frac{t + a}{1 + \frac{1}{k}} \right),$$

where $t = (t_1, t_2, \ldots, t_n)'$ and the logarithm is defined component wise. In the next example we consider the binomial response model with logit link.
4.2 Binomial Response Model With Logit Link

Suppose \( y_1, \ldots, y_n \) are \( n \) independent observations from Binomial(\( m_i, p_i \)) model with probability density

\[
f(y_i \mid p_i) = (m_i) p_i^{y_i} (1 - p_i)^{m_i - y_i},
\]

for all \( i = 1, \ldots, n \).

In other words, we can express the likelihood as

\[
f(y_i \mid p_i) \propto \exp \{ y_i \theta_i - \psi(\theta_i) \},
\]

where

\[
\theta_i = \log \left( \frac{p_i}{1 - p_i} \right),
\]

and

\[
\psi(\theta_i) = m_i \log(1 + e^{\theta_i}).
\]

Here \( \theta_i \) is the natural parameter of the natural exponential family with systematic component given under the logit link as \( \log \left( \frac{p_i}{1 - p_i} \right) = \eta_i = x'_i \beta, \ i = 1, \ldots, n \). We choose the conjugate prior for \( p_i \) as Beta(\( a, b \)). Therefore the conjugate prior for the natural parameters \( \theta_i \) is \( \pi(\theta_i) \propto \exp \{ a \theta_i - \left( \frac{a + b}{m_i} \right) m_i \log(1 + e^{\theta_i}) \}, \) where \( m_0 = a, m = \frac{a+b}{m_i} \). Now by the Theorem 1, we know the solution of Equation (7) gives the posterior mode of \( \eta_i \). For our case, \( g(\eta_i) = \eta_i, \) and thus \( g'(\eta_i) = 1 \) and \( g''(\eta_i) = 0, \) we have

\[
\begin{align*}
\hat{\eta}_i &= \frac{y_i + a}{m_i + b} \\
\end{align*}
\]

Hence the posterior mode will be \( \hat{\eta}_i = \log \left( \frac{y_i + a}{m_i + b} \right) \) and the required estimator for \( \beta \) is

\[
\hat{\beta} = (X'X)^{-1}X'y = (X'X)^{-1}X' \log \left( \frac{y + a}{b + m - y} \right),
\]

where \( y = (y_1, y_2, \ldots, y_n)' \) and the logarithm is defined component wise. Note that if the response variable has the Bernoulli distribution, then \( m_i \) will be equal to unity for all \( i \). Setting \( a = 1 \) and \( b = 1 \) for flat prior, the resulting \( \hat{\beta} \) simplifies to

\[
\hat{\beta} = (X'X)^{-1}X' \log \left( \frac{y + 1}{2 - y} \right).
\]

In the next example, we will consider the binomial response model, where we use some link function other than logit link.

4.3 Binomial Response Model When Covariates are All Positive

This example considers an extension of the standard GLIM, which demonstrates the application of Theorem 1. Suppose \( y_1, \ldots, y_n \) are \( n \) independent observations from Binomial(\( m_i, p_i \)) as defined in section 4.2. Consider the link function as

\[
\theta_i = \log(\eta_i) = \log(x'_i \beta + \alpha).
\]

Note that we can express more concisely as

\[
\eta_i = x'_i \beta^*, \text{ where } x'_i = (1, x_i) \text{ and } \beta^* = (\alpha, \beta)',
\]

Again consider the prior for \( p_i \) as Beta(\( a, b \)), where \( \theta_i = \log \left( \frac{p_i}{1 - p_i} \right) \). Therefore the conjugate prior for the natural parameter \( \theta_i \) is

\[
\pi(\theta_i) \propto \exp \left\{ a \theta_i - \left( \frac{a + b}{m_i} \right) m_i \log(1 + e^{\theta_i}) \right\}.
\]

Following Theorem 1, and setting \( m_0 = a, m = \frac{a+b}{m_i} \) and \( g(\eta_i) = \log(\eta_i) \), we have,

\[
\frac{m_i e^{\log(\eta_i)}}{1 + e^{\log(\eta_i)}} = \frac{(y_i + a)m_i}{m_i + a + b} - \frac{m_i}{m_i + a + b},
\]

which is equivalent to

\[
e^{\log(\eta_i)} \left( \frac{m_i}{m_i + a + b} - 1 \right) = \frac{y_i + a - 1}{m_i + a + b}.
\]

Hence the posterior mode for \( \eta_i \) is

\[
\hat{\eta}_i = \frac{y_i + a - 1}{m_i + y_i + b + 1}.
\]

Therefore the required estimator for the regression parameter is

\[
\hat{\beta} = (X'X)^{-1}X'y,
\]

\[
= (X'X)^{-1}X' \log \left( \frac{y + a - 1}{m - y + b + 1} \right),
\]

where \( y = (y_1, \ldots, y_n)' \), \( m = (m_1, \ldots, m_n)' \), and logarithm is defined component wise.

