International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 09 Issue 12 ǁ 2021 ǁ PP. 43-49

Bayesian quantile regression based on ALD

WEI Xue-yong[1]，JINLiang-qiong[*1]，SHEN-ting[1]，ZHU Wei-yecheng[1]
[1]*School of Data Science and Information Engineering, Guizhou Minzu University, Guiyang 550025, China*
Corresponding Author: LiangqiongJin

Abstract
This article assumes that the random error term obeys the asymmetric Laplace distribution, and realizes the quantile regression of the likelihood function of the asymmetric Laplace distribution specified by the negative binomial distribution for Bayesian analysis, and uses the Gibbs sampling algorithm to obtain the parameters. The statistical properties of the posterior distribution are compared and the influence of whether the scale parameter is parameterized on the statistical properties of the model estimated coefficients is compared. The experimental results show that the statistical properties of the estimator obtained after the scale parameter is parameterized are better.

Keywords: *Quantile regression, asymmetric Laplace distribution, scale parameter, negative binomial distribution, Gibbs sampling.*

Date of Submission: 08-12-2021 Date of acceptance: 23-12-2021

I. INTRODUCTION

Since the least-squares estimation was proposed, it has been widely used in the field of social economy, but sometimes it can't fully describe the relationship between dependent variables and independent variables. Therefore, quantile regression was proposed to supplement the deficiency of focusing only on conditional mean estimation based on least-squares estimation. Koenker first proposed the concept of "Quantile Regression" in 1978 [1]; Koenker and Machado studied the relationship between quantile and asymmetric Laplace distribution in 1999 [2]; Yu used Bayesian quantile regression based on asymmetric Laplace distribution to estimate parameters in 2001 [3]. At the same time, it is confirmed that even if the prior distribution is not appropriate, the posterior distribution obtained is also appropriate; with the development of the times, the combination of quantile regression and Bayesian analysis method is applied to more and more fields. In 2005, Yu et al. studied the distribution of wages in Britain based on this method [4].

In previous studies, researchers have always set the scale parameter of asymmetric Laplace distribution as 1. In 2009, Wang Xinyu and other researchers proved that the scale parameter should be parameterized [5].In 2012, Zeng Ping and others studied the Bayesian analysis of non-standard distribution in WinBUGS software [6]. Based on this idea, this paper realizes the parameter estimation of Bayesian quantile regression of asymmetric Laplace distribution specified by negative binomial distribution in WinBUGS.

II. ASYMMETRIC LAPLACE DISTRIBUTION (ALD)

When using the Bayesian analysis method to estimate the quantile regression model, to make the model more robust, it is assumed that the random error term obeys ALD, so the maximum likelihood function is obtained according to its probability density function. The posterior distribution of parameters can be obtained from the prior distribution of parameters by the Bayesian theorem.

2.1 LAPLACE DISTRIBUTION (LD)

Since the probability density function of Laplace distribution (LD) is composed of two exponential functions, it is also called double exponential distribution. The probability density function of LD is

$$f(y;\mu,\sigma)=\frac{1}{2\sigma}\exp\left\{-\frac{|y-\mu|}{\sigma}\right\} \tag{2.1}$$

It is said that the random variable y obeys the Laplace distribution, the mean is μ , and the variance is $2\sigma^2$. Among them, the location parameter $-\infty<\mu<+\infty$ and the scale parameter $\sigma>0$; this distribution is also called the one-variable Laplace distribution, and the normal distribution is the binary Laplace distribution.

2.2 PROBABILITY DENSITY FUNCTION OF ASYMMETRIC LAPLACE DISTRIBUTION

The probability density function of ALD is

$$f(y;\mu,\sigma,q)=\frac{q(1-q)}{\sigma}\exp\left\{-\frac{y-\mu}{\sigma}\left[q-I\left(y\le\mu\right)\right]\right\} \qquad y\in(-\infty,+\infty) \tag{2.2}$$

where the location parameter $-\infty<\mu<+\infty$, the scale parameter $\sigma>0$, the skewness parameter $0<q<1$, and the loss function is $\rho_q\left(\frac{y-\mu}{\sigma}\right)=\frac{1}{2}\left[\left|\frac{y-\mu}{\sigma}\right|+(2q-1)\frac{y-\mu}{\sigma}\right]$.

