© Statistical Research and Training Center

مجله‌ی پژوهش‌های آماری ایران
دوره‌ی ۱۷، شماره‌ی ۱، بهار و تابستان ۱۳۹۹، صص ۱۹–۴۴
پژوهشکده‌ی آمار

J. Statist. Res. Iran **17** (2020): 19–44
DOI: 10.52547/jsri.17.1.19
DOR: 20.1001.1.25385771.2020.17.1.2.3

Semi-parametric Quantile Regression for Analysing Continuous Longitudinal Responses

Omid Khazaei, Mojtaba Ganjali* and Mojtaba Khazaei

Shahid Beheshti University

Received: 2021/01/03 Approved: 2022/30/05

Abstract. Recently, quantile regression (QR) models are often applied for longitudinal data analysis. When the distribution of responses seems to be skew and asymmetric due to outliers and heavy-tails, QR models may work suitably. In this paper, a semi-parametric quantile regression model is developed for analysing continuous longitudinal responses. The error term's distribution is assumed to be Asymmetric Laplace (AL) distribution for modeling the continuous responses. The correlation of longitudinal responses belong to the same individual is taken into account by using a random-effects approach. We use the local polynomial kernel to approximate the non-parametric part of the model. The parameter estimation procedure is performed under a Bayesian paradigm using the Gibbs sampling method. The performance of the model is evaluated in a simulation study. To show the proposed model's application, a Peabody Individual Achievement Test (PIAT) dataset is analyzed.

Keywords. Semi-parametric Quantile regression, continuous longitudinal data, local polynomial kernel, asymmetric Laplace distribution, semi-parametric model, Gibbs sampling.

MSC 2010: 62G08, 62F15.

* Corresponding author

1 Introduction

Linear regression models are statistical tools used to model the association between variables. Many researchers used Classic Mean Regression (CMR) as a general tool to analyze the data. The principal purpose of CMR is to model the mean of the response variable under a set of predictor variables. Quantile regression (QR) is a type of regression analysis used for data analysis in statistics, biomedical, medicine and econometrics. Koenker and Bassett (1978) used QR for the first time, which does not follow the Gaussian assumption for the model error distribution, but formalised asymptotic properties of the least absolute deviations estimator for independent observations. Many researchers used QR models for longitudinal data , i.e., data collected on the same individuals over time (Fairclough, 2005), in recent years. Diggle *et al.* (2002) introduced three approaches to analyzing longitudinal data using extensions of generalized linear models (GLMs): marginal, random effects, and transition (Markov) models.

QR models are the quantiles of the response variable given covariates. Quantiles are less sensitive to skewed distributions and outliers. A comprehensive guide of the subject can be found in Koenker (2005) and Hao and Naiman (2007). QR estimators determine a linear plan and can be measured efficiently. In the last few years, QR models became a more generally utilised method to define a response variable's distribution given covariates. Karlsson (2007) used a quantile method to analyse the nonlinear longitudinal responses. Hong and He (2010) recommended the transformed ordinal QR for single-index models. Alhamzawi *et al.* (2011) suggested a method for regularisation in mixed QR models. Aghamohammadi and Meshkani (2017) introduced Bayesian quantile regression for skew-normal linear mixed models. Matteo and Emmanuel (2020) compared two conformal inference methods for using quantile regression approach.

Kottas and Krnjajic (2009) considered a semi-parametric method for the quantile regression model using a Bayesian approach. They suggested the Dirichlet process for the error term in an additive quantile regression construction. Zhou *et al.* (2012) used the semi-parametric QR model to decrease the covariates' dimension and keep non-parametric regression flexibility. Cai and Xiao (2015) proposed semi-parametric quantile regression estimation for dynamic models with partially varying coefficients so that the values of some coefficients may be functions of informative covariates. They recommended the estimation of both parametric and non-parametric

© 2020, SRTC Iran

functional coefficients. Zhou (2010) used the QR model with a latent variable approach for the ordinal response variable. Hong and Zhou (2013) introduced a multi-index semi-parametric QR model for ordinal data. For more information about semi-parametric QR models, see Roca and Ordonez (2019), Iren *et al.* (2020) and Bravo (2020). Chen *et al.* (2015) suggested semi-parametric partial linear quantile regression of longitudinal data with time-varying coefficient and informative observation of time. Huang (2016) proposed quantile regression semi-parametric nonlinear mixed-effects models to analyze longitudinal data using a Bayesian approach. Yang et al. (2019) proposed the Bayesian nonlinear QR model for analyzing longitudinal ordinal data. Wang *et al.* (2019) used a new semi-parametric quantile regression approach based on copula for analyzing longitudinal data.

This paper introduces a semi-parametric QR model for analysing longitudinal data using a Bayesian framework where the posterior distribution can be evaluated under various priors. In the proposed approach, we use the Local Polynomial Kernel (LPK) method to approximate the model's non-parametric part. We extended a Metropolis-Hastings algorithm within the Gibbs sampling to sample from the posterior distribution.

This paper is organized as follows: In Section 2, we first introduce the simple semi-parametric QR (SQR) model, and then an SQR model for continuous longitudinal responses is presented. Section 3 specifies prior distributions for unknown parameters and constructs a Bayesian hierarchical model for the SQR model defined in Section 2. Section 4 provides some simulation studies to evaluate the performance of the model. The proposed model is used for analysing a data set (PIAT), and its sensitivity analysis with respect to using different priors is considered in Section 5. Some results and discussion are presented in Section 6.

2 Semi-parametric Quantile Regression Models

2.1 Simple Semi-parametric QR Model

Consider the following semi-parametric model for the responses y_i,

$$y_i = \boldsymbol{x}_i^T \boldsymbol{\beta}_\tau + \eta_{i,\tau}(.) + \varepsilon_{i,\tau}, \qquad i = 1, ..., n \tag{1}$$

where $\boldsymbol{x}_i$ is the $k \times 1$ fixed vector of covariates, $\boldsymbol{\beta}_\tau$ is the $k \times 1$ vector of unknown parameters, $\eta_{i,\tau}(.)$ is the non-parametric part of the model and $\varepsilon_{i,\tau}$ is the error of the model. The error distribution is supposed to be unknown

in the QR model, and it is restricted to have the τ-th quantile to be equal to zero for $0 < \tau < 1$.

A quick approach to estimate $\eta_{i,\tau}(.)$ in (1) is the local polynomial estimating method (Fan , 1992, 1993; Fan and Gijbels , 1992). By Taylor expansion, $\eta_{i,\tau}(.)$ can be approximated with a polynomial of the degree of p. Suppose $\eta_{i,\tau}(.)$ has a $(p+1)$-st continuous derivative for some integer $p \geq 0$ at, for example, t_0; thus we will have

$$\begin{aligned} \eta_{i,\tau}(t_i) &\simeq \eta_{t_0} + (t_i - t_0)\eta_{t_0}^{(1)} + ... + \frac{(t_i - t_0)^p}{p!}\eta_{t_0}^{(p)} \\ &\simeq c_i^T \boldsymbol{\alpha}_\tau \end{aligned} \tag{2}$$

where $\boldsymbol{c}_i = (1, t_i - t_0, (t_i - t_0)^2, ..., (t - t_0)^p)^T$ and $\boldsymbol{\alpha}_\tau = (\eta_{t_0}, \eta_{t_0}^{(1)}, \frac{\eta_{t_0}^{(2)}}{2!} ..., \frac{\eta_{t_0}^{(p)}}{p!})^T$. Consequently in (1) we will have

$$\begin{aligned} y_i &= \boldsymbol{x}_i^T \boldsymbol{\beta}_\tau + \boldsymbol{c}_i^T \boldsymbol{\alpha}_\tau + \varepsilon_{i,\tau} \\ &= \boldsymbol{x}_i^{*T} \boldsymbol{\beta}_\tau^* + \varepsilon_{i,\tau} \end{aligned} \tag{3}$$

where $\boldsymbol{x}_i^* = (\boldsymbol{x}_i^T, \boldsymbol{c}_i^T)^T$ is the new covariate vector and $\boldsymbol{\beta}_\tau^* = (\boldsymbol{\beta}_\tau^T, \boldsymbol{\alpha}_\tau^T)^T$ is the new fixed regression coefficient.
One of the methods to estimate the fixed regression coefficient in the semi-parametric QR model is to proceed with minimizing the sum of the asymmetrically weighted absolute errors

$$\min_{\beta_\tau^*} \sum_{i=1}^{n} \rho_\tau(y_i - \boldsymbol{x}_i^{*T} \boldsymbol{\beta}_\tau^*) w_i \tag{4}$$

where $\rho(.)$ is the loss function (Koenker , 2005) determined by

$$\begin{aligned} \rho_\tau(s) &= s(\tau - I(s < 0)) \\ &= |s|((1 - \tau)I(s < 0) + \tau I(s \geq 0)), \qquad 0 < \tau < 1 \end{aligned} \tag{5}$$

such that $I(.)$ is the indicator function and $w_i > 0$ is the weight allocated to the ith individual. Considering that (4) is not recognizable at the origin, there is no closed-form solution for β^* and a linear programming algorithm can perform the minimization of (4) (Kotz *et al.* , 2001; Koenker , 2004; Fu and Wang , 2012).

