Asian Journal of Scientific Research

ISSN 1992-1454

OPEN ACCESS Asian Journal of Scientific Research

ISSN 1992-1454
DOI: 10.3923/ajsr.2019.346.351

Research Article
Bayesian Quantile Regression Methods in Handling Non-normal and Heterogeneous Error Term

[1]Ferra Yanuar, [1]Hazmira Yozza and [2]Aidinil Zetra

[1]Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Andalas, Kampus Unand Limau Manis, 25163 Padang, Indonesia
[2]Department of Political Science, Faculty of Social and Political Sciences, Universitas Andalas, Kampus Unand Limau Manis, 25163 Padang, Indonesia

Abstract

Background and Objective: Quantile regression is a developing statistical tool which is used to explain the relationship between response and predictor variables. Quantile approach has ability to model the data which non-normal distributed and non-constant variance assumption. This study presented the ability of the quantile and Bayesian quantile method in overcoming the problem of violation of normality and homogenous assumption for error terms and compare the results. **Materials and Methods:** This research implemented the simulation study to explore the performance of the asymmetric Laplace distribution for working likelihood in posterior estimation process. Markov Chain Monte Carlo method using Gibbs sampling algorithm was then applied to estimate the parameter in quantile regression model. This study designed distributions for error term; normal, non-normal and heterogeneous variability, then compare the bias and Monte Carlo standard error as the results of classical quantile and Bayesian quantile method. Convergency of parameter estimated were also checked. **Results:** Bayesian quantile estimation method resulted lower biases and lower Monte Carlo standard error than the classical quantile method for all selected conditions of error term. **Conclusion:** This study proved that Bayesian quantile regression method produced better proposed model then classical quantile method in the case of non-normal and heterogenous error term.

Key words: Asymmetric laplace distribution, bayesian quantile, gibbs sampling, monte carlo standard error, quantile regression, posterior estimation

Received: October 13, 2018 Accepted: December 29, 2018 Published: June 15, 2019

Citation: Ferra Yanuar, Hazmira Yozza and Aidinil Zetra, 2019. Bayesian quantile regression methods in handling non-normal and heterogeneous error term. Asian J. Sci. Res., 12: 346-351.

Corresponding Author: Ferra Yanuar, Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Andalas, Kampus Unand Limau Manis, 25163 Padang, Indonesia Tel: +6281364643808

Competing Interest: The authors have declared that no competing interest exists.

Data Availability: All relevant data are within the paper and its supporting information files.

INTRODUCTION

Quantile regression is gradually emerging as a comprehensive approach to estimate the relationship between response variable y and the explanatory variables x. A number of papers have recently published that related to quantile regression, such as Yu and Moyeed[1], Kozumi and Kobayashi[2], Feng *et al.*[3] and Hong and Zhou[4]. Quantile regression is a statistical procedure based on minimizing sums of asymmetrically weighted absolute residuals and can be used to explore the relationship between quantiles of linear or non-linear response models distribution and available covariates. Quantile method provides more complete description of the response distribution than the mean regression.

Let y_i is the response variable, x_i is a $p\times 1$ vector of p indicator variables for the ith observation. Let $q_\tau(x_i)$ denote the τth (0<τ<1) quantile regression function of y_i given x_i. It can be modeled as:

$$q_\tau(x_i) = x_i^T\beta(\tau)$$

where, β(τ) is a $p\times 1$ vector of coefficients for indicator variables at specified τ. Therefore, the quantile regression model is considered as following:

$$y_i = x_i^T\beta(\tau) + u_i, \quad i = 1,2,...,n \tag{1}$$

where, u_i is the error term with mean zero and constant variance. Then, quantile regression estimation for β(τ) is obtained by minimizing:

$$\min\sum_i \rho_\tau\left(y_i - x_i^T\beta(\tau)\right) \tag{2}$$

where, $\rho_\tau(u)$ is the loss function defined by:

$$\rho_\tau(u) = u(\tau\text{-}I(u<0)) \tag{3}$$

It also may write (Eq. 3) as:

$$\rho_\tau(u) = u(\tau I(u>0) - (1-)I(u<0)) \text{ or } \rho_\tau(u) = \frac{|u| + (2\tau-1)u}{2}$$

where, I(.) denotes the indicator function. However, these indicator function is not differentiable at zero, thus this study cannot obtain explicit solutions in minimizing (Eq. 2). In quantile regression methods, it is commonly implement linear programming methods to derive parameter estimated[2].

