Quarterly Journal of Econometrics Research
2020 Vol. 6, No. 1, pp. 12-17.
ISSN(e): 2411-0523
ISSN(p): 2518-2536
DOI: 10.18488/journal.88.2020.61.12.17

ASSESSMENT OF MENTAL HEALTH OF UNDERGRADUATE STUDENTS BASED ON AGE: A BAYESIAN ORDINAL QUANTILE REGRESSION APPROACH

Nwakuya M. T.
Department of Mathematics/Statistics University of Port Harcourt, Choba, Rivers State, Nigeria.
Email: maureen.nwakuya@uniport.edu.ng Tel: +2348033167003

Article History
Received: 27 October 2020
Revised: 12 November 2020
Accepted: 30 November 2020
Published: 14 December 2020

Keywords
Quantile regression
Gibbs sampler
Posterior distribution
Latent variable
Bayesian ordinal quantile Regression.

JEL Classification:
C11 & C40.

ABSTRACT

The traditional frequentist quantile regression makes minimal assumptions that accommodate errors that are not normal given that the response variable (y) is continuous even in Bayesian framework. However inference on these models where y is not continuous proves to be challenging particularly when the response variable is an ordinal data. This paper portrays the idea of Bayesian quantile estimation on ordinal data. This method utilizes the latent variable inferential framework. Estimation was done using Markov chain Monte Carlo simulation with Gibbs sampler where the cut points were set ahead of time and remained fixed all through the analysis. The method was applied in a mental health study of University undergraduate students. Investigations of the model exemplify the practical utility of Bayesian ordinal quantile models. In this paper we were able to investigate the mental health state of undergraduate students at different points in the distribution of their ages. Our findings show that the age of the students has a significant effect on their mental health. The results revealed that at 25th, 50th and 75th quantiles the ages had a negative effect on their mental health while at the 95th quantile the effect was positive. This study was able to show that older undergraduate students are more mentally equipped to withstand the stress of higher learning in the University.

Contribution/Originality: The paper's primary contribution is to apply Bayesian ordinal quantile regression to mental health analysis. The study utilized the Gibbs sampler with fixed cut-points. It portrayed insight to the effect of age on the mental health of undergraduate students at different points on the age distribution.

1. INTRODUCTION

The traditional frequentist quantile regression as proposed by Koenker (2004) makes minimal assumptions that accommodates continuous response variables with errors that are not normal. The Bayesian quantile framework also assumes the response variable to be continuous. However inference on these models where the response variable is not continuous proves to be challenging particularly for ordinal data. Ordinal models arise when the response variable is discrete and inherently ordered or ranked with the characteristic that values assigned to outcomes have an ordinal meaning, but no cardinal interpretation. For example, in a survey regarding the performance of the economy, responses may be recorded as follows: 1 for 'bad', 2 for 'average' and 3 for 'good'. The responses in such a case have ordinal meaning but no cardinal interpretation, so one cannot say a value of 2 is twice as good as a value of 1, (Rahman, 2016). The ordinal ranking of the responses differentiates these data from unordered choice outcomes. Quantile regression allows us to uncover interesting structures that might be present in the tails of the distribution, including heavy-tailed or skewed distributions, that would otherwise be masked in standard regression and distort inference. Quantile regression allows us to quantify a more complex relationship between the covariates and the distribution of the response variable by modeling the conditional quantile function of response variable y, that is $Q_y(\tau \mid x)$, where τ is the quantile with interval $0 < \tau < 1$. Here, the τ^{th} quantile of y given covariates x is defined as inf{y : F(y) ≥ τ } for the cumulative distribution function F(y). As a result, quantile regression estimates both the variable effects of covariates across conditional quantiles and the shape of response distributions conditional on x. This is useful in any situation where the mean might not adequately describe the conditional response distribution, such as in the presence of non-Gaussian residuals. The Bayesian approach to ordinal quantile regression is of interest in this paper.

