New Bayesian Single Index Quantile Regression Based on Uniform Scale Mixture

Samer Alalaqa, Taha Alshaybawee, b
a General Directorate of Education, Misan, Iraq
b Department of Statistics, College of Administration and Economics, University of Al-Qadisiyah, Al Diwaniyah, Iraq
E-mail addresses: taha.alshaybawee@qu.edu.iq.
sam371979371979@qmail.com

Abstract:
To scale back the dimensionality while holding a lot of flexibility of a nonparametric model Wu, et al. (2010) proposed a single index conditional quantile regression model. In this paper, a new Bayesian lasso for single index quantile regression model is proposed based on a scale mixture uniform. In addition, we construct an efficient and sampling Gibbs algorithm for posterior inference based on a uniform scale mixture representation for Laplace distribution. Simulation study have considered to evaluate our proposed method compare to the existing methods. The results of simulations indicate that the new Bayesian algorithm performs well.

Keywords: Bayesian inference, single index, quantile regression, prior distribution, uniform scale mixture.

1. Introduction

In real applications, the exact response variable cannot be predicted from the prediction variables. Classical regression focuses on predicting a modal response variable Y for a single or set of predictor variables X, prediction of the response variable is given by prediction variables E(Y|X), called regression function (Gujarati 2003; Weisberg 2005). Nevertheless, many researchers claim that the fitting of the conditional mean is unable to give a complete picture of underlying interrelationships. For example, Mosteller and Tukey (1977) indicate “… The regression curve gives a large summary of the averages corresponding distributions to the set x”s. We could go further and compute several different regression curves corresponding to the various percentage points of the distributions and thus get a
more complete picture of the set. Ordinarily, this is not done, and so regression often gives a rather incomplete picture. Just as the mean gives an incomplete picture of a single distribution, so the regression curve gives a correspondingly incomplete picture for a set of distributions”. Quantile regression (QR) suggested by Koenker and Basset (1978), allowing the person to thoroughly review the conditional distribution of $Y$ over $X$ in completely different locations, thus giving a complete view of the interrelationships between $Y$ and $X$. Subsequently, quantile regression has become a unite statistical methodology for estimating models of conditional quantile functions Wu & Liu (2009). QR usually provides a more complete picture description of the distribution of the response over the mean, it offers a practically necessary alternative to classical mean regression. Quantile regression is additionally a valuable alternative to the usually utilized Cox proportional risk model and the accelerated failure time model (AFT) utilized in survival analysis Koenker, R., & Geling, O. (2001). For instance, in a cancer study, it's known that treatment can cause different impacts among patients in lower or higher quantiles of the survival distribution Wu, et al. (2010). Nonparametric quantile regression was studied by Koenker (2005). Kernel primarily based methods and spline smoothing ways, have been used to accommodate nonlinear relationships; see for more details, Yu and Jones (1998), He and Shi (1994) and Koenker et al. (1994). There are problems in the non-parameter estimation of the existence of high dimensions that can be solved using the single index model and avert the “curse of dimensionality” (Bellman et al., 1966) in nonparametric problems by that the response is just associated with a single linear set of the covariates. In addition, the main problem of non-parameter quantile regression that the estimated function is often hard to see and interpret with multivariable variables. To scale back the dimensionality while holding a lot of flexibility of a nonparametric model Wu, et al. (2010) proposed a single index conditional quantile regression model. They are the most common and necessary models in statistics as well as applied quantitative sciences such as econometrics and psychology due to their ability to reduce dimensions (Ichimura, 1993). Whereas this model is conditional on the quantile regression so it becomes more immune to outliers and includes a heavier tail errors distribution than the mean regression. The mathematical type for the $\tau^{th}$ quantile level SIQR model is shows by:

$$y = m_\tau(x_\tau^T \beta_\tau) + \varepsilon_\tau$$

(1)

where $\tau \in (0,1)$, $y$ is real response variable, $x$ is that the p-dimensional vector of explanatory variables $\beta_\tau$, is that the index parameter vector for the $\tau^{th}$
quantile, and \( g_\tau(.) \) is the unknown nonparametric link function for the \( \tau^{th} \) quantile, that is estimated for a single variable \((x_i^T \beta_\tau)\).

