Al-Qadisiyah Journal for Administrative and Economic Sciences ISSN$_{Online}$: 2312-9883
QJAE, Volume 21, Issue 4(2019) ISSN$_{Print}$: 1816-9171

Bayesian Estimation for Single Index Quantile Regression with Normal Scale Mixture

Taha Hussein Alshaybawee
Samer Jassim Alalaq
Department of Statistics , College of Administration and Economics, University of Al-Qadisiyah, Al Diwaniyah, Iraq
Corresponding Author: Taha Hussein Alshaybawee Samer Jassim Alalaq

***Abstract* :** *In this paper, the new Bayesian approach based on normal scale mixture is proposed for estimating the quantile regression of the single index model and selecting variables. The Gaussian process prior have been considered to the nonparametric link function. Bayesian hierarchical model construct and Gibbs sampler algorithm using for posterior inference. Simulation study was addressed to compare our suggested method with three other existing methods. The simulation studies significant that the new suggested method offers substantial improvement over the other three methods.*
Keyword: Quantile regression, single index model, Bayesian inference, normal scale mixture.

1- INTRODUCTION

Classical regression concentrate on the expectation of a variable (Y) conditional the single or a group of variables X, 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 instance, Mosteller and Tukey (1977) indicate "that the mean often gives a fairly incomplete picture of a single distribution, and the regression curve gives an incomplete picture of a set of distributions". Briollais and Durrieu (2014) say "There is no practical or theoretical reason to assume that the effect of the covariates is the same at different quantiles of the distribution". Such a function may be more or less complicated, however it restricts completely on a specified location of the (Y) conditional distribution. Quantile regression (QR) extends this approach, permitting one to comprehensive review the conditional distribution of Y on X at totally different, locations and therefore giving a global view on the interrelationships between Y and X. (QR) models became quite common since the research's Koenker, R. and Bassett, G. (1978) proposed it. Quantile regression has become a unite statistical methodology for estimating models of conditional quantile functions Wu, Y., & Liu, Y. (2009). QR usually provides more complete description picture of the response distribution than the mean and offers a practically necessary alternative to classical mean regression. There exists a large literature on quantile regression ways see for more detail Yu, J., & Zhang, L. (2003) and Koenker (2005). Nonparametric models are important in all areas of scientific data, on some occasions the inevitable the use of nonparametric models in quantile regression Koenker (2005). Nevertheless, the nonparametric quantile regression with common multivariate variables is a difficult estimation problem because of the "curse of dimensionality". To scale back the dimensionality while holding a lot of flexibility of a nonparametric model Wu et al. (2010). The single index quantile regression model SIQR suggested by Wu et al. (2010), it appeals to many researchers as a result of its features: it combines the power of parametric and the flexibility of nonparametric modelling. The SIQR is an attractive model to use with high dimensional data that are lessen for a single index, so allowing well approximation of the nonparametric function ,see for additional details, Ichimura (1993), Yatchew (2003) and Härdle et al. (2012). In addition, 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 τ^{th} quantile level SIQR model is shows by:

$$y = m_\tau(x_i^T \beta_\tau) + \varepsilon_\tau , \qquad (1)$$

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

Offer a process algorithm for estimating single index vector based on linear quantile regression and with the unknown correlation function estimated by the local linear method by Wu et al. (2010). This model has been given a great deal of interact from many researchers on both applied and theoretical side. Some of the drawbacks experienced the frequentist approach are mentioned by Benoit and Van den Poel (2012); they include the difficult optimization of the estimated parameters and the structure of the confidence interval. Because of these problems,

Benoit and Van den Poel (2012) relied a new estimation method based on the Bayesian approach. The Bayesian analysis method has become vastly applied, outcome of its ability to interest from all ready information in the analysis. Many studies presented to develop and expand the single index quantile regression model, for instance; Completely Bayesian method was developed by Hu et al. (2013). A similar Bayesian approach was used by Zhao and Lian (2015) for the controlled data in the single index quantile regression model. Computational algorithm for estimation and variable selection was advanced by Alkenani and Yu (2013), Lv et al. (2014) and Kuruwita (2016). Alshaybawee (2017) suggest a Bayesian approach to elastic net penalty to estimate and select variables in the single index quantile regression model.

2. The SIQR model and prior assumption

The model of single index quantile regression:-

$$y_i = m_\tau(\, x_i^\tau \beta_\tau\,) + \epsilon_{i\tau}\ , \qquad i = 1,2,\dots,n \qquad (2)$$

At τth quantile, where the quantile error term $\epsilon_{i\tau}$ distributed as an asymmetric Laplace distribution with the density:

$$\pi\,(\epsilon|\sigma,\tau) = \frac{\tau(1-\tau)}{\sigma}\ exp\ \left\{-\frac{1}{\sigma}\ \rho_\tau(\epsilon)\right\}, \qquad (3)$$

where σ is that the scale parameter . $\rho_{\,\tau}(\cdot)$ 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 $\epsilon_{i\tau} = y_i - m_\tau(x_i^T\beta_\tau)$ $i = 1,2,\dots,n.$ there n be a sample size, is shown as:

$$\pi(\, y|\beta, m\ , \sigma) = \frac{\tau^n(1-\tau)^n}{\sigma} exp\ \left\{-\frac{1}{\sigma}\sum_{i=1}^{n}\ \rho_\tau\Big(y_i - m(x_i^T\beta)\Big)\right\}, \qquad (4)$$

quantile regression is often based on minimization of the check function. However, direct use of the probability top on convenient for Bayesian inference. A location-scale mixture illustration of the ALD (Kozumi and Kobayashi 2011) is useful here . We can write the notes appropriate (4) instead as:

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

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

$$\pi(\, y_i|\beta, m_\tau, v_i) = \prod_{i=1}^{n}\ (2\pi\sigma\theta v_i)^{-\frac{1}{2}}\,.\,exp\{-\frac{1}{2\sigma\theta v_i}\ (y_i - m_\tau(x_i^T\beta_\tau) - \delta v_i)^2\} \qquad (5)$$

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

Here $y_i = (y_1\ , y_2\ , \dots, y_n)^T,$ $v_n = (v_1, v_2, \dots, v_n)^T,$ $D^{-1} = \delta\ \sigma\ diag(v_1, v_2, \dots, v_n)$, and $m_n = (m_1, m_2, \dots, m_n)^T = (m(x_1^T\beta), m(x_2^T\beta), \dots, m(x_n^T\beta))^T$.

