Int. J. Nonlinear Anal. Appl. 13 (2022) 1, 3059-5071
ISSN: 2008-6822 (electronic)
http://dx.doi.org/10.22075/ijnaa.2022.6042

Bayesian parameter estimation in addiction model

Najla A. AL-Khairullah[a,*], Tasnim Hasan Kadhim AlBaldawi[a]

[a] *Department of Mathematics, College of Science, University of Baghdad, Iraq*

(Communicated by Madjid Eshaghi Gordji)

Abstract

In this paper, we investigated the performance of Bayesian Computational methods for estimating the parameters of the multinomial Logistic regression model. We discussed two of the most common Bayesian computational algorithms: the Random walk Metropolis-Hastings (RWM) and Slice algorithms and their application to estimating the parameters of the addiction model as well as comparing the performance of these algorithms using the mean square error (MSE) criterion. The results revealed that the performance of the algorithms is excellent, with a slight superiority to the RWM algorithm.

Keywords: Multinomial Logistic Regression, MCMC, Random Walk Metropolis-Hasting Algorithm, Slice Sampling.

1. Introduction

Bayesian statistical approaches have grown increasingly prominent in modern scientific research since examining any inferences that may be drawn from complicated models requires the use of Bayesian computational methods, the most popular of which being Markov chain Monte Carlo (MCMC) methods.

These methods have become fundamental in Bayesian computation. The evolution of these methods has had a significant impact on Bayesian statistics. This power was clearly exhibited in [9, 6].

One of the most essential strategies for categorical data analysis is the Multinomial Logistic Regression model. The fundamental idea for the Multinomial Logistic Regression was based upon binary logistic regression where the response variable has more than two categories or levels also there is no natural order of the categories.

*Corresponding author

Email addresses: nalkhairalla5@gmail.com (Najla A. AL-Khairullah), tasnim.h@sc.uobaghdad.edu.iq (Tasnim Hasan Kadhim AlBaldawi)

Received: December 2021 *Accepted:* January 2022

Multiple binary indicator variables indicate the multinomial response variable in regression analysis. There will be j-1 binary indicator variables for every j categories of the response variable.

As a result, j-1 binary logistic regression models are produced. Following that, the multinomial logistic regression generates a separate binary logistic regression model for each of indicator variables. Each one compares the effect of the predictors on the probability of success in that category to reference category [1].

There are many studies that indicated multinomial logistic regression. For example, [4] proposed two new data augmentation strategies for sampling the parameters, inside a Bayesian framework of the binary or multinomial logit model from the posterior distribution. Another study [2] showed that the regularization parameter used in sparse multinomial logistic regression with a Laplace prior may be integrated analytically. Also, [7] proved that the suitable data augmentation technique offers faster posterior sampling than alternatives in the research and also that the calculation time of the posterior sample technique increases linearly as the number of categories increases.

In this paper, we discussed two of the most common Bayesian computational algorithms, namely: Random walk Metropolis-Hastings (RWM) and Slice algorithms. We investigated the performance of Bayesian Computational methods for estimating the parameters of the multinomial Logistic regression model, as well as comparing the performance of Random walk Metropolis-Hastings (RWM) and Slice algorithms using the MSE criterion. Section two presents a brief overview of multinomial logistic regression. Section three introduces the Random walk Metropolis-Hastings (RWM) and Slice algorithms in the process of obtaining Bayesian parameter estimation in multinomial logistic regression model. Section Four includes application of algorithms presented in section three on real data concerns the problem of addiction which recently witnessed a notable spread in our society. Finally, section five includes some basic conclusions.

2. Multinomial Logistic Regression

2.1. Logistic regression model

The Logistic regression model has grown popular in data analysis, which includes binary or binomial response and many explanatory variables. The logistic regression model is a special case of the statistical models known as generalized linear models.

For an explanatory variable X and a response variable Y with two measurement levels.

$\pi(x) \equiv P\left(Y=1|X=x\right)$

For logit of this probability, the logistic regression model has a linear form

$\text{logit}\left(\pi\right) = log\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_1 x_1$

Where

$odds = \left(\frac{\pi}{1-\pi}\right) = e^{\beta_0+\beta_1 x_1}$

3. Multinomial Logistic Regression

The multinomial logistic regression model is a general form of the binary model that can be extended to include multiple explanatory variables and both models reliant on logit analysis or logistic regression.

To construct the logits in the multinomial case, we need n independent observations with p-explanatory variables and a qualitative response variable with k categories.

One of categories must always be considered the basic level, and all logits are constructed around that one. Since any category can be used as the base level, we will use category k. Because there is no ordering, it is obvious that any category can be designated k. [3].