© 2020, SRTC Iran

One of the popular methods in quantile regression analysis is to use asymmetric Laplace distribution for the model's errors. Suppose the errors in (1) given a known τ, are independent and $\varepsilon_{i,\tau} \sim ALD(0, \sigma/w_i, \tau)$. Then a continuous random variable y_i given $\boldsymbol{x}_i^*$ has AL distribution with the location parameter $\mu_i = \boldsymbol{x}_i^{*T}\boldsymbol{\beta}_\tau^*$, the scale parameter σ/w_i, and the skewness parameter τ, $ALD(\mu_i, \sigma/w_i, \tau)$. This density function is given by

$$f_\tau(y_i) = \frac{\tau(1-\tau)}{\sigma} exp\left\{-\rho_\tau(\frac{y_i-\mu_i}{\sigma})w_i\right\}, \tag{6}$$

where $y_i, \mu_i \in R, \sigma/w_i > 0, 0 < \tau < 1$. For a given $\tau \in (0,1)$, $Q_\tau(y_i|\boldsymbol{x}_i^*) = F^{-1}_{y_i|\boldsymbol{x}_i^*}(\tau)$ is the τth quantile function of the response variable. It means that

$$F^{ALD}_{y_i|\boldsymbol{x}_i^*}(\mu_i) = p(y_i < \mu_i) = \tau. \tag{7}$$

Equation (7) includes the way to model quantiles of the response variable in the QR analysis by the Bayesian approach. Some other properties of the ALD can be found in Yu and Zhang (2005). The likelihood function of $\boldsymbol{\beta}_\tau^*$ and σ can be composed based on n independent components of $\boldsymbol{y} = (y_1, y_2, ..., y_n)$ as follows:

$$L(\boldsymbol{\beta}_\tau^*, \sigma|\boldsymbol{y}) = \prod_{i=1}^{n} f_\tau(y_i) = \left(\frac{\tau(1-\tau)}{\sigma}\right)^n exp\left\{-\sum_{i=1}^{n}\rho_\tau(\frac{y_i-\mu_i}{\sigma})w_i\right\}. \tag{8}$$

Maximization of Equation (8) with respect to $\boldsymbol{\beta}_\tau^*$ for obtaining maximum likelihood estimate (MLE) of $\boldsymbol{\beta}_\tau^*$ and σ, is equivalent to minimizing the loss function in (4) with respect to β_τ^*. The connection between Bayesian QR model and ALD first proceeded in Yu and Moyeed (2001). The ALD has different mixture representations as shown by Kotz *et al.* (2001). We use a mixture demonstration of the standard exponential distribution and standard normal distribution to generate a sampling algorithm for the QR model. This structure can be found in Kotz *et al.* (2001).

Proposition 1. Let U be a standard exponential variable with mean σ/w, Z be a standard normal variable and it is assumed that U and Z are independent of each other. If a random variable Y follows the ALD with density (6), then we can represent Y as a location-scale mixture of normals given by

$$Y = \sqrt{a\,U}Z + b\,U + \mu \tag{9}$$

where $a = \dfrac{2\sigma/w}{\tau(1-\tau)}$, $b = \dfrac{1-2\tau}{\tau(1-\tau)}$ and $Y \sim ALD(\mu, \sigma/w, \tau)$.

A proof for Proposition 1 can be found in Kozumi and Kobayashi (2011). The equation (9) can be formulated in the hierarchical structure as follows

$$\begin{aligned} Y|U &\sim N(bU + \mu, a\,U), \\ U &\sim exp(\sigma/w). \end{aligned}$$

This arrangement can be used in performing the Bayesian approach.

2.2 Monotone Equivariance Property of QR Models

Authors often transform the variable's scale to get an appropriate model fit or normalize the data. One of the quantiles' features is the monotone equivariance property. See more in ,Hao and Naiman (2007). It means that if we use a monotone transformation g to a random variable, then quantiles of the new random variable are collected using a similar transformation with the quantile of the untransformed random variable. For example, if Q_τ is the τth quantile of the random variable Y then $g(Q_\tau)$ is the τth quantile of $g(Y)$. QR model contains monotone equivariance property as well. It means that if g is a monotone function, then we have

$$Q_\tau\left(g(Y_i)|x_i\right) = g\left(Q_\tau(Y_i|x_i)\right).$$

CMR does not hold the above condition. The QR model can reinterpret by using this feature. For more details, see Hao and Naiman (2007) and Koenker (2005).

2.3 Notation and Model Specification for Longitudinal Responses

Let $\boldsymbol{y} = (y_{ij}, j = 1, ..., n_i, i = 1, ..., n)$, where y_{ij} denotes the continuous response for the ith subject measured at the jth time. For modeling continuous responses, we consider the following semi-parametric mixed-effect model:

$$\begin{aligned} y_{ij} &= \boldsymbol{x_{ij}}^T \boldsymbol{\beta_\tau} + \boldsymbol{Z_i}^T \boldsymbol{b_i} + \eta_\tau(t_{ij}) + \varepsilon_{ij}, \qquad j = 1, ..., n_i, i = 1, ..., n \\ &\boldsymbol{b}_i \sim N_q(\boldsymbol{0}, \boldsymbol{\Sigma_b}), \varepsilon_{ij} \sim ALD(0, \sigma/w_{ij}, \tau) \end{aligned} \tag{10}$$

where $\boldsymbol{x_{ij}} = (x_{ij1}, ..., x_{ijp})^T$ denotes the vector of p covariates observed at time t_{ij}, $\boldsymbol{\beta_\tau} = (\beta_1, ..., \beta_p)^T$ is a $p \times 1$ vector of fixed-effect regression co-

© 2020, SRTC Iran

efficients, $\boldsymbol{Z_i} = (z_{i1}, ..., z_{iq})^T$ is a $q \times 1$ vector of covariate connected with random effects, $\boldsymbol{b_i} = (b_{i1}, ..., b_{iq})^T$ is a $q \times 1$ vector of random-effect regression coefficients, $\eta_\tau(t_{ij})$ is a non-parametric part of the model and ε_{ij} is the error of the model. In model (10) we assumed $\boldsymbol{b_i}$ and ε_{ij} are independent. Then y_{ij} given $\boldsymbol{b_i}$'s has $ALD(\mu_{ij}, \sigma/w_{ij}, \tau)$ where μ_{ij}'s is the mean responses:

$$\mu_{ij} = E(y_{ij}) = \boldsymbol{x_{ij}}^T \boldsymbol{\beta_\tau} + \boldsymbol{Z_i}^T + \eta_\tau(t_{ij}).$$

The covariance structure of (y_{ij}, y_{ik}) by the above assumption is as follows:

$$\begin{aligned} Cov(y_{ij}, y_{ik}) &= Cov[(\boldsymbol{x_{ij}}^T \boldsymbol{\beta_\tau} + \boldsymbol{Z_i}^T \boldsymbol{b_i} + \eta_\tau(t_{ij}) + \varepsilon_{ij}) \\ &\qquad , (\boldsymbol{x_{ik}}^T \boldsymbol{\beta_\tau} + \boldsymbol{Z_i}^T \boldsymbol{b_i} + \eta_\tau(t_{ik}) + \varepsilon_{ik})] \\ &= Cov(\boldsymbol{Z_i}^T \boldsymbol{b_i}, \boldsymbol{Z_i}^T \boldsymbol{b_i}) = \boldsymbol{Z_i}^T Var(\boldsymbol{b_i}) \boldsymbol{Z_i} \\ &= \boldsymbol{Z_i}^T \boldsymbol{\Sigma_b} \boldsymbol{Z_i}, \end{aligned}$$

Therefore, the covariance structure of longitudinal responses for each individual is taken into account by using random-effect models.