The density function diagram of ALD is shown in figure 1. When the position parameter μ and the scale parameter σ are fixed, with different values of the skewness parameter q, the density function diagram of ALD is also different; when $q=0.2$, the density diagram of ALD is biased to the right. When $q=0.5$, the density diagram of ALD is symmetrical on the left and right sides of the position parameter. When $q=0.8$, the density diagram of ALD is biased to the left.

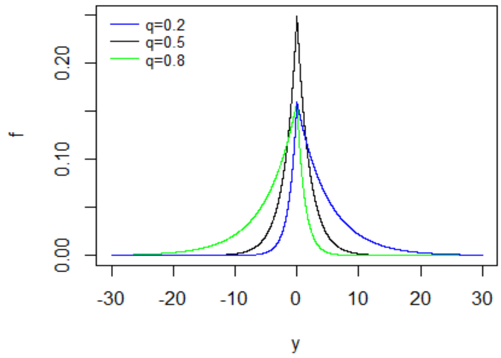

Figure 1:Density function of asymmetric Laplace distribution (ALD) ($\mu=0,\sigma=1,q$)

III. BAYESIAN QUANTILE REGRESSION

When using the Bayesian analysis method to estimate the quantile regression model, we first need to build a linear regression modeland then assume that the random error term obeys ALD. Then, based on the Bayesian method, we use the Markov chain Monte Carlo (MCMC) method to sample to obtain the posterior distribution of parameters. Finally, we test the estimation effect of parameters through data simulation analysis.

3.1 LINEAR MODEL

The linear regression model is:

$$y=\beta_0+\beta_1 x+u, \tag{3.1}$$

where let $Z(x;\beta)=\beta_0+\beta_1 x$, so $y=Z(x;\beta)+u$, where y is the dependent variable, x is the independent variable, β_0 is the constant parameter to be estimated, β_1 is the coefficient parameter to be estimated, and the random error term $u\sim ALD(0,\sigma,q)$, that is, $y\sim ALD(Z(x;\beta),\sigma,q)$.

3.2 ASYMMETRIC LAPLACE DISTRIBUTION LIKELIHOOD FUNCTION

The likelihood function of ALD probability density function is:

$$\mathrm{L}(y_i;Z(x_i;\beta),\sigma,q)=\frac{q^n(1-q)^n}{\sigma^n}\exp\left\{-\frac{1}{\sigma}\sum_{i=1}^{n}\rho_q(u_i)\right\} \qquad y_i\in(-\infty,+\infty), \tag{3.2}$$

where $\rho_q(u_i)=u_i\left[q-I(u_i\le 0)\right]$ is loss function,and $u_i=y_i-Z(x_i;\beta)$.The core idea of Bayesian quantile regression based on asymmetric Laplacian distribution is: transform the optimal solution problem of quantile regression into solving an asymmetric Laplace distribution probability density maximization of the likelihood function.

3.3 SPECIFYING ASYMMETRIC LAPLACE DISTRIBUTION WITH NEGATIVE BINOMIAL DISTRIBUTION

WinBUGS (Bayesian inference Using Gibbs Sampling, BUGS)is a software dedicated to Bayesian statistical analysis. WinBUGS is a version under the Windows operating system. The biggest advantage of this software is that it is free for users and flexible in operation, so it is widely used in Bayesian analysis; but in

WinBUGS software only more than 20 standard distributions are provided, and the non-standard distribution of ALD involved in this article is not included. Therefore, it is necessary to specify ALD with a negative binomial distribution. First, compile the program that specifies the likelihood function of ALD in the WinBUGS software for negative binomial distribution; then use the "bugs()" function in the Rstudio software to call the WinBUGS compiled and saved program, and perform Gibbs sampling based on the WinBUGS software to achieve Simulation of Bayesian quantile regression.

