Science and Technology Indonesia
e-ISSN:2580-4391 *p*-ISSN:2580-4405
Vol. 5, No. 3, July 2020

Research Paper

Simulation Study of Autocorrelated Error Using Bayesian Quantile Regression

Nayla Desviona[1], Ferra Yanuar[2]*

[1,2] *Department of Mathematics, Andalas University, Limau Manis Campus, Padang 25163, Indonesia*
***Corresponding author**: ferrayanuar@sci.unand.ac.id

Abstract
The purpose of this study is to compare the ability of the Classical Quantile Regression method and the Bayesian Quantile Regression method in estimating models that contain autocorrelated error problems using simulation studies. In the quantile regression approach, the data response is divided into several pieces or quantiles condition on indicator variables. Then, The parameter model is estimated for each selected quantiles. The parameters are estimated using conditional quantile functions obtained by minimizing absolute asymmetric errors. In the Bayesian quantile regression method, the data error is assumed to be asymmetric Laplace distribution. The Bayesian approach for quantile regression uses the Markov Chain Monte Carlo Method with the Gibbs sample algorithm to produce a converging posterior mean. The best method for estimating of parameter is the method that produces the smallest absolute value of bias and the smallest confidence interval. This study resulted that the Bayesian Quantile method produces smaller absolute bias values and confidence interval than the quantile regression method. These results proved that the Bayesian Quantile Regression method tends to produce better estimate values than the Quantile Regression method in the case of autocorrelation errors.

Keywords
Quantile Regression Method, Bayesian Quantile Regression Method, Confidence Interval, Autocorrelation.

Received: 20 May 2020, Accepted: 16 July 2020
https://doi.org/10.26554/sti.2020.5.3.70-74

1. INTRODUCTION

Parameter estimation in linear regression analysis is performed using the Ordinary Least Squares Method. This method has several assumptions that must be approved in order to get the Best Linear Unbias Estimator (BLUE). In some empirical data, not all assumptions can be fulfilled, such as autocorrelation among errors.

Then, the quantile regression method appears to overcome weaknesses in the Ordinary Least Square (OLS) method. This method uses a parameter estimation approach by separating or dividing data into quantiles by assuming conditional quantile functions in data distribution and minimizing the absolute symmetric weighted errors. This quantile regression analysis is used to overcome assumptions that are not met, including the existence of autocorrelation, normality assumptions, no multicollinearity and homogeneity of variances (Yanuar et al., 2019a)

Large data size is usually needed in the quantile regression method. A sampling of large data requires a lot of time and a lot of energy. Therefore it is used the Bayes method to evaluate parameters with quantile regression. Bayesian related to variable selection in quantile regression has received great attention in the literature because the Bayes method is able to get good models with small data (Oh et al., 2016), (Yanuar et al., 2013), (Yanuar et al., 2019d).

In Bayes's views, the unknown parameter is assumpted as a random variable and has distribution. Distribution related to the parameter can be obtained from corresponding previous research or based on expert opinion, this distribution known as the prior distribution. Then, the prior distribution is combined with information from data obtained from sampling (known as likelihood function). The combination of both distribution then results in a posterior distribution of parameters. Averages and variations of this posterior distribution are made estimators for the regression parameters by the Bayesian method (Yanuar et al., 2019b). In case of difficulties to identify the distribution of posterior distribution or because of complex formulation, Bayesian method uses the MCMC (Markov Chain Monte Carlo) algorithm to estimate the mean posterior and variance posterior of the parameter model.

In previous studies (Muharisa et al., 2018), the Bayesian Quantile Regression Method with Abnormal Error has been discussed in the case of Low Birth Weight (BBLR) in West Sumatra in the data of 2016 to 2018. Furthermore (Delviyanti et al., 2018) has been examined the application of the Quantile Regression with the Bootstrap Method to the autocorrelated error in the case of the interest rate on Indonesia's inflation rate. This article will compare the ability of the Quantile Regression method and the Bayesian quantile Regression method in overcoming the autocorrelated error problem using a simulation study.

2. EXPERIMENTAL SECTION

2.1 Materials

2.1.1 Quantile Regression

The quantile method is one of the methods of regression modeling by dividing data sets into several equal parts with data sorted from the smallest or largest. Quantile regression, in theory, is able to overcome the existence of autocorrelation, normality assumptions, heteroskedasticity, multicollinearity problems and etc. Quantile regression minimizes asymmetric weighted data and agrees with the data function on the data distribution, (Muharisa et al., 2018). Linear equations for quantile θ can be written with the equation :

$$Y = X_i^{'}\beta(\theta) + \varepsilon \tag{1}$$

The general estimation for it to θ can be written as follows:

$$argmin_{x_i^{'}\beta\theta\epsilon R}\sum_{i=1}^{n}\rho_\theta(y_i - Q_\theta(Y|X)) \tag{2}$$

Where :

θ : Show as quantile index $\epsilon(0,1\ )$

ρ_θ : is an asymmetric loss function

$Q_\theta(Y|X) = X_i^{'}\beta$ the quantile function to θ from Y on X condition, (Yanuar et al., 2019a).