Let π_j indicate the multinomial probability of an observation falling into the j^{th} category, to identify the relation between of probability and the p explanatory variables, $X_1, X_2, X_3 \ldots, X_p$ then the model of multinomial logistic regression is

$\text{logit}\left(\frac{\pi_j(x_i)}{\pi_k(x_i)}\right) = \beta_{0i} + \beta_{1j}x_{1i} + \ldots + \beta_{\text{pj}}x_{\text{pi}}$

Where $j = 1, 2, 3, \ldots, k-1$

$i = 1, 2, 3, \ldots, n$

This is reduced to $\text{logit}\left(\pi_j(x_i)\right) = \frac{e^{\beta_{0i}+\beta_{1j}x_{1i}+\ldots+\beta_{\text{pj}}x_{\text{pi}}}}{1+e^{\beta_{0i}+\beta_{1j}x_{1i}+\ldots+\beta_{\text{pj}}x_{\text{pi}}}}$

4. Methodology

This section describes the Bayesian techniques that we suggest for estimating the parameters in multinomial logistic regression models.

4.1. Metropolis-Hastings algorithm

Metropolis-Hastings is one of the most used MCMC techniques. This algorithm's main goal is to generate a candidate value $x_t^{'}$ from that of arbitrary conditional distribution q $\left(.|x_t^{'}\right)$ and then accept or reject it, with probability $\alpha\left(x_t, x_t^{'}\right) = min(\frac{q(x_t|x_t^{'})p(x_t^{'})}{q(x_t^{'}|x_t)p(x_t)}, 1)$ proposal $x_t^{'}$ is accepted.

Algorithm: Metropolis-Hastings

Initialize x_0

For $t = 1$ *to* N *do*

Genarate $x_t^{'} \sim$ q $\left(.|x_t^{'}\right)$

Genarate $u \sim U[0,1]$

Calculate $c = \frac{q(x_t|x_t^{'})p(x_t^{'})}{q(x_t^{'}|x_t)p(x_t)}$

If u $< \min\left(c\left(x_t, x_t^{'}\right), 1\right)$

then $x_t = x_t^{'}$

else

$x_t = x_{t-1}$

end

There are a number of Metropolis-Hastings extensions available. Only symmetric ideas would be evaluated in the Metropolis algorithm $q\left(x_t|x_t^{'}\right) = q(x_t^{'}|x_t)$.

Random Walk Metropolis (RWM) is a special case of Metropolis-Hastings algorithm where $q\left(x_t, x_t^{'}\right) = q\left(x_t^{'} - x_t\right)$. Independent Metropolis–Hastings is also special case, as in RWM, proposal distribution do not depend on previous state of chain.

5. Slice Sampler

Slice sampler introduced by [8]. The general method is to take a sample of $f(x)$ from which the sample is usually difficult to generate. For this purpose, one may define the auxiliary variables u and the conditional $f(u|x)$ to construct the joint distribution $f(u, x) = f(u|x)\,f(x)$.Then uses an MCMC algorithm sample of $(x, u)$and marginalize this with computing over u to obtain samples from $f(x)$. Although selecting auxiliary variables is typically a difficult problem that is dependent on the specific problem and its physical characteristics, uniform distribution $U(0, f(x))$is the usual choice of $f(u|x)$.

Algorithm: Slice Sampler

- Generate $u \sim f(u|x) = U(0, f(x))$

- Generate an interval $(x_L, x_R), x \in (x_L, x_R)$

Generate x $\sim f(u|x)\, f(x) = U(x_L, x_R)$
If f$(x) > u$ *then*
Stop anner loop
else
Take the interval (x_L, x_R) shorter

6. Application and Results

Addiction is one of the biggest problems facing any society. This is clearly evident with the increase in the number of addicts every year. It can be defined as chronic medical condition. It involves intricate connections between brain circuits, heredity, the environment, and an individual's life experience. Addicts utilize substances or participate in obsessive behaviors, which they frequently continue despite negative consequences. There are multiple kinds of addictions. Alcohol and drug addiction are the most common.

The study sample includes 72 persons suffering from drugs and alcohol addiction or both, Samples were collected from Ibn Rushd psychiatric training hospital, and the study variables included independent variables, which are as follows:
X_1 : Age
X_2 : Job
X_3 : Weak family supervision
X_4 : Weakness of the rule of law
X_5 : Absence of a religious deterrent
X_6 : Co-occurring mental conditions
X_7 : Experience using narcotic substances at an early age. And y dependent variable that included four categories as follows:
Alcohol addiction, drug addiction, medical drug addiction (Captagon, Artane, Tramadol and Somadril) and addictive to more than one substance.