The algorithm for estimating single index vector, based on linear quantile regression and with the unknown nonparametric function estimated by the local linear method was constructed by Wu et al. (2010). Computational algorithm for estimating and variable selection was advanced by Alkenani, Yu (2013), Lv et al. (2014) and Kuruwita (2016). This model has been given a great deal of interact from many researchers on both the applied and theoretical side. Some of the drawbacks experienced the frequentist approach are mentioned by Benoit and Van den Poel (2012); the difficult optimization of the estimated parameters and the structure of the confidence interval. Hu et al. (2013) were developed new Bayesian hierarchical model to estimate nonparametric link function in the single index model (SIM) conditional to QR. The Bayesian analysis method has become vastly applied, outcome of its ability to interest from all ready information in the analysis. Many studies newly presented to develop and expand the single index quantile regression model, for instance; Completely Bayesian method was sophisticated by Hu et al. (2013). A similar Bayesian approach was utilized by Zhao and Lian (2013) for the censored data in the single index quantile regression model. Alshaybawee et al. (2017) proposed a Bayesian approach with elastic net penalty to estimate and selected variables at the single index quantile regression model. In this paper, a new Bayesian lasso for single index quantile regression model is proposed based on a scale mixture uniform to the prior parameters. In addition, we construct an efficient and simple Gibbs sampling algorithm for posterior inference based on a uniform scale mixture representation for Laplace distribution.

The reset of this paper is orderly as follows: In Section 2, we briefly review the single index model and the prior distribution. The Bayesian hierarchical model and the posterior inference for updating the parameters is described in section 3. Simulation study include in section 4 and some conclusions are given in section 5.

2. Single index quantile regression model and prior assumptions

The model of single index quantile regression was given as:-

\[ y_i = m_\tau(x_i^T \beta_\tau) + \epsilon_\tau \quad i = 1,2,\ldots,n \quad (2) \]

At \( \tau^{th} \) quantile, the quantile error term \( \epsilon_\tau \) distributed as an asymmetric Laplace distribution with the density:

\[ \pi(\epsilon | \sigma, \tau) = \frac{\tau(1-\tau)}{\sigma} \exp \{ -\frac{1}{\sigma} \rho_\tau(\epsilon) \} \quad (3) \]

where \( \sigma \) is the scale parameter, \( \rho_\tau(.) \) is the check loss function defined by \( \rho_\tau(\epsilon) = \epsilon(\tau - I(\epsilon < 0)) \) for quantile level \( \tau \in (0,1) \).

The likelihood function for
$$
\epsilon_i = y_i - m_\tau(x_i^T \beta), \quad i = 1, 2, ..., n.
$$

where $n$ is a sample size, shown as:

$$
\pi(y|m, \beta, \sigma)
= \frac{\tau^n(1 - \tau)^n}{\sigma} \exp \left\{ - \frac{1}{\sigma} \sum_{i=1}^{n} \rho_\tau(y_i - m(x_i^T \beta)) \right\}
$$

Quantile regression is often based on minimizing of the check function. However, direct use of the likelihood function is not convenient for Bayesian inference. A location-scale mixture (Kozumi and Kobayashi 2011) have used to the ALD which is useful here. We can rewrite Eq. (2) based on location – scale representation as follows:

$$
y_i = m(x_i^T \beta) + \delta v_i + \sqrt{\theta \sigma v_i} \quad u_i
$$

where $v_i$ is a standard exponential random variable, $u_i$ is a standard normal random, \(\delta = \frac{1 - 2\tau}{\tau(1 - \tau)}\) and \(\theta = \frac{2}{\tau(1 - \tau)}\). Therefore, the likelihood function are rewritten as Kozumi, H., & Kobayashi, G. (2011).