2.1.2 Bayesian Methods

Bayes introduced a parameter estimation method where we need to know the form of the initial distribution (prior) of parameters to find the estimated θ parameter of the population, known as the Bayesian method. The Bayesian method uses the prior $f\ (\theta)$ distribution together with the likelihood function to determine the posterior $f(\theta|x_1, x_2, ..., x_n)$ distribution (Yanuar et al., 2019c)

Given $y = (y_1, y_2, \ldots y_n)$ where the prior distribution of β is ρ (β). The prior distribution taken in this research is prior informative which is originating from previous research. Determination of prior distribution parameters are very subjective, depending on the researcher's intuition.

2.1.3 Bayesian Quantile Regression

A random variable of Y is said to follow Asymmetric Laplace Distribution (ALD)(μ, σ, ρ) with location μ parameter, scale of σ>0 parameter and skewness p parameter in (0,1), the density function of the probability of ALD as follows, (Alhamzawi and Yu, 2012) ; (Feng and He, 2015).

$$f(y|\mu, \sigma, \rho) = \frac{p(1-p)}{\sigma}exp(-\rho_p(\frac{y-\mu}{\sigma})) \tag{3}$$

Where ρ_p is loss function that is defined as follows :

$$\rho_p(u) = u(p - 1_{u<0}) \tag{4}$$

With I shows the indicator function. (Benoit and Van den Poel, 2010) ; (Yanuar et al. (2019).

2.2. Methods

The data used in this study was data generated using R software version 3.6.1 (R Development Core Team, 2011). The data used in this study consisted of two variables X_1 and X_2 each generated from $N(0.1)$. While the dependent Y variable was set with the value $Y_t = 0.5X_{1t} + 2X_{2t} + \varepsilon_t.\varepsilon_t = Sin(S_{eq}(0.1\pi, 15.0\pi, 0.1\pi)) + Z_t$ with ~ $N(0, 0.1)$ to $t = 1, 2, ..., 150$.

3. RESULTS AND DISCUSSION

In this section, We will describe the results of parameter estimations and comparisons for the Quantile regression and Bayesian Quantile regression methods.

3.1. Durbin Watson (DW) test on Simulation Data

In this study, to see the existence of autocorrelation in error from simulation data used The Dubin Watson test.

Based on the results of Durbin Watson (DW) using R version 3.6.1 software (R Development Core Team, 2011) the statistical DW value was 0.124904. To find out whether the error of the simulation data was free from autocorrelation, We have compared the statistical DW values with DW table values. With the number of independent K variables was 2 and the number of n observations was 150.

Then the values for d_L=1.7062 and d_u=1.7602 were obtained from the DW Table values. In other words, Based on the Durbin Watson test, if the DW value lies between 0 and d_L stated that 0<DW<d_L Or 0<0.124904<1.7062. Then H_0 is rejected, the error contains autocorrelation. Thus, in this case, the simulation data contains autocorrelated errors.

3.2. Parameter Estimation by using the Quantile Regression and Bayesian Quantile Regression Model

At this step estimation parameters for quantile regression and Bayesian quantile regression will be performed. The data consists of two variables X_1 and X_2 each generated from N (0,1). While the dependent Y variable was set with the value Y_t = $0.5X_{1t} + 2X_{2t} + \varepsilon_t$

The quantile regression method is done by dividing or splitting the data into groups that have different estimations of the results in the quantiles.

In the Bayesian quantile regression method, the results of estimation are also seen in each quantiles. The Bayesian method uses the MCMC to estimate its posterior distribution with the help of R version 3.6.1 software (R Development Core Team, 2011). The results of parameter estimation by the quantile method and the Bayesian quantile method are presented in Table 1.

Table 1. Estimated Parameters by Both Methods.

