Scholars Journal of Physics, Mathematics and Statistics

OPEN ACCESS

Abbreviated Key Title: Sch J Phys Math Stat
ISSN 2393-8056 (Print) | ISSN 2393-8064 (Online)
Journal homepage: https://saspublishers.com

Parameters Estimation Sensitivity of the Linear Mixed Model To Alternative Prior Distribution Specifications

Nasir Elmesmari[1*], Farag Hamad[1], Abdelbaset Abdalla[2]

[1]Department of Statistics, Faculty of arts and Science/ Al Maraj, University of Benghazi, Benghazi, Libya
[2]Department of Statistics, Faculty of Science, University of Benghazi, Benghazi, Libya

DOI: 10.36347/sjpms.2021.v08i09.001 | **Received:** 29.09.2021 | **Accepted:** 06.11.2021 | **Published:** 11.11.2021

***Corresponding author:** Nasir Elmesmari

Abstract **Original Research Article**

Markov chain Monte Carlo (MCMC) is the most widely used method for estimating joint posterior distributions in Bayesian analysis. The Markov chain Monte Carlo technique has been used in order to estimate the model parameters based on the different prior distributions. MCMC simulations were carried out in order to evaluate the linear mixed model using different parameters of the prior distribution. In this paper, we established the linear mixed model with different types of variables. The proposed parameters of the prior distribution are different from the traditional parameters of the prior distribution. We assumed special parameters for the prior distribution based on some background or information about the data science. This work aims to estimate the parameters using a point estimator or find a confidence interval (credible interval) for the unknown parameters. Also, a specific hypothesis about these parameters can be tested using a random sample from the posterior distribution. The performance of each prior is measured based on the effective sample size (ESS) for the estimated model. The results showed that the estimated linear mixed model with proposed parameters of the prior distribution performed very well in comparison with the standard or traditional prior (inverse-Wishart prior for random effect component). Based on the scale reduction factors, the estimated model with proposed parameters performed better in comparison with scale reduction factors for the traditional model.

Keywords: Markov chain Monte Carlo (MCMC), traditional prior, standard prior, posterior distribution, effective sample size, linear mixed model.

1. INTRODUCTION

In Bayesian data analysis, it's possible to compute posterior distribution when the problems are simple. But as soon as we start to get into some complex real data, the analysis gets very complicated (Andrieu *et al.*, 2003). Many different methods were developed to analyze the complicated data based on the Bayesian technique. Markov chain Monte Carlo (MCMC) algorithms were implemented to estimate the posterior distribution. In the last two decades, the Markov chain Monte Carlo has played a significant role in statistics, econometrics, physics, and computing science (Levy, 2009). The MCMC is a general approach that we can use for providing a solution for the complex problem within some reasonable time. Often, the Markov chain Monte Carlo technique is used to solve integration and optimization problems in high-dimensional spaces (Jackman, 1999). The idea of MCMC is to iteratively produce parameter values that are representative samples from the joint posterior. After, the samples have been selected the conditional distribution is used to summarize the posterior distribution of the parameters (Smith, 2007).

In this work, we applied Markov chain Monte Carlo (MCMC) algorithms to estimate the linear mixed model parameters. We used MCMC to fit a linear mixed model with a different prior distribution of the parameters (Alkhamisi & Shukur, 2005). Moreover, the algorithm was used to see whether the different parameters of the prior distribution have an impact on the estimated effective sample size (ESS). In the linear mixed model, we have two types of effects (fixed and random effects) (Harrison *et al.*, 2018). The experiments were implemented with inverse-Wishart priors, by assuming a normal posterior distribution with very large variance for the fixed effects and a flat improper (weakly informative) prior. Since the inverse-Wishart contains two variances components V and nu (Walker, 2016). Usually, V and nu are unknown parameters, must be estimated from the data. The ultimate goal of this work is to have an optimal variance component V while we fix the component nu using the Markov chain Monte Carlo (MCMC) technique.