$$
\pi(y_i|\beta, m_\tau, v_i)
= \prod_{i=1}^{n} \left(2\pi \sigma \theta v_i \right)^{-\frac{1}{2}} \exp\left\{ - \frac{1}{2 \sigma \theta v_i} (y_i - m_\tau(x_i^T \beta) - \delta v_i)^2 \right\}
$$

$$
\alpha \left\{ (\text{det}[D^{-1}])^{-\frac{1}{2}} \exp\left\{ - \frac{(y-m_n-\delta v_n)^T D^{-1} (y-m_n-\delta v_n)}{2} \right\} \right\}
$$

$$
y_i = (y_1, y_2, ..., y_n)^T, v_n = (v_1, v_2, ..., v_n)^T
$$

$$
m_n = (m_1, m_2, ..., m_n)^T = (m(x_1^T \beta), m(x_2^T \beta), ..., m(x_n^T \beta))^T
$$

$$
D^{-1} = \delta \sigma \text{diag}(v_1, v_2, ..., v_n), \text{and}
$$

As in Choi et al. (2011) and Gramacy and Lian (2012), the Gaussian method distribution is put as a prior distribution for the unknown nonparametric link function $m_\tau(.)$. More specially, the prior distribution of $m_\tau(.)$ is a Gaussian process, with the mean zero and the square covariance exponential function as follows:

$$
m_\tau(.) \sim GP(0, E(\ldots)), E(x_i, x_j)
= \vartheta \exp(-\frac{(x_i - x_j)^2}{\omega})
$$

Where $\vartheta$ and $\omega$ are hyperparameters.

Therefore, The prior distribution for unknown function can be shown as follows:

$$
\pi(m_\tau|\beta, \vartheta)
= \text{det}[E_n]^{-\frac{1}{2}} \exp\left\{ - \frac{m_n^T E_n^{-1} m_n}{2} \right\}
$$

where $E_n$ is the covariance matrix with dimension $(n \times n)$ and it can be written as

$$
E(x_i^T \beta, x_j^T \beta) = \vartheta \exp \left\{ - (x_i - x_j)^T \beta \beta^T (x_i - x_j) \right\}.
$$

We have seen that the variance function $E(x_i, x_j) = \vartheta \exp(-\frac{(x_i^T \beta - x_j^T \beta)^2}{\omega})$, follow Gramacy and
Lian (2012), when the Gaussian process is employed as the prior for the nonparametric function, \( \frac{\beta}{\sqrt{\omega}} \) is identifiable without the necessity for the constraint \( ||\beta|| = 1 \). Therefore, the parameter \( \beta \) will be instead of \( \frac{\beta}{\sqrt{\omega}} \).

\[
E(x_i, x_j) = \partial \exp \left\{ -(x_i^T \beta - x_j^T \beta)^2 \right\}
\]

The inverse gamma distribution is considered as a hyper prior for \( \partial \), which implies that \( \partial \sim IG(a_\gamma, b_\gamma) \) where \( a_\gamma \) and \( b_\gamma \) are the hyperparameters.

Scale mixture of uniform distribution (SMUD) for regression models has been utilized in few of occasions in literature. Walker et al. (1997) utilized SMU distribution in natural regression models in non-Bayesian framework. Damien and Qin (2000), & Walker, S. G. (1997) provided Gibbs sampler by utilized SMU in variance regression models and additionally to derive Gibbs sampler for autocorrelated heteroscedastic regression models. Nevertheless, its use has been specified in penalized regression setting. Choy et al. (2009) consider it in random volatility model by employing a two-stage scale mixture representation of the t-student distribution. It is important to note that the potential density function of Laplace distribution can be expressed in different forms of mixture representation. This representation is much importance in research and studies, especially using the pyramidal method. This representation is important in facilitating the estimation process and increasing its efficiency. Some examples of this representation are important in this research. See for instance, Mallick and Yi (2014) they were developed the former methods by utilizing method called a new Bayesian Lasso, such suggested procedure for variable selection(VS), estimation of linear regression coefficient. Husseini (2017) proposed the Bayesian Lasso quantile regression method for (VS) to allocate an independent scale mixture of uniform distributions for regression coefficients. Alhamzawi & Ali (2018) offered a complete Bayesian method for tobit QR to reach the \( L_1 \) based on the representation of a mixture of the deviant Laplace distribution. A simple and effective Bayesian MCMC algorithm was introduced. The results of Mallick in his research were ideal and proved their competence from the previous Bayesian processes utilized.