Yu and Moyeed[1] proposed combination of Bayesian approach to quantile regression method in the minimizing problem. They used asymmetric Laplace error distribution to maximize likelihood distribution as equivalent way in minimizing[5-8] (Eq. 1-4). They assumed that error term follows an independent asymmetric Laplace distribution:

$$f_\tau(u) = \tau(1\text{-}\tau)e^{-\rho\tau(u)}, \quad u\in R \tag{4}$$

The mode of $f_\tau(u)$ is the solution to Eq. 2, thus the asymmetric Laplace distribution is closely related to quantile regression. However, the posterior density for parameter estimated β(τ) is not simple to obtained due to the complexity of the likelihood function, then Markov Chain Monte Carlo (MCMC) method is applied to sample from the approximate posterior distribution. Dunson and Taylor[8] used a random walk Metropolis algorithm with a Gaussian density centred at the current parameter value. Meanwhile Kozumi and Kobayashi[2] developed a Gibbs sampling algorithm based on a location-scale mixture representation of the asymmetric Laplace distribution.

This study considers to implement the asymmetric Laplace distribution for the error terms in the framework of quantile regression from a Bayesian perspective. It is assumed that error term violates of normality and homogenous assumption. The main objective of this study was to implement the superiority of Gibbs sampler to quantile method in the case of non-normal distribution and heteroscedastic variance of error and then compare the result to quantile regression. Data generated will be used to implement the proposed methods.

MATERIALS AND METHODS

Asymmetric laplace distribution: Let U is random variable assumed follow the asymmetric Laplace distribution with its probability distribution is formed by Eq. 4. If parameter τ = 0.5, which determines the skewness of distribution, Eq. 4 changes into standard symmetric Laplace distribution, with its density function is 0.25exp(-0.5|u|). Meanwhile, the density in Eq. 4 is asymmetric for all other values of τ. It can be proven that the mean and variance of this asymmetric Laplace distribution are respectively given by:

$$E(u_i) = \frac{1-2\tau}{\tau(1-\tau)} \text{ and } Var(u_i) = \frac{1-2\tau+2\tau^2}{\tau^2(1-\tau)^2} \tag{5}$$

Then incorporate the location (μ) and scale (σ) inside (Eq. 4), its density changes into following form Yang *et al.*[9] and Yu and Zhang[10]:

$$f_\tau(u;\mu,\sigma) = \frac{\tau(1-\tau)}{\sigma}\exp\left\{-\rho_\tau\left(\frac{u-\mu}{\sigma}\right)\right\}, \quad u \in \mathbb{R}$$

where, $\tau \in (0,1)$ is the skewness parameter, $-\infty<\mu<\infty$ is the location parameter which reflecting both the mode and τth quantile and $\sigma>0$ is the scale parameter[11].

Bayesian quantile regression: Yu and Zhang[10] wrote some other properties of this asymmetric Laplace distribution. Let p and ε are identic and independent standard exponential distributions, thus $\frac{p}{\tau}-\frac{\varepsilon}{(1-\tau)}$ is assumed has the asymmetric Laplace distribution as well. Hong and Zhou[4] applied Gibbs sampling algorithm for estimating parameter in quantile regression model.

It has defined above that random variable U assumed follow the asymmetric Laplace distribution with density as presented in Eq. 4. Let z be an standard exponential variable and ϵ a standard normal variable. Therefore, it can represent u as a location-scale mixture of normals given by:

$$u = \theta z + \sigma\sqrt{z}\,\epsilon$$

Where:

$$\theta = \frac{1-2\tau}{\tau(1-\tau)} \text{ and } \sigma^2 = \frac{2}{\tau(1-\tau)}$$

Based on this result, the y_i variable can equivalently represented as follows:

$$y_i = x_i^T\beta(\tau) + \theta z_i + \sigma\sqrt{z_i}\,\epsilon_i \quad (6)$$

where, z_i~exp(1) denotes an exponential distribution with mean 1 and ϵ_i~N(0,1)z_i and ϵ_i are mutually independent. Meanwhile conditional distribution of y_i given z_i is assumed follow normal distribution with mean $x_i^T\beta(\tau)+\theta z_i$ and variance $\sigma^2 z_i$. The mixture representation in Eq. 6 expands the likelihood specification into such a hierarchical structure which normal linear model framework then can be transferred to the quantile approach.