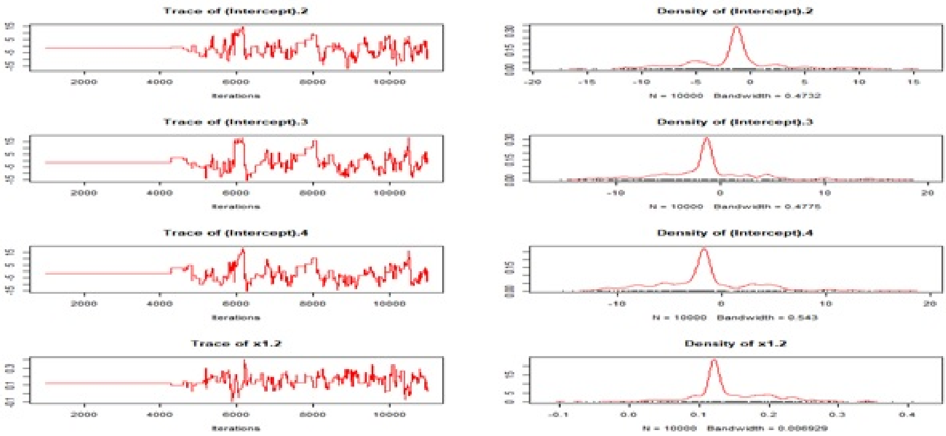
For Bayesian analysis, we use the R statistical software, which contains a suitable framework for MCMC simulation; we use RWM and slice algorithms to simulate the posterior density of multinomial logistic regression model.
Using the RWM and Slice algorithms, we simulate the multinomial logistic regression model's posterior density. First, we calculated the posterior simulation of the RWM method where the baseline or reference category is alcohol addiction. The sample is summarized in the following tables:

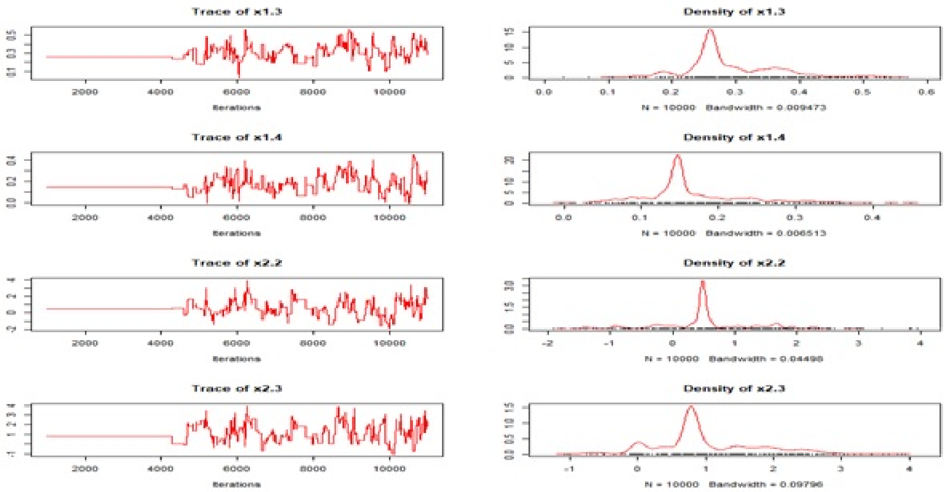
Table 1: Empirical mean and standard deviation for each variable, standard error of the mean by RWM algorithm of the multinomial logistic regression

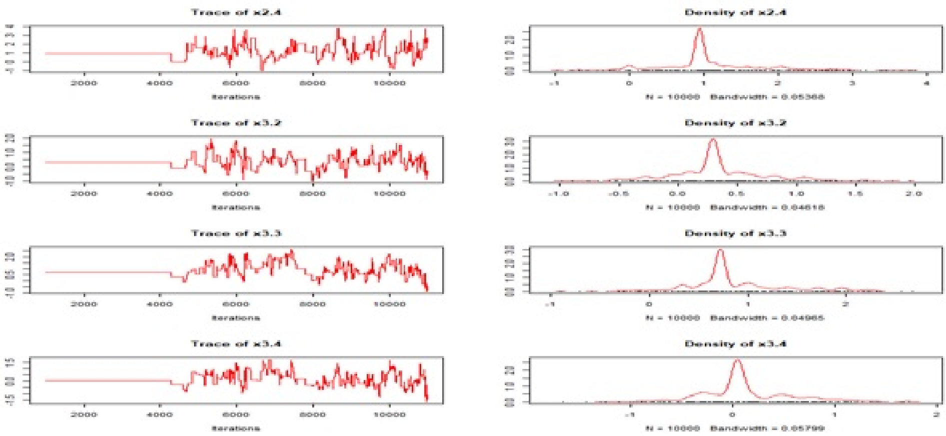
variables	Mean	SD	Naive SE	Time-series SE
(Intercept).2	-1.74290	4.43176	0.0443176	0.607722
(Intercept).3	-1.17814	5.22458	0.0522458	0.815206
(Intercept).4	-1.66597	4.76421	0.0476421	0.738226
x1.2	0.14240	0.05943	0.0005943	0.005922
x1.3	0.29169	0.07534	0.0007534	0.009578
x1.4	0.16591	0.06841	0.0006841	0.009220
x2.2	0.53417	0.80610	0.0080610	0.106048
x2.3	1.01196	0.79896	0.0079896	0.105658
x2.4	1.09812	0.72870	0.0072870	0.091637
x3.2	0.33234	0.40973	0.0040973	0.050419
x3.3	0.87528	0.52010	0.0052010	0.073339
x3.4	0.13444	0.46060	0.0046060	0.059887
x4.2	-0.03852	0.42824	0.0042824	0.053261
x4.3	-0.85961	0.48686	0.0048686	0.065099
x4.4	-1.02733	0.43251	0.0043251	0.056038
X5.2	0.12555	0.35290	0.0035290	0.042000
X5.3	0.70711	0.42778	0.0042778	0.049479
X5.4	0.75757	0.45647	0.0045647	0.052952
X6.2	0.06106	0.35711	0.0035711	0.042997
X6.3	-0.13806	0.41877	0.0041877	0.046485
X6.4	0.30728	0.35657	0.0035657	0.034874
X7.2	-1.09444	0.81894	0.0081894	0.149750
X7.3	-2.40954	0.84464	0.0084464	0.136247
X7.4	-1.34378	0.78086	0.0078086	0.119453