In the next example, we will consider the negative binomial response model, where we use the corresponding canonical link function. This model incorporates over dispersion.

4.4 Negative Binomial Response Model With Canonical Link Function

In real life there are various data which suffer from overdispersion. Negative-binomial response model is one of the many models which take care of overdispersion in the data and also this model is a special case of our general set-up. Suppose \( y_1, \ldots, y_n \) are \( n \) independent observations from Negative binomial(\( r, p_i \)) with probability distribution

\[
f(y_i \mid p_i) = \binom{y_i + r - 1}{y_i} p_i^r (1 - p_i)^{y_i},
\]

for all \( i = 1, \ldots, n \).

Consider the canonical link \( \eta_i = \log \left( \frac{r(y_i + b)}{a + r} \right) \). Under the conjugate prior, \( p_i \sim \text{Beta}(a, b) \), the posterior distribution of \( p_i \) is Beta(\( a + r, y_i + b \)). Hence by using Theorem 1, the posterior mode for \( \eta_i \) is

\[
\hat{\eta}_i = \log \left( \frac{r(y_i + b)}{a + r} \right).
\]

Therefore our required estimator for \( \beta \) is

\[
\hat{\beta} = (X'X)^{-1}X'y = (X'X)^{-1}X' \log \left( \frac{r(y + b)}{a + r} \right),
\]

where \( y = (y_1, \ldots, y_n)' \) and the logarithm is defined component wise. Note that if the response variable has geometric distribution, then \( r \) will be equal to unity. Then choose \( a = 1, b = 1, \)

\[\text{The American Statistician, August 2006, Vol. 60, No. 3} \]
since our motivation is to choose a flat prior and as a result the required estimator for $\beta$ is

$$\hat{\beta} = (X'X)^{-1}X' \log \left( \frac{y + 1}{2} \right),$$

where $y = (y_1, \ldots, y_n)'$ and the logarithm is defined component wise. All the results are summarized in Table 1.

In the next section, we develop the credible interval for the regression estimator for general case, using the Bernstein-von Mises theorem.

5. CREDIBLE INTERVAL FOR THE REGRESSION PARAMETERS

In this section, we develop a credible interval for the regression parameters using the asymptotic normality of the posterior distribution. In order to find the standard error of the estimator, we use the Bernstein-von Mises theorem, to obtain the asymptotic normality. From Ferguson (1996) suppose $X_1, X_2, \ldots$ are iid random variables having likelihood function $L(\theta) = \prod_{i=1}^{n} f(X_i \mid \theta)$ and $\pi(\theta)$ is the prior pdf which is continuous over the parameter space $\Theta$ and the posterior pdf of $\theta$ is $\pi(\theta \mid X)$. Suppose further $\hat{\theta}_n$ is the maximum likelihood estimator of $\theta$ and $I(\hat{\theta}_0)$ is the Fisher’s information, when $\theta_0$ is the true value. Then the posterior distribution of $\sqrt{n}(\theta - \hat{\theta}_n)$ approaches the density of $N(0, I(\theta_0)^{-1})$ as $n \to \infty$. Further this limiting posterior distribution is independent of the prior distribution $\pi(\theta)$. For our case, using the Bernstein-von Mises theorem, the posterior distribution of

$$\sqrt{n}(\eta - \hat{\eta}_n) \to N(0, I(\eta_0)^{-1}),$$

as $n \to \infty$, when $\hat{\eta}_n$ is the MLE and $\eta_0$ is the true value. Defining $\Sigma = \text{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2)$, with $\sigma_i^2 = (nI(\eta_0))^{-1}$ and $\hat{\beta} = A\hat{\eta}$, where $A = (X'X)^{-1}X'$, it follows that the posterior distribution of

$$\sqrt{n}(\beta - \hat{\beta}_n) \to N(0, A\Sigma A')$$

as $n \to \infty$, where $\hat{\beta}_n = A\hat{\eta}_n$ and the true value of $\beta$ is embedded in $\Sigma$. Therefore, for large $n$, we can compute the $100 \times (1 - \alpha)$% credible interval for $\beta$.

In the next section we consider a real life dataset on the Poisson response model with certain covariates.

6. A REAL DATA EXAMPLE

In this section we consider a Poisson model from a real life dataset. The study investigated factors affecting whether the female crab had any other males, called satellites, residing nearby. Explanatory variables are the female crab’s color, spine condi-