2. Method: Markov Chain Monte Carlo (MCMC)

Monte Carlo simulation is considered as one of the most potential statistic tools in many fields such as engineering and science. Monte Carlo refers to the methods used to generate random numbers. In the Monte Carlo method, we generate sets of random numbers from different distributions (Fong *et al*., 2010). For instance, we generate a random sample from a normal distribution with mean μ and variance σ^2, called proposal distribution. The proposal distribution will be used to accept or reject samples. Using the random samples generated from the proposal distribution, we can estimate the distribution of θ (Robert & Casella, 2011). Estimating the parameter θ using Monte Carlo simulation is a straightforward approach. We can simulate independent and identically distributed (i.i.d.) observations from a proposal distribution and use them to estimate the sample mean. Although with non-Gaussian response variables the likelihood cannot be obtained in closed form, the generalized linear mixed models provide a flexible framework for modeling a range of data (De la Cruz *et al*., 2016). Markov chain Monte Carlo methods solve this problem by sampling from a series of simpler conditional distributions that can be evaluated (Hadfield, 2010). For the linear mixed model, we have:

$Y = X\beta + Zd + \varepsilon,$

Where X and Z are design matrices relating to fixed and random effects, respectively. These matrices have associated parameter vectors β and d, while ε is a residual vector (Archila, 2016). The distribution of vectors d and ε are assumed to be multivariate normal distribution as d~N(0, G) and ε~N(0, R). The matrices G and R are (co)variances of the random effects and residuals, respectively (Fan et al., 2008). The structure form of matrix G is:

$$G = V_1 \otimes A_1 \oplus V_2 \otimes A_2 \oplus \dots$$

Typically, the (co)variance matrices V are low-dimensional, and the structured matrices are (A) high dimensional (Wang et al., 1994). The terms are separated by a direct sum (⊕) as component terms and each component term is formed through the Kronecker product (⊗) (Mathew et al., 2012). So, if we have two-component terms, we can write matrix G as:

$$\boldsymbol{G} = \begin{bmatrix} (\boldsymbol{V}_1 \otimes \boldsymbol{A}_1) & \boldsymbol{0} \\ \boldsymbol{0} & (\boldsymbol{V}_2 \otimes \boldsymbol{A}_2) \end{bmatrix}$$

By the same manner, we can get R: $= (\boldsymbol{V}_{e1} \otimes \boldsymbol{I}_1) \oplus (\boldsymbol{V}_{e2} \otimes \boldsymbol{I}_2) \oplus \dots$, where Ve_1, Ve_2 are residual variances and I_1, I_2 are identity matrices.

The objective of this work is to show how different prior distribution parameters affect on the posterior distribution. We used the Markov chain Monte Carlo approach to fit a linear mixed model with different parameters for the prior distribution. We used data named (l51). This data is available in an *R package gap*. The data contains 51 individuals in a pedigree with the quantitative trait (qt) as a dependent variable.

3. RESULTS & DISCUSSION

In Bayesian statistics, the goal is to find the posterior distribution by combining the information from the data and the prior distribution. To assess the performance of our estimation model, we conducted a simulation study for the linear mixed model with different parameters of the prior distribution. The MCMC experiment was replicated several thousand times to evaluate the estimated model. We performed 250000 iterations of the Markov chain Monte Carlo technic. After the first 40000 iterations (burn-in) samples were selected, we obtain the first independent samples at the spacing 40 iterations. Indeed, we ended up with (5625) samples were used to estimate the posterior distribution of the model parameters.

Table 1: Summary statistic for the MCMC of the linear mixed model fitted by deferent prior parameters

	Effect	post.mean	l-95% CI	u-95% CI	eff.samp	pMCMC
Model1	(Intercept)	2.17	1.06	3.42	437.67	0.00018***
	Independent variable	-0.56	-1.26	0.16	1449.42	0.12480
Model2	(Intercept)	1.98	0.88	3.13	3434.95	0.00107**
	Independent variable	-0.63	-1.39	0.06	3395.26	0.08284

The results illustrate fitted linear mixed model with different prior parameters. We can see that the estimated posterior means by both models are identical. Moreover, the model 2 has better effective sample size (ESS) since the prior distribution parameters coming from the data information. PMCMC value for model2 shows that the model significant at 0.1 level while for the model1 do not. We can get the posterior distribution using the posterior mean and the variance.