Table 2: Posterior Quantiles of multinomial logistic regression by RWM algorithm

variables	2.5%	25%	50%	75%	97.5%
(Intercept).2	-11.195582	-4.27962	-1.29680	-0.5053	9.36390
(Intercept).3	-11.526499	-3.36821	-1.30583	0.4401	13.85018
(Intercept).4	-10.896478	-4.19765	-1.71869	0.1331	10.09733
x1.2	0.028102	0.12149	0.12149	0.1768	0.27463
x1.3	0.158505	0.25988	0.25988	0.3354	0.49382
x1.4	0.054499	0.13697	0.14741	0.1889	0.34231
x2.2	-1.372136	0.32857	0.48426	0.6873	2.28765
x2.3	-0.513304	0.69858	0.79111	1.4800	2.72409
x2.4	-0.206621	0.91799	0.94152	1.3462	2.78272
x3.2	-0.518150	0.13381	0.29300	0.5022	1.27302
x3.3	-0.118278	0.67083	0.72365	1.0669	2.16543
x3.4	-0.705120	-0.06898	0.04264	0.3935	1.16008
x4.2	-0.825022	-0.22687	-0.10590	0.2253	0.88222
x4.3	-1.962758	-1.12527	-0.77285	-0.6355	0.15971
x4.4	-2.046243	-1.22972	-0.90515	-0.8545	-0.33769
X5.2	-0.511489	-0.10537	0.13726	0.2065	0.90128
X5.3	-0.090201	0.54006	0.54006	0.9919	1.62010
X5.4	-0.003999	0.52819	0.65324	0.9462	1.80124
X6.2	-0.573281	-0.11378	0.04606	0.1242	0.90689
X6.3	-0.951068	-0.34028	-0.12539	-0.1077	0.84279
X6.4	-0.372212	0.23530	0.27913	0.4710	1.02454
X7.2	-3.317743	-1.32783	-0.95224	-0.7720	0.54506
X7.3	-4.698336	-2.91090	-2.01511	-1.9315	-1.18779
X7.4	-3.487843	-1.59095	-1.10496	-1.1050	-0.03872


Trace of (Intercept).2
Density of (Intercept).2
Trace of (Intercept).3
Density of (Intercept).3
Trace of (Intercept).4
Density of (Intercept).4
Trace of x1.2
Density of x1.2


Trace of x1.3
Density of x1.3
Trace of x1.4
Density of x1.4
Trace of x2.2
Density of x2.2
Trace of x2.3
Density of x2.3


Trace of x2.4
Density of x2.4
Trace of x3.2
Density of x3.2
Trace of x3.3
Density of x3.3
Trace of x3.4
Density of x3.4