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

$$m_\tau(.) \sim GP\big(0\ , E(.,.)\big), \qquad E\big(x_i, x_j\big) = \partial\ \ exp\left(-\ \frac{(x_i -\ x_j)^2}{\omega}\right),$$

where ∂ and ω are Hyperparameters.

Writing this out in the single-index model framework using the observed covariates $\boldsymbol{x_i}$, we have,

$$\pi(\, m_\tau|\beta_\tau, \partial) = det[\ E_n\ ]^{-\frac{1}{2}}\ \ exp\{-\frac{m_n^T\, E_n^{-1} m_n}{2}\}\ ,$$

E_n is the covariance matrix with dimension $(n \times n\,)$ and elements $E(.,.)$ given in Equation $E(x_1^T\beta, x_1^T\beta) = \partial$ $exp\ exp\ \{-\big(x_i - x_j\big)^T\beta\ \beta^T\big(x_i - x_j\big)\}$.

Follow Gramacy and Lian (2012), and when we use the Gaussian process as a prior distribution to the nonparimatric link function, then $\frac{\beta}{\sqrt{\omega}}$ is identifiable without the necessity for the constraint $\|\beta\| = 1$. Therefore we will instead $\frac{\beta}{\sqrt{\omega}}$ by β and the covariance function is reformulated as follows:

$$E\big(x_i, x_j\big) = \partial\quad exp\ exp\ \{-(x_i^T\beta -\ x_i^T\beta)^2\}\ , \qquad (6)$$

the inverse gamma distribution is ready as a hyper prior for which implies that $\partial \sim IG\ (a_\gamma\ , b_\gamma\ )$ Where$a_\gamma$and b_γ are the hyperparameters.

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 great 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, Andrews & Mallows (1974). Their properties have been helpful in several areas. In theoretical studies distributional properties like moments are often readily derived by exploiting the special structure. Practically, robustness studies have often utilized these distribution for simulation and in the analysis of external models. Park and Casella (2008) introduced the Bayesian lasso regression, using a conditional Laplace prior distribution represented as a scale mixture of normal with an exponential mixing distribution. As well, Alhamzawi, Yu, & Benoit, (2012) Proposed adaptive Lasso quantile regression (BALQR) from a Bayesian perspective .The method extends the Bayesian Lasso quantile regression by allowing different penalization parameters for different regression coefficients. Inversed gamma prior distributions are placed on the penalty parameters. Benoit, D. F., Alhamzawi, R., & Yu, K. (2013) propose Bayesian lasso binary quantile regression. Alhamzawi, R., & Taha Mohammad Ali, H. (2018) proposed Bayesian treatment for the adaptive lasso and related estimators by used a scale mixture of truncated normal representation of the Laplace density with exponential mixing densities.

3. Hierarchical model and MCMC sampler

Laplace distribution will be written as a scale mixture of truncated normal distribution (with the exponential density), Andrews and Mallows, 1974. i.e

$$\left(\frac{\gamma}{2}\right) exp\{-(\gamma|\beta_{j\,\tau}|)\} = \int_0^\infty \frac{1}{\sqrt{2\pi\sigma_j}}\ exp\ exp\left(\frac{\beta_j^2}{2\sigma_j}\right)\ \frac{\gamma^2}{2} exp\ exp\left(-\frac{\gamma^2}{2}\sigma_j\right) d\sigma_j,$$

hierarchical model Based on what's mentioned in Section (2), the hierarchic model for the Bayesian single index quantile regression will be written as:

$$y_i = m_\tau(\, x_i^\tau \beta_\tau\,) + \epsilon_{i\tau}\ \ ,\quad i = 1,2,\dots,n$$

$$\beta_\tau \Big/ \sigma\,, \gamma\ \sim\ \prod_{j=1}^{p} \int_0^\infty \frac{1}{2\pi\sigma_j}\ exp\ exp\,(\frac{\beta_j^2}{2\sigma_j})\quad \frac{\gamma^2}{2} exp\ exp\,(-\frac{\gamma^2}{2}\sigma_j)\ d\sigma_j\ \ ,$$

$$y/\beta_\tau\,, \sigma, m_\tau\,, v \sim N(m_\tau(x_i^T\beta) +\ \delta v\,, \theta\sigma\, u\,),$$

$$m_\tau|x_i, \beta_\tau, \partial \sim GP(0\,, E_n), \qquad (7)$$

$$u_i \sim N(\,0\,,1\,)\quad ,\quad v_i \sim \frac{1}{\sigma} exp\ exp\left(-\frac{v_i}{\sigma}\right)\quad ,\quad \gamma^2 \sim Gamma\,(\,a_\gamma\,, b_\gamma\,)$$

$$\sigma \sim Inv\,.Gamma\,(\,a_\sigma\,, b_\sigma\,)\quad ,\quad \partial \sim Inv\,.Gamma\,(\,a_\partial\,, b_\partial\,)$$