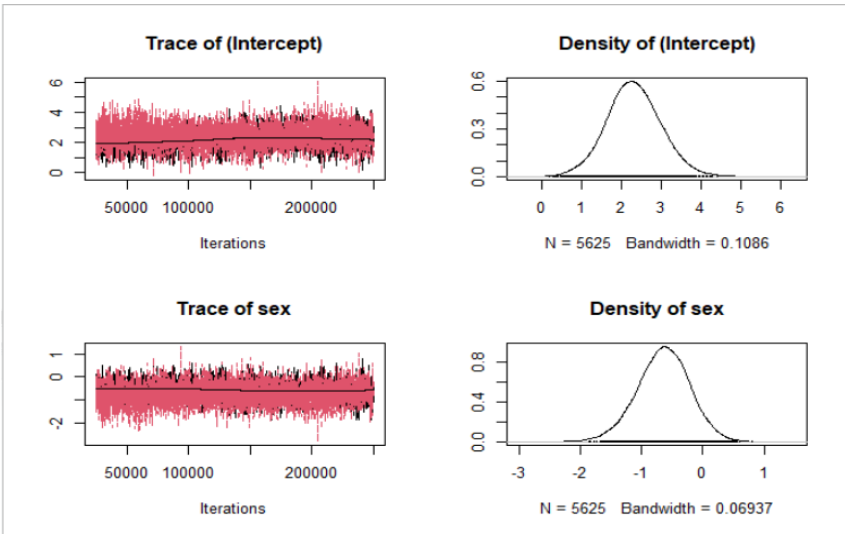

Figure 1: The MCMC summary for the intercept and sex coefficients from the linear mixed model in which the id assumed as random effects

The above figures show the MCMC output for different models with different prior parameters. Model 1 with traditional prior which is assumed to be flat normal with a short tail. Model 2 with prior distribution in which the parameters have been determined based on the data background. As we can see, the outputs are very close which means that the prior with parameters coming from the data information provides an estimate for the model equivalent to the estimate of the model with traditional prior.

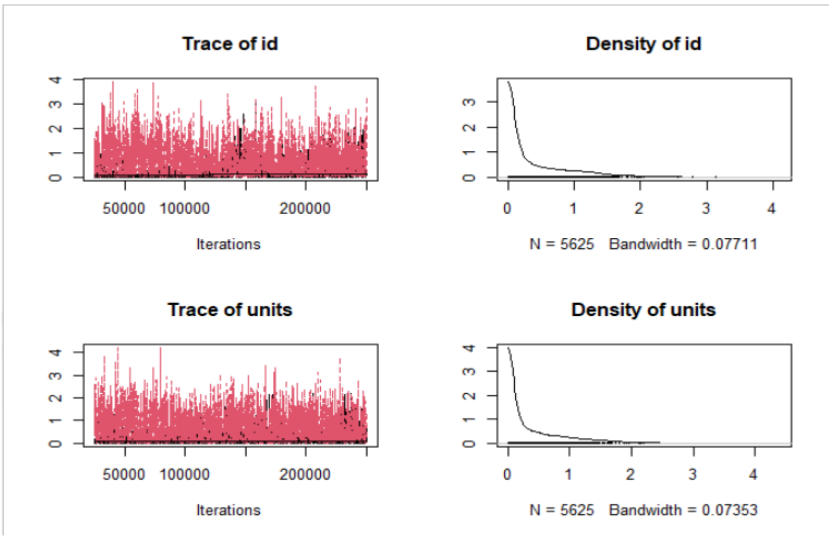

Figure 1: MCMC plot of the variance component associated with (id) and the residual variance component (units)

There are two components of the random effects in the linear mixed model, one is associated with the "id" factor, and one is associated with the "units". Figure 1 shows the trace and density plot for the two random effects. From the Figure, we can see that the posterior distributions are virtually identical assuming the correct but with traditional prior the autocorrelation exists. Autocorrelation between samples computed in Table *2*.

Table 2: Autocorrelation between the samples for different models based on different prior parameters

		Samples of Model 1				**Samples of Model 2**			
		$\hat{B}_0$	$\hat{B}_1$	$\hat{V}_e$	$\hat{V}_d$	$\hat{B}_0$	$\hat{B}_1$	$\hat{V}_e$	$\hat{V}_d$
Lags	0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	40	0.474	0.435	0.781	0.676	0.206	0.206	0.516	0.292
	200	0.146	0.098	0.475	0.283	0.009	0.001	0.113	0.001
	400	0.113	0.053	0.267	0.154	0.009	0.011	0.026	0.009
	2000	0.089	0.047	0.160	0.121	-0.008	-0.015	-0.010	-0.003

Even though the posterior distribution of the parameters is unknown, we can set up a Markov chain that sampled directly from the posterior distribution. The true posterior distribution can be accurate only when the chain has converged. As we can see, based on the model output there is a strong dependence between each successive sample. From the correlation results between the successive samples, we can see that model 1 has a higher correlation in comparison with the correlation in model 2 for both components. When autocorrelation is high the chain needs to be run for longer iterations to reach stability. To get asymmetric trend using the prior of the first model (no dependence), we need 2000 samples while 200 samples using the prior of the second model. Clearly that a strong dependency in the first model makes the information about the posterior distribution meaningless.