The likelihood function of the unknown parameters in model (10) for given τ and $Y_i = (y_{i1}, y_{i2}, ..., y_{in_i})^T$ can be formulated as

$$\begin{aligned} L(\boldsymbol{\beta}, \sigma | \boldsymbol{Y_i}, \boldsymbol{b_i}) = f(\boldsymbol{Y_i} | \boldsymbol{\beta}, \sigma, \boldsymbol{b_i}) &= \prod_{j=1}^{n_i} f(y_{ij} | \boldsymbol{\beta}, \sigma, \boldsymbol{b_i}), \\ &= \prod_{j=1}^{n_i} \left(\frac{\tau(1-\tau)}{\sigma/w_{ij}} \right) \exp\left\{ -\rho_\tau(\frac{y_{ij} - \mu_{ij}}{\sigma} w_{ij}) \right\}, \\ &= \left(\prod_{j=1}^{n_i} \frac{\tau(1-\tau)}{\sigma/w_{ij}} \right) \exp\left\{ -\sum_{j=1}^{n_i} \rho_\tau(\frac{y_{ij} - \mu_{ij}}{\sigma}) w_{ij} \right\}. \end{aligned}$$

And

$$\begin{aligned} L(\boldsymbol{\beta}, \sigma | \boldsymbol{Y}, \boldsymbol{b_i}) = f(\boldsymbol{Y} | \boldsymbol{\beta}, \sigma, \boldsymbol{b}) &= \prod_{i=1}^{n} f(y_i | \boldsymbol{\beta}, \sigma, b_i) \\ &\propto \exp\left\{ -\sum_{i=1}^{n} \sum_{j=1}^{n_i} \rho_\tau(\frac{y_{ij} - \mu_{ij}}{\sigma}) w_{ij} \right\}, \end{aligned}$$

such that $\boldsymbol{Y} = \left(\boldsymbol{Y_1^T}, \boldsymbol{Y_2^T}, ..., \boldsymbol{Y_n^T}\right)^T$ and $\boldsymbol{b} = \left(\boldsymbol{b_1^T}, \boldsymbol{b_2^T}, ..., \boldsymbol{b_n^T}\right)^T$.

Before estimating the parameters, we specify the approximation method of the non-parametric part of the model ($\eta_\tau(t_{ij})$). One of these methods is a Local Polynomial estimation. In this method $\eta_\tau(t_{ij})$ can be approximated with a polynomial of degrees p by Taylor Expansion. Suppose $\eta_\tau(t_{ij})$ has a $(p+1)$-st continuous derivative for some integer $p \geq 0$ at, for example, t_0; then we will have

$$\begin{aligned}\eta_\tau(t_{ij}) &\simeq \eta_{t_0}(\tau) + (t_{ij}-t_0)\eta_{t_0}^{(')}(\tau) + ... + \frac{(t_{ij}-t_0)^p}{p!}\eta_{t_0}^{(p)}(\tau),\\ &\simeq \boldsymbol{c}_{ij}^T\boldsymbol{\alpha}_\tau \end{aligned} \tag{11}$$

where

$$\begin{aligned}\boldsymbol{c_{ij}} &= (1, t_{ij}-t_0, (t_{ij}-t_0)^2, ..., (t_{ij}-t_0)^p)^T,\\ \boldsymbol{\alpha_\tau} &= (\eta_{t_0}(\tau), \eta_{t_0}^{(')}(\tau), \frac{\eta_{t_0}^{('')}(\tau)}{2!}, ..., \frac{\eta_{t_0}^{(p)}(\tau)}{p!})^T.\end{aligned}$$

consequently we will have

$$\begin{aligned}\mu_{ij} = E(y_{ij}) &\cong \boldsymbol{x_{ij}}^T\boldsymbol{\beta_\tau} + \boldsymbol{Z_i}^T\boldsymbol{b_i} + \boldsymbol{c}_{ij}^T\boldsymbol{\alpha_\tau},\\ &\cong \boldsymbol{x_{ij}^*}^T\boldsymbol{\beta_\tau^*} + \boldsymbol{Z_i}^T\boldsymbol{b_i},\\ &\cong \tilde{\mu}_{ij}\end{aligned} \tag{12}$$

where $\boldsymbol{x}_{ij}^* = (\boldsymbol{x}_{ij}^T, \boldsymbol{c}_{ij}^T)^T$ is the new covariate vector and $\boldsymbol{\beta}_\tau^* = (\boldsymbol{\beta}_\tau^T, \boldsymbol{\alpha}_\tau^T)^T$ is the new regression coefficient that can be estimated by solving

$$\min_{\beta_\tau^*} \sum_{i=1}^{n}\sum_{j=1}^{n_i} \rho_\tau(y_{ij} - \boldsymbol{x}_{ij}^{*T}\boldsymbol{\beta}_\tau^* - \boldsymbol{Z_i}^T\boldsymbol{b_i})w_{ij} \tag{13}$$

where $\rho(.)$ is the check loss function defined in (5), $0 < \tau < 1$ and w_{ij} is the kernel function allocated to the i-th individual for time t_{ij} such that $w_{ij} = K_h(t_{ij}-t_0) = \frac{K(.|h)}{h}$ and $h > 0$ is called a bandwdith or smothing parameter, in which the kernel K and the bandwidth h are assumed to be fixed. In practice, h needs to be carefully chosen. When h is too small, the resulting estimators $\widehat{\eta}_{ij}$ are usually very noisy, and when h is too large, $\widehat{\eta}_{ij}$ may oversmooth the data so that some important information in the data is not sufficiently captured. In this paper, the "leave- one-subject-out" cross-validation (SCV) criterion is used to select a proper bandwidth for estimating

© 2020, SRTC Iran

η_{ij}. Some different methods can be found in Wu and Zhang (2006). When fitting the local polynomial kernel, the selection of the p is not as important as the h. So, after choosing h, an important choice is a polynomial degree. Boundary bias considerations indicate that one should, at the least, fit local lines. But local line estimators can have a substantial bias to estimate regions of high curvature of the function, such as at peaks and valleys. This problem can be reduced by fitting higher-degree polynomials such as quadratics and cubics. However, there are costs in terms of increased computational complexity that need to be considered. Nevertheless, we have noticed that gains can often be made using higher-degree fits in many examples that we have tried. This is our primary motivation for developing beyond linear fits (Fan *et al.* , 1995).

We use a mixture demonstration of the standard exponential distribution and standard normal distribution to generate a sampling algorithm for the QR model. This structure can be found in Kotz *et al.* (2001). Since the structure of the ALD is not simple to trade with, then we do its random mixture representation (see equation (9)) in model (10) to facilitate some scientific and statistical computations. By studying the subsequent illustration for error in model (10), we have

$$\varepsilon_{ij} = \sqrt{au_{ij}}\, z_{ij} + bu_{ij}$$

where $a = \dfrac{2\sigma/w_{ij}}{\tau(1-\tau)}$, $b = \dfrac{1-2\tau}{\tau(1-\tau)}$, u_{ij}s have standard exponential distribution with mean σ/w_{ij} and z_{ij}s have standard normal distribution. Therefore, the conditional distribution of ε_{ij} given u_{ij} is

$$\varepsilon_{ij}\,|u_{ij} \sim N\left(bu_{ij}, au_{ij}\right)$$

then

$$y_{ij}\,|u_{ij}, b_i \sim N\left(\mu^*_{ij}, au_{ij}\right) \tag{14}$$

where $\mu^*_{ij} = bu_{ij} + \mu_{ij}$.