For all parameters the posterior can be given as:

$$(\,m_n, \beta, v, \gamma, \partial, \sigma|y)\alpha\ \{\,(det[D^{-1}])^{-\frac{1}{2}}\, exp\left\{-\frac{(y - m_n - \delta v_i)^T D^{-1}(y - m_n - \delta v_i}{2}\right\}$$

$$\times\, det[\,E_n\,]^{-\frac{1}{2}}\ \ exp\{-\frac{m_n^T E_n^{-1} m_n}{2}\}\ \times \prod_{j=1}^{p} \int_0^\infty \frac{1}{2\pi\sigma_j}\ exp\ exp\,(\frac{\beta_j^2}{2\sigma_j})\quad \frac{\gamma^2}{2} exp\ exp\,(-\frac{\gamma^2}{2}\sigma_j)$$

$$\times \prod_{i=1}^{n} \frac{1}{\sigma} exp\ exp\left(-\frac{v_i}{\sigma}\right)\ \times \prod_{j=1}^{p} (\gamma_j)^{a_\gamma - 1}\, exp\ exp\left(-b_\gamma \gamma_j\right)\ \times (\partial)^{-a_\partial - 1}\, exp\ exp\left(-\frac{b_\partial}{\partial}\right)$$

$$\times\, (\sigma)^{-a_\sigma - 1}\, exp\ exp\left(-\frac{b_\sigma}{\sigma}\right),$$

where $a, b, c\ and\ d$ are hyperparameters, Here $y_i = (y_1\,, y_2\,, \dots, y_n)^T$, $v_n = (v_1, v_2, \dots, v_n)^T$, $D^{-1} = \delta\ \sigma\ diag(v_1, v_2, \dots, v_n)$, and $m_n = (m_1, m_2, \dots, m_n)^T = m(x_1^T\beta), m(x_2^T\beta), \dots, m(x_n^T\beta))^T$, E is that the covariance matrix with dimension $(n \times n\,)$and components $E(.,.)$given in Equation $E(x_1^T\beta, x_1^T\beta) = \partial\ exp\ \{-(x_i - x_j)^T\beta\ \beta^T(x_i - x_j)\}$.

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 several cases in the simulation study the kernel matrix E_n is almost singular, in order that m_τ is integrated to avert the singularity problem and therefore the inverse is calculated for $(E_n + D\,)$ the matrix (Hu et al. (2013) and Zhao and Lian (2015)). The conditional posterior densities of all the parameters and variables, excepting for β and , 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 β and ∂. For the $\tau\ th$ quantile an easy and efficient Gibbs sampler can be shown, works as follows:

Sampling $\boldsymbol{\beta}$

$$1-\pi(\beta|m_n,\sigma,\gamma,\partial,y)\propto\ \pi(y|v_n,\sigma,m_n)\times\ \pi(m_n|\beta,\partial\ )\times\pi(\beta|\sigma,\gamma)$$

Full conditional posterior distributions

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

To β sample the posterior, the Metropolis algorithm is used within the Gibbs sample.

Sampling $\boldsymbol{\partial}$

$$2-\pi(\partial\ |\beta,v_n,\sigma,y)\propto\ \pi(y|v_n,\sigma,m_n)\times\ \pi(m_n|\beta,\partial)\ \times\pi(\partial\ )$$

Full conditional posterior distributions

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

we use the Metropolis algorithm for sampling .

Sampling $\boldsymbol{m_n}$

$$3-\ \pi(m_n|\beta,v_i,\sigma,\gamma,\partial,y)=\pi(y|v_n,\sigma,m_n)\times\ \pi(m_n|\beta,\partial)$$

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

$$\pi(m_n|\beta,v_i,\sigma,\partial,y)\ \sim\ N(\mu_n,\sum_n),$$

Full conditional posterior distributions

The link function has (ND) with mean.

$\mu_n=\ E_n(E_n+D)^{-1}(y-\delta v_i)$, variance $,\sum_n\ =\ E_n(E_n\ +D)^{-1}D$

Sampling $\boldsymbol{\sigma}$

$$4-\ \pi(\sigma|\beta,m_n,v_i,\gamma,\partial,y)=\pi(y|\sigma,m_n,v_i\ )\times\pi(\beta|\sigma,\gamma\ )\times\pi(1/\sigma)\times\pi(v_i)$$

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

$$\times\prod_{j=1}^{p}\int_0^\infty\frac{1}{2\pi\sigma_j}\exp\left(\frac{\beta_j^2}{2\sigma_j}\right)\ \frac{\gamma^2}{2}\exp\left(-\frac{\gamma^2}{2}\sigma_j\right)\times\prod_{i=1}^{n}\frac{1}{\sigma}\exp\left(-\frac{v_i}{\sigma}\right)\times(\sigma)^{-a_\sigma-1}\exp\left(-\frac{b_\sigma}{\sigma}\right)$$

$$\pi(\sigma|,m_n,\beta,v_i,\gamma,y)\ \sim\ IG(\alpha_\sigma\ ,\vartheta_\sigma\ )$$

Full conditional posterior distributions

As we can see, the conditional distribution of $\boldsymbol{\sigma}$ is inverse Gamma with the shape parameter

$\alpha_\sigma\ =\ \frac{3n+p}{2}+a_\sigma+1,$

and scale parameter

$$\vartheta_\sigma=\left\{\sum_{i=1}^{n}\frac{(y-m_n-\delta v_i)^T(y-m_n-\delta v_i)}{2\theta v_i}+v_i+\sum_{j=1}^{p}\frac{\beta^2}{2}+b_\sigma\right\}-\frac{\gamma_\sigma^2}{2}$$

Sampling $\boldsymbol{\gamma}$

$$5\text{-}\ \pi(\gamma|\sigma,\beta,m_n,v_i,\partial,y)=\pi(\beta|\sigma,\gamma)\times\pi(\gamma)$$

$$\alpha\ \ \frac{\gamma^2}{2}\exp\left(-\frac{\gamma^2}{2}\sigma_j\right)\ \times\ \prod_{j=1}^{p}(\gamma_j)^{a_\gamma-1}\exp\left(-b_\gamma\gamma_j\right)$$

Full conditional posterior distributions

the conditional distribution of γ is Gamma inverse.

$$\pi(\gamma|\sigma,\beta)\ \sim\ Ga(\ \ a_\gamma\ +2\ p\ \ ,\ \ \ b_\gamma\ +\frac{\gamma_\sigma}{2}\ \ )$$