Quantile	Quantile Reg		Bayesian Quantile Reg	
to θ	β_1	β_2	β_1	β_2
0.1	0.5230*	1.9899	0.4308*	1.9356*
0.25	0.4722*	1.9996*	0.4325*	1.9357*
0.5	0.3787*	1.9632	0.4340*	1.9391*
0.75	0.4282*	2.0619*	0.4329*	1.9296*
0.9	0.4779*	2.0418	0.4230*	1.9366*

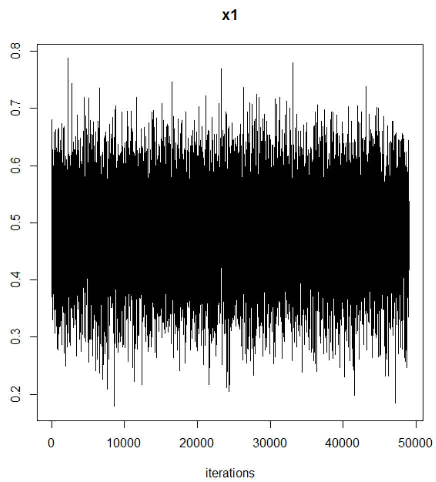

Figure 1. Trace plot β_1 with quantile 0.25

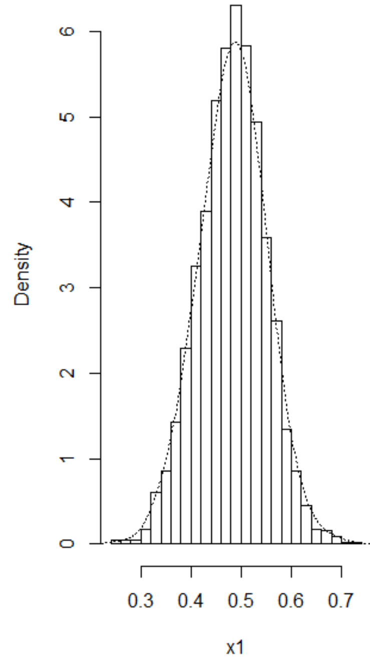

Figure 2. Plot Density β_1 with quantile 0.25

In Table 1, it can be seen that the parameter estimation for β_1 in quantile regression has been significant for all quantiles by selecting the real level $\alpha = 0.05$. Whereas β_2 is only significant at the quantiles 0.25 and 0.75. While the estimated parameters for Bayesian quantile regression for β_1 and β_2 have been significant for all each quantiles with a real level $\alpha = 0.05$.

The analysis continued with the Durbin Watson (DW) test, to estimate the two methods by looking at the DW values in each quantiles, this aimed to find out whether the model meets the Best Linear Unbias Estimator (BLUE) by seeing no more assumptions that are violated. For the Quantile Regression Method in the simulation data, the DW value was 2.0829 and the Bayesian Quantile Regression had DW value 2.1132 with a p-value <0.05 . In the DW table with n=150 and $K = 2$ and the values, D_L = 1.7249 and D_L = 1.7915.

Based on the Durbin Watson test, if the DW value is located at D_U < 4-D_L, then H_0 is accepted.

Thus, the error does not contain autocorrelation for both methods. Likewise for quantile 0.25, 0.5, 0.75, 0.9.

3.3 Comparison between the Quantile Regression Method and the Bayesian Quantile Regression

In this section, We compared the estimated results of the model parameters with the two methods by estimating the value of the Absolute Bias and the width of the Bayes confidence interval. Estimated results for both criteria are presented in Table 2.

Based on Table 2, it is known that the Absolute bias value for β_1 results of Bayesian quantile regression tends to produce a smaller value than the quantile regression results. Smaller values are marked in bold. While the Absolute Bias value for β_2 results of the quantile regression tends to produce a smaller value than the Bayesian quantile regression results. This is acceptable because the Bayes estimator is a biased estimator so that the results of estimated posterior mean may slightly deviate from the true values.

Then the analysis results from Table 2, also show that the confidence interval is 95% for β_1 and β_2 with the Bayes approach tends to produce a smaller value than the classical quantile method. This result is acceptable because the Bayes method tends to produce a smaller variance so that in the calculation of the confidence interval will also produce a smaller value than the classical quantile method.

Thus, it could be concluded that the Bayes estimator method tends to produce better guess values than the quantile method in autocorrelated error cases.

Furthermore, the convergence test would be conducted for the parameters of the model that has been obtained. The convergence test used was by choosing a trace plot, density plot and autocorelation plot of the selected parameters.

Figure 1 presents a trace plot of the β_1 parameter at the quantile 0.25. This figure shows that the estimated values of the parameter with the iteration process has spread around two

Table 2. Comparison of Absolute Bias & Interval of Confidence Method of Quantile Regression and Bayesian Quantile Regression.