In the Bernoulli experiment sequence, the probability of occurrence of event A in each experiment is p. If X is the number of experiments when event A occurs for the r th time, then the possible value of X is $r, r+1, \cdots, r+m, \cdots$, and X is said to obey negative binomial distribution or Baska distribution, its distribution is listed as follows:

$$f(X=k) = C_{k-1}^{r-1} p^r (1-p)^{k-r} \qquad k = r, r+1, \cdots, \tag{3.3}$$

where $X \sim Nb(r, p)$, mathematical expectation as $\frac{1}{p} r$, and variance as $\frac{1}{p^2} r(1-p)$ [7].

3.3.2 The negative binomial distribution specifies the asymmetric Laplace distribution

Suppose the logarithmic density function of the data is $l_i = \log f(y_i; Z(x_i;\beta), \sigma, q)$, the probability density function is $f(y_i; Z(x_i;\beta), \sigma, q) = \exp(l_i)$, and the negative binomial distribution data are $h_i = 0, i = 1, 2, \cdots, n$, then the likelihood function is $L(y_i; Z(x_i;\beta), \sigma, q) = \prod_{i=1}^{n} \exp(l_i)$, the parameter

$$p_i = \exp(l_i - c), \tag{3.4}$$

in order to ensure $p_i \in [0,1]$, so a larger $c = 600$ is reduced, where $l_i = \log(q(1-q)) - \log\sigma - \frac{1}{2\sigma}\left[|u_i| + (2q-1)u_i\right]$, $u_i = y_i - Z(x_i;\beta)$, then the likelihood function is:

$$\begin{aligned} L(y_i; Z(x_i;\beta), \sigma, q) &\propto \prod_{i=1}^{n} \frac{(0+1-1)}{0!(1-1)!}\left[\exp(l_i - c)\right]^1 \left[1 - \exp(l_i - c)\right]^0 \\ &\propto \prod_{i=1}^{n} \exp(l_i - c) \end{aligned} \tag{3.5}$$

IV. NUMERICAL SIMULATION ANALYSIS

The model for generating random data is

$$Y = 1 + 2X + \varepsilon, \tag{4.1}$$

where $X \sim U(0,10)$ and $\varepsilon \sim ALD(0,1,q)$ use the generated random data to simulate model(3.1).

4.1 THE SCALE PARAMETER σ IS SET TO 1

When the scale parameter σ is set to 1, and the parameters to be estimated are β_0 and β_1, it is assumed that the prior distribution of the parameters to be estimated obeys the normal distribution. According to the Bayesian theorem, the joint posterior distribution of the parameters is

$$P(y_i; \beta, \sigma) \propto L(y_i; Z(x_i;\beta), 1, q)\, f(\beta), \tag{4.2}$$

where $f(\beta)$ is the prior distribution of the coefficient β to be estimated, so

$$l_i = \log(q(1-q)) - \frac{1}{2}\left[|u_i| + (2q-1)u_i\right], \tag{4.3}$$

substituting formula (4.3) into formula (3.4), the expression of parameter p_i can be obtained:

$$\begin{aligned} p_i &= \exp(l_i - \mathrm{c}) \\ &= \exp\left\{\log\left[q(1-q)\right] - \frac{1}{2}\left[|u_i| + (2q-1)u_i\right] - 600\right\}, \\ &= q(1-q)\exp\{-\rho_q(u_i) - 600\} \end{aligned} \tag{4.4}$$

substituting formula (4.4) into formula (3.5), the joint posterior distribution of the parameters can be expressed as:

$$P\left(y_i;\beta,\sigma\right)\propto L\left(y_i;Z\left(x_i;\beta\right),1,q\right)f\left(\beta\right)$$
$$\propto q^n(1-q)^n\exp\{-nc\}\exp\left\{-\sum_{i=1}^{n}\rho_q(u_i)\right\}. \tag{4.5}$$

Figure 2 shows the likelihood function of ALD specified by the negative binomial distribution, the prior distribution is set to $\beta\sim N(0,100)$, the parameters β_0 and β_1 are at the 0.75 quantile and the sample size is 75, the iteration is 10000 times, and the burn-in period is 5000 times. Sampling trajectory graph, density graph and autocorrelation graph, the results show that the 5000 sampling values of parameter parameters β_0 and β_1 fluctuate up and down the set value, and the autocorrelation coefficient tends to zero as the lag period increases. Therefore, the Markov chains formed by sampling all converge.