Thus the joint density of y = $(y_1, y_2,..., y_n)'$ is written as:

$$f\left(y\,|\,\beta(\tau),z\right) \propto \left(\prod_{i=1}^{n} z_i^{-1/2}\right)\exp\left\{-\sum_{i=1}^{n}\frac{\left(y_i - x_i^T\beta(\tau)-\theta z_i\right)^2}{2\sigma^2 z_i}\right\} \quad (7)$$

where, z = $(z_1, z_2,..., z_n)'$.

In Bayesian analysis, the posterior distribution of selected parameter are estimated by multiplying likelihood distribution and prior distribution. Equation 7 is equivalent to likelihood distribution, thus it needs prior distribution then. In this study, the prior distribution used here is assumed as follows:

$$\beta(\tau) \sim N(\beta(\tau 0), B(\tau 0)) \quad (8)$$

where, $\beta(\tau 0)$ is prior mean of $\beta(\tau)$ and $B(\tau 0)$ is covariance of $\beta(\tau)$. Therefore prior distribution of $\beta(\tau)$, known as conditional density of $\beta(\tau)$ given y and z, given by:

$$\beta(\tau)|y, z \sim N(\hat{\beta}(\tau), \hat{\beta}(\tau)) \quad (9)$$

Where:

$$\hat{B}^{-1}(\tau) = \sum_{i=1}^{n}\frac{x_i x_i'}{\sigma^2 z_i} + B^{-1}(\tau 0)$$

and:

$$\hat{\beta}(\tau) = \hat{B}(\tau)\left\{\sum_{i=1}^{n}\frac{x_i\left(y_i-\theta z_i\right)}{\sigma^2 z_i} + B^{-1}(\tau 0)\hat{\beta}(\tau 0)\right\}$$

Since it is assumed that z_i follows standard exponential distribution, the full conditional distribution of z_i is given by:

$$f\left(z_i\,|\,y,\beta(\tau)\right) \propto z_i^{-1/2}\exp\left\{-\frac{1}{2}\left(\hat{\delta}_i^2 z_i^{-1} + \hat{\gamma}_i^2 z_i\right)\right\} \quad (10)$$

Where:

$$\hat{\delta}_i^2 = \frac{\left(y_i - x_i^T\right)^2}{\sigma^2}$$

and $\hat{\gamma}_i^2 = 2+\theta^2/\sigma^2$. Equation 10 is known as the kernel of a generalized inverse Gaussian distribution and it can be rewritten as follows[2]:

$$f\left(z_i\,|\,y,\beta(\tau)\right) \propto \mathcal{GIG}\left(\frac{1}{2}, \hat{\delta}_i^2, \hat{\gamma}_i^2\right) \quad (11)$$

Where:

$$\mathcal{GIG}\left(\frac{1}{2}, \hat{\delta}_i^2, \hat{\gamma}_i^2\right)$$

denotes a generalized inverse Gaussian and its probability density function is given by:

$$f\left(z_i\,|\,y,\beta(\tau)\right) = \frac{\left(\hat{\gamma}_i^2/\hat{\delta}_i^2\right)^{1/2}}{2K_{1/2}\left(\hat{\delta}_i^2\,\hat{\gamma}_i^2\right)} z_i^{-1/2}\exp\left\{-\frac{1}{2}\left(\hat{\delta}_i^2 z_i^{-1} + \hat{\gamma}_i^2 z_i\right)\right\},\ z_i>0,\ \hat{\delta}_i^2,\ \hat{\gamma}_i^2 \geq 0 \quad (12)$$

and $K_{1/2}(.)$ is a modified Bessel function of the third kind.

So, the priors that used in this Bayesian analysis are as defined in Eq. 9 and 11. Based on empirical studies, this asymmetric Laplace for likelihood distribution can be computed directly from the MCMC chains. Many studies have shown that the Gibbs sampler then can be easily applied to quantile regression estimation[5,2]. Choi and Hobert[7] and Khare and Hobert[12] have proven that Markov chain underlying Gibbs sampling algorithm are geometrically ergodic.

Alhamzawi and Yu[13] utilized efficient Gibbs sampling algorithms for Bayesian quantile regression using R packages; brq, meanwhile Benoit used bayesQR[5]. Kozumi and Kobayashi[2] implemented the quantreg package for the R language. This study implement bayesQR and MCMCquantreg in MCMCpack package for R language to estimate the parameter model and result the Traceplot and Marginal posterior distribution of each parameters[14]. Traceplot is used to monitor the convergency of the algorithm based on Gibbs sampling method[1,15] meanwhile marginal posterior distribution is used for check the distribution of generated sample. While Sriram *et al.*[16] constructed sufficient condition for the posterior consistency of model parameters in Bayesian quantile regression with mis-specified asymmetric Laplace density.

