ANÁLISIS COMPARATIVO DE LAS TÉCNICAS DE SIMULACIÓN POSTERIOR EN LA ESTIMACIÓN DEL MODELO DE REGRESIÓN BAYESIANA

ADEDAYO A. ADEPOJU
Department of Statistics, University of Ibadan, Nigeria.
e-mail: pojuday@yahoo.com
OLUWADARE O. OJO
Department of Statistics, The Federal University of Technology Akure, Nigeria.
e-mail: ojooo@futa.edu.ng

RESUMEN

En los últimos años, el desarrollo de varias técnicas de simulación posterior ha impulsado el campo de la econometría bayesiana, especialmente en trabajos aplicados. La distribución previa juega un papel dominante en el análisis bayesiano; los anteriores están destinados a reflejar la información que el investigador tiene antes de ver los datos. Esta investigación examinó la sensibilidad de los métodos de muestreo de Gibbs (GS) e Integración de Monte Carlo (MCI) a tres niveles diferentes de correlación en covarianza previa para conocer los efectos de la correlación variable en los métodos de simulación posterior al estimar los parámetros en un modelo de regresión lineal. Los tres niveles diferentes de correlación son: Correlación negativa (NC), correlación positiva (PC) y correlación cero (ZC). Los resultados mostraron que el MCI superó al GS en la mayoría de los casos y la precisión del MCI no depende del nivel de correlación, ya sea positivo o negativo, mientras que el GS se desempeñó mejor cuando se usó el nivel de correlación positivo como información en la covarianza previa que el uso de un nivel negativo de correlación. El uso de MCI en la inferencia bayesiana podría ser de importancia práctica para los profesionales.

Palabras Clave: Simulación posterior, regresión, covarianza previa, bayesiana.

A comparative analysis of posterior simulation techniques in the estimation of bayesian regression model

ABSTRACT

In recent years, the development of several posterior simulation techniques has boosted the field of Bayesian econometrics especially in applied works. Prior distribution plays a dominant role in Bayesian analysis; priors are meant to reflect information the researcher has before seeing the data. This research examined the sensitivity of the Gibbs sampler (GS) and Monte Carlo Integration (MCI) methods to three different levels of correlation in prior covariance to know the effects of varying correlation on posterior simulation methods in estimating the parameters in a linear regression model. The three different levels of correlation are; Negative Correlation (NC), Positive Correlation (PC) and Zero correlation (ZC). The results showed that MCI outperformed the GS in most cases and the accuracy of MCI does not depend on the level of correlation either positive or negative while GS performed better when positive level of correlation was used as information in the prior covariance than using negative level of correlation. The use of MCI in Bayesian inference might be of practical importance to practitioners.

Keywords: Posterior simulation, Regression, Prior Covariance, Bayesian.

Clasificación JEL: C01, C12

Artículo recibido el 6 de octubre de 2018 y aceptado el 5 de enero de 2019
Artículo disponible en versión electrónica en la página www.revista-eea.net

ISSN 1697-5731 (*online*) – ISSN 1133-3197 (*print*)

1. INTRODUCCIÓN

In many areas of research and applications, regression model is a tool that has been applied to study wide range of real phenomenon and is also the workhouse of econometrics.

Recently, there has been a long time arguments on the kind of posterior simulation techniques that will yields more efficient posterior simulation algorithms for classes of models.

Posterior simulation techniques are techniques used in computation of range of integrals that are necessary in Bayesian Posterior Analysis (Chib (2001), Geweke and Keane (2001)). There are various posterior simulation techniques meant for evaluating such integrals they are; Monte Carlo Integration (MCI), Laplace's method for asymptotic normality of posterior (Gelman et al (2004), Carlin and Louis (2008)); Gibbs sampler by Gelfand and Smith (1990), Metropolis –Hastings algorithms by Metropolis et al (1953), Hastings (1970) that can handle non-linear regression appropriately, Slice sampler that uses auxiliary variable to generate a random sample from a given distribution (Neal (2003) and also the importance sampling method (Robert and Casella (2004) that involves taking draws from an important function etc. The two common posterior simulation methods in linear regression analysis are Monte Carlo Integration (MCI) and Gibbs sampler (GS).