Sampling $\boldsymbol{v_i}$

6- The total conditional distribution for v_i is a generalized inverse Gaussian distribution *(GIG)*,

$$\pi(v_i\ |\sigma,\beta,m_n,\gamma,\partial,y)=\pi(y|\sigma,m_n,v_i\ )\times\pi(v_i)$$

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

Full conditional posterior distributions

The conditional distribution for v_i is (GIG)

$$\pi(v_i\,|\sigma, m_n, y) \sim GIG\left(\frac{1}{2}, \sqrt{\frac{(y-m_n)^2}{\theta\sigma}}, \sqrt{\frac{\delta^2}{\theta\sigma}+\sqrt{\frac{2}{\sigma}}}\right)$$

where the probability density function of $GIG(\rho, w, z)$ is

$$f\,(x|\rho, w, z) = \frac{(z/w)^{\rho}}{2K\rho(wz)}\, x^{\rho-1}\, exp\left\{-\frac{1}{2}(w^2\, x^{-1} + z^2 x)\right\}$$

$x > 0,\ -\infty < \rho < \infty,\ w \geq 0,\ z \geq 0,$

and $K\rho$ is the modified Bessel function of the third kind (Barndorff-Nielsen and Shephard 2001).

4. Simulation

The simulation study adopted in this section to evaluate our suggested method Bayesian single index quantile regression based on normal mixture (BNSIQR) compared to some other existing methods; Bayesian quantile regression (BQR), Bayesian single index quantile regression (BSIQR) and Bayesian lasso quantile regression (BLQR). Simulation examples is utilized by Wu et al. (2010) and it has become popular, whereas it was employed by the most researchers whose study single index quantile regression model (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, in the number of covariates and the values of the parameters. The performance of the methods are evaluated based on the median of mean absolute deviations denoted as MMAD, bias and standard deviation. These measurements are calculated based on 100 replications. The Bayesian algorithm was run for 10000 iteration and the first 2000 iterations was burn in. Computations were done using R package.

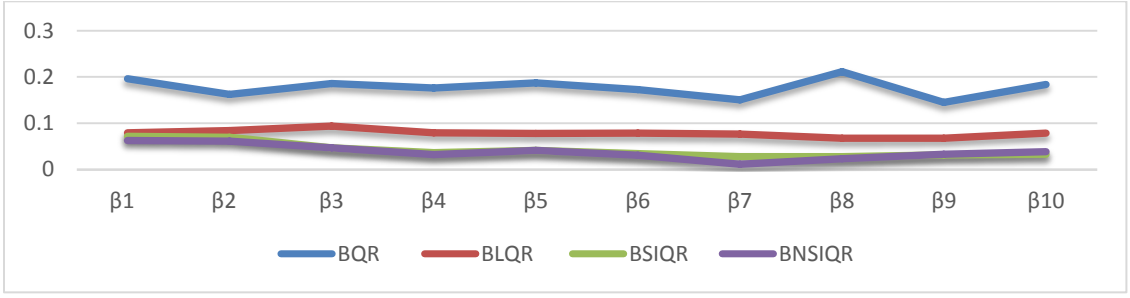
Simulation Example

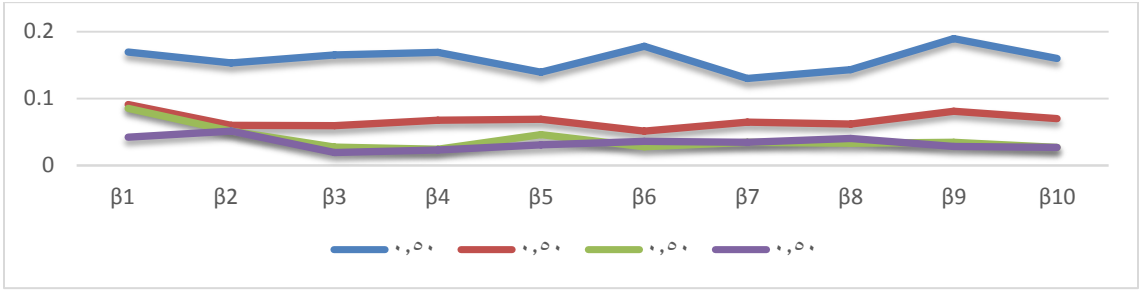
In this example three samples size are considered 50, 100 and 200, these samples were generated with homoscedastic errors according to the following SIM:

$$y = m(x_i^T\beta) + 0.1\epsilon\,, m(h) = sin\left\{\frac{\pi(h-\mathrm{C})}{R-\mathrm{C}}\right\},$$

where $x_i\,, i = 1, \ldots, 10,$ is uniformly distributed (0,1), $\mathrm{C} = \frac{\sqrt{3}}{2} - \frac{1{,}645}{\sqrt{12}}$, $R = \frac{\sqrt{3}}{2} + \frac{1{,}645}{\sqrt{12}}$, $\beta = (\beta_1, \ldots, \beta_{10})^T = \frac{1}{\sqrt{12.25}}\ (3,1,5,0,0,2,0,0,0,0)^T$ and ϵ_i is generated from N (0,1). The results were summarized in the following tables and figures based on 100 replication for three quantiles $\tau = \{0.25, 0.50, 0.75\}$.

In Figures 1,2 and 3, we summarize the standard deviation results of the parameter estimates to the four methods with different sample size 50,100 and 200 respectively . A good estimator that one get the smallest SD.