This study generate data to apply proposed methods, quantile regression and Bayes quantile regression. Response variable, y_i is generated from the model:

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + u_i, \qquad i = 1,..., 200 \qquad (13)$$

where, covariate x_{i1} and x_{i2} are generated from a standard normal distribution, while β_1 and β_2 are set to one. This study considers five different distribution for u_i: (i) Standard normal distribution, N(0,1), (ii) Exponential distribution with one degrees of freedom minus one, Exp(1)-1, (iii) The Student's t distribution with five degree of freedom, t_5, (iv) Heteroscedastic normal, $N\left(0, \sqrt{0.01*(X\beta)^2}\right)$ and (v) Heteroscedastic error, $\left(1+0.2x_1^2\right)t_5$. All five different distribution for u_i, each is independent of x_{i1} and x_{i2}. Case (i) represents a normal and homoscedastic error model, cases (ii) and (iii) represent non normal error model, meanwhile case (iv) and (v) represent heteroscedastic error models. For classical method, the parameters are estimated by minimizing Eq. 2 using quantreg package in R language[17]. The normal prior for $\beta(\tau)$ is as given by Eq. 8 and the hyper parameter are chosen as $\beta(\tau 0) = 0$ and $B(\tau 0) = 100I$.

RESULTS AND DISCUSSION

For Bayesian estimation, this study implemented Gibbs sampling method to the quantile regression approach which posterior mean is calculated from a sample of 5000 draws with the first 1000 of which are discarded to mitigate the impact of start up effects. The result are summarized in Table 1, presenting the Bias and Monte

Table 1: Simulation study for non-normal and heteroscedastic error

		Parameter of $X_1(\beta_1)$		Parameter of $X_2(\beta_2)$	
Conditions of error and methods	Quantile τ	Bias	MCSE*	Bias	MCSE*
u_i~N(0,1) Classical quantile	0.25	0.063	0.113	0.029	0.103
	0.50	-0.051	0.102	0.078	0.093
	0.75	0.022	0.100	0.048	0.092
Bayesian quantile	0.25	0.050	0.002	0.050	0.001
	0.50	-0.030	0.001	0.056	0.001
	0.75	-0.003	0.001	0.024	0.001
u_i~Exp(1)-1 Classical quantile	0.25	-0.067	0.052	-0.013	0.047
	0.50	-0.022	0.075	-0.099	0.068
	0.75	0.071	0.091	-0.086	0.083
Bayesian quantile	0.25	-0.041	0.001	-0.004	0.001
	0.50	-0.011	0.001	-0.076	0.001
	0.75	0.065	0.002	-0.078	0.001
u_i~t_5 Classical quantile	0.25	0.000	0.071	-0.026	0.065
	0.50	-0.079	0.088	0.050	0.081
	0.75	-0.240	0.102	0.209	0.094
Bayesian quantile	0.25	0.015	0.002	0.028	0.001
	0.50	-0.073	0.001	0.049	0.001
	0.75	0.003	0.000	-0.002	0.000
$u_i \sim N\left(0, \sqrt{0.01*(X\beta)^2}\right)$ Classical quantile	0.25	0.001	0.000	0.002	0.000
	0.50	0.001	0.000	0.001	0.000
	0.75	0.004	0.001	0.001	0.001
Bayesian quantile	0.25	0.002	0.000	-0.001	0.000
	0.50	0.002	0.000	-0.002	0.000
	0.75	0.003	0.000	-0.002	0.000
$u_i \sim (1+0.2x_1^2)t_5$ Classical quantile	0.25	0.002	0.010	-0.004	0.009
	0.50	0.002	0.001	-0.002	0.001
	0.75	0.013	0.004	-0.002	0.004
Bayesian quantile	0.25	0.002	0.000	-0.007	0.000
	0.50	0.004	0.001	0.000	0.000
	0.75	0.002	0.001	0.000	0.000