2. BAYESIAN QUANTILE REGRESSION

Given a linear model:

$$y_i = x_i^T\beta + e_i y_i = x_i^T\beta + e_i, \tag{1}$$

where $x_i \epsilon \mathbb{R}^j \, and \, \beta \epsilon \mathbb{R}^j$ are column vectors of size j and $y_i \epsilon \, \mathbb{R}$ is a scalar variable while e_i is the residual. To obtain the conditional mean regression it is assumed that $E(e/x) = 0$, in the same vein the conditional median regression is obtain under the assumption that med$(e/x) = 0$. We can deduce from Equation 1 that the residual is given as;

$$e_i = y_i - x_i^T\beta \tag{2}$$

In mean regression the coefficient β can be obtained by OLS method by minimizing the sum of squared residuals;

$$\hat{\beta} = \underset{\beta \epsilon \mathbb{R}^j}{\text{argmin}} \; \sum_{i=1}^{n} \rho_\tau (y_i - x_i^T\beta)^2 \tag{3}$$

In median regression the regression coefficient $\beta\beta$ is obtained by minimizing the absolute deviations;

$$\hat{\beta} = \underset{\beta \epsilon \mathbb{R}^j}{\text{argmin}} \; \sum_{i=1}^{n} |y_i - x_i^T\beta| \tag{4}$$

The extension of the median regression to all other quantiles gave rise to quantile regression analysis. The quantile regression estimation proceeds by minimizing, with respect to β_τ, the following objective function.

$$\underset{\beta \epsilon \mathbb{R}^j}{\text{argmin}} \left[\; \sum_{e_i<0} (1-\tau) \, |y_i - x_i^T\beta| + \sum_{e_i>0} \tau |y_i - x_i^T\beta| \right] \tag{5}$$

This objective function, Equation 5 can be written as a sum of check functions or piecewise linear functions as follows;

$$\hat{\beta}_\tau = \underset{\beta \epsilon \mathbb{R}^j}{\text{argmin}} \; \sum_{i=1}^{n} \rho_\tau (y_i - x_i^T\beta) \tag{6}$$

Given that $\hat{\beta}_\tau$ is the τ^{th} regression quantile, where the check function $\rho_\tau(e) = e\big(\tau - \mathrm{I}(e < 0)\big) \; given \; that \; \mathrm{I}(.)$ is the indicator function which equals 1 if the condition inside the parentheses is true and 0 otherwise. This check function assigns τ to positive residuals and 1-τ to negative residuals. Note $\tau = 0.5$ is the median regression. Because the check function is not differentiable at the origin, computational techniques such as the simplex algorithm, the interior point algorithm or the smoothing algorithm are often applied.

Koenker. and Machado (1999) were the first to show that likelihood-based inference using independently distributed asymmetric Laplace densities (ALD) is directly related to the minimization problem in Equation 6. Yu and Zhang (2005) proposed a three-parameter ALD with a skewness parameter that can be used directly to model the quantile of interest. The Bayesian version of quantile regression still uses the check function and its method follows the proposed three-parameter ALD that assumes that the residual follows an Asymmetric Laplace distribution (ALD) i.e $e \sim ALD(\theta, \delta, \tau)$ where θ is the location, δ is the scale parameter and τ is the skewness parameter and it also relies on the observation that the asymmetric Laplace distribution (ALD) contains a check function within its probability density function (PDF). The Bayesian framework, when adopting an ALD prior on the residuals e_i with θ = 0 and δ = 1, has the probability density function given as;

$$f_\tau(e_i) = (1-\tau)\begin{cases}\exp(-e_i(\tau-1)) & e < 0\\ \exp(-e_i\tau) & e \geq 0\end{cases} \tag{7}$$

And the likelihood function of the model $y_i = x_i^T\beta + e_i$ becomes

$$L(y|\beta) = \tau^n(1-\tau)^n \exp(-\sum_{i=1}^n \rho_\tau(e_i))$$
$$= \tau^n(1-\tau)^n \exp(-\sum_{i=1}^n \rho_\tau(y_i - x_i^T\beta)) \tag{8}$$

The likelihood Function in equation 8 can be seen to contain the objective function for the τ^{th} regression quantile in Equation 6. The mean and variance of the PDF in Equation 7 is given by;

$$E(e_i) = \frac{1-2\tau}{\tau(1-\tau)} \; and \; V(e_i) = \frac{1-2\tau+2\tau^2}{\tau^2(1-\tau)^2} \tag{9}$$

where E(ei) is the mean and V(ei) is the variance.