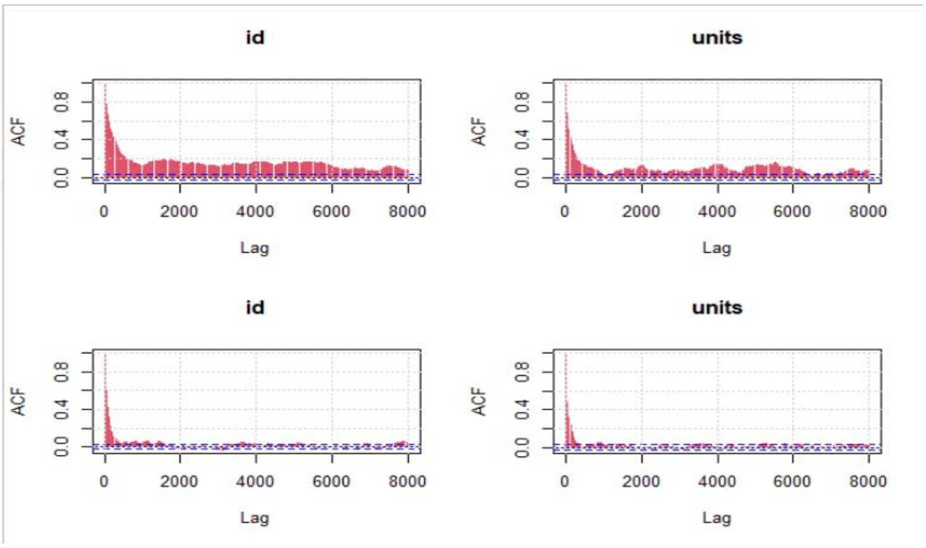

Figure 3: Autocorrelation output of the random effects for different models using different prior distribution parameters (Top: ACF of the residual component using traditional prior and Bottom: ACF of the residual component using proposed prior)

From these plots, we can see that the ACF of the random effects is reduced for the proposed model in comparison with the ACF of the random effects of the original model. The comparison showed that the chain of the model with traditional prior has high dependency than the chain of the model with proposed prior. More results of the autocorrelations based on different priors are presented in Table 2.

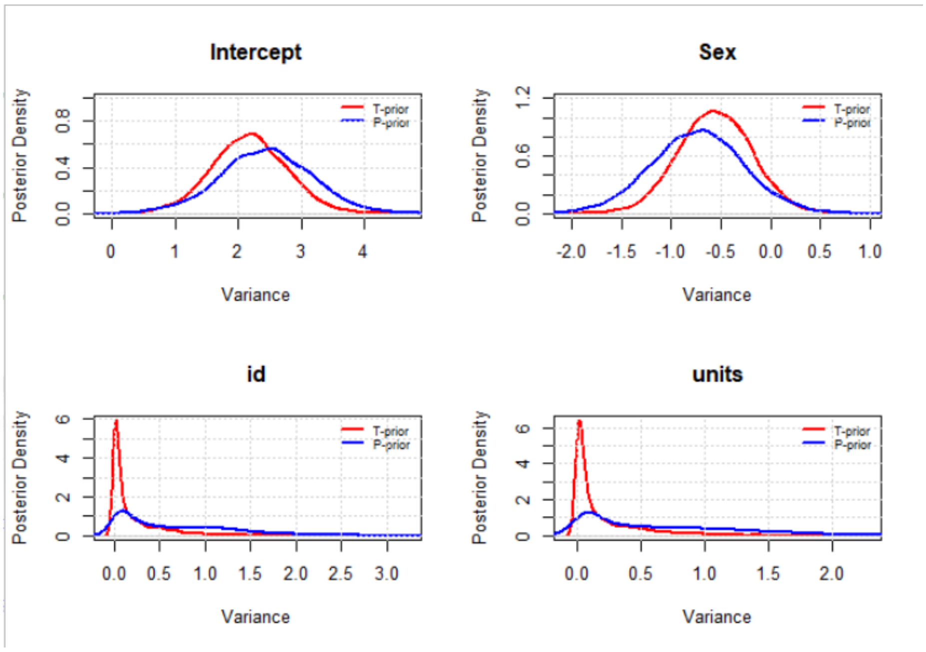

Figure 4: Estimated density of the variance components using the proposed prior

From the four plots, we can say that the posterior distribution of the components (mean and variance) is sensitive to the prior distribution parameters. Clearly, different values for the prior distribution parameters show many differences between the estimated posterior distributions. The estimated densities for the posterior distribution of the model components show that the prior distribution of the parameters influences on the posterior mode as well.