MCI is a method that is known to straightforwardly evaluate a complex integrals and it involves drawing from a joint posterior distribution while GS on the other hand draws sequentially from a full conditional posterior distributions and that kind of draw is treated as if they came from a joint posterior distribution. Dijk and Kloek (1980) employed a MCI for a nine dimensional parameter space of Klein's model while Geweke (1993) demonstrated how GS algorithms can provide an alternative to inference subject to linear inequality constraints. Fosdisck and Raftery (2012) considered the problem of estimating the correlation between bivariate normal regression model when the means and variances were assumed to be known with emphasis on a small sample case. Koop et al (2007) estimated regression model parameters with a known bivariate normal posterior, MCI performed well than the GS.

This work is a follow-up on Koop et al (2007) where a bivariate normal model was used to study the effect of degree of correlation between two parameters on GS and MCI; a high positive correlation between the parameters was considered in their work.

This study investigates the sensitivity of prior covariance on GS and MCI methods given three different levels of correlation coefficients in a linear regression model. The performance of the methods was judged using the mean parameter estimates in terms of closeness of their estimated parameter to the true values and Mean Squared Error (MSE).

The three different levels of correlation are; Negative Correlation (NC), Positive Correlation (PC) and Zero correlation (ZC).

The remainder of this paper is structured as follows:- section 2 considers the specification of model and description of methods of posterior simulation; GS and MCI. Simulation studies are conducted in section 3; section 4 presents the analysis, results and discussion. Section 5 concludes.

2. MATERIALS AND METHODS

Given the following model:

$$y = x\theta + \varepsilon \qquad (1)$$

Where y and x are the observed data on the $n \, x \, 1$ vector of dependent and $n \, x \, m$ matrix of explanatory variables of the regression respectively. θ is the $m \, x \, 1$ vector of parameters to be estimated and ε is an error term which is normally distributed with mean zero and constant $\sigma^2$and x values are independent of the error term.

2.1 Bayesian Posterior Simulation techniques

The Bayesian estimation method in Regression model involves three steps:

a) Obtain the likelihood function of the model

b) Specify the prior density function for the model
c) Obtain the Posterior density function.

The relationship between the three steps can be written as:

$$P(\theta|y) \propto P(\theta)\, P(y|\theta) \quad (2)$$

$P(\theta|y)$ is referred to as Posterior density function, $P(\theta)$ is the prior density function, $P(y|\theta)$ is the likelihood function and $\propto$ means proportionality

The likelihood is written as follows;

$$P(y|\theta, h) = \frac{h^{N/2}}{(2\pi)^{N/2}} \{\exp[-\frac{h}{2}(y - x\theta)'(y - x\theta)]\} \quad (3)$$

where $h = \frac{1}{\sigma^2}$

Priors play a defining role in Bayesian inference which can take any form and are also meant to reflect any information the researcher has before seeing the data. However, it is common to choose particular classes of priors that are easy to interpret or which would make computation easier (Koop (2003), an independent Normal-gamma and Natural conjugate priors typically belong to such class.

Using an independent Normal-gamma:

While the independent normal gamma prior will be:

$$f(\theta) = \frac{1}{(2\pi)^{k/2}} |V^0|^{-1/2} \exp[-\frac{1}{2}(\theta - \theta^0)' V^{0^{-1}}(\theta - \theta^0)] \quad (4)$$

$$f(h) = C_G^{-1} h^{\frac{v^0-2}{2}} \exp(-\frac{hv^0}{2(S^0)^{-2}}) \quad (5)$$

θ^0denotes the prior mean for parameter θ, $2(S^0)^{-2}$ is the prior mean of gamma distribution for model precision h, V^0 is the un-scaled variance matrix for parameter θ and v^0 is the prior sample size.

Note that the symbols 0 over the parameters denote parameters of a prior density and * over parameters denote parameters of a posterior density.

Where C_G is the integrating constant for the gamma distribution.