Both the mean and variance, as shown above, depend on the skewness parameter τ, but are fixed for a given value of τ. Interestingly, τ also defines the quantile of an AL distribution. This feature becomes useful in quantile regression since estimation of a model at different quantiles basically requires a change in the value of τ.

3. BAYESIAN ORDINAL QUANTILE REGRESSION

The standard approach to regression with ordinal response variables is to use the ordinal probit model. Fitting this ordinal probit model only captures the mean of the conditional distribution of the continuous latent variable underlying each response but quantile regression will study the full conditional distributions of such outcomes without assuming Gaussianity. The Bayesian method of estimating quantile regression stems from the fact that maximization of the likelihood, where the error follows an AL distribution, is equivalent to minimization of the quantile objective function in Equation 6. Bayesian implementation of quantile regression begins by forming a likelihood based on the AL distribution, thus the posterior distribution is proportional to the product of the likelihood and the prior distribution of the parameters and it can be represented as follows;

$$\phi(\beta,\delta|\tau,y,x) \propto \xi(\beta,\delta)\prod_{i=1}^n ALD(y|\beta,\delta,\tau) \tag{10}$$

Where $\phi(\beta,\delta|\tau,y,x)$ is the full posterior distribution, $\xi(\beta,\delta)$ is the joint prior on the regression parameters and $\prod_{i=1}^n ALD(y|\beta,\delta,\tau)$ is the product of the likelihoods. Bayesian inference for the estimation of β at any quantile is done by querying the posterior distribution with the skewness parameter set to that quantile. Unfortunately the posterior distributional form cannot be tracked; therefore an appropriate Markov chain Monte Carlo (MCMC) method is usually used to perform a Bayesian analysis to estimate the full posterior distribution which is then queried to obtain β at any desired quantile. Previous approaches that have been developed for Bayesian quantile regression include random walk Metropolis-Hastings method and Gibbs sampler methods which are more computationally efficient and require less parameter tuning. Hideo and Genya (2012) showed that Gibbs sampling can be used for Bayesian quantile regression provided the AL distribution is represented as a mixture of normal–exponential distributions. There is also a partially collapsed Gibbs sampler (Reed & Yu, 2009). In particular, there is no widely accepted quantile regression method for ordinal variables. Ordinal variables are especially common in medical contexts, where many health outcomes are expressed as ordered categories rather than as strictly numerical measures, like the case we are considering in this paper, where the mental health status of undergraduate students are assessed based on the ages of the students.

The Bayesian quantile regression methods thus discussed so far assumes that the response variable y is continuous. Since y is assumed continuous in Bayesian Quantile regression and its residuals are modeled directly as Asymptotic Laplace Distributed (ALD) variables then this direct approach is very meaningful but the situation where y the response variable is not continuous but ordinal, the direct approach becomes meaningless. To handle the situation where y is ordinal, Rahim and Haithem (2017) introduced a continuous latent variables z_i corresponding to each y_i, the variable z_i is unobserved and relates to the observed response y_i, which has J categories or outcomes gotten by implementing cut-points σ. Thus a quantile regression ordinal model can be represented using a continuous latent random variable z_i as;

$$z_i = x_i'\beta_\tau + \delta e_i, \;\; i = 1,\dots,n \tag{11}$$

Where the residual $e_i \sim ALD(\theta,\delta,\tau)$.

Let's assume y can take on D possible ordered values, which can be coded as y ∈ {, d_1, d_2 . . . , d_D }. Then these coded responses relate to z as follows:

$$y_i \begin{cases} d_1 \; if \; \sigma_0 \leq z_i < \sigma_1 \\ d_j \; if \; \sigma_{j-1} \leq z_i < \sigma_j, \qquad j = 2, \dots, D-1 \\ D \; if \; \sigma_{D-1} \leq z_i < \sigma_D \end{cases} \tag{12}$$

Grabski, Vito, and Engelhardt (2019) this formulation allows us to transform the ordinal response into continuous responses so that we can apply the above mentioned method. The Bayesian quantile regression models $y_i = x_i^T \beta + e_i$ given observed and continuous y, while the transformed Bayesian model with the unobserved but continuous z, models $z_i = x_i' \beta_\tau + e_i$, which is coded as an ordinal response correlated with y as in Equation 12. In both cases, e_i is drawn from an Asymptotic Laplace distribution. The coefficients have Laplace priors, and the cutpoint vector's prior is an order statistics from a uniform distribution. Gibbs sampling can then be used to perform inference on this model, leveraging the key observation that the ALD can be written as a conditionally conjugate normal-exponential mixture. Thus, the Bayesian regression quantiles for Equation 11 are easily estimated.