The likelihood function of the refreshed parameters for continuous QR

model is given by

$$L_y\left(\boldsymbol{\beta}_\tau, \boldsymbol{\alpha}_\tau, \sigma \left|\boldsymbol{Y}, \boldsymbol{b}, \boldsymbol{u}\right.\right) = f\left(\boldsymbol{Y}\left|\boldsymbol{\beta_\tau}, \boldsymbol{\alpha}_\tau, \sigma, \boldsymbol{b}, \boldsymbol{u}\right.\right) = \prod_{i=1}^{n}\prod_{j=1}^{n_i} f\left(y_{ij}\left|\boldsymbol{\beta}_\tau, \boldsymbol{\alpha}_\tau, \sigma, \boldsymbol{b_i}, u_{ij}\right.\right) \tag{15}$$

where $\boldsymbol{Y} = \left(\boldsymbol{Y_1^T}, \boldsymbol{Y_2^T}, ..., \boldsymbol{Y_n^T}\right)^T$, $\boldsymbol{b} = \left(\boldsymbol{b_1^T}, \boldsymbol{b_2^T}, ..., \boldsymbol{b_n^T}\right)^T$,$\boldsymbol{u} = \left(\boldsymbol{u_1^T}, \boldsymbol{u_2^T}, ..., \boldsymbol{u_n^T}\right)^T$,and

$$f\left(y_{ij}\left|\boldsymbol{\beta}_\tau, \boldsymbol{\alpha}_\tau, \sigma, \boldsymbol{b_i}, u_{ij}\right.\right) \propto \exp\left\{\sum_{i,j} -\rho\left(\frac{y_{ij}-\mu_{ij}^*}{\sigma}\right) K_h(t_{ij}-t_0)\right\}.$$

3 Bayesian Inferences

This section performs the Bayesian approach using MCMC methods to estimate the previous section's model parameters. Bayesian approaches provide convenient alternative inference tools for quantile regression. Even though conventional quantile regression does not need any parametric distributional assumptions, a working likelihood is required to carry out Bayesian analysis. Suppose $\boldsymbol{\Theta} = (\boldsymbol{\beta}_\tau, \boldsymbol{\alpha}_\tau, \sigma, \boldsymbol{b}, \boldsymbol{u})$ is the vector of unknown parameters in (15). We allocate independent prior distributions for each element of $\boldsymbol{\Theta}$. We consider the following priors in the Bayesian hierarchical framework to the SQR model parameters for $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, n_i$ as

$$\begin{aligned}
&y_{ij}|b_i, u_{ij}, \boldsymbol{\beta}_\tau, \sigma \sim N\left(\mu_{ij}^*, au_{ij}\right),\\
&\boldsymbol{b}_i|\boldsymbol{\Sigma_b} \sim N\left(\boldsymbol{0}, \boldsymbol{\Sigma_b}\right),\\
&u_{ij} \sim \exp(\sigma/w_{ij})\\
&\boldsymbol{\alpha}_\tau \sim N\left(\boldsymbol{\mu}_\alpha, \boldsymbol{\Sigma}_\alpha\right), \quad \boldsymbol{\beta}_\tau \sim N\left(\boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta\right),\\
&\boldsymbol{\Sigma_b} \sim IW\left(\boldsymbol{\Omega_b}, \boldsymbol{\nu_b}\right), \quad \sigma \sim IG\left(c, d\right),
\end{aligned} \tag{16}$$

such that μ_{ij}^*, w_{ij} and a are defined in the previous sections and c and d are constants. IG and IW also indicate inverse gamma and inverse Wishart distributions, respectively. An MCMC method, such as the Gibbs sampling algorithm, can perform Bayesian analysis as a numerical method. We need to determine the full conditional distributions of parameters to receive this goal. Since the posterior distributions are analytically intractable, we can use

© 2020, SRTC Iran

a Metropolis-Hastings algorithm to generate the parameters' full-conditional distributions.

3.1 Convergence Diagnostic

One of the most critical steps of the Bayesian MCMC algorithm is to determine the chain's convergence and the burn-in period's length, typically assessed using diagnostics tools. For this aim, we use several chains with several initial values and then analyse their behaviours. Therefore chains will be converged when they begin to generate the same values. The Gelman–Rubin diagnostic (Gelman and Rubin , 1992) assesses MCMC convergence by analysing the difference between multiple Markov chains. This method evaluated the convergence by comparing the estimated Inter-chain and Intra-chains variances for each model parameter. Significant differences between these variances indicate non-convergence. Let M be the number of chains with length N, although the chains may be of different lengths. The same-length opinion is used for convenience and simplifies the formulas. For a model parameter θ, let $\{\theta_{mt}\}_{t=1}^{N}$ be the mth simulated chain, $m = 1, \ldots, M$. Let $\hat{\theta}_m$ and $\hat{\sigma}_m^2$ be the sample posterior mean and variance of the mth chain, and let the overall sample posterior mean be $\hat{\theta} = (1/M)\sum_{m=1}^{M}\hat{\theta}_m$. The between-chains and within-chain variances are given by

$$B = \frac{N}{M-1}\sum_{m=1}^{M}(\hat{\theta}_m - \hat{\theta})^2$$

$$W = \frac{1}{M}\sum_{m=1}^{M}\hat{\sigma}_m^2$$

Under certain stationarity conditions, the pooled variance is as follows

$$\widehat{V} = \frac{N-1}{N}W + \frac{M+1}{MN}B$$

then

$$\widehat{R} = \frac{\widehat{V}}{W}$$

The chains will be converging if the produced data-size is large. The Intra-chain variance should be relatively small than the Inter-chain variance; thus, R goes to one. In other words, R values greater than one show non-convergence.

3.2 Deviation Information Criteria (DIC)

DIC is a metric used for model comparison in the Bayesian approach. This criterion is a reduced version of the DIC introduced by Spiegelhalter *et al.* (2002). DIC consists of two parts based on the posterior distribution. The first part, the posterior mean, measures the goodness of fit and the second part, the effective number of parameters, is the penalty for increasing model complexity. The DIC formulation is:

$$DIC = \bar{D}(\theta) + p_D,$$

where $p_D = \bar{D}(\theta) - D(\bar{\theta})$, and $D(\theta) = E_{\theta|y} - 2\left[log\left(f(\boldsymbol{y}|\theta)\right)\right]$. $\bar{D}(\theta)$ and $\bar{\theta}$ are the posterior mean estimates of $D(\theta)$ and θ, respectively. Additionally $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$ is the complete data with density $f(\boldsymbol{y}\,|\theta)$ and θ is the vector of unknown parameters. We can obtain a Monte Carlo approximation of DIC utilizing the MCMC samples of the posterior distribution. The lowest value of DIC intends a better fit of a model to the observations.

4 Simulation

To specify the proposed SQR model's performance, we carry out some MCMC simulation studies. Three scenarios for simulation study are considered such that the continuous response distribution is ALD with skew parameters $\tau = 0.25, 0.50$, and 0.75. Each generated dataset is fitted by a semi-parametric Quantile Regression Model (SQRM) for $\tau = 0.25$, $\tau = 0.5$ and $\tau = 0.75$ and later we compare performance of our method with that of a standard method. These datasets are analyzed by a Classic Mean Regression Model (CMRM). The models' performances are evaluated using two criteria: the relative bias (R.Bias) and mean-squared errors (MSE). These are

$$R.Bias = \frac{1}{N}\sum_{i=1}^{N}\frac{(\hat{\theta}_i - \theta)}{\theta}$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{\theta}_i - \theta\right)^2$$

where $\hat{\theta}_i$ is the estimate of θ for the ith sample, and N is the number of simulated samples ($N = 100$ in this section). We run the Gibbs sampler algorithm for 10,000 iterations and disregarding the first 5000 iterations to

© 2020, SRTC Iran

reduce the initial values' effect. The last 5000 iterations are considered to get the Bayesian estimation. To fit our suggested model to produce data in a Bayesian structure, we use non-informative prior distributions to estimate parameters in Equation (16) as follows

$$\begin{aligned}(\boldsymbol{\mu_\alpha}, \boldsymbol{\Sigma_\alpha}) &= (\mathbf{0}_3, 100\boldsymbol{I}_3)\\ (\boldsymbol{\mu_\beta}, \boldsymbol{\Sigma_\beta}) &= (\mathbf{0}_3, 100\boldsymbol{I}_3)\\ (c, d) &= (0.01, 0.01)\\ \boldsymbol{\Sigma}_b = \sigma_b &\sim IG(0.01, 0.01).\end{aligned}$$

The MCMC method is performed using OpenBUGS software which associates with R software by the R2OpenBUGS package. This software is accessible online. See R core Team (2013).