Then the joint posterior distribution when (3), (4) and (5) are combined will be:

$$P(\theta, h|y) \propto \{\exp[-\frac{1}{2}\{h\left(y - x\theta\right)'(y - x\theta) + \left(\theta - \theta^0\right)' V^{0^{-1}}(\theta - \theta^0)\}]\} \quad (6)$$

$$h^{\frac{N+v^0-2}{2}} \exp[-\frac{hv^0}{2(S^0)^{-2}}]$$

Formally, Equation (6) has posterior form as given by Joyce (2009)

Thus, the conditional posteriors are:

$$P(\theta|y, h) \propto \exp[-\frac{1}{2}(\theta - \theta^*)' V^{*^{-1}}(\theta - \theta^*)] \quad (7)$$

Where $V^* = (V^{0^{-1}} + h\, x' x)^{-1}$

while the conditional for h is of the form:

$$h|y, \theta \sim G(S^{*-2}, v^*) \quad (8)$$

Where

$$v^* = N + v^0$$

And

$$S^{*2} = \frac{(y - x\theta)'(y - x\theta) + (vS^2)^*}{v^*}$$

The conditional Posterior obtained in (7) will be used for GS.

Using Natural Conjugate prior (Normal-Gamma):

We assume elicit a prior for θ conditional on h which is of the form:

$$\theta| h \sim N(\theta^0, h^{-1}V^0)$$

And a prior for h of the form:

$$h \sim G((S^0)^{-2}, v^0) \tag{9}$$

Then the prior is Normal-Gamma distribution which will be:

$$\theta, h \sim NG (\theta^0, V^0 , (S^0)^{-2}, v^0) \tag{10}$$

Hence, combining the likelihood in (3) by the prior in (10) and collecting the like terms will yields a posterior of the form:

$$\theta, h|y \sim NG (\theta^*, V^* , S^{-2\,*} , v^*) \tag{11}$$

Where $V^* = ({V^0}^{-1} + x'x)^{-1}$

$\theta^* = V^* ({V^0}^{-1}\theta^0 + x'x\,\hat{\theta})$

And $S^{-2\,*}$ defined implicitly through

$$(vS^2)^* = (vS^2)^0 + vS^2 + (\hat{\theta} - \theta^0)' [V^0 + (x'x)^{-1}]^{-1} (\hat{\theta} - \theta^0) \tag{12}$$

Where $\hat{\theta}$ is an estimator for θ

Since our interest is on θ, we integrate out h, and then we have:

$$\theta|y \sim t (\theta^* , S^{2\,*} V^*, v^*) \tag{13}$$

t - follows a t-distribution

The marginal Posterior obtained in (13) will be used for MCI.

3. SIMULATION STUDY

The data experiment is set up using the Data Generating Process (DGP) below:

$$y = 0.4\, x_1 + 2\, x_2 + 4.5\, x_3 + \varepsilon \tag{14}$$

The error term, $\varepsilon \sim$ N (0, 1) and the explanatory variables, $x_i \sim$ U (0, 1), i = 1, 2, 3 generate the dependent variable, y. The sample sizes are: 10, 30, 100, and 1000 with 10000 replication. The Posteriors obtained in equations (7) and (13) for both GS and MCI respectively will be used to perform Posterior analysis while the values for levels of correlation for prior covariance are specified as:

Negative Correlation (NC): -0.01, -0.7 and -0.9

Positive Correlation (PC): 0.01, 0.7 and 0.9

Zero Correlation (ZC): 0

Prior parameter values are given as: $\theta_0^0 = 0$, $\theta_1^0 = 0.4$, $\theta_2^0 = 1.5$, $\theta_3^0 = 3.8$, $S^{0-2} = 1$and prior degree of freedom, v^0 is 4.

4. ANALYSIS, RESULTS AND DISCUSSION

The results for three different levels of correlation PC NC and ZC for both the GS and MCI using Mean of parameter estimates and MSE are presented in this section with the true parameter values in parenthesis.