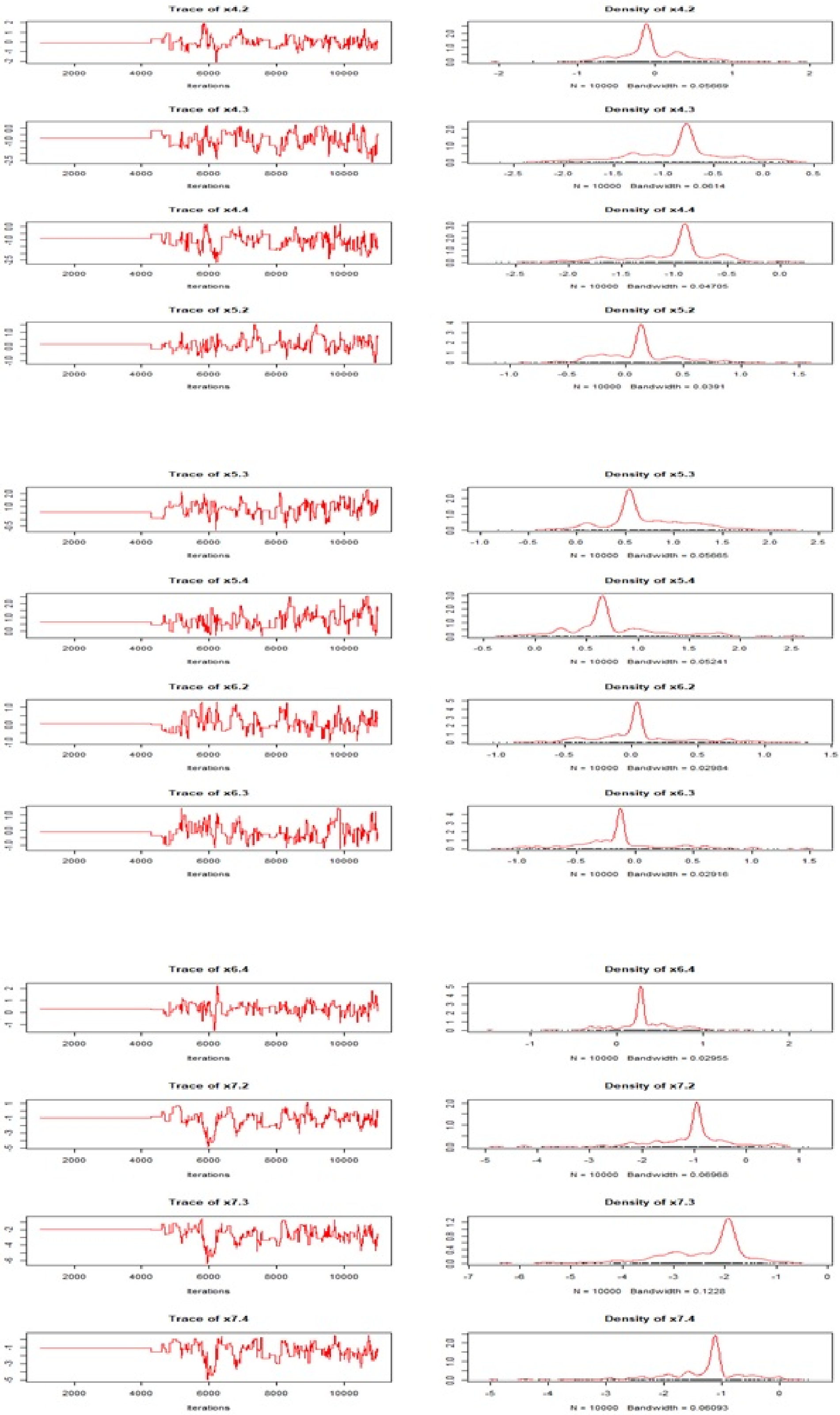


Figure 1: Trace plots and the density distribution of corresponding posterior estimates for variables in addiction Model by RWM algorithm

We employed geweke diagnostics [5] to determine the convergence of Markov chains to stationarity and given its statistics in Table 3.

Table 3: Geweke diagnostic statistics for multinomial logistic regression model variables that estimated by RWM algorithm.

variables	Z statistic
(Intercept).2	0.5234
(Intercept).3	-0.4300
(Intercept).4	-0.09698
x1.2	-3.4210
x1.3	-3.2010
x1.4	-2.2430
x2.2	-0.3233
x2.3	-2.0340
x2.4	-2.0080
x3.2	-0.6452
x3.3	-2.0760
x3.4	-1.3570
x4.2	-0.8741
x4.3	1.2930
x4.4	2.8220
X5.2	-0.5154
X5.3	-3.3910
X5.4	-2.2190
X6.2	-0.4531
X6.3	0.5595
X6.4	-0.5205
X7.2	1.4700
X7.3	5.1970
X7.4	2.6840

Variables with $|z| > 2$ are shown in Table 3. As a result, the age variables in three models have not converged. Job, Absence of a religious deterrent and Experience using narcotic substances at an early age variables in two models have not converged. Also Weakness of the rule of law, Weak family supervision variables in one model have not converged. According to Geweke diagnosis, all other variables of the multinomial logistic regression models have converged via RWM algorithm, and the models RWM can be represented as:

$$\widehat{\mathbf{y}}_{\mathbf{2(RWM)}} = -1.74290 + 0.53417x_2 + 0.33234x_3 - 0.03852x_4 + 0.12555x_5 + 0.06106x_6 - 1.09444x_7$$
$$\widehat{\mathbf{y}}_{\mathbf{3(RWM)}} = -1.17814 - 0.85961x_4 - 0.13806x_6$$
$$\widehat{\mathbf{y}}_{\mathbf{4(RWM)}} = -1.66597 + 0.13444x_3 + 0.30728x_6$$

We calculated the posterior simulation of the Slice method. The sample is summarized in the following tables:

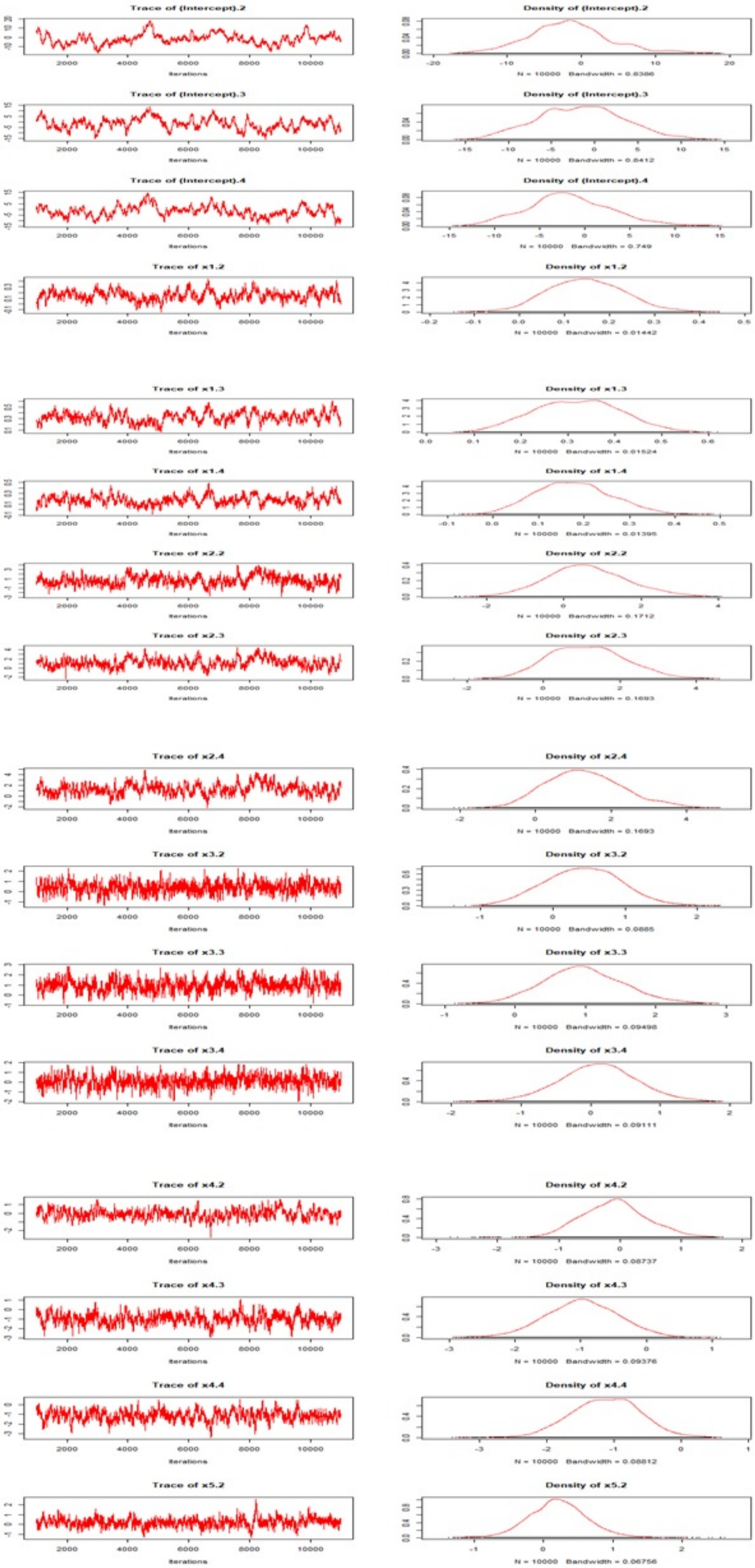
Table 4: Empirical mean and standard deviation for each variable, standard error of the mean by Slice algorithm of the multinomial logistic regression

variables	Mean	SD	Naive SE	Time-series SE
(Intercept).2	-1.61546	5.60349	0.0560349	1.221739
(Intercept).3	-1.63257	5.00719	0.0500719	0.945274
(Intercept).4	-1.67682	4.64399	0.0464399	0.871068
x1.2	0.14705	0.08584	0.0008584	0.009059
x1.3	0.319187	0.09073	0.0009073	0.010677
x1.4	0.17707	0.08303	0.0008303	0.008880
x2.2	0.60068	1.02936	0.0102936	0.086346
x2.3	1.15145	1.00802	0.0100802	0.094102
x2.4	1.30023	1.00784	0.0100784	0.096335
x3.2	0.41607	0.52680	0.0052680	0.027939
x3.3	0.96971	0.56863	0.0056863	0.033198
x3.4	0.10523	0.56305	0.0056305	0.031037
x4.2	-0.08558	0.53557	0.0053557	0.043453
x4.3	-0.97905	0.55973	0.0055973	0.040740
x4.4	-1.13091	0.53115	0.0053115	0.039644
X5.2	0.21391	0.43696	0.0043696	0.029043
X5.3	0.77467	0.54435	0.0054435	0.043350
X5.4	0.84311	0.53745	0.0053745	0.045760
X6.2	0.11353	0.42023	0.0042023	0.030035
X6.3	-0.08735	0.50211	0.0050211	0.040171
X6.4	0.35439	0.48286	0.0048286	0.043482
X7.2	-1.29274	0.86700	0.0086700	0.136351
X7.3	-2.59563	0.76227	0.0076227	0.095308
X7.4	-1.51944	0.71644	0.0071644	0.095304