Malik's results in his research were exemplary and proved efficient from previous Bayesian processes utilized. Hierarchical model and MCMC sampler

A Laplace distribution function can be written as a scale mixture of uniform
distribution and a specific gamma distribution Mallick, H., Yi, N., 2014, i.e.

\[
\pi(\beta_j | \lambda_j, \sigma) = \Pi_{j=1}^p \left( \frac{\lambda_j}{2\sigma} \right) \exp\left(-\left(\frac{\lambda_j | \beta_j |}{\sigma} \right)\right), \lambda > 0 \quad (6)
\]

\[
\frac{\gamma^2}{2} \exp\{- (\gamma | \beta_j, \tau)\} = \int_{s > |\beta|} \frac{1}{2s} \frac{\gamma^2}{\Gamma_2} s^{2-1} e^{-\gamma s} ds \quad (7)
\]

Based on the formulas above (6) and (7) we can rewrite the hierarchical representation model as follows:

\[
y | \beta, \sigma, m, \tau, v \sim N(m(x_i^T \beta) + \delta \nu, \theta \sigma u)
\]

\[
\beta | s \sim \Pi_{j=1}^p \text{uniform } (-\sqrt{s_j}, \sqrt{s_j}) \quad (8)
\]

\[
u_i \sim N(0, 1) \quad v_i \sim \frac{1}{\sigma} \exp\left(-\frac{v_i}{\sigma}\right)
\]

\[
s_j | \gamma \sim \Pi_{j=1}^p \Gamma(2, \gamma)
\]

\[
\gamma^2 \sim \Gamma(a_\gamma, b_\gamma)
\]

\[
\sigma \sim \text{Inv}.\Gamma(a_\sigma, b_\sigma)
\]

\[
\delta \sim \text{Inv}.\Gamma(a_\delta, b_\delta)
\]

For all parameters the posterior can be given as:

\[
p(m_n, \beta, v, s, \gamma, \delta, \sigma | y) \alpha \left\{ (\det[D^{-1}])^{-\frac{1}{2}} \exp\left\{-\frac{(y - m_n - \delta \nu)^T D^{-1} (y - m_n - \delta \nu)}{2}\right\} \times \det[E_n]^{-\frac{1}{2}} \exp\left\{-\frac{m_n^T E_n^{-1} m_n}{2}\right\} \times \Pi_{j=1}^p \frac{1}{2s} \frac{\gamma^2}{\Gamma_2} s^{2-1} e^{-\gamma s} \times \Pi_{j=1}^p \frac{1}{\sigma} \exp\left(-\frac{v_i}{\sigma}\right) \times \Pi_{j=1}^p \frac{ba}{\Gamma_a} s^{a-1} e^{-bs} \times \right.
\]

\[
\Pi_{j=1}^p (\gamma_j)^{a_j-1} \exp\left(-\frac{b_j \gamma_j}{\sigma}\right) \times (\delta)^{-a_\delta-1} \exp\left(-\frac{ba}{\delta}\right) \times (\sigma)^{-a_\sigma-1} \times \exp\left(-\frac{ba}{\sigma}\right) \quad (9)
\]

Where \(a, b, c\ and \ d\ are \ hyperparameters, \ here\ \ \ y_i = (y_1, y_2, ..., y_n)^T \quad v_n = (v_1, v_2, ..., v_n)^T \quad D^{-1} = \delta \sigma \text{ diag}(v_1, v_2, ..., v_n) \quad and \ m_n = (m_1, m_2, ..., m_n)^T \quad \text{(m(x_i^T \beta), m(x_i^T \beta), ..., m(x_i^T \beta))^T} \quad \text{E is the covariance matrix with dimension } (n \times n) \quad \text{and elements } E(., .) \text{ given in Equation} \quad E(x_i^T \beta, x_i^T \beta) = \delta \exp \exp \{-(x_i - x_j)^T \beta \beta^T (x_i - x_j)\}. \text{For the parameters and potential variables, the conditional posterior distributions will simply be derived via a Gibbs sampler algorithm for the Bayesian single index quantile regression. In the simulation study the kernel matrix } E \text{ is almost be singular, in order that } m \text{ is integrated to avert the singularity problem and therefore the inverse is calculated for } (E + D) \text{ the matrix (Hu et al. (2013) and Zhao and Lian (2015)). The conditional posterior densities of all the parameters and variables, excepting for } \beta \text{ and } \partial \text{, are}
common distributions. The conditional distributions utilized in the sampling are given below. \( m_n \) is marginalized when look considering the posterior conditional distribution of \( \beta \) and \( \vartheta \). For the \( T \) quantile an easy and efficient Gibbs sampler can be shown, works as follows:

**Sampling \( \beta \)**

\[
1 - \frac{1}{\pi(\beta|m_n, v_i, y, \sigma, s, \vartheta, y)} \propto \pi(y|v_i, \sigma, m_n) \times \pi(m_n|\beta, \vartheta) \times \pi(\beta|s)
\]

Full conditional posterior distributions

\[
\propto \exp\left\{-\frac{(y-v_i)^T(D+E_n)^{-1}(y-v_i)}{2}\right\} \times
\left((\text{det}[E_n+D])\right)^{-\frac{1}{2}} \times \prod_{j=1}^{p} \frac{1}{2s}
\]

Updating \( \beta \) The full conditional distribution of \( \beta_j \) is truncated normal With mean \( \bar{\beta}_j \) and variance \( S^2_\beta \) where

\[
\bar{\beta}_j = S^2_\beta \sum_{i=1}^{n} \frac{\alpha}{2\theta v_i} (y_i - m_T(x_i^T \beta_T) - \delta v_i)
\]

\[
\mathbb{I}(\beta_j < -\sqrt{s}), \quad \text{and} \quad S^2_\beta = \left(\sum_{i=1}^{n} \frac{\alpha}{2\theta v_i}\right)^{-1}
\]

**Sampling \( m_n \)**

\[
2 - \frac{1}{\pi(m_n|\beta, v_i, \sigma, s, y, \vartheta, y)} \propto \pi(y|v_i, m_n, \sigma, \beta) \times \pi(m_n|\beta, \vartheta)
\]

Full conditional posterior distributions

\[
\pi(m_n|\beta, v_i, \sigma, \vartheta, y) \sim N(\mu_n, \Sigma_n)
\]

**Sampling \( \sigma \)**

\[
3 - \frac{1}{\pi(\sigma|\beta, m_n, v_i, y, \vartheta, y)} \propto \pi(y|v_i, m_n, \sigma) \times \pi(v_i/\sigma) \times \pi(\sigma)
\]

Full conditional posterior distributions

\[
\alpha_\sigma^\sigma = -\frac{3n}{2} - a_\sigma
\]

\[
\sigma_\sigma = \left\{\sum_{i=1}^{n} (\frac{y-m_n-\delta v_i)^T(y-m_n-\delta v_i)}{2\theta v_i} + v_i + b_\sigma\right\}
\]

\[
\pi(\sigma|m_n, \beta_n, v_i, y, y) \sim IG(\alpha_\sigma, \sigma_\sigma)
\]

**Sampling \( y \)**

\[
4 - \frac{1}{\pi(y|\sigma, s, m_n, v_i, \vartheta, y)} \propto \pi(\beta/s) \times \pi(y)
\]

Full conditional posterior distributions

\[
\pi(y|s, \beta) \sim IG(\alpha_y + 2p, b_y + \sum_{j=1}^{p} s_j)
\]

**Sampling \( s \)**

\[
5 - \frac{1}{\pi(s|m_T, v_i, y, \sigma, \beta, \vartheta, y)} \pi(\beta/s) \times \pi(s)
\]

Full conditional posterior distributions

\[
\pi(s|m_T, y, \beta, y) \sim Gamma(p - pa, \sum_{j=1}^{p} (y + b))
\]

**Sampling \( \vartheta \)**

\[
6 - \frac{1}{\pi(\vartheta|\beta, m_n, y, \vartheta, y)} \propto \pi(y|v_n, \sigma, m_n) \times \pi(m_n|\beta, \vartheta) \times \pi(\vartheta)
\]

\[
\propto \left((\text{det}[D^{-1}])\right)^{-\frac{1}{2}} \exp\left\{-\frac{(y-m_n-\delta v_i)^T(y-m_n-\delta v_i)}{2}\right\}
\]
\[
\times \det \left[ E_n \right]^{-\frac{1}{2}} \exp \left\{ -\frac{m_n^T E_n^{-1} m_n}{2} \right\} \times \\
(\partial)^{-a_\partial - 1} \exp \left( -\frac{b_\partial}{\partial} \right)
\]

Full conditional posterior distributions
We use the Metropolis algorithm for sampling \( \partial \).