Tables 1, 2, 4 and 5 present the Mean estimates of the GS and MCI for PC, NC and ZC for samples 10, 30, 100 and 1000 while Tables 3 and 6 give the MSE estimates of the GS and MCI for PC, NC and ZC for samples 10, 30, 100 and 1000.

Table 1

Mean Criteria for sensitivity of GS and MCI when the level of correlation is Positive Correlation (PC) for samples 10 and 30

		Sample size =10			Sample size =30		
Parameter	**Positive Correlation (PC)**	**0.01**	**0.7**	**0.9**	**0.01**	**0.7**	**0.9**
$\boldsymbol{\theta_0}$ (0)	**GIBBS**	0.5256	0.3326	0.2621	0.3926	0.3027	0.2745
	MCI	0.178	0.2619	0.2672	0.1871	0.0257	0.1859
$\boldsymbol{\theta_1}$ (0.4)	**GIBBS**	-0.3355	0.3314	0.529	-0.0667	0.2684	0.4886
	MCI	0.5349	0.5727	0.5815	0.4257	0.6849	0.4286
$\boldsymbol{\theta_2}$ (2.0)	**GIBBS**	0.7972	1.4045	1.6167	1.7312	1.7735	1.7726
	MCI	0.9337	1.3500	1.5887	1.716	1.947	1.7235
$\boldsymbol{\theta_3}$(4.5)	**GIBBS**	5.4893	4.489	4.1751	4.5293	4.347	4.19
	MCI	5.2546	4.5845	4.2488	4.5344	4.4219	4.5301

Table 2

Mean Criteria for sensitivity of GS and MCI when the level of correlation is Positive Correlation (PC) for samples 100 and 1000

		Sample size =100			Sample size =1000		
Parameter	**Positive Correlation (PC)**	**0.01**	**0.7**	**0.9**	**0.01**	**0.7**	**0.9**
$\boldsymbol{\theta_0}$ (0)	**GIBBS**	0.1078	0.1407	0.1869	0.1209	0.0865	0.1041
	MCI	0.1147	0.1469	0.1935	0.0798	0.086	0.1043
$\boldsymbol{\theta_1}$ (0.4)	**GIBBS**	0.9297	0.8763	0.7896	0.1546	0.3195	0.3401
	MCI	0.9243	0.8649	0.7759	0.311	0.3196	0.3402
$\boldsymbol{\theta_2}$ (2.0)	**GIBBS**	1.4132	1.4668	1.5536	1.9266	1.9875	1.9661
	MCI	1.4162	1.4789	1.5691	1.9963	1.9884	1.9659
$\boldsymbol{\theta_3}$(4.5)	**GIBBS**	4.376	4.3126	4.2156	4.5344	4.4438	4.4107
	MCI	4.3642	4.2984	4.1992	4.4555	4.4434	4.41

In Table 1, the MCI performs better since its estimates are close to the true parameter values in most cases for all the parameters considered for sample sizes 10 and 30. As the correlation increases the estimates of the GS move closer to the true parameter values while the estimates of MCI move farther away from the true parameter values. Also, in Table 2, MCI provides better estimates using Mean of parameter as criterion to judge the performance of the posterior simulation methods.

Table 3

MSE Criteria for sensitivity of GS and MCI when the level of correlation is Positive Correlation (PC) for samples 10, 30, 100 and 1000

		Sample size			
Levels of correlation		**10**	**30**	**100**	**1000**
0.01	**GIBBS**	0.8107	0.1113	0.163	0.0204
	MCI	0.4391	0.0294	0.1618	0.0041
0.7	**GIBBS**	0.1415	0.0459	0.1415	0.0043
	MCI	0.1375	0.0227	0.1375	0.0043
0.9	**GIBBS**	0.1167	0.1127	0.1167	0.0059
	MCI	0.1137	0.0282	0.1137	0.0059

Results obtained from Table 3 shows that MCI has the least MSE for all the sample sizes in most cases for Positive Correlation (PC) which also supported the results obtained using the Mean of estimate as criteria.