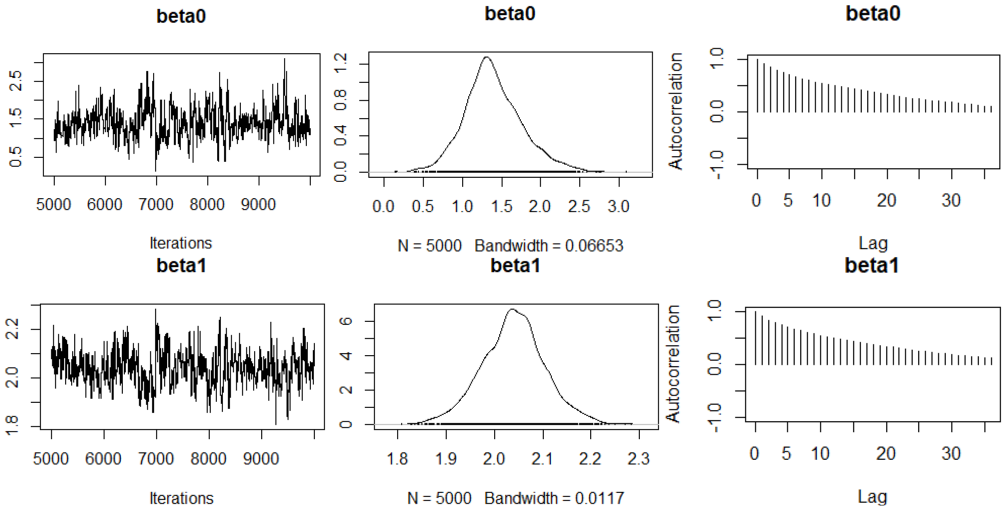

Figure 2:Trajectory graph, density graph, and autocorrelation graph of β_0 and β_1 sampled values

Table 1:The mean, standard deviations, and MC errors of parameters β_0 and β_1 when σ is set to 1.

prior information	sample size	quantile	0.25			0.5			0.75		
		parameter	mean	sd	MC errors	mean	sd	MC errors	mean	sd	MC errors
$\beta\sim N(0,100)$	25	β_0	-0.35840	0.63150	0.04100	0.18450	0.48930	0.05223	1.79900	0.69480	0.05175
		β_1	2.08100	0.09703	0.00629	2.09700	0.07615	0.00811	1.98100	0.10740	0.00796
	75	β_0	-0.32360	0.55320	0.04079	44.99000	10.32000	1.22800	1.39600	0.37280	0.02616
		β_1	1.99400	0.08611	0.00640	-34.29000	1.13100	0.13460	2.03900	0.06557	0.00462
	100	β_0	-0.30170	0.43700	0.03132	84.40000	12.04000	1.43300	18.37000	8.46500	0.95450
		β_1	1.91100	0.08437	0.00593	-40.34000	1.37400	0.16360	-30.82000	1.36600	0.15430
$\beta\sim N(0,10)$	25	β_0	-0.29870	0.65390	0.04001	0.21630	0.53060	0.04218	1.77600	0.65420	0.04501
		β_1	2.07000	0.10390	0.00630	2.09200	0.08788	0.00705	1.98300	0.10100	0.00693
	75	β_0	-0.32150	0.52270	0.03719	0.58650	0.33750	0.02382	1.37800	0.38330	0.02871
		β_1	1.99300	0.08427	0.00607	2.02700	0.05553	0.00397	2.04300	0.06915	0.00511
	100	β_0	-0.22660	0.39260	0.02421	0.55010	0.32010	0.02214	1.31600	0.28910	0.01687
		β_1	1.90000	0.07868	0.00454	1.99000	0.06267	0.00430	2.01600	0.05881	0.00330
$\beta\sim N(0,1)$	25	β_0	-0.11460	0.54310	0.03086	0.23480	0.46300	0.03398	1.31800	0.55460	0.03778
		β_1	2.04200	0.09239	0.00530	2.08500	0.07854	0.00577	2.03900	0.09213	0.00603
	75	β_0	-0.17590	0.47580	0.03267	0.57610	0.32780	0.02268	1.28700	0.32700	0.02252
		β_1	1.96800	0.07960	0.00542	2.02600	0.05278	0.00357	2.05400	0.06142	0.00417
	100	β_0	-0.17960	0.37620	0.02335	0.53360	0.31380	0.02143	1.22900	0.26450	0.01510
		β_1	1.88800	0.07653	0.00455	1.99100	0.06314	0.00419	2.02700	0.05690	0.00324