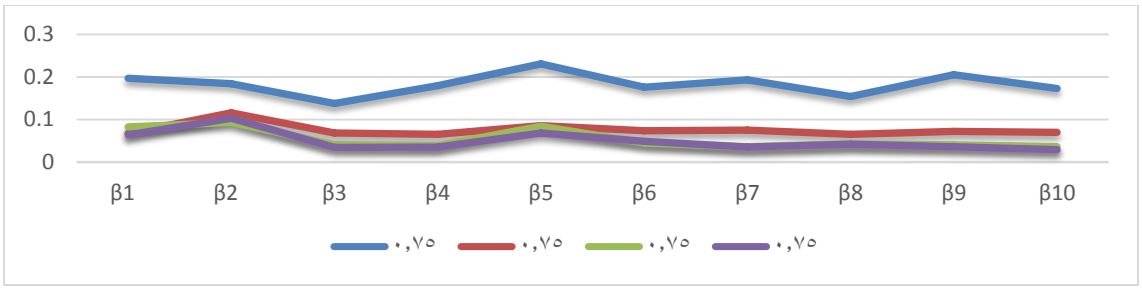
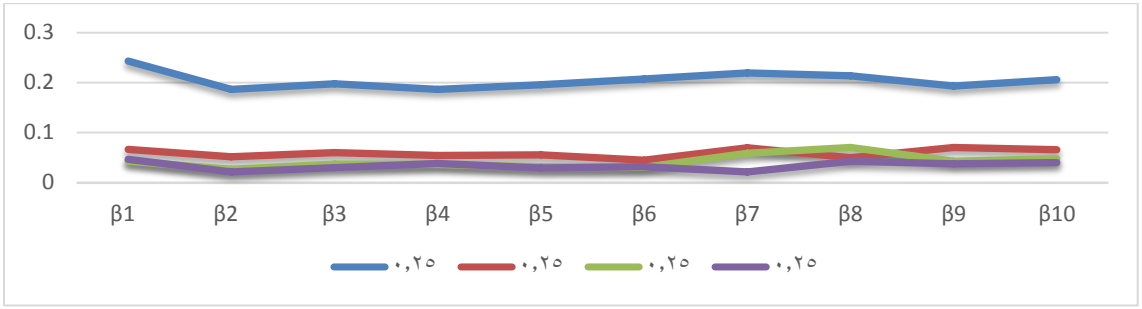
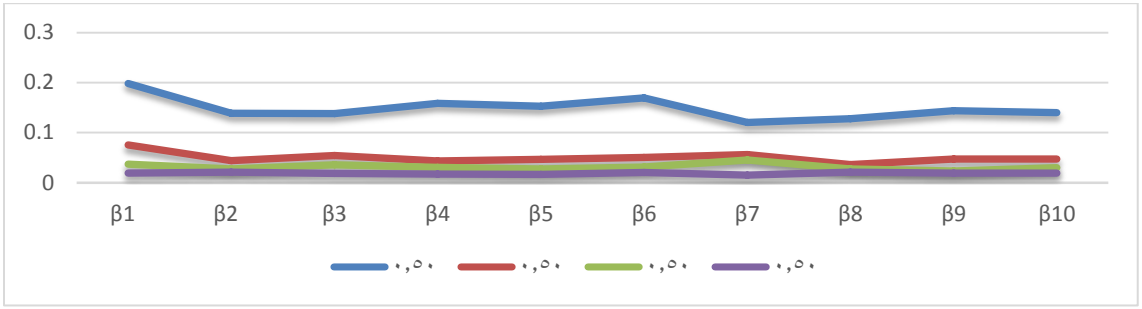


Figure (1) shows the SD of the estimated parameters for the quantiles (0.25, 0.50, 0.75) in 100 replications when the sample size is 50.

From Figure 1, we can see that the parameter estimates for the proposed BNSIQR method is better than the other methods, whereas this method have the smallest SD compare to the others. On the contrary, the BQR method get the largest values of SD. Also we can see that the BSIQR method are get SD smaller than the BLQR method.





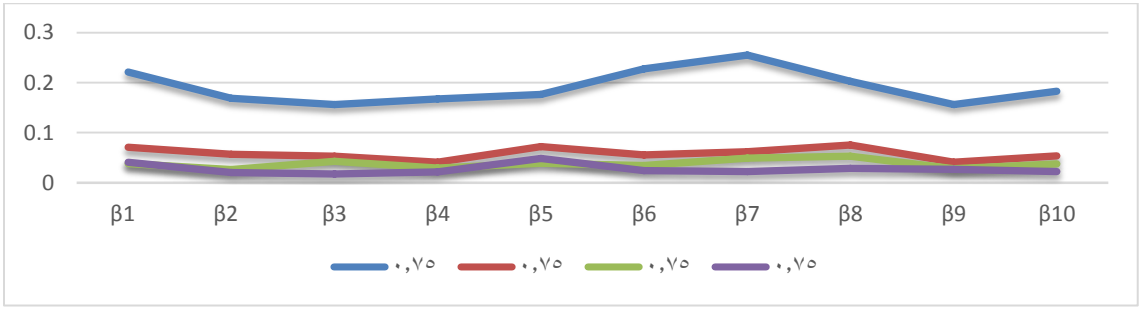


Figure (2) shows the SD of the estimated parameters for the quantiles (0.25, 0.50, 0.75) in 100 replications when the sample size is 100.