4. METHODOLOGY

Our data is an ordinal data on the mental health state of university students. Information gotten from students was classified based on their age and answers to questions that point to their mental health wellbeing. Their mental health state was ordered into 3 categories; stable, mildly-unstable and unstable. In this paper we consider the ordered responses yi (mental health state) and the corresponding covariate x_i (age), for i = 1, . . . ,n. Using the latent variable inferential framework of Albert and Chib (1993) we employed the Gibbs sampler, leveraging on the assumption that the Asymptotic Laplace distribution of the residual is a mixture of normal and exponential distribution where the residual is given as;

$$e_i = \delta\theta\omega_i + \delta\psi\sqrt{\omega_i\mu_i} \quad \forall\, i = 1, \dots, n \tag{13}$$

Where δ is the scale parameter, $\omega_i \sim Exponential(1) \;\&\; \mu_i \sim Normal(0,1)$ and are mutually independent. And the constants θ and ψ are defined as;

$$\theta = \frac{1-2\tau}{\tau(1-\tau)} \; and \; \psi = \sqrt{\frac{2}{\tau(1-\tau)}} \tag{14}$$

Given Equation 13, the Bayesian Ordinal quantile model thus incorporates the normal-exponential Al distribution into the latent variable model of Equation 11, thus the τ^{th} ordinal quantile model can therefore be expressed as;

$$z_i = x_i'\beta_\tau + \delta\theta\omega_i + \delta\psi\sqrt{\omega_i\mu_i} \quad \forall\, i = 1, \dots, n \tag{15}$$

However, the scale parameter can be removed from the model through a simple transformation as follows:

$$z_i = x_i'\beta_\tau + \theta v_i + \psi^2\delta v_i \tag{16}$$

Where $v_i = \delta\omega_i$, Equation 16, makes it clear that $z_i|\beta_\tau, \, z_i \sim normal(x_i'\beta_\tau + \theta v_i, \psi^2\delta v_i)$ and this gives the access to the properties of the normal distribution in the estimation procedure.

5. RESULTS

Our data has a sample size of 707 undergraduate students. The number of observations corresponding to each category are; 138(19.52%) for stable state, 533 (75.39%) for mildly-stable state and 36 (5.15%) for unstable state. Using data on the above variables, the application studied the effect of age on the mental health state of undergraduate students. In the analysis we considered the 25^{th}, 50^{th}, 75^{th} and 95^{th} quantiles. The ages of the students range from 15-42. Where we classified ages (15-19) years to fall within the 25^{th} quantile, (20-21) years falls within the 50^{th} quantile, (22-23) years falls within the 75^{th} quartile and ages >23 falls within the 95^{th} quantile.

The posterior estimates of the Bayesian quantile ordinal models with the inefficiency values were obtained and the results are shown below;

Table-1. Posterior mean, posterior standard deviation and inefficiency values.

Quantiles	Parameters Estimates		
		Intercept	Age
25th quantile	Mean	-1.4284	-0.7301
	Standard Deviation	0.8282	0.0388
	Inefficiency	1.1156	1.1667
50th quantile	Mean	-1.0017	-0.1995
	Standard Deviation	0.1894	0.0637
	Inefficiency	1.4307	1.2034
75th quantile	Mean	-0.2227	-0.0036
	Standard Deviation	0.0347	0.0466
	Inefficiency	1.3620	1.3075
90th quantile	Mean	0.0591	0.3261
	Standard Deviation	0.1662	0.2124
	Inefficiency	1.1567	1.2864

The effect of age at the 25th, 50th and 75th quartile has a negative effect on the probability of supporting stable mental health. This shows that a significant number of students between the ages of 15 − 23 are not in a stable mental health state. At the 90th quantile, we see a positive effect, this shows that significantly students from ages 23 and above are in a state of stable mental health.

Table-2. Deviance Information Criterion (DIC) for all Quantiles.