The above table uses the Gibbs sampling algorithm to sample the model parameters and the number of samples is 10,000. To eliminate the influence of the initial value of the parameters on the sampling distribution, the first 5000 sampling values are removed. Finally, the mean, standard deviations, and MC errors of the posterior distribution of the parameters can be obtained, as shown in table 1:

(1)When the prior distribution and quantile are the same, as the sample size n increases, the standard deviations and MC errors of the parameters β_0 and β_1 to be estimated gradually decrease.

(2)Under the same condition of sample size and quantile, with the enhancement of the prior distribution, the standard deviation and MC error of the parameters β_0 and β_1 will be smaller, that is, when the prior information is enhanced, the estimation accuracy of the parameters can be improved.

(3) When the prior information, sample size, and quantile are the same, the standard deviation and MC errors of β_1 are always smaller than the standard deviation and MC errors of β_0, that is, the estimation accuracy of the coefficient term obtained in the WinBUGS software is often higher than that of the constant term estimated accuracy.

(4)In these three prior settings, when the sample size and prior information are the same, except that the prior information is $\beta \sim N(0,100)$ and the sample size is 100, the standard deviations and MC errors ratio of parameters β_0 and β_1 at the 0.25 quantile are in the standard deviations and MC errors at 0.5 quantile and 0.75 quantile are small. Under other conditions, the standard deviations and MC errors of parameters β_0 and β_1 at 0.25 quantile are both at 0.5 quantile and 0.75 quantile. The lower standard deviations and MC errors are large, that is, the estimation accuracy of the parameters β_0 and β_1 in the median and high quantile are higher.

4.2 PARAMETERIZE THE SCALE PARAMETER σ

When the scale parameter in ALD is set to 1, although the estimated accuracy of the parameter is high, it is not appropriate in practical applications. Therefore, the scale parameter in ALD should be parameterized [5].

When σ, β_0, and β_1 are all to be estimated, assuming that the prior distribution of the parameters β_0 and β_1 to be estimated is a normal distribution, the prior distribution of the scale parameter σ is a chi-square distribution, and the degree of freedom is 4, the smoothness of the three parameters to be estimated is higher.

when the three parameters to be estimated have higher smoothness. According to the Bayesian theorem, the joint posterior distribution of the parameters is

$$P(y_i;\beta,\sigma) \propto L(y_i;Z(x_i;\beta),\sigma,q)\, f(\beta)\, g(\sigma), \tag{4.6}$$

where $f(\beta)$ is the prior distribution of the coefficient β to be estimated, $g(\sigma)$ is the prior distribution of the scale parameter σ, and $\sigma \sim \chi^2(4)$,

$$l_i = \log[q(1-q)] - \log\sigma - \frac{1}{2}\left[\left|\frac{u_i}{\sigma}\right| + (2q-1)\frac{u_i}{\sigma}\right], \tag{4.7}$$