Table 5: Posterior Quantiles of multinomial logistic regression by Slice algorithm

variables	2.5%	25%	50%	75%	97.5%
(Intercept).2	-12.11860	-5.24194	-1.82539	1.4473	11.58190
(Intercept).3	-11.18803	-5.17579	-1.60670	1.6920	8.45397
(Intercept).4	-10.23661	-4.65181	-1.96149	1.3221	8.20518
x1.2	-0.01626	0.08708	0.14612	0.2057	0.32289
x1.3	0.14756	0.25292	0.32031	0.3836	0.49562
x1.4	0.02543	0.11804	0.17390	0.2306	0.34664
x2.2	-1.32496	-0.09441	0.55035	1.2712	2.75302
x2.3	-0.66074	0.41730	1.11267	1.8171	3.21573
x2.4	-0.48578	0.58756	1.24408	1.9430	3.47414
x3.2	-0.59969	0.05022	0.42004	0.7766	1.44965
x3.3	-0.10198	0.58664	0.94767	1.3442	2.14913
x3.4	-1.03182	-0.25949	0.11155	0.4672	1.22666
x4.2	-1.06796	-0.44842	-0.08765	0.2485	1.00147
x4.3	-2.10098	-1.34631	-0.97630	-0.5984	0.10711
x4.4	-2.17624	-1.47562	-1.11414	-0.7727	-0.12659
X5.2	-0.62787	-0.06434	0.20750	0.4746	1.10951
X5.3	-0.28275	0.41429	0.76745	1.1333	1.82744
X5.4	-0.16064	0.49031	0.82363	1.1881	1.94415
X6.2	-0.65797	-0.18057	0.10038	0.3880	0.97361
X6.3	-1.04589	-0.42548	-0.11525	0.2324	0.97273
X6.4	-0.58574	0.03604	0.34578	0.6652	1.36369
X7.2	-3.15520	-1.85329	-1.22644	-0.7136	0.31717
X7.3	-4.21304	-3.08198	-2.56117	-2.0913	-1.17063
X7.4	-3.01330	-1.97258	-1.52412	-1.0676	-0.05915

Figure 2 shows trace plots and density plots of the posterior distributions of the Markov chain where trace plots indicate good Markov chain mixing for all variables.


Trace of (Intercept).2
Density of (Intercept).2
Trace of (Intercept).3
Density of (Intercept).3
Trace of (Intercept).4
Density of (Intercept).4
Trace of x1.2
Density of x1.2
Trace of x1.3
Density of x1.3
Trace of x1.4
Density of x1.4
Trace of x2.2
Density of x2.2
Trace of x2.3
Density of x2.3
Trace of x2.4
Density of x2.4
Trace of x3.2
Density of x3.2
Trace of x3.3
Density of x3.3
Trace of x3.4
Density of x3.4
Trace of x4.2
Density of x4.2
Trace of x4.3
Density of x4.3
Trace of x4.4
Density of x4.4
Trace of x5.2
Density of x5.2

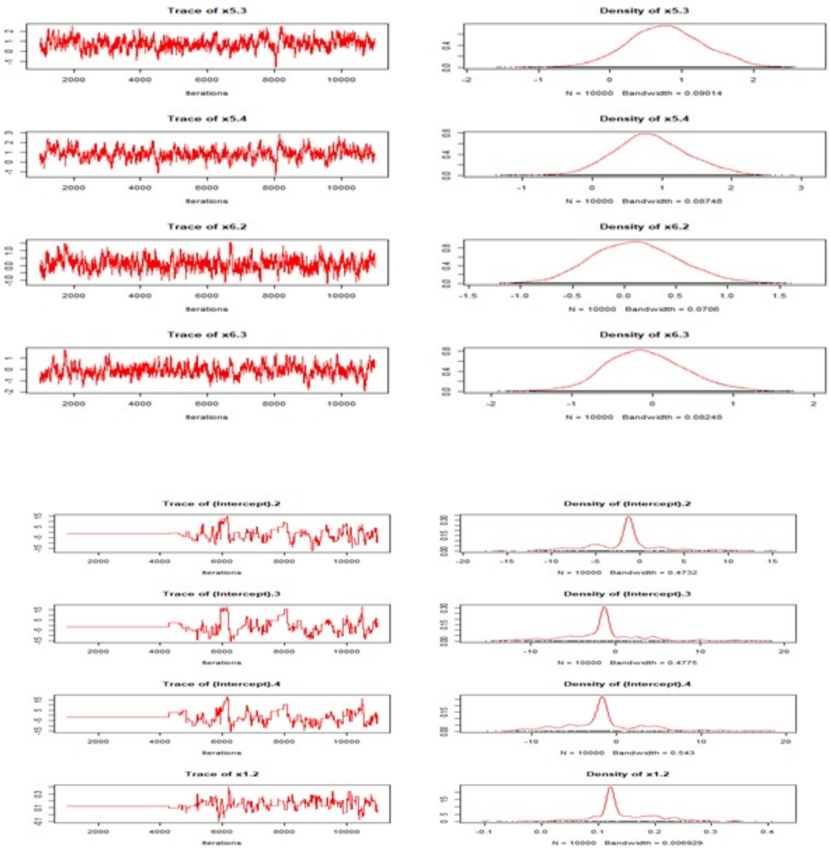


Figure 2: Trace plots and the density distribution of corresponding posterior estimates for variables in addiction Model by Slice algorithm

We used geweke diagnostics to determine the convergence of Markov chains to stationarity and given its statistics in Table 6.