**Sampling \( v_i \)**
The full conditional distribution for \( v_i \) is a generalized inverse Gaussian distribution (\( GIG \)),
\[
7 - \pi(v_i | \sigma, \beta, m_n, \gamma, \partial, y) \alpha \\
\pi(y | \sigma, m_n, v_i) \times \pi(v_i)
\]

Full conditional posterior distributions
\[
\pi(v_i | \sigma, \beta, m_n, \gamma, \partial, y) = \\
\pi(v_i | \sigma, m_n, y) \sim GIG \left( \frac{1}{2}, \sqrt{\frac{(y - m_n)^2}{\theta \sigma}}, \sqrt{\frac{\delta^2}{\theta \sigma}} + \sqrt{\frac{\gamma^2}{\sigma}} \right)
\]

where the likelihood density function of \( GIG(\rho, w, z) \) is
\[
\left( \frac{w}{2K\rho(wz)} \right)^{p-1} x^{p-1} \exp \left\{ -\frac{1}{2} (w x^{-1} + z^2 x) \right\}
\]

\( x > 0, -\infty < \rho < \infty, w \geq 0, z \geq 0, \) and \( K \rho \) is the changed Bessel function of the third kind (Barndorff-Nielsen and Shephard 2001).

3. **Simulation**

The simulation study considered in this section to evaluate our suggested method Bayesian single index quantile regression with a uniform scale mixture (BUSIQR) and compared it\’s with other existing methods Bayesian quantile regression (BQR) and Bayesian single index quantile regression (BSIQR) and Bayesian lasso quantile regression (BLQR). Here, simulation example is considered in this study. This example was utilized by Wu et al. (2010) and have been very popular often it employed by most researchers whose study single index model conditional to quantile regression [See, (Hu et al. (2013), Alkenani and Yu (2013), Lv et al. (2014), Zhao and Lian (2015), Kuruwita (2015) and Alshaybawee et al. (2016))], but with some differences. Methods performance is evaluated based on the median mean absolute deviations referred to as (MMAD), bias and standard deviation. These measurements are calculated based on 100 replications. We have run the proposed Bayesian algorithm for 10000 iteration and the first 2000 iterations was burn in. Computations were done based on R package.

**Simulation Example**

In this example we have considered three samples with different size 50, 100 and 200 these samples are generated based on the following model:
\[ y = m(x^T \beta) + \sqrt{(\sin (x^T \beta) + 1)} \epsilon \]

where \( m(h) = 10 \sin (0.75 h) \),
where \( x_i, i = 1, \ldots, 6 \) are i.i.d normal distribution with \( N \sim [0, (0.25)^2] \), \( \beta = (\beta_1, \ldots, \beta_6)^T = \frac{1}{\sqrt{2}} (0, 1, 0, 0, 1, 0)^T \), and the error term (\( \epsilon \)) is standard normal distribution (SND), the models were supply with three various quantiles, \( T = \{ 0.25, 0.50, 0.75 \} \).

The standard deviation parameter estimates of this simulation example for the samples 50, 100 and 200 and based on 100 replications are summarized in Figures 1, 2 and 3 of the parameter estimates of each quantile.

![Figure 1](image1.png)
Figure (1) shows the trace plots for the estimated parameters at quantile 0.25 for simulation Example with 100 sample size.

![Figure 2](image2.png)
Figure (2) shows the trace plots for the estimated parameters at quantile 0.75 for simulation Example with 100 sample size.

We have display the trace plots in Figures (1) and (2), to check the rapprochement of the parameter estimates for the chain of the parameters \( \beta \) at two quantiles (0.25, 0.75). Based on the trace plots, we have seen clear that the chains of our suggested algorithm given a good performance.