substituting formula (4.7) into formula (3.4), the expression of parameter p_i can be obtained as:

$$\begin{aligned} p_i &= \exp(l_i - \mathrm{c}) \\ &= \exp\left\{\log[q(1-q)] - \log\sigma - \frac{1}{2}\left[\left|\frac{u_i}{\sigma}\right| + (2q-1)\frac{u_i}{\sigma}\right] - 600\right\} \\ &= \frac{q(1-q)}{\sigma}\exp\left\{-\rho_q\left(\frac{u_i}{\sigma}\right) - 600\right\}, \end{aligned} \tag{4.8}$$

substituting formula (4.8) into formula (3.5), the joint posterior distribution of the parameters can be expressed as:

$$\begin{aligned} P(y_i;\beta,\sigma) &\propto L(y_i;Z(x_i;\beta),\sigma,q)\, f(\beta)\, g(\sigma) \\ &\propto \frac{q^n(1-q)^n}{\sigma^n}\exp\{-nc\}\exp\left\{-\frac{1}{\sigma}\sum_{i=1}^{n}\rho_q(u_i)\right\} \end{aligned} \tag{4.9}$$

Figure 3 shows the likelihood function of the ALD specified by the negative binomial distribution, and the prior distribution is set to $\beta \sim N(0,100)$, the parameters σ, β_0, and β_1 are at the 0.75 quantile, the sample size is 75, the variance is 100, iteration is 10,000 times, and the burn-in is the trajectory graph, density graph and autocorrelation graph of Gibbs sampling values with a period of 5000 times. The results show that the 5000 sampling values of parameters σ, β_0, and β_1 fluctuate up and down the set value, and the autocorrelation coefficient tends to zero with the increase of the lag period, so the Markov chain is all convergesby sampling.

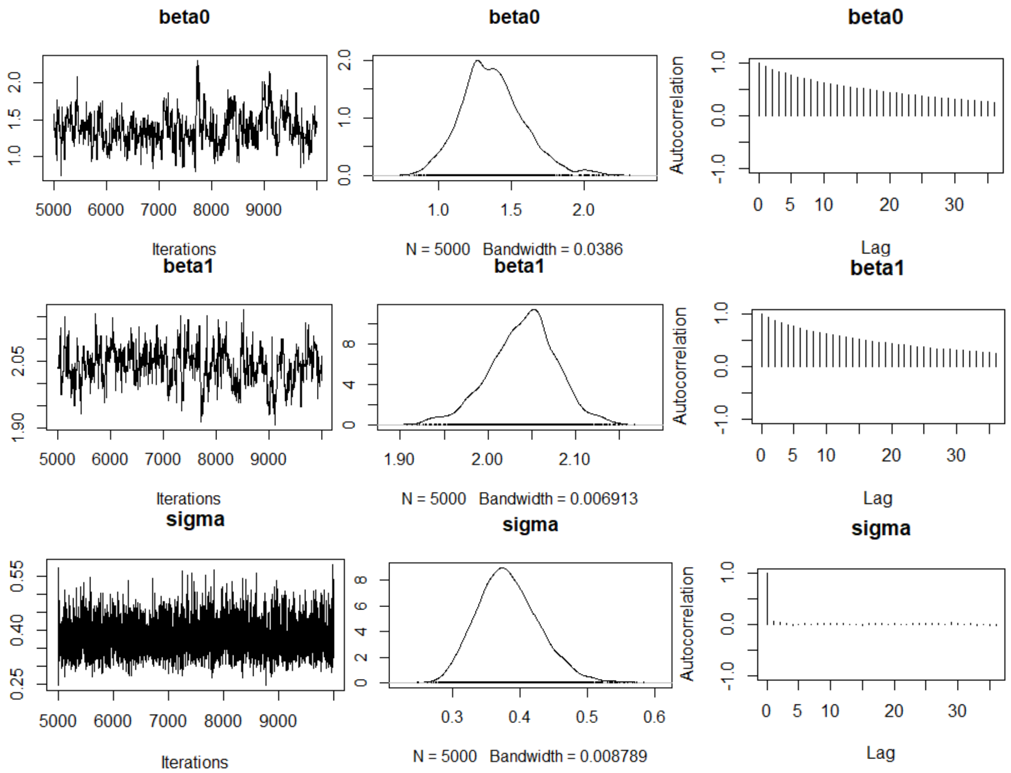

Figure 3:.The trajectory graph, density graph, and autocorrelation graph of the sampled values of β_0, β_1, and σ