Table 6: Geweke diagnostic statistics for multinomial logistic regression model variables that estimated by Slice algorithm.

variables	Z statistic
(Intercept).2	0.4018
(Intercept).4	-0.0552
x1.2	0.1098
x1.3	-0.3767
x1.4	-0.1871
x2.2	-1.2420
x2.3	-1.4460
x2.4	-1.2570
x3.2	-0.2572
x3.3	-0.8504
x3.4	-1.3340
x4.2	1.3000
x4.3	0.5787
x4.4	0.5459
X5.2	1.2190
X5.3	0.6756
X5.4	1.1030
X6.2	1.2220
X6.3	-0.2802
X6.4	0.2591
X7.2	0.2273
X7.3	-0.5639
X7.4	-0.0643

According to Geweke diagnosis, all variables of the multinomial logistic regression models have converged via Slice algorithm, and the models Slice can be represented as:

$$\begin{aligned}
\widehat{\mathbf{y}}_{\mathbf{2(Slice)}} &= -1.61546 + 0.14705x_1 - 0.60068\ x_2 + 0.41607\ x_3 - 0.08558\ x_4 + 0.21391\ x_5 \\
&+ 0.11353\ x_6 - 1.29274\ x_7 \\
\widehat{\mathbf{y}}_{\mathbf{3(Slice)}} &= -1.63257 + 0.31918\ x_1 + 1.15145\ \ x_2 + 0.96971\ \ x_3 - 0.97905\ \ x_4 + 0.77467x_5 \\
&- 0.08735\ x_6 - 2.59563\ x_7 \\
\widehat{\mathbf{y}}_{\mathbf{4(Slice)}} &= -1.67682\ + 0.17707\ x_1 + 1.30023\ \ x_2 + 0.10523\ \ x_3 - 1.13091\ \ x_4 + 0.84311x_5 \\
&+ 0.35439x_6 - 1.51944x_7
\end{aligned}$$

The Mean Square Error (MSE) that given by $MSE = \frac{\sum_{i=1}^{n}(y-\widehat{y})^2}{n}$ used to compare the algorithms. The smaller MSE, better the overall performance of the estimates. Table 7 shows the MSE value for each simulated data set using the RWM and Slice algorithms.

Table 7: Values of MSE for the RWM and Slice algorithms

Model	MSE(RWM)	MSE(Slice)
y_2	9.124997e-31	1.998345e-29
y_3	9.546360e-30	3.201050e-29
y_4	2.495559e-30	6.067450e-30

The MSE findings showed that the algorithms are successful and work excellently. It can also be configured as follows based on performance preferences: The RWM algorithm and the Slice algorithm.

7. Conclusion

The multinomial logistic regression technique would be used to describe categorical response variables. In this paper we covered two Bayesian approaches: the RWM algorithm and the slice algorithm in the process of obtaining Bayesian parameter estimation in a multinomial logistic regression model. We investigated the effectiveness of RWM and slice algorithms in estimating the parameters in addiction model. A series of plots is provided and will be plotted to determine whether the chain has reached convergence, the trace plot depicts sampled values to every variable and follows the iterations. The probability density function is generated by the density plot using the simulated data and our analysis showed that both algorithms seem to be efficient in estimating multinomial logistic regression model parameters.

References

[1] N.A. . Aziz, Z. Ali, N.M. Nor, A. Baharum and M. Omar, *Modeling multinomial logistic regression on characteristics of smokers after the smoke-free campaign in the area of Melaka*, AIP Conf. Proc. 1750(1) (2016) 60020.

[2] G.C. Cawley, N.L.C. Talbot and M. Girolami, *Sparse multinomial logistic regression via bayesian l1 regularisation*, Adv. Neural Inf. Process. Syst. 19 (2007) 209.

[3] A.M. El-Habil, *An application on multinomial logistic regression model*, Pakistan J. Statist. Oper. Res. 8(2) (2012) 271–291.

[4] S. Frühwirth-Schnatter and R. Frühwirth, *Data augmentation and MCMC for binary and multinomial logit models*, Statist. Modell. Regression Structures (2010) 111–132.

[5] J. Geweke, *Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments*, Federal Reserve Bank of Minneapolis, Research Department Minneapolis, 148 (1991).

[6] P.J. Green, K. Łatuszyński, M. Pereyra and C.P. Robert, *Bayesian computation: a summary of the current state, and samples backwards and forwards*, Stat. Comput., 25(4) (2015) 835–862.

[7] S. Kwon, D. Kim and S. Lee, *An efficient algorithm for the non-convex penalized multinomial logistic regression*, Commun. Statist. Appl. Methods 27(1) (2020) 129-–140.

[8] R.M. Neal, *Slice sampling*, Ann. Statist. 13(3) (2003) 705–741.

[9] C. Robert and G. Casella, *A short history of Markov chain Monte Carlo: Subjective recollections from incomplete data*, Statist. Sci. 26(1) (2011) 102–115.