![Figure 3](image3.png)
Figure (3) Shows the SD of the estimated parameters for the quantiles (0.25, 0.50, 0.75) in 100 replications when the sample size is 50.

From Figure (3), We can view that the proposed parameter estimates the BUSIQR method better than other methods, whereas this method have the smallest SD compare to the others. The small values of SD mean that the parameter...
estimates are homogeneous. On the contrary, the BQR and BLQR methods are got the largest values of SD. Also we can see that the BSIQR method are got SD smaller than the BLQR and BQR method.

Figure (4) Shows the SD of the estimated parameters for the quantiles (0.25, 0.50, 0.75) in 100 replications when the sample size is 100.

Figure (5) Shows the SD of the estimated parameters for the quantiles (0.25, 0.50, 0.75) in 100 replications when the sample size is 200.

As same as Figure (3), Figures (4) and (5), shows that the proposed method BUSIQR get the smallest SD for most parameter estimates at the different quantiles. In addition, the BQR method get the largest SD for all parameter estimates at all quantiles and in the different sample size 50,100 and 200. Also we can see that when the sample size are increase the SD for the BLQR and BSIQR are converge.

Table (1) shows the bias for the different quantiles when the sample size is 50

<table>
<thead>
<tr>
<th>τ</th>
<th>Methods</th>
<th>$\hat{\beta}_1$</th>
<th>$\hat{\beta}_2$</th>
<th>$\hat{\beta}_3$</th>
<th>$\hat{\beta}_4$</th>
<th>$\hat{\beta}_5$</th>
<th>$\hat{\beta}_6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>BQR</td>
<td>0.898</td>
<td>0.939</td>
<td>0.73</td>
<td>0.557</td>
<td>0.923</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>BSIQR</td>
<td>0.096</td>
<td>0.033</td>
<td>0.255</td>
<td>0.184</td>
<td>0.377</td>
<td>0.213</td>
</tr>
<tr>
<td></td>
<td>BLQR</td>
<td>0.161</td>
<td>0.877</td>
<td>0.623</td>
<td>0.459</td>
<td>0.888</td>
<td>0.743</td>
</tr>
<tr>
<td></td>
<td>BUSIQR</td>
<td>0.021</td>
<td>0.01</td>
<td>0.054</td>
<td>0.072</td>
<td>0.105</td>
<td>0.09</td>
</tr>
<tr>
<td>0.50</td>
<td>BQR</td>
<td>1.018</td>
<td>1.875</td>
<td>0.687</td>
<td>0.821</td>
<td>2.829</td>
<td>0.725</td>
</tr>
<tr>
<td></td>
<td>BSIQR</td>
<td>0.829</td>
<td>1.047</td>
<td>0.277</td>
<td>0.552</td>
<td>0.097</td>
<td>0.261</td>
</tr>
</tbody>
</table>
The proposed BUSIQR method and the existing methods, namely BLQR and BSIQR, show improved performance compared to the existing methods. Table (2) and (3) show the bias for the different quantiles when the sample size is 100 and 200, respectively. In both tables, the BUSIQR method consistently exhibits lower bias than the BLQR and BSIQR methods. Furthermore, Table (4) compares the Mean Modified Absolute Deviation (MMAD) for the SQR, BSIQR, BLQR, and BUSIQR methods based on 100 replications. The results indicate that the BUSIQR method generally outperforms the other methods, particularly for larger sample sizes.

In conclusion, the BUSIQR method shows promising results, especially in terms of bias reduction across different quantiles and sample sizes. Further research could explore the method's effectiveness in various applications and compare it with other state-of-the-art methods.
5. Conclusion

Figure (6) shows the MMAD for SQR, BSIQR, BLQR and BUSIQR methods based on 100 replications.

Table (4) and Figure (6) show that the median of mean absolute deviations (MMAD) of the new proposed method BUSIQR method is got the smallest values of MMAD compare to the other three existing methods at all the quantiles and the different sample size. BQR method is got the largest MMAD values, which indicate not good performance. In addition, the MMAD of the BSIQR method is smaller than BLQR and its better.

References