Table 2:The mean, standard deviations, and MC errors when parameters β_0, β_1, and σ are all to be estimated

prior information	sample size	quantile	0.25			0.5			0.75		
		parameter	mean	sd	MC errors	mean	sd	MC errors	mean	sd	MC errors
$\beta \sim N(0,100)$	25	β_0	-0.33670	0.40060	0.03062	0.35630	0.20250	0.01542	1.83000	0.36700	0.02749
		β_1	2.09500	0.05705	0.00431	2.06900	0.03493	0.00265	1.97100	0.05173	0.00382
		σ	0.26650	0.05981	0.00097	0.27850	0.06252	0.00135	0.28530	0.06190	0.00100
	75	β_0	-0.30430	0.37670	0.02906	0.67030	0.25230	0.01818	1.37000	0.21980	0.01712
		β_1	1.99400	0.05685	0.00435	2.01600	0.03935	0.00282	2.04000	0.03849	0.00303
		σ	0.36110	0.04300	0.00069	0.44200	0.05199	0.00077	0.38260	0.04594	0.00085
	100	β_0	-0.22370	0.30070	0.02114	0.59590	0.24460	0.01644	1.29900	0.17460	0.00969
		β_1	1.90100	0.05651	0.00377	1.98500	0.04770	0.00317	2.01300	0.03728	0.00209
		σ	0.43600	0.04554	0.00068	0.50060	0.05226	0.00081	0.41040	0.04221	0.00051
$\beta \sim N(0,10)$	25	β_0	-0.32930	0.40150	0.02860	0.37600	0.20600	0.01616	1.74600	0.42000	0.03560
		β_1	2.09300	0.05813	0.00405	2.06500	0.03475	0.00270	1.98100	0.06092	0.00512
		σ	0.26570	0.05918	0.00102	0.27930	0.06258	0.00118	0.28690	0.06471	0.00131
	75	β_0	-0.23050	0.36190	0.02608	0.60700	0.22840	0.01558	1.37600	0.21680	0.01600
		β_1	1.98200	0.05572	0.00397	2.02500	0.03634	0.00251	2.03800	0.03760	0.00277
		σ	0.36020	0.04188	0.00067	0.44150	0.05192	0.00082	0.38270	0.04496	0.00070
	100	β_0	-0.19130	0.28030	0.01765	0.57280	0.25850	0.01950	1.30100	0.18850	0.00996

		β_1	1.89500	0.05317	0.00330	1.98800	0.05010	0.00375	2.01300	0.04022	0.00216
		σ	0.43480	0.04560	0.00073	0.50180	0.05292	0.00076	0.41110	0.04203	0.00069
		β_0	-0.19100	0.39360	0.02783	0.35350	0.19120	0.01292	1.64100	0.36190	0.02686
	25	β_1	2.07400	0.05814	0.00418	2.06800	0.03273	0.00219	1.99300	0.05299	0.00392
		σ	0.26610	0.05888	0.00123	0.27820	0.06085	0.00122	0.28750	0.06408	0.00123
		β_0	-0.23810	0.33500	0.02487	0.58420	0.22110	0.01404	1.31800	0.21150	0.01673
$\beta \sim N(0,1)$	75	β_1	1.98200	0.05384	0.00402	2.02700	0.03471	0.00226	2.04700	0.03851	0.00306
		σ	0.36030	0.04248	0.00064	0.44260	0.05299	0.00071	0.38260	0.04450	0.00061
		β_0	-0.16190	0.26430	0.01769	0.59740	0.25150	0.02053	1.26700	0.17340	0.01124
	100	β_1	1.88800	0.05178	0.00335	1.98200	0.04916	0.00402	2.01900	0.03739	0.00236
		σ	0.43530	0.04439	0.00061	0.50080	0.05209	0.00082	0.40940	0.04283	0.00064