*MCSE: Monte Carlo Standard Error

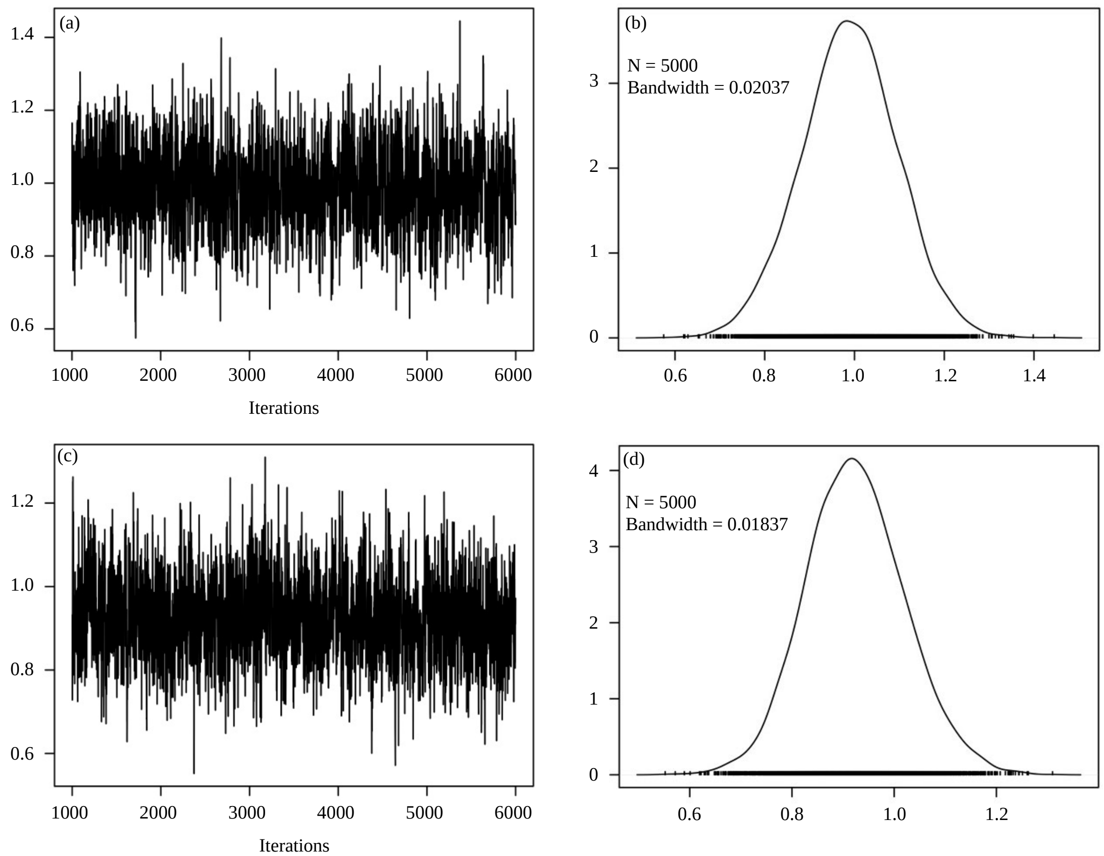

Fig. 1(a-d): Trace plot and marginal posterior distribution for corresponding X_1 and X_2 at $\tau = 0.50$, u_i~Exp(1)-1, (a) Trace of x1, (b) Density of x1, (c) Trace of x2 and (d) Density of x2

Carlo standard error (MCSE) for each parameter model in selected different distributions for error.

The data in Table 1 presented the simulation result for $\tau = 0.25$, 0.50 and 0.75 and results for the bias and Monte Carlo standard error (MCSE) for β_1 and β_2 in each conditions of error. Table 1 shows that Bayesian quantile estimation method yields lower biases than the classical method for all five different distributions for error. Table 1 also informs us that Monte Carlo standard errors of the Gibbs sampler are lower than classical quantile standard errors for all conditions of error. These results indicated the advantages of Gibbs sampler method to the classical method. Gibbs sampler methods could result parameter estimated with lower bias and lower Monte Carlo standard error than classical quantiles at any selected quantiles in respective conditions of error.

The trace plots and marginal posterior distribution for all conditions of error at all selected quantiles are checked. Due to limited space, this study reported only when $\tau = 0.50$ for non-normal error distributions, u_i~Exp(1)-1, which reported in Fig. 1. For other quantiles, similar plots and marginal posterior distributions are available by author.

Based on Fig. 1, it can be seen that the convergency of all parameter estimated can be reached easily in only 5,000 iteration with burn in 1000. This also reported that all marginal posterior density at $\tau = 0.50$ are very similar for the different distributions of error. Based on these empirical results, this study informed that Gibbs sampling method produces well estimated parameters in all three different condition for error terms.