4.1 Simulation 1

Three scenarios are considered in this simulation study, i.e. the continuous response is assumed to have the AL distribution for $\tau = 0.25, 0.50$, and 0.75. We generate 100 data sets (iterations) based on the above mentioned three cases. Each produced dataset is fitted by $SQRM(\tau = 0.25), SQRM(\tau = 0.5), SQRM(\tau = 0.75)$ and $CMRM$. The longitudinal studies are considered with $n = 100, n = 200$ and $n = 500$ subjects in which the continuous response has measured four times. The continuous longitudinal responses are sampled from the following SQRM:

$$y_{ij} = \beta_0 + \beta_1 x_{ij1} + \beta_2 x_{ij2} + \eta_{ij}(t) + b_i + \varepsilon_{ij}^{ALD}, i = 1, 2, ..., n, j = 1, 2, 3, 4$$

where $(\beta_0, \beta_1, \beta_2) = (-1, 3, 2), t_j = j$ and $\eta_{ij}(t)$ is approximated with polynomial, by Taylor expansion, to be $\alpha_1 t_j + 0.5\alpha_2 t_j^2 + \frac{1}{6}\alpha_3 t_j^3$. Also we assume $(\alpha_1, \alpha_2, \alpha_3) = (-1, 1, 2)$. x_{ij1} and x_{ij2} are generated from Bernoulli distribution with $p = 0.5$ and standard normal distribution, respectively. ε_{ij}^{ALD}s sampled from AL distribution. Simultaneously, two MCMC chains are conducted with different initial values for 10,000 iterations, disregarded the first 5000 iterations, the Bayesian inference is done. For each iteration, the convergence of chains is controlled by using the Gelman-Rubin diagnostic test. If the chains are not converged based on this test, in each iteration, then the sample is discarded.

Table 1 contains estimates of parameters, standard deviations, relative bias and root of MSE (RMSE) for the SQRM(0.25) and CMRM with $n =$

Table 1. Results of simulation study from fitting SQRM(0.25) and CMRM to generated data by SQRM(0.25).

			SQRM(0.25)				CMRM			
n	Par	Real	Est.	SD	R.Bias	RMSE	Est.	SD	R.Bias	RMSE
100	β_0	-1	-0.818	0.824	-0.182	0.844	-1.273	0.836	0.273	0.880
	β_1	3	2.781	0.522	-0.073	0.566	2.185	0.362	-0.272	0.891
	β_2	2	2.001	0.214	$5e^{-4}$	0.214	1.510	0.180	-0.245	0.522
	α_1	-1	-0.847	0.927	-0.153	0.940	-0.641	0.884	-0.359	0.954
	α_2	1	1.177	1.205	0.177	1.218	-0.334	0.981	-1.334	1.656
	α_3	2	2.282	0.326	0.141	0.431	1.521	0.805	-0.240	0.937
	σ_b	1	0.879	0.886	-0.121	0.894	0.877	0.563	-0.123	0.576
	σ	1	0.720	0.010	-0.280	0.280	1.927	0.153	0.927	0.940
200	β_0	-1	-0.894	0.637	-0.106	0.646	-1.572	0.857	0.572	1.030
	β_1	3	2.952	0.388	-0.016	0.391	3.081	0.329	0.027	0.339
	β_2	2	1.976	0.197	-0.012	0.199	2.002	0.165	0.001	0.165
	α_1	-1	-0.774	0.848	-0.226	0.877	-0.882	0.994	-0.118	1.001
	α_2	1	1.806	0.832	0.806	1.158	-0.139	1.102	-1.139	1.585
	α_3	2	2.131	0.388	0.066	0.409	2.050	0.346	0.025	0.350
	σ_b	1	0.856	0.850	-0.144	0.862	0.891	0.512	-0.109	0.524
	σ	1	0.910	0.088	-0.090	0.125	0.496	0.140	-0.504	0.524
500	β_0	-1	-0.845	0.497	-0.155	0.521	-1.714	0.680	0.714	0.986
	β_1	3	3.018	0.237	0.006	0.237	3.033	0.212	0.011	0.214
	β_2	2	2.000	0.118	0.000	0.118	2.010	0.105	0.005	0.106
	α_1	-1	-1.096	0.717	0.096	0.723	-1.021	0.907	0.021	0.907
	α_2	1	1.841	0.754	0.841	1.129	1.131	0.799	0.131	0.809
	α_3	2	2.046	0.316	0.023	0.319	1.995	0.322	$-2.5e^{-3}$	0.322
	σ_b	1	1.113	0.376	0.113	0.393	1.100	0.394	0.100	0.406
	σ	1	0.898	0.005	-0.102	0.102	1.487	0.092	0.487	0.496

Par: parameters, Est.: estimate of parameter, S.D.: posterior standard deviation

R.bias: relative biase, RMSE: root of mean-square errors.

© 2020, SRTC Iran

Table 2. Results of simulation study from fitting SQRM(0.5) and CMRM to generated data by SQRM(0.5).

			SQRM(0.50)				CMRM			
n	Par	Real	Est.	SD	R.Bias	RMSE	Est.	SD	R.Bias	RMSE
100	β_0	-1	-1.093	0.677	0.093	0.683	-1.107	0.860	0.107	0.867
	β_1	3	3.127	0.297	0.042	0.323	3.003	0.344	0.001	0.344
	β_2	2	2.006	0.221	0.003	0.221	1.973	0.170	-0.013	0.172
	α_1	-1	-1.026	0.895	0.026	0.895	-0.821	1.035	-0.180	1.050
	α_2	1	0.866	0.836	-0.134	0.847	-0.664	0.594	-1.664	1.767
	α_3	2	2.028	0.361	0.014	0.362	2.068	0.362	0.034	0.369
	σ_b	1	1.281	0.867	0.281	0.911	0.967	0.448	-0.033	0.449
	σ	1	0.951	0.155	-0.049	0.162	1.147	0.090	0.147	0.172
200	β_0	-1	-1.018	0.521	0.018	0.521	-1.110	0.704	0.110	0.712
	β_1	3	3.059	0.308	0.020	0.313	3.079	0.245	0.026	0.257
	β_2	2	2.000	0.155	0.000	0.155	2.002	0.122	0.001	0.122
	α_1	-1	-1.001	0.732	0.001	0.732	-0.874	0.922	-0.127	0.930
	α_2	1	0.986	0.793	-0.014	0.794	-0.137	0.893	-1.137	1.446
	α_3	2	1.999	0.336	$5e^{-4}$	0.336	2.047	0.327	0.023	0.330
	σ_b	1	0.833	0.570	-0.167	0.594	0.973	0.330	-0.027	0.331
	σ	1	0.856	0.016	-0.144	0.145	1.115	0.065	0.115	0.132
500	β_0	-1	-0.858	0.381	-0.142	0.407	-0.882	0.524	-0.118	0.537
	β_1	3	3.018	0.194	0.006	0.195	3.028	0.157	$9e^{-3}$	0.159
	β_2	2	1.995	0.098	$-2.5e^{-3}$	0.098	2.006	0.115	0.003	0.115
	α_1	-1	-1.169	0.564	0.169	0.589	-1.148	0.727	0.148	0.742
	α_2	1	1.275	0.482	0.275	0.555	2.311	0.336	1.311	1.353
	α_3	2	1.958	0.259	-0.021	0.262	1.947	0.259	-0.027	0.265
	σ_b	1	1.071	0.211	0.071	0.222	1.073	0.211	0.073	0.223
	σ	1	1.101	0.021	0.101	0.103	1.101	0.041	0.101	0.108

Par: parameters, Est.: estimate of parameter, S.D.: posterior standard deviation
R.bias: relative biase, RMSE: root of mean-square errors.

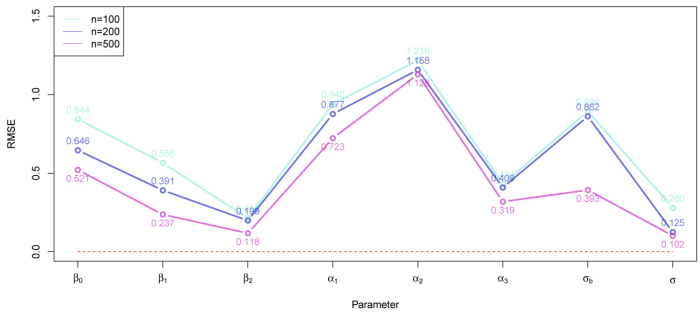

Figure 1. RMSE of parameter estimates in the SQRM(0.25) model for n = 100, 200, and 500, (parameter estimates are given inside the plot with the sign "∘").