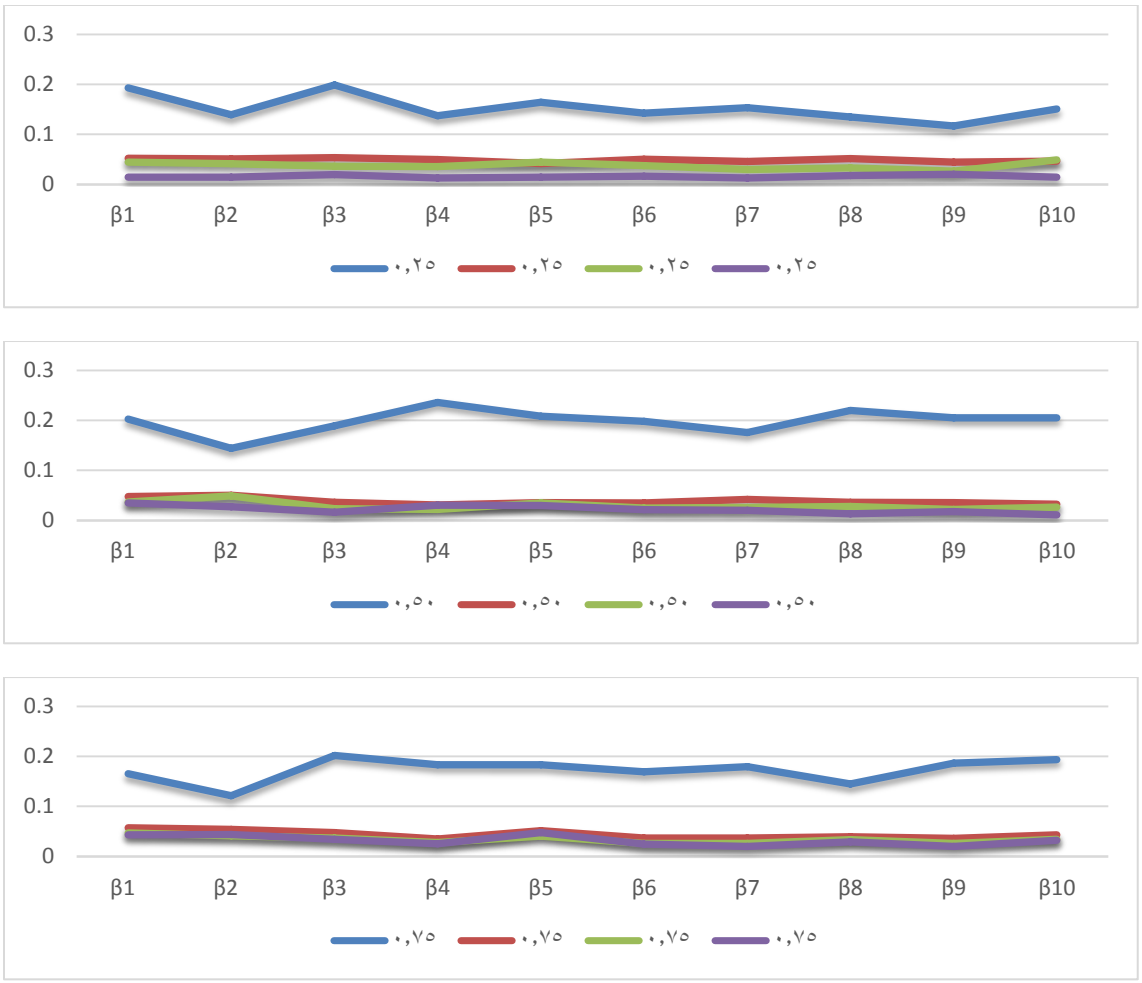


Figure (3) shows the SD of the estimated parameters for the quantiles (0.25, 0.50, 0.75) in 100 replications when the sample size is 200.

Figures 2 and 3, shows that the proposed method BNSIQR gets 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. The bias of parameter estimates for BQR, BSIQR, BLQR and BNSIQR based on 100 replications when the sample size is 50.

τ	Methods	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_6$	$\hat{\beta}_7$	$\hat{\beta}_8$	$\hat{\beta}_9$	$\hat{\beta}_{10}$
0.25	BQR	0.2946	0.2662	0.2466	0.1873	0.2665	0.2080	0.1801	0.2155	0.1940	0.1918
	BSIQR	0.1546	0.0854	0.0906	0.0379	0.3265	0.2374	0.0208	0.0412	0.0385	0.0515
	BLQR	0.1217	0.2329	0.1693	0.0322	0.4574	0.0448	0.0201	0.0522	0.0285	0.0291
	BNSIQR	0.1236	0.0817	0.0520	0.0251	0.2241	0.0201	0.0020	0.0213	0.0615	0.0056
0.50	BQR	0.2662	0.2220	0.2220	0.2131	0.2438	0.2239	0.2248	0.2253	0.2159	0.2461
	BSIQR	0.1919	0.1703	0.0805	0.0674	0.0432	0.3113	0.1492	0.0581	0.0514	0.0369
	BLQR	0.0789	0.2122	0.0316	0.0165	0.3299	0.1046	0.0455	0.0112	0.0065	0.0031
	BNSIQR	0.0301	0.1036	0.0477	0.0114	0.0426	0.1044	0.0258	0.0111	0.0058	0.0055

0.75	BQR	0.3367	0.1682	0.1201	0.1817	0.2560	0.1772	0.1995	0.1364	0.2101	0.1665
	BSIQR	0.1495	0.1036	0.1606	0.0682	0.1052	0.2835	0.1871	0.0127	0.0851	0.0420
	BLQR	0.1083	0.2325	0.0557	0.0029	0.2090	0.1427	0.0563	0.0116	0.0042	0.0218
	BNSIQR	0.1048	0.1264	0.0221	0.0024	0.0840	0.0697	0.0214	0.0080	0.0036	0.0123

Table 2. The bias of parameter estimates for BQR, BSIQR, BLQR and BNSIQR based on 100 replications when the sample size is 100.

τ	Methods	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_6$	$\hat{\beta}_7$	$\hat{\beta}_8$	$\hat{\beta}_9$	$\hat{\beta}_{10}$
0.25	BQR	0.8350	0.5829	0.3101	0.2477	0.7119	0.2741	0.4401	0.4519	0.0597	0.4610
	BSIQR	0.3518	0.2857	0.1730	0.2005	0.3432	0.1496	0.1689	0.1114	0.1641	0.1571
	BLQR	0.6750	0.5868	0.1768	0.1431	0.4936	0.1040	0.1767	0.3595	0.0779	0.1665
	BNSIQR	0.0948	0.0722	0.0007	0.0144	0.0855	0.0047	0.0102	0.0052	0.0029	0.0058
0.50	BQR	0.7747	0.4943	0.0935	0.1397	0.6629	0.1573	0.3045	0.2453	0.0694	0.1608
	BSIQR	0.1891	0.2320	0.1845	0.2461	0.3055	0.2345	0.2014	0.1989	0.2211	0.1292
	BLQR	0.6198	0.5105	0.0444	0.0833	0.5071	0.0226	0.1300	0.0974	0.0051	0.0051
	BNSIQR	0.0059	0.0126	0.0012	0.0030	0.0138	0.0090	0.0072	0.0005	0.0053	0.0085
0.75	BQR	0.6660	0.6220	0.0353	0.0107	0.6619	0.0483	0.1387	0.0843	0.0190	0.0063
	BSIQR	0.3069	0.2415	0.1103	0.1943	0.3726	0.2010	0.2147	0.1898	0.0997	0.1415
	BLQR	0.5663	0.5477	0.0227	0.0057	0.5498	0.0003	0.0633	0.0331	0.0306	0.0277
	BNSIQR	0.0256	0.0221	0.0009	0.0009	0.0313	0.0016	0.0070	0.0039	0.0131	0.0045