Quantile to θ	*Absolute Bias*				*Bayes Credible Interval*			
	QR		BQR		QR		BQR	
	β_1	β_2	β_1	β_2	β_1	β_2	β_1	β_2
0.1	0.023	0.01	0.0692	0.0643	0.5578	0.4438	0.3005	0.2792
0.25	0.0578	0.046	0.0475	0.0643	0.4381	0.3744	0.3205	0.2921
0.5	0.1212	0.037	0.0666	0.0609	0.304	0.3246	0.3231	0.2746
0.75	0.0718	1.794	0.0711	0.0704	0.2996	0.2891	0.2354	0.2912
0.9	0.052	0.072	0.0772	0.0634	0.4443	0.2955	0.3111	0.2962

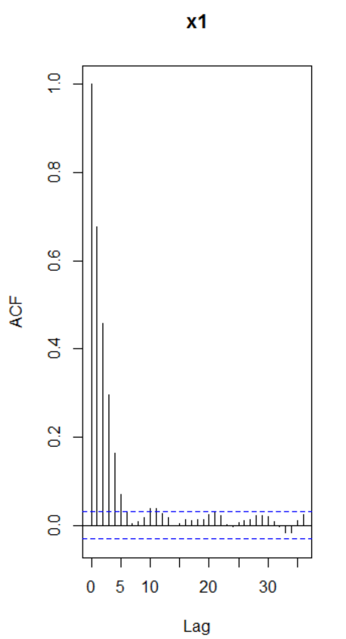

Figure 3. ACF plot β_1 with quantile 0.25

linear horizontal lines so the parameter is said to be convergent.

The Density Plot model parameters for quantile 0.25 with β_1 can be seen in Figure 2 and for the ACF plot can be seen in Figure 3. In Figure 3, the ACF plot shows stationary data on average. It can be seen that the ACF plot on the lag line above does not cross the 1.0 ACF line.

In this case, it can be said that there is no error containing autocorrelation for the selected quartile, the quartile 0.25 with the parameter β_1.

4. CONCLUSIONS

This study used a simulation study to apply the quantile regression and Bayesian quantile regression methods for autocorrelated errors. By comparing the quantile values in each parameters, and looking at the Absolute Bias values and the smallest confidence interval for each quantiles in the parameter. The quantile regression method is done by dividing or splitting the data into groups that have different estimates of the results in the quantiles. In the Bayesian quantile regression method the parameter estimation results are also seen in each quantiles. The Bayesian method uses MCMC to estimate its posterior distribution. The obtained results proved that the best method for estimating model parameter is the Bayesian quantile regression method. By looking at the quantile value of each parameter, the smallest quantile 0.25 was obtained based on the absolute bias value and the smallest confidence interval 95%.

REFERENCES

Alhamzawi, R. and K. Yu (2012). Variable selection in quantile regression via Gibbs sampling. *Journal of Applied Statistics*, **39**(4); 799–813

Benoit, D. F. and D. Van den Poel (2010). Binary quantile regression: A Bayesian approach based on the asymmetric Laplace density. *Journal of Applied Econometrics*

Delviyanti, O. S., F. Yanuar, and D. Devianto (2018). Simulation Study The Implementation of Quantile Bootstrap Method on Autocorrelated Error. *CAUCHY: Jurnal Matematika Murni dan Aplikasi*, **5**(3); 95–101

Feng, Y., Yang ; Chen and X. He (2015). Bayesian quantile regression with approximate likelihood. *Bernoulli*, **21**(2); 832–850

Muharisa, C., F. Yanuar, and D. Devianto (2018). Simulation Study The Using of Bayesian Quantile Regression in Nonnormal Error. *CAUCHY: Jurnal Matematika Murni dan Aplikasi*, **5**(3); 121–126

Oh, M., J. Choi, and E. Sug (2016). Bayesian variable selection in quantile regression using the Savage–Dickey density ratio. *Journal of the Korean Statistical Society*, **45**(3); 466–476

Yanuar, F., K. Ibrahim, and A. A. Jemain (2013). Bayesian structural equation modeling for the health index. *Journal of Applied Statistics*, **40**(6); 1254–1269

Yanuar, F., H. Yozza, F. Firdawati, I. Rahmi, and A. Zetra (2019a). Applying bootstrap quantile regression for the construction of a low birth weight model. *Makara Journal of Health Research*, **23**(2); 5

Yanuar, F., H. Yozza, and R. V. Rescha (2019b). Comparison of Two Priors in Bayesian Estimation for Parameter of Weibull Distribution. *Science and Technology Indonesia*, **4**(3); 82–87

Yanuar, F., H. Yozza, and A. Zetra (2019c). Bayesian Quantile Regression Methods in Handling Non-normal and Heterogeneous Error Term. *Asian Journal of Scientific Research*, **12**(3); 346–351

Yanuar, F., A. Zetra, C. Muharisa, D. Devianto, A. R. Putri, and Y. Asdi (2019d). Bayesian Quantile Regression Method to Construct the Low Birth Weight Model. In *Journal of Physics: Conference Series*, volume 1245. IOP Publishing, page 012044