$100, n = 200$ and $n = 500$ subjects. Despite the low standard deviation of the estimated parameters in CMRM, by increasing the value of n, the performance of the SQRM is better than that of the CMRM because it has lower relative biases and RMSE among the estimated parameters. Table 2 and Table 3 contain estimates of parameters, standard deviations, relative bias and RMSE for SQRM(0.5) and CMRM and SQRM(0.75) and CMRM, respectively. The results from Tables 2 and 3 give interpretations similar to those in Table 1. However, Table 2 shows that CMRM performance is very close to that of the SQRM(0.50) when data are generated from SQRM(0.50). Figure 1 shows how the RMSE varies with the variation of the sample size n in the SQRM(0.25) model when data are generated by SQRM(0.25).

4.2 Simulation 2

To assess the performance of the proposed SQRM models, we use another approach to generate the data. The structure for this simulation study is the same as that of the simulation 1, but ε_{ij}^{ALD} in model of Eq.(10) is generated from two other distributions i.e. skew-normal (SN) distribution and mixture of normal (MN) distribution as follows

$$\varepsilon_{ij}^{SN} \sim SN(0, 1, 3)$$
$$\varepsilon_{ij}^{MN} \sim 0.4N(0, 1) + 0.6N(0, 1)$$

© 2020, SRTC Iran

Table 3. Results of simulation study from fitting SQRM(0.75) and CMRM to generated data by SQRM(0.75).

			SQRM(0.75)				CMRM			
n	Par	Real	Est.	SD	R.Bias	RMSE	Est.	SD	R.Bias	RMSE
100	β_0	-1	-1.259	0.860	0.259	0.898	-1.747	1.102	0.747	1.331
	β_1	3	3.070	0.527	0.023	0.532	2.956	0.669	-0.015	0.670
	β_2	2	2.054	0.275	0.027	0.280	2.051	0.235	0.025	0.240
	α_1	-1	-1.091	1.066	0.091	1.070	-0.782	1.175	-0.218	1.195
	α_2	1	0.738	0.974	-0.262	1.017	-0.260	0.185	-1.260	1.274
	α_3	2	1.982	0.455	-0.009	0.456	2.012	0.406	0.006	0.406
	σ_b	1	0.931	0.854	-0.069	0.856	1.063	0.701	0.063	0.704
	σ	1	0.911	0.091	-0.089	0.127	1.422	0.192	0.422	0.463
200	β_0	-1	-1.430	0.637	0.430	0.769	-1.713	0.907	0.713	1.154
	β_1	3	3.054	0.375	0.018	0.379	3.107	0.427	0.036	0.440
	β_2	2	2.009	0.187	$4.5e^{-3}$	0.187	2.006	0.194	0.003	0.194
	α_1	-1	-0.885	0.849	-0.115	0.857	-0.978	1.098	-0.022	1.098
	α_2	1	0.799	0.902	-0.201	0.924	0.714	0.903	-0.286	0.947
	α_3	2	1.995	0.395	$-2.5e^{-3}$	0.395	2.015	0.384	$7.5e^{-3}$	0.384
	σ_b	1	0.821	0.817	-0.179	0.836	0.930	0.523	-0.070	0.528
	σ	1	0.888	0.006	-0.112	0.113	1.435	0.109	0.435	0.448
500	β_0	-1	-1.349	0.479	0.349	0.592	-1.693	0.666	0.693	0.961
	β_1	3	3.037	0.209	0.012	0.212	3.036	0.233	0.012	0.236
	β_2	2	1.992	0.104	-0.004	0.104	2.002	0.117	0.001	0.117
	α_1	-1	-0.938	0.710	-0.062	0.713	-0.924	0.877	-0.076	0.880
	α_2	1	0.699	0.411	-0.301	0.510	0.438	0.500	-0.562	0.752
	α_3	2	2.010	0.310	0.005	0.310	2.017	0.322	$8.5e^{-3}$	0.322
	σ_b	1	1.146	0.203	0.146	0.250	1.034	0.387	0.034	0.389
	σ	1	0.916	0.005	-0.084	0.085	1.426	0.090	0.426	0.435

Par: parameters, Est.: estimate of parameter, S.D.: posterior standard deviation
R.bias: relative biase, RMSE: root of mean-square errors.

Table 4. Results (relative biases and RMSEs) of simulation study from fitting SQRM(τ), $\tau = 0.25, 0.50$ and 0.75 and CMRM to generated data in simulation 2.

	SQRM(0.25)		SQRM(0.50)		SQRM(0.75)		CMRM	
parameters	R.Bias	RMSE	R.Bias	RMSE	R.Bias	RMSE	R.Bias	RMSE
data set is generated from skew normal Model								
β_0	-0.673	0.722	-0.711	0.769	-0.676	0.763	-0.687	0.771
β_1	0.016	0.214	0.023	0.228	0.026	0.239	0.026	0.238
β_2	0.004	0.109	0.003	0.114	0.004	0.112	0.003	0.112
α_1	0.113	0.372	-0.017	0.406	-0.263	0.579	-0.241	0.567
α_2	0.710	1.402	-0.218	1.023	-0.176	1.032	-0.967	1.364
α_3	-0.001	0.162	0.006	0.183	0.033	0.229	0.028	0.232
σ_b	0.170	0.247	0.212	0.283	0.226	0.294	0.215	0.285
σ	-0.970	0.970	-0.964	0.964	-0.969	0.969	-0.970	0.970
data set is generated from mixture of normal Model								
β_0	-0.926	1.360	-0.179	1.167	-0.254	1.262	-0.217	1.627
β_1	-0.105	0.960	-0.079	0.856	-0.084	0.962	-0.045	0.754
β_2	0.032	0.493	0.006	0.435	0.004	0.486	-0.045	0.384
α_1	0.490	1.294	0.215	1.234	0.030	1.272	0.246	1.443
α_2	-1.175	1.461	-1.077	1.607	-1.602	1.890	-0.775	1.076
α_3	0.111	0.575	0.011	0.568	-0.039	0.579	-0.019	0.482
σ_b	0.228	0.515	0.222	0.875	0.226	0.876	0.137	0.869
σ	-0.272	0.272	-0.298	0.299	-0.370	0.371	1.427	1.451

R.bias: relative biase, RMSE: root of mean-square errors.

where $SN(\xi, \omega, \lambda)$ denotes the skew-normal distribution with location papameter ξ, scale parameter ω and skew parameter λ.

Table4 shows the results of simulation studies for the models SQRM(0.25), SQRM(0.50), SQRM(0.75), and CMRM. The relative biases and root of MSEs using the suggested models for generating the data sets are obtained. It can be seen that there is some τ, which gives good results in each of the above cases. In both data sets generated from the skew-normal model and the mixture of normal model, the maximum of absolute value of R.Bias of parameters and the maximum of RMSE of parameters in SQRM(0.50) are fewer than other models. Therefore, some τ (here $\tau = 0.5$) works better than others. This model gives lower R.Biases and RMSEs for the estimates of parameters than that of CMRM, even though the model's error does not follow from AL distribution, except for some parameters, e.g., the estimate of α_2.

© 2020, SRTC Iran

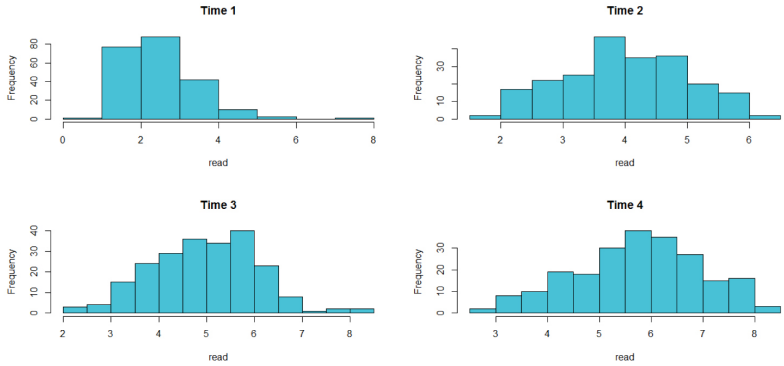

Figure 2. Histogram of child's reading recognition skill (read) at four time-points of PIAT data set .