The above table uses the Gibbs sampling algorithm to sample the model parameters and the number of samples is 10,000. To eliminate the influence of the initial value of the parameter on the sampling distribution, the first 5000 sampling values are removed. Finally, the mean, standard deviations and MC errors of the posterior distribution of the parameters can be obtained, as shown in table 2:

(1)When the prior distribution and quantile are the same, as the sample size increases, the standard deviations and MC errors of parameters σ , β_0 , and β_1 gradually become smaller.

(2)Under the same conditions of prior information, sample size, and quantile, for the size of the standard deviations: $\beta_1 < \sigma < \beta_0$, but for the size of the MC errors: $\sigma < \beta_1 < \beta_0$.

(3)Under the same condition of sample size and quantile, with the enhancement of the prior distribution, the standard deviations and MC errors of parameters σ , β_0 , and β_1 are smaller, that is, when the prior information is enhanced, the estimation accuracy of the parameters can be improved.

(4)When the prior information is the same and the sample size is 100, the standard deviations and MC errors of the parameters σ , β_0 , and β_1 at the 0.75 quantile are smaller than the standard deviations and the MC errors at the 0.25 quantile and 0.5 quantile.

Comparing table 1 and table 2, we can see that when the prior information, sample size, parameters and quantile are the same, the standard deviations and MC errors of parameters β_0 and β_1 in table 2 are smaller than those of parameters β_0 and β_1 in table 1 MC errors. Therefore, parameterizing the scale parameter can improve the estimation accuracy of parameters β_0 and β_1 , that is, the scale parameter should be parameterized instead of being set to 1.

V. CONCLUSION

In this paper, in the Bayesian quantile regression analysis, the relevant statistical properties of the parameter posterior distribution are obtained through the Gibbs sampling algorithm. The experimental results show that the likelihood functionof ALD is specified based on the negative binomial distribution. After the scale parameter is parameterized, the statistical property of the estimator of the posterior distribution of the parameter is better than the statistical property of the estimator when the scale parameter is assumed to be 1, so the scale parameter should be parameterized; even if the prior distribution is improper, the obtained posterior distribution is also suitable; a proper prior distribution can improve the accuracy of Gibbs sampling estimates.

REFERENCES

[1]. Koenker R, Bassett G W. (1978)"Regression quantiles" J. Econometrica, Vol. 46,No.1, pp. 33 - 50.
[2]. Roger Koenker,José A. F. Machado. (1999)"Goodness of Fit and Related Inference Processes for Quantile Regression" J. Journal of the American Statistical Association,Vol. 94,No.448,pp. 1296 - 1310.
[3]. Keming Yu,Rana A. Moyeed (2001)"Bayesian quantile regression" J. Statistics and Probability Letters, Vol. 54,No.4, 437 - 447.
[4]. Yu Keming,van Kerm Philippe,Zhang Jin (2005),"Bayesian Quantile Regression: An Application to the Wage Distribution in 1990s Britain"J. Sankhyā: The Indian Journal of Statistics, Vol. 67,No.2, pp.359 - 377.
[5]. Xinyu Wang, Xuefeng Song (2009),"Measure Market Risk Based on Bayesian QuantileRegression Model with an Application" J.Journal of Systems Management, Vol. 18,No.1, pp. 40-48.
[6]. Ping Zeng, Ting Wang, Peng He (2012),"WinBUGS implementation of Nonstandard distribution Bayesian analysis" J.China Health Statistics, Vol. 29,No.4, pp. 614–615+617.
[7]. Sisong Mao, JinglongWang, XiaolongPu (2006), "Advanced Mathematical Statistics" M. Higher Education Press.