Table 1. Regression Estimators of the Standard Models

<table>
<thead>
<tr>
<th>Model</th>
<th>Link function</th>
<th>Prior over the natural parameter $\theta_i$</th>
<th>Posterior estimator of regression parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poisson</td>
<td>Log link</td>
<td>Gamma($a, b$)</td>
<td>$\hat{\beta} = (X'X)^{-1}X' \log \left( \frac{t + a}{1 + \frac{a}{b}} \right)$</td>
</tr>
<tr>
<td>Binomial</td>
<td>Logit link</td>
<td>Beta($a, b$)</td>
<td>$\hat{\beta} = (X'X)^{-1}X' \log \left( \frac{y + a}{b + m - y} \right)$</td>
</tr>
<tr>
<td>Binomial with all positive covariates</td>
<td>$\theta_i = \log \eta_i$</td>
<td>Beta($a, b$)</td>
<td>$\hat{\beta}^* = (X'X)^{-1}X' \log \left( \frac{y + a - 1}{m - y + b + 1} \right)$,</td>
</tr>
<tr>
<td>Negative binomial</td>
<td>Canonical link</td>
<td>Beta($a, b$)</td>
<td>$\hat{\beta} = (X'X)^{-1}X' \log \left( \frac{r(y + b)}{a + r} \right)$</td>
</tr>
</tbody>
</table>

Table 2. Estimates of the Regression Parameters Using the MLE and the Jacobian Method, Under Two Different Choices of Prior

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$\hat{\beta}_0$</th>
<th>$\hat{\beta}_1$</th>
<th>$\hat{\beta}_2$</th>
<th>$\hat{\beta}_3$</th>
<th>$\hat{\beta}_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MLE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estimates</td>
<td>$-0.4541$</td>
<td>$-0.1916$</td>
<td>$0.0372$</td>
<td>$0.0254$</td>
<td>$0.4728$</td>
</tr>
<tr>
<td>s.e.</td>
<td>$0.9466$</td>
<td>$0.0665$</td>
<td>$0.0568$</td>
<td>$0.0479$</td>
<td>$0.1653$</td>
</tr>
<tr>
<td><strong>Posterior mode</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estimates</td>
<td>$0.2098$</td>
<td>$-0.1050$</td>
<td>$0.0178$</td>
<td>$0.0143$</td>
<td>$0.3079$</td>
</tr>
<tr>
<td>s.e.</td>
<td>$0.8087$</td>
<td>$0.0543$</td>
<td>$0.0489$</td>
<td>$0.0419$</td>
<td>$0.1517$</td>
</tr>
<tr>
<td><strong>Gamma(5/2, 2)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posterior Mode</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estimates</td>
<td>$-0.5145$</td>
<td>$-0.1590$</td>
<td>$0.0346$</td>
<td>$0.0264$</td>
<td>$0.4227$</td>
</tr>
<tr>
<td>s.e.</td>
<td>$1.0775$</td>
<td>$0.0723$</td>
<td>$0.0636$</td>
<td>$0.0562$</td>
<td>$0.2034$</td>
</tr>
</tbody>
</table>
tion, weight, and carapace width. The response outcome for each female crab is her number of satellites. Color has four levels (1 = light medium; 2 = medium; 3 = dark medium; 4 = dark). Spine condition has three levels (1 = both good; 2 = one worn or broken; 3 = both worn and/or broken). Carapace width (cm) and weight (kg) are two other explanatory variables other than spine condition and color. The dataset, which was described by Agresti (2002), resulted from a study conducted by Jane Brockmann of the Zoology Department at the University of Florida.

We consider a log link for the model with four covariates with \( \beta_1, \beta_2, \beta_3, \) and \( \beta_4 \), representing the regression coefficients for the corresponding covariates, namely color, spine condition, carapace width (cm), weight (kg). An intercept \( \beta_0 \) is also included in the model. We estimate the parameters by MLE using S-Plus software, which uses the Fisher scoring method. We also use S-Plus to calculate the same regression parameters using our method. The calculations are based on two separate two-parameter Gamma priors with the first prior having mean 5, and a variance of 10, while the second prior having the same mean but a variance of 20 to represent more non-informative prior. The results are summarized in Table 2.

It follows from Table 2 that when we have the prior with mean 5 and variance 10—that is, when we have a more informative prior—the estimates are farther away from the maximum likelihood estimates, but the posterior standard errors are lower than the standard error of the maximum likelihood estimates. On the other hand, when we have the same prior with the same mean 5, but variance as 20, which is a more noninformative prior, then the estimates are very close to the maximum likelihood estimates, but the posterior standard errors are higher than the standard error of the maximum likelihood estimates. It seems to be that there is a trade-off between bias and variance. However, since our main objective is to estimate the regression parameters, one should choose the more noninformative prior to get a good estimate of the regression parameter.

7. CONCLUSION

We have developed a closed form estimator of the regression parameter, from the Bayesian point of view, for most of the standard GLIM, such as the binomial response model with logit link function, or the Poisson model with log link function. Because this is a closed form estimator, we avoid any issue regarding convergence problems, which often appear when using other approaches. From the example discussed in Section 6, this method seems to work reasonably well, if one chooses a reasonably flat prior. Estimates from the study are quite close to the MLE, but posterior standard errors for our method are higher than that of the MLE, when the prior is more noninformative. However, for a more informative prior, posterior standard errors are lower than the standard error of the MLE’s, but the estimates we get from our technique, are farther away from the likelihood estimates. Since our objective is to estimate the regression parameter, we should choose a flatter prior for this purpose.

[Received July 2005. Revised June 2006.]

REFERENCES