5 Application

In this section, to specify the performance of the proposed SQRM, we analyse a set of real data i.e. the Peabody Individual Achievement Test (PIAT) data set. The PIAT data set was previously analyzed by Samani and Ganjali (2011), Noorian *et al.* (2016), and Ghasemzadeh *et al.* (2020). The PIAT data were prepared at two-year intervals between 1986 and 1992 and contains both the mother and her child's interviews. This paper has selected one important continuous response from the dataset as the primary variable for modelling. The continuous response is the child's reading recognition skill (read). Figure 2 shows that the distributions of continuous response at four time- points are almost asymmetric. Then, it is reasonable to use quantile regression for the analysis of this data.

The Peabody Individual Achievement Test (PIAT) Reading Recognition subtest, one of five in the PIAT series, measures word recognition and pronunciation ability, essential reading achievement components. Children read a word silently, then say it aloud. PIAT Reading Recognition contains 84 items, each with four options, increasing difficulty from preschool to high school levels. Skills assessed include matching letters, naming names, and reading single words aloud. The reading recognition measure was computed by summing the total number of correct items for the 84 items subtest, and scores could range in value from 0 to 84. The final reading recognition scores were divided by 10.

These data consist of 221 children and mothers who had perfect answers at all times. These individuals data have been studied as complete data in this paper. Dataset also includes some covariates which will be used in our analysis: child's gender (*Gender*, girl=0 and boy=1), mother's age at first time-point (*Mage*, from age 21 up to age 29), child's age at first time-point (*Kage*, from age six up to age 8) and time (*Time*=$1, 2, 3, 4$ corresponds to the first, second, third and fourth time-points, respectively).

5.1 Model Implementation and Results of Data Analysis

We assume the τth quantile regression model has the following semi-parametric model, **for** $i = 1, 2, \ldots, 221$ **and** $j = 1, 2, 3, 4$

$$Q_\tau\left(read_{ij}|\boldsymbol{x}, t_{ij}\right) = \beta_0 + \beta_1 Gender_i + \beta_2 Kage_i + \beta_3 Mage_i + \eta(t_{ij}) + b_i$$

where $\eta(t_{ij})$ is a non-parametric part of the model approximated by the degree 3 local polynomials. b_i is a random effect in the model that has a normal distribution with parameters $(0, \sigma_b^2)$. We fit four different models to PIAT data. In order to compare the models, the DIC criterion is used. Fitted models are: CMR and the τth SQR model for responses by taking $\tau = 0.25, 0.50$ and 0.75. To implement the Bayesian method, The non-informative priors are employed as follows

$$\begin{aligned}
&\boldsymbol{\beta} \sim N_4(\mathbf{0}_4, 100\boldsymbol{I}_4), \qquad \boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_3)^{'},\\
&\boldsymbol{\alpha} \sim N_3(\mathbf{0}_3, 100\boldsymbol{I}_3), \qquad \boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)^{'},\\
&\sigma_b \sim IGamma(0,01, 0,01),\\
&\sigma \sim IGamma(0,01, 0,01).
\end{aligned}$$

The prior distributions used in the Bayesian implementation were similar to the simulation study's prior distributions. We run two parallel MCMC chains with distinctive initial values for 60,000 iterations, and ignore the first 30,000 iterations to do the Bayesian inference. Table 5 gives posterior mean, standard deviation, and DIC. According to the results of Table 5, SQRM(0.25) has the lowest DIC, and by this model, gender has no significant effects on reading ability. The age of the mother and her child have meaningful effects on reading ability. It means that the older mothers have

© 2020, SRTC Iran

Table 5. Results of real data from fitting SQRM($\tau = 0.25, 0.5$ and 0.75) and CMRM to PIAT data.

	SQRM($\tau = 0.25$)			SQRM($\tau = 0.50$)		
parameter	Est.	S.D.	95% PCI	Est.	S.D.	95% PCI
β_0	-6.015	0.805	(-7.657,-4.469)	-5.549	1.000	(-7.637,-3.841)
β_1	-0.115	0.106	(-0.192,0.094)	-0.107	0.106	(-0.313,0.103)
β_2	0.097	0.023	(0.051,0.098)	0.100	0.031	(0.047,0.169)
β_3	0.499	0.078	(0.343,0.648)	0.476	0.085	(0.296,0.627)
α_1	3.040	0.464	(2.131,3.852)	2.776	0.403	(2.043,3.540)
α_2	-12.590	0.205	(-13.01,-11.32)	-11.968	0.635	(-13.99,-9.627)
α_3	0.346	0.172	(0.005,0.643)	0.221	0.147	(0.005,0.569)
σ	0.056	0.002	(0.052,0.061)	0.067	0.002	(0.062,0.072)
σ_b	0.5363	0.062	(0.427,0.668)	0.530	0.062	(0.419,0.663)
DIC		1291			1321	
	SQRM($\tau = 0.75$)			CMRM		
parameter	Est.	S.D.	95% PCI	Est.	S.D.	95% PCI
β_0	-4.84	0.828	(-6.423,-3.318)	-5.651	0.646	(-6.17,-4.385)
β_1	-0.089	0.106	(-0.296,0.119)	-0.083	0.065	(-0.210,0.044)
β_2	0.089	0.025	(0.041,0.141)	0.528	0.052	(0.426,0.631)
β_3	0.465	0.073	(0.324,0.628)	0.528	0.046	(0.431,0.612)
α_1	2.522	0.403	(1.675,3.218)	2.116	0.423	(1.342,2.627)
α_2	-8.968	0.635	(-10.34,-6.648)	-7.476	0.585	(-8.696,-6.649)
α_3	0.221	0.144	(0.010,0.243)	0.241	0.152	(0.001,0.486)
σ	0.057	0.002	(0.053,0.061)	0.012	0.003	(0.006,0.018)
σ_b	0.528	0.061	(0.420,0.657)	0.056	0.004	(0.048,0.063)
DIC		1322			2349	

Est.: posterior mean, S.D.: posterior standard deviation, PCI: posterior credible interval.

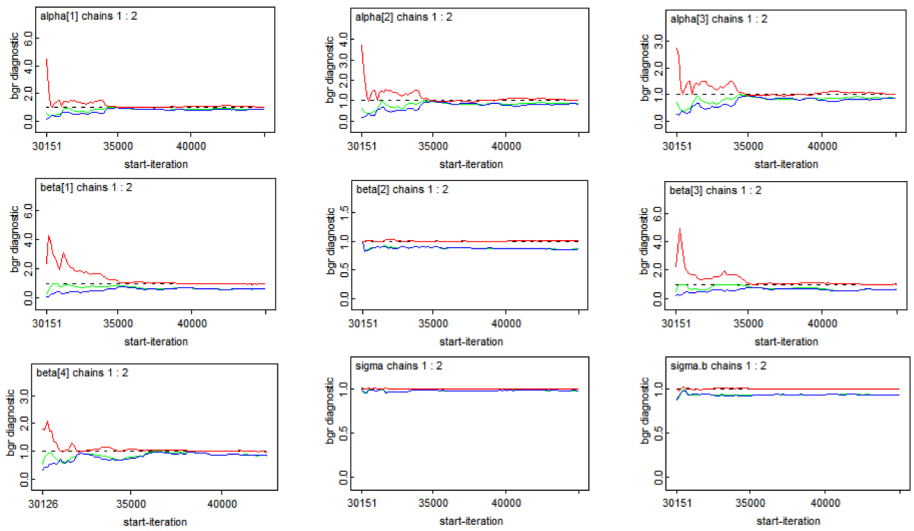

Figure 3. Gelman-Rubin diagnostic plot for each parameter in the SQRM(0.50).

a more significant effect on children's reading skills than others. Also, older children have a better reading ability than younger children. In Table 5, estimates of σ_bs in all models are significant, indicating the importance of the correlation structure in the continuous longitudinal responses of PIAT data.

Figure 3 shows the output diagrams of the Gelman-Rubin diagnostics for SQRM(0.5) of different parameters. Three curves exist in each plot of which one is a horizontal line, which shows the value of one. The curves under the horizontal line display $\hat{V}$ in green color and W in blue color, and the top curve above the horizontal line display the ratio of $\hat{V}$ and W, i.e. $\hat{R}$. Figure 3 shows the convergence of the MCMC chains. As it is seen, when the number of replication increases, $\hat{V}$ and W are steadying, and $\hat{R}$ is conducing to one.