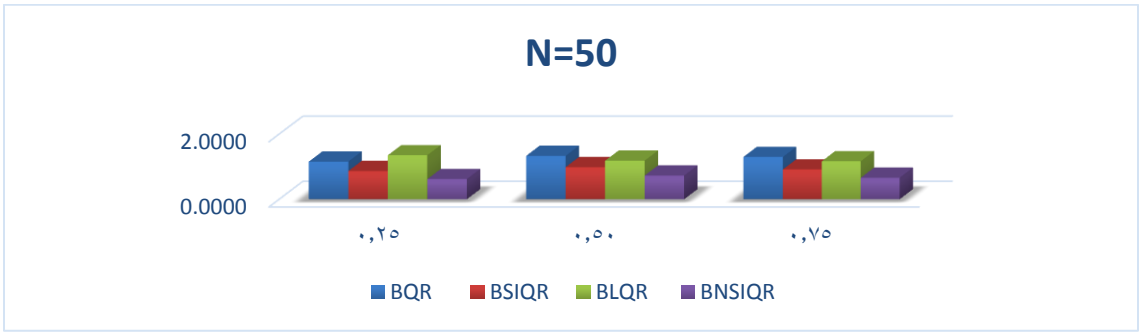
Table 3. The bias of parameter estimates for BQR, BSIQR, BLQR and BNSIQR based on 100 replications when the sample size is 200.

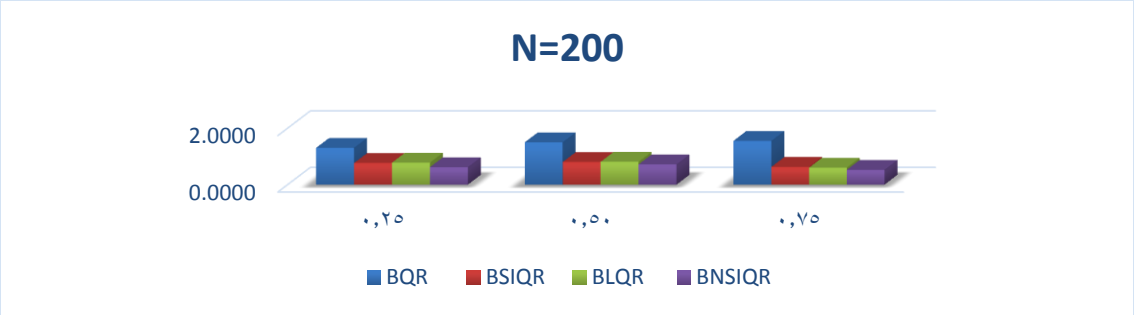
τ	Methods	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_6$	$\hat{\beta}_7$	$\hat{\beta}_8$	$\hat{\beta}_9$	$\hat{\beta}_{10}$
0.25	BQR	0.3806	0.3687	0.1685	0.2529	0.2805	0.2133	0.0575	0.1237	0.0705	0.3449
	BSIQR	0.2107	0.1770	0.0956	0.1535	0.1195	0.1107	0.0461	0.0669	0.0182	0.2145
	BLQR	0.2823	0.2745	0.2155	0.1326	0.1924	0.1631	0.1971	0.1523	0.1994	0.1607
	BNSIQR	0.0011	0.0042	0.0029	0.0061	0.0061	0.0001	0.0040	0.0012	0.0063	0.0037
0.50	BQR	0.4125	0.3744	0.0630	0.0510	0.3355	0.0793	0.0031	0.1089	0.0100	0.1737
	BSIQR	0.2724	0.2200	0.0207	0.0609	0.1919	0.0701	0.0176	0.0864	0.0256	0.0957
	BLQR	0.3566	0.3309	0.1742	0.1678	0.2913	0.1474	0.1522	0.1466	0.1748	0.1852
	BNSIQR	0.0176	0.0143	0.0047	0.0049	0.0115	0.0050	0.0013	0.0023	0.0058	0.0037
0.75	BQR	0.4314	0.4785	0.0019	0.0190	0.4027	0.0269	0.0299	0.0771	0.0379	0.0613
	BSIQR	0.2943	0.2377	0.0112	0.0224	0.2552	0.0286	0.0113	0.0668	0.0409	0.0325
	BLQR	0.3860	0.4375	0.1932	0.1942	0.3309	0.1986	0.2538	0.1883	0.2485	0.1942
	BNSIQR	0.0316	0.0177	0.0061	0.0058	0.0200	0.0065	0.0033	0.0025	0.0081	0.0046

Tables 1, 2 and 3 show that the bias of the proposed method are the smallest at all the quantiles levels and the different size of samples 50, 100 and 200. On the contrary, the BQR methods are getting the largest bias at the different quantiles and sample size. The bias of the BSIQR method is better than the bias of BQR when the sample size 50, whereas when the sample size are increase the bias for these methods are converge.