Table 4

Mean Criteria for sensitivity of GS and MCI when the level of correlation is Negative Correlation (NC) and Zero Correlation (ZC) for samples 10 and 30

		Sample size = 10				Sample size = 30			
Parameter	**Negative Correlation (NC) & Zero Correlation (ZC)**	**-0.01**	**-0.7**	**-0.9**	**0**	**-0.01**	**-0.7**	**-0.9**	**0**
$\boldsymbol{\theta_0}$ (0)	**GIBBS**	0.5314	0.5949	0.6015	0.6184	0.3945	0.3451	0.3695	0.3969
	MCI	0.1753	-0.0106	-0.0289	0.1768	-0.1787	-0.2835	-0.2908	-0.1769
$\boldsymbol{\theta_1}$ (0.4)	**GIBBS**	-1.1293	-1.3385	-1.4115	-1.0738	-0.0733	-0.1014	-0.1412	-0.0755
	MCI	0.533	0.5534	0.5367	0.53439	0.7181	0.7560	0.7545	0.7183
$\boldsymbol{\theta_2}$ (2.0)	**GIBBS**	0.7835	0.3906	0.3941	0.3963	1.7318	1.7699	1.7534	1.732
	MCI	0.927	0.9062	0.8823	0.9299	2.0765	2.1588	2.1626	2.0756
$\boldsymbol{\theta_3}$(4.5)	**GIBBS**	5.5102	6.9923	7.0415	6.6128	4.5282	4.6723	4.6696	4.5293
	MCI	5.2637	5.9214	5.9604	5.2589	4.6086	4.7624	4.7589	4.6081

Table 5

Mean Criteria for sensitivity of GS and MCI when the level of correlation is Negative Correlation (NC) and Zero Correlation (ZC) for samples 100 and 1000

		Sample size = 100				Sample size = 1000			
Parameter	**Negative Correlation (NC) & Zero Correlation (ZC)**	**-0.01**	**-0.7**	**-0.9**	**0**	**-0.01**	**-0.7**	**-0.9**	**0**
$\boldsymbol{\theta_0}$ (0)	**GIBBS**	0.1089	0.0568	0.0568	0.1082	0.0803	0.0747	0.0758	0.0805
	MCI	0.1155	0.0561	0.0618	-0.0804	0.0798	0.0740	0.0745	0.0798
$\boldsymbol{\theta_1}$ (0.4)	**GIBBS**	0.9314	0.975	0.975	0.9296	0.3101	0.3122	0.3105	0.3105
	MCI	0.9241	0.9740	0.9711	0.9379	0.3109	0.3127	0.3119	0.3109
$\boldsymbol{\theta_2}$ (2.0)	**GIBBS**	1.4115	1.4289	1.4289	1.4126	1.9958	2.0006	2.0007	1.9956
	MCI	1.4148	1.4347	1.4278	1.7122	1.9963	2.0014	2.001	1.9963
$\boldsymbol{\theta_3}$(4.5)	**GIBBS**	4.3748	4.4256	4.4256	4.3759	4.4559	4.462	4.4617	4.4561
	MCI	4.364	4.4222	4.4188	4.4168	4.4555	4.4614	4.4612	4.4555

From tables 4 and 5 which present the Mean Criteria for performance of the posterior simulation techniques revealed that MCI has the best performance for all the parameters considered for all sample cases. As the sample size increases, the estimates of both the GS and MCI become better and tend toward the true parameters for each values of NC across the parameters. Also for ZC, MCI also performed better than MCI.

Table 6

MSE Criteria for sensitivity of GS and MCI when the level of correlation is Negative Correlation (NC) and Zero Correlation (ZC) for samples 10, 30, 100 and 1000

Levels of correlation		Sample sizes 10	30	100	1000
-0.01	**GIBBS**	1.2804	0.1131	0.1641	0.0041
	MCI	0.4457	0.0377	0.1622	0.0041
-0.7	**GIBBS**	3.0445	0.1133	0.1664	0.0037
	MCI	0.8101	0.0753	0.1646	0.0036
-0.9	**GIBBS**	3.1704	0.1298	0.1664	0.0038
	MCI	0.8504	0.0759	0.1660	0.0037
0	**GIBBS**	2.3976	0.1141	0.1632	0.0041
	MCI	0.4426	0.0375	0.0964	0.0041

From Table 6, the MCI also has a better performance that GS for each values of Negative correlation (NC) across the parameters for all the samples sizes with the least MSE.