According to DIC values in Table 5, the SQRM(τ) model is an excellent fit compared to that of the CMRM model. SQRM(0.25) provides the lowest DIC value, then SQRM(0.25) gives the most suitable model in existing models based on DIC values.

© 2020, SRTC Iran

6 Conclusion

This paper aimed to use a semi-parametric quantile regression model for longitudinal data using random effects. An asymmetric Laplace distribution was assigned to the mixed effect model's error for continuous data. As shown, the suggested model can be fitted under a Bayesian structure via MCMC methods. Two simulations were carried out to study the performance of the proposed method. We earned consistent and desirable results and analysed the PIAT data. The results specified that SQRM has a better performance than the CMRM. According to this, the performance of the suggested model was good. Furthermore, ongoing research is considering the proposed model to analyse data with missing responses.

References

Aghamohammadi, A. and Meshkani, M.R. (2017). Bayesian quantile regression for skew-normal linear mixed models. *Communications in Statistics - Theory and Methods*, **46**, 10953-10972. DOI: 10.1080/03610926.2016.1257713.

Alhamzawi, R., Yu, K., Vinciotti, V., and Tucker, A. (2011). Prior elicitation for mixed quantile regression with an allometric model. *Environmetrics*, **22**, 911–920.

Bravo, F. (2020). Semiparametric quantile regression with random censoring. *Ann. Inst. Stat. Math.*, **72**, 265-295. https://doi.org/10.1007/s10463-018-0688-3.

Cai, Z., and Xiao, Z. (2015). Semiparametric quantile regression estimation in dynamic models with partially varying coefficients. *Journal of Econometrics*, **167**, 413-425.

Chen, X., Sun, J., and Liu, L. (2015). Semiparametric partial linear quantile regression of longitudinal data with time varying coefficients and informative observation times. *Statistica Sinica*, **25**, 1437-1458.

Diggle, P., Heagerty, P., Liang, K.Y., and Zeger, S. (2002). *Analysis of Longitudinal Data*, Oxford University Press.

Fairclough, N. (2005). Peripheral Vision: Discourse Analysis in Organization Studies: The Case for Critical Realism, *Organization Studies*, **26**, 915-939.

Fan, J. (1992). Design-adaptive nonparametric regression. *Journal of American Statistical Association*, **87**, 998- 1004.

Fan, J. (1993). Local linear regression smoothers and their minimax efficiency. *Annals of Statistics*, **21**, 196-216.

Fan, J., and Gijbels, I. (1992). Variable bandwidth and local linear regression smoothers. *Annals of Statistics*, **20**, 2008-2036.

Fan, J., Heckman, N.E., and Wand, M.P. (1995). Local Polynomial Kernel Regression for Generalized Linear Models and Quasi-Likelihood Functions. *Journal of the American Statistical Association*, **90**, 141-150. doi:10.2307/2291137.

Fitzmaurice, G., Davidian, M., Verbeke, G., and Molenberghs, G. (2008). *Longitudinal Data Analysis*. Boston, CRC Press.

Fu, L., and Wang, Y.G. (2012). Quantile regression for longitudinal data with a working correlation model. *Comput. Stat. Data Anal.*, **56**, 2526–2538.

Gelman, and Rubin, D.B. (1992). Inference from iterative simulation using sequences. *Statistical Science*, **7**, 457-511.

Geraci, M., and Bottai, M. (2007). Quantile regression for longitudinal data using the asymmetric Laplace distribution. *Biostatistics*, **8**, 140-154.

Ghasemzadeh, S., Ganjali, M., and Baghfalaki, T. (2020). Bayesian quantile regression for joint modeling of longitudinal mixed ordinal and continuous data. *Communications in Statistics-Simulation and Computation*, **49**, 375-395.

Hao, L., and Naiman, D.Q. (2007). *Quantile Regression*. New York, Sage.

Hong, H.G., He, X. (2010). Prediction of functional status for the elderly based on a new ordinal regression model. *Journal of American Statistical Association*, **105**, 930–941.

Hong, H.G., and Zhou, J. (2013). A multi-index model for quantile regression with ordinal data. *Journal of Applied Statistics*. **40**, 1231–1245.

Huang, Y. (2016). Quantile regression-based Bayesian semiparametric mixed-effects models for longitudinal data with non-normal, missing and mismeasured covariate, *Journal of Statistical Computation and Simulation*, **86**, 1183-1202, DOI: 10.1080/00949655.2015.1057732

Irene, G., Rezaul, K., and Anneleen, V. (2020). Semiparametric quantile regression using family of quantile-based asymmetric densities. *Computational Statistics & Data Analysis*, **157**, 107-129. https://doi.org/10.1016/j.csda.2020.107129.

Karlsson, A. (2007). Nonlinear quantile regression estimation of longitudinal data. *Communications in Statistics-Simulation and Computation*. **37**, 114–131.

Koenker, R., and Bassett, G. (1978). Regression quantiles. *Econometrica*, **46**, 33-50.

Koenker, R. (2004). Quantile regression for longitudinal data. *J. Multivar. Anal.*, **91**, 74–89.

Koenker, R. (2005). *Quantile regression*. Cambridge, Cambridge university press.

© 2020, SRTC Iran

Kottas, A., and Krnjajic, M. (2009) Bayesian Semiparametric Modelling in Quantile Regression. *Scandinavian Journal of Statistics*, **36**, 297–319.

Kotz, S., Kozubowski, T.J., and Podgorski, K. (2001). *The Laplace distribution and generalizations: a revisit with applications to communications, economics, engineering, and finance.* Boston, Birkhauser.

Kozumi, H., and Kobayashi, G. (2011). Gibbs sampling methods for Bayesian quantile regression. *Journal of Statistical Computation and Simulation*, **81**, 1565-1578.

Matteo, S.L., and Emmanuel, J.C. (2020). A comparison of some conformal quantile regression methods. *Stat.*, **9**.

Noorian, S., Ganjali, M., and Bahrami Samani, E. (2016). A Bayesian test of homogeneity of association parameter using transition modeling of longitudinal mixed responses. *Journal of Applied Statistics*, **43**, 1850-1863.

R Core Team (2013). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. Available at http://www.R-project.org/.

Roca, P., and Ordonez, C.P. (2019). Predicting pollution incidents through semiparametric quantile regression models. *Stochastic Environmental Research and Risk Assessment*, **33**, 673-685. DOI:10.1007/s00477-019-01653-7.

Samani, E.B., and Ganjali, M. (2011). Sensitivity analysis for non-ignorable missing responses with application to multivariate Random effect model. *Metron*, **69**, 309-322.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P., and Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **64**, 583-639.

Wang, Huixia, J., Xingdong, F., and Chen, D. (2019). Copula-based quantile regression for longitudinal data. *Statistica Sinica*, **29**, 245-264.

Wu, H., and Zhang, J.T. (2006). *Nonparametric regression methods for longitudinal data analysis: mixed-effects modeling approaches* (Vol. 515). John Wiley and Sons.

Yang, H., Chen, Z., and Zhang, W. (2019). Bayesian Nonlinear Quantile Regression Approach for Longitudinal Ordinal Data. *Communications in Mathematics and Statistics*, **7**, 123-140. DOI:10.1007/s40304-018-0148-7.

Yu, K., Moyeed, R.A. (2001). Bayesian quantile regression. *Stat. Probab.Lett*, **54**, 437–447.

Yu, K., and Zhang, J. (2005). A three-parameter asymmetric Laplace distribution and its extension. *Communications in Statistics-Theory and Methods*, **34**, 1867-1879.

Zhou, L.H., and M. Li, R. (2012). Semiparametric quantile regression with high-dimensional covariates. *Statistica Sinica*, **22**, 1379-1401

Zhou, L. (2010). *Conditional quantile estimation with ordinal data* , Ph.D. thesis, University of South Carolina.

Omid Khazaei
Department of Statistics,
Faculty of Mathematical Sciences,
Shahid Beheshti University,
Tehran, Iran.
email: *o-khazaei@sbu.ac.ir*

Mojtaba Ganjali
Department of Statistics,
Faculty of Mathematical Sciences,
Shahid Beheshti University,
Tehren, Iran.
email: *m-ganjali@sbu.ac.ir*

M. Khazaei
Department of Statistics,
Faculty of Mathematical Sciences,
Shahid Beheshti University,
Tehran, Iran.
email: *m-khazaei@sbu.ac.ir*

© 2020, SRTC Iran