DISCUSSION

Bayesian quantile method implemented the asymmetric Laplace distribution to form the likelihood function and normal for prior distribution in posterior estimation method. Simulation study was implemented to show all three conditions of error distributions, they are normal error, non-normal error and heteroscedastic error. The MCMC method with Gibbs sampling is implemented to obtain the value of parameter estimated. The convergency of parameter estimated are also tested using traceplot and marginal posterior distribution. This results that the use of asymmetric

Laplace working likelihood makes the Bayesian method produces well estimated parameters to the classical quantile regression method.

The results found in this study were consistent with the study of Alhamzawi *et al.*[18], Oh *et al.*[19], Mollica and Petrella[11]. Meanwhile Alhamzawi and Yu[20] also proved that Bayesian approach in quantile regression can result better model though in non-normal condition.

CONCLUSIONS

This study showed that Bayesian asymmetric Laplace distribution method for quantile regression is a viable strategy to model conditional quantiles when non normal and heteroskedasticity are available in the data.

SIGNIFICANCE STATEMENT

This study discovered the new method that have ability to model the data which do not fulfill the normality and homogenous assumption of error term. This study hope can beneficial for researcher especially who will construct a model using simulation or empirical data.

ACKNOWLEDGMENTS

The authors are grateful for any anonymous referees whose gave comments on our earlier draft of the paper and helped us improve presentation. This research was financially supported by Higher Education of Indonesia.

REFERENCES

1. Yu, K. and R.A. Moyeed, 2001. Bayesian quantile regression. Stat. Probab. Lett., 54: 437-447.
2. Kozumi, H. and G. Kobayashi, 2011. Gibbs sampling methods for Bayesian quantile regression. J. Stat. Comput. Simul., 81: 1565-1578.
3. Feng, Y., Y. Chen and X. He, 2015. Bayesian quantile regression with approximate likelihood. Bernoulli, 21: 832-850.
4. Hong, H.G. and J. Zhou, 2013. A multi-index model for quantile regression with ordinal data. J. Applied Stat., 40: 1231-1245.
5. Benoit, D.F. and D. Van den Poel, 2017. Bayes QR: A bayesian approach to quantile regression. J. Stat. Software, Vol. 76.
6. Chib, S., 2001. Markov chain monte carlo methods: Computation and inference. Handbook Econometr., 5: 3569-3649.
7. Choi, H.M. and J.P. Hobert, 2013. Analysis of MCMC algorithms for Bayesian linear regression with Laplace errors. J. Multivariate Anal., 117: 32-40.
8. Dunson, D.B. and J.A. Taylor, 2005. Approximate Bayesian inference for quantiles. J. Nonparamet. Stat., 17: 385-400.
9. Yang, Y., H.J. Wang and X. He, 2016. Posterior inference in Bayesian quantile regression with asymmetric Laplace likelihood. Int. Stat. Rev., 84: 327-344.
10. Yu, K. and J. Zhang, 2005. A three-parameter asymmetric Laplace distribution and its extension. Commun. Statistics-Theory Meth., 34: 1867-1879.
11. Mollica, C. and L. Petrella, 2017. Bayesian binary quantile regression for the analysis of bachelor-to-master transition. J. Applied Stat., 44: 2791-2812.
12. Khare, K. and J.P. Hobert, 2012. Geometric ergodicity of the Gibbs sampler for Bayesian quantile regression. J. Multivariate Anal., 112: 108-116.
13. Alhamzawi, R. and K. Yu, 2014. Bayesian lasso-mixed quantile regression. J. Stat. Comput. Simul., 84: 868-880.
14. Lancaster, T. and S.J. Jun, 2010. Bayesian quantile regression methods. J. Applied Econometr., 25: 287-307.
15. Yanuar, F., 2014. The estimation process in Bayesian structural equation modeling approach. J. Phys.: Conf. Ser., Vol. 495.
16. Sriram, K., R.V. Ramamoorthi and P. Ghosh, 2013. Posterior consistency of Bayesian quantile regression based on the misspecified asymmetric Laplace density. Bayesian Anal., 8: 479-504.
17. Koenker, R., 2017. Quantreg: Quantile regression. R package version 5.33.
18. Alhamzawi, R., K. Yu and D.F. Benoit, 2012. Bayesian adaptive Lasso quantile regression. Stat. Model., 12: 279-297.
19. Oh, M.S., E.S. Park and B.S. So, 2016. Bayesian variable selection in binary quantile regression. Stat. Probab. Lett., 118: 177-181.
20. Alhamzawi, R. and K. Yu, 2012. Variable selection in quantile regression via Gibbs sampling. J. Applied Stat., 39: 799-813.