Table(4) Comparison of MMAD for SQR, BSIQR, BLQR and BNSIQR methods based on 100 replications for the different sample size.

n	τ	BQR	BSIQR	BLQR	BNSIQR
50	0.25	0.4540	0.3655	0.3486	0.1163
	0.50	0.3293	0.3844	0.1807	0.1248
	0.75	0.2950	0.3140	0.1653	0.1579
100	0.25	1.8701	0.4861	1.2560	0.2327
	0.50	1.4272	0.4234	0.9321	0.3206
	0.75	1.1259	0.4499	0.8627	0.3405
200	0.25	0.3022	0.3641	0.4780	0.0176
	0.50	0.4997	0.3997	0.5073	0.0159
	0.75	0.5499	0.4270	0.6014	0.0235





Figure(4) shows the MMAD for SQR, BSIQR, BLQR and BNSIQR methods based on 100 replications for the different sample size.

From Table 4 and Figure 4 we can note that clearly that the median of mean absolute deviations (MMAD) of the BNSIQR method is the smallest compare to the other three methods. Whereas the BQR method get the largest MMAD values. As same as that in SD and bias we can see that the MMAD for BSIQR is smaller than MMAD for BLQR. Whereas these values are converged when the sample size are increased.

5. Conclusion

In this article, the Bayesian approach based on normal scale mixture is proposed for estimating a single index conditional quantile regression model and selecting variables. The Gaussian process have been considered as a prior to the nonparametric link function. We construct a Bayesian hierarchical model and Gibbs sampler algorithm was used for posterior inference.

Simulation was adopted to compare our suggested method, BNSIQR, with three other methods, BQR ,BSIQR and BLQR . The simulation study signify that the BNSIQR offers substantial improvement over the other three methods. In our simulation study MCMC algorithm exhibited good mixing, so that BNSIQR consistently has the smallest standard deviations, bias and MMAD and good convergent properties. Therefore, we can concluded that the suggested method BNSIQR performs better than the other existing methods.

References

Alhamzawi, R., & Taha Mohammad Ali, H. (2018). A new Gibbs sampler for Bayesian lasso. *Communications in Statistics-Simulation and Computation*, 1-17.

Alhamzawi, R., Yu, K., & Benoit, D. F. (2012). Bayesian adaptive Lasso quantile regression. Statistical Modelling, 12(3), 279-297.

Alkenani, A., & Yu, K. (2013). Penalized single-index quantile regression. *International Journal of Statistics and Probability*, *2*(3), 12.

Alshaybawee, T., Midi, H., & Alhamzawi, R. (2017). Bayesian elastic net single index quantile regression. *Journal of Applied Statistics*, *44*(5), 853-871.

Andrews, D. F., & Mallows, C. L. (1974). Scale mixtures of normal distributions. *Journal of the Royal Statistical Society: Series B (Methodological)*, *36*(1), 99-102.

Barndorff-Nielsen, O. E., & Shephard, N. (2001). Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *63*(2), 167-241.

Benoit, D. F., & Van den Poel, D. (2012). Binary quantile regression: a Bayesian approach based on the asymmetric Laplace distribution. *Journal of Applied Econometrics*, *27*(7), 1174-1188.

Benoit, D. F., Alhamzawi, R., & Yu, K. (2013). Bayesian lasso binary quantile regression. *Computational Statistics*, *28*(6), 2861-2873.

Briollais, L., & Durrieu, G. (2014). Application of quantile regression to recent genetic and-omic studies. *Human genetics*, *133*(8), 951-966.

Choi, T., Shi, J. Q., & Wang, B. (2011). A Gaussian process regression approach to a single-index model. *Journal of Nonparametric Statistics*, *23*(1), 21-36.

Gramacy, R. B., & Lian, H. (2012). Gaussian process single-index models as emulators for computer experiments. *Technometrics*, *54*(1), 30-41.

Härdle, W. K., Müller, M., Sperlich, S., & Werwatz, A. (2012). *Nonparametric and semiparametric models*. Springer Science & Business Media.

Hu, Y., Gramacy, R. B., & Lian, H. (2013). Bayesian quantile regression for single-index models. *Statistics and Computing*, *23*(4), 437-454.
Ichimura, H. (1993). Semiparametric least squares (SLS) and weighted SLS estimation of single-index models. *Journal of Econometrics*, *58*(1-2), 71-120.
Koenker, R., & Basset, G. (1978). Asymptotic theory of least absolute error regression. *Journal of the American Statistical Association*, *73*(363), 618-22.
Koenker. (2005). *Quantile Regression (Econometric Society monographs; no. 38)*. Cambridge university press.
Kozumi, H., & Kobayashi, G. (2011). Gibbs sampling methods for Bayesian quantile regression. *Journal of statistical computation and simulation*, *81*(11), 1565-1578.
Kuruwita, C. N. (2016). Non-iterative Estimation and Variable Selection in the Single-index Quantile Regression Model. *Communications in Statistics-Simulation and Computation*, *45*(10), 3615-3628.
Lv, Y., Zhang, R., Zhao, W., & Liu, J. (2014). Quantile regression and variable selection for the single-index model. *Journal of Applied Statistics*, *41*(7), 1565-1577.
Mosteller, F., & Tukey, J. W. (1977). Data analysis and regression: a second course in statistics. *Addison-Wesley Series in Behavioral Science: Quantitative Methods*.
Park, T., & Casella, G. (2008). The bayesian lasso. *Journal of the American Statistical Association*, *103*(482), 681-686.
Wu, T. Z., Yu, K., & Yu, Y. (2010). Single-index quantile regression. *Journal of Multivariate Analysis*, *101*(7), 1607-1621.
Wu, Y., & Liu, Y. (2009). Variable selection in quantile regression. *Statistica Sinica*, *19*(2), 801.
Yatchew, A. (2003). *Semiparametric regression for the applied econometrician*. Cambridge University Press.
Yu, K., Lu, Z., and Stander, J. (2003). "Quantile regression: applications and current researchv areas," *The Statistician* 52, 331–350.
Zhao, K., & Lian, H. (2015). Bayesian Tobit quantile regression with single-index models. *Journal of Statistical Computation and Simulation*, *85*(6), 1247-1263.