4. CONCLUSION

We studied the impact of the prior distribution parameters on the fix and random effects of the linear mixed model. We consider the traditional Bayesian approaches to model development and parameters estimation. Markov chain Monte Carlo (MCMC) algorithms have been used to fit the linear mixed models. Also, two different models with different parameters were established in order to evaluate the model's improvement. The first model was implemented based on the traditional prior distribution parameters which is assumed to be normal with large variance for fixed effect and inverse-Wishart for a random effect. The second model (propose model) is implemented based on the prior distribution that its parameters are estimated from the data science. To compare the estimated posterior distribution of each model component, we used effective sample size (ESS) and chain dependency (Autocorrelation) to evaluate the model's improvement.

The study showed that for different prior distribution parameters, there are different posterior distributions for the model components. In the traditional prior case, the posterior mode 2.17 and -0.43 for the intercept and sex component respectively. For the proposed prior the posterior mode 2.53 and -0.75 for the intercept and sex component respectively. Moreover, there is a positive impact of different prior distribution parameters on the random components. In the proposed prior the dependency of the chain was decreased compared with the dependency of the chain based on the traditional prior.

REFERENCES

- Alkhamisi, M., & Shukur, G. (2005). Bayesian analysis of a linear mixed model with AR (p) errors via MCMC. *Journal of Applied Statistics*, *32*(7), 741–755.
- Andrieu, C., De Freitas, N., Doucet, A., & Jordan, M. I. (2003). An introduction to MCMC for machine learning. *Machine Learning*, *50*(1), 5–43.
- Archila, F. H. A. (2016). *Markov chain Monte Carlo for linear mixed models*.
- De la Cruz, R., Meza, C., Arribas-Gil, A., & Carroll, R. J. (2016). Bayesian regression analysis of data with random effects covariates from nonlinear longitudinal measurements. *Journal of Multivariate Analysis*, *143*, 94–106.
- Fan, Y., Leslie, D. S., & Wand, M. P. (2008). Generalised linear mixed model analysis via sequential Monte Carlo sampling. *Electronic Journal of Statistics*, *2*, 916–938.
- Fong, Y., Rue, H., & Wakefield, J. (2010). Bayesian inference for generalized linear mixed models. *Biostatistics*, *11*(3), 397–412.
- Hadfield, J. D. (2010). MCMC methods for multi-response generalized linear mixed models: The MCMCglmm R package. *Journal of Statistical Software*, *33*, 1–22.
- Harrison, X. A., Donaldson, L., Correa-Cano, M. E., Evans, J., Fisher, D. N., Goodwin, C. E., Robinson, B. S., Hodgson, D. J., & Inger, R. (2018). A brief introduction to mixed effects modelling and multi-model inference in ecology. *PeerJ*, *6*, e4794.
- Jackman, S. (1999). Bayesian modeling in the social sciences: An introduction to Markov-chain Monte Carlo.
- Levy, R. (2009). The rise of Markov chain Monte Carlo estimation for psychometric modeling. *Journal of Probability and Statistics*, *2009*.
- Mathew, B., Bauer, A., Koistinen, P., Reetz, T., Léon, J., & Sillanpää, M. (2012). Bayesian adaptive Markov chain Monte Carlo estimation of genetic parameters. *Heredity*, *109*(4), 235–245.
- Robert, C., & Casella, G. (2011). A short history of Markov chain Monte Carlo: Subjective recollections from incomplete data. *Statistical Science*, 102–115.
- Smith, B. J. (2007). boa: An R package for MCMC output convergence assessment and posterior inference. *Journal of Statistical Software*, *21*(1), 1–37.
- Walker, J. A. (2016). Monte Carlo simulation of OLS and linear mixed model inference of phenotypic effects on gene expression. *PeerJ*, *4*, e2575.
- Wang, C., Rutledge, J., & Gianola, D. (1994). Bayesian analysis of mixed linear models via Gibbs sampling with an application to litter size in Iberian pigs. *Genetics Selection Evolution*, *26*(2), 91–115.