5. CONCLUSION

This study investigated the sensitivity of prior covariance by using three different levels of correlation; Negative Correlation (NC), Positive Correlation (PC) and Zero correlation (ZC) on two Posterior simulation techniques namely; GS and MCI to examine the method more suitable for estimation of parameters in a linear regression model. Two kinds of priors known as independent Normal–Gamma and a natural conjugate (Normal Gamma) were used. The data for the experiment were generated using a varying sample sizes while Mean of parameter estimates and MSE were used to judge the performances of the methods; GS and MCI.

Going by the various results obtained, it can be concluded that MCI performed that GS in all cases of the experiment which means that accuracy of MCI does not depend on the any level of correlation either positive or negative, but however GS is much better using a Positive level of correlation as an information in the prior covariance than using negative level of correlation.

The parameter estimates of the both GS and MCI are also consistent with the samples sizes, the mean are not far from the true parameter values for all the levels of correlation considered.

The work therefore concluded that Monte Carlo Integration (MCI) is a good method of Bayesian estimation of parameters for linear regression model especially when the researcher has a choice of selecting a particular Posterior simulator.

REFERENCIAS BIBLIOGRÁFICAS

CARLIN, B.P. AND LOUIS, T.A. (2008): “Bayesian Methods for data Analysis”, third edition. *CRC Press, Boca Ratcon, FL.*

CHIB, S. (2001): “Markov chain Monte Carlo methods: Computation and inference”. in eds. J. J. Heckman and E. Leamer, *Handbook of Econometrics* Volume 5, 3569–3649, Elsevier Amsterdam,.

DIJK, H.K. AND KLOEK, T (1980): “Further Experience in Bayesian Analysis Using Monte Carlo Integration” *Journal of Econometrics* 14,307-328. North-Holland Publishing Company.

GELFAND, A.E AND SMITH, A.F.M (1990): “Sampling Based Approaches to Calculating Marginal Densities”, *Journal of the American Statistical Association* 85,398-409.

GELMAN, A., CARLIN, J.B., STERN, H.S., AND RUBIN, D. B. (2004): “Bayesian data analysis”, second edition. *Chapman & Hall/CRC, London, UK*.

GEWEKE, J. (1993): “Evaluating the accuracy of sampling-based approaches to calculating posterior moments”, In eds. J.M. Bernado, *Bayesian Statistics* 4.

GEWEKE, J. AND M. KEANE (2001): “Computationally intensive methods for integration in econometrics”, in (eds. J. J. Heckman and E.Leamer), Handbook of *Econometrics* Volume 5 Elsevier, Amsterdam, 3463–3568.

FOSDICK, B. AND RAFTERY, A (2012): “Estimating the Correlation in Bivariate Normal Data with Known Variances and small sample sizes”, *The American Statistician* 66(1), 34-41.
HASTINGS, W.K. (1970): “Monte Carlo Sampling-based methods using Markov chains and their applications”. *Biometrika*, 57, 97-109.
JOYCE, D. (2009): “A short introduction to Bayesian Statistics”.
KOOP, G. (2003): “Bayesian Econometrics”. *John Wiley & Sons Ltd., UK.*
KOOP, G., POIRIER, D. AND TOBIAS, J (2007): “Bayesian Econometric Methods”, *Cambridge University Press*
METROPOLIS, N., ROSENBLUTH, A.W., ROSENBLUTH, M.N., TELLER, A.H., AND TELLER,E. (1953): “ Equations of state calculations by fast computing machines”. *Journal of Chemical Physics*, 21, 1087-91
NEAL, R. M. (2003):. “Slice sampling”, Annals of Statistics 31, 705–767.
ROBERT, C., AND CASELLA, G. (2004): “Monte Carlo Statistical methods”, 2nd edn, *Springer-verlag*, London, UK.