Quantiles	Deviance Information Criterion (DIC)
25th Quantile	1004.16
50th Quantile	839.93
75th Quantile	789.25
90th Quantile	987.53

The model selection criterion such as deviance information criterion (DIC) (Celeux, Forbes, Robert, & Titterington, 2006; Spiegelhalter, Best, Carlin, & van der Linde, 2002) was utilized to choose a value of τ that is most consistent with the data. To show this, DIC was computed for the 25th, 50th, 75th and 90th quantile models and the values were 1004.16, 839.93, 789.25 and 987.53, respectively as shown in Table 2. Hence, amongst all the models considered, the 75th quantile model provides the best fit.

6. CONCLUSION

The paper considers the Bayesian analysis of quantile regression models for univariate ordinal data. The method exploits the latent variable inferential framework of Albert and Chib (1993) and capitalizes on the normal–exponential mixture representation of the AL distribution. Estimation utilizes Gibbs sampling with fixed cut-points. Posterior means, standard deviations and inefficiency factors are calculated for $(\beta_\tau, \sigma_\tau)$. The Posterior estimates of β_τ are statistically significant at all quantiles, the standard deviations are small and inefficiency factors are all less than 6 at all quantiles. The explanation is that the age of university undergraduate students has an effect on their mental health state. Considering the signs on the β_τ estimates, the 25th, 50th and 75th quantile effects were all negative while the 90th quantile effect was positive. This shows that ages of students between 15 and 23 showed a negative effect on the mental health state while the ages of students above 23 had a positive effect on the mental health state. Based on the findings we can say that the older an undergraduate student is, the better mentally equipped he/she is in coping with the stress associated with higher learning in the university. The model comparison based on DIC selects the 75th quantile model to be the best fitting model.

Funding: This study received no specific financial support.
Competing Interests: The author declares that there are no conflicts of interests regarding the publication of this paper.

REFERENCES

Albert, J., & Chib, S. (1993). Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association*, *88*(422), 669–679.Available at: https://doi.org/10.1080/01621459.1993.10476321.

Celeux, G., Forbes, F., Robert, C. P., & Titterington, D. M. (2006). Deviance information criteria for missing data models. *Bayesian Analysis*, *1*(4), 651–674.Available at: https://doi.org/10.1214/06-ba122.

Grabski, I. N., Vito, R. D., & Engelhardt, B. E. (2019). *Bayesian ordinal quantile regression with a partially collapsed gibbs sampler*. Paper presented at the Joint Statistical Meeting online Program, July 27th - Aug 1st 2019; Colorado USA. Cited as: arXiv:1911.07099 [stat.ME].

Hideo, K., & Genya, K. (2012). Gibbs sampling methods for bayesian quantile regression. *Journal of Statistical Computation and Simulation*, *81*(11), 1565–1578.Available at: https://doi.org/10.1080/00949655.2010.496117.

Koenker, R. (2004). Quantile regression for longitudinal data. *Journal of Multivariate Analysis*, *91*(1), 74-89.

Koenker., R. W., & Machado, J. A. F. (1999). Goodness of fit and related inference processes for quantile regression. *Journal of the American Statistical Association*, *94*(448), 1296-1310.Available at: https://doi.org/10.1080/01621459.1999.10473882.

Rahim, A., & Haithem, T. M. A. (2017). Bayesian quantile regression for ordinal longitudinal data. *Journal of Applied Statistics*, *45*(5), 815-828.Available at: 10.1080/02664763.2017.1315059.

Rahman, M. A. (2016). Bayesian quantile regression for ordinal models. *Bayesian Analysis*, *11*(1), 1-24.Available at: https://doi.org/10.1214/15-ba939.

Reed, C., & Yu, K. (2009). A partially collapsed gibbs sampler for bayesian quantile regression. Computing and Mathematics Working Papers, Brunel University. Retrieved from: http://bura.brunel.ac.uk/handle/2438/3593.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society – Series B Statistical Methodology*, *64*(4), 583–639.

Yu, K., & Zhang, J. (2005). A three paramter asymmetric laplace distribution and its extensions. *Communications in Statistics – Theory and Methods*, *34*(9-10), 1867– 1879.Available at: https://doi.org/10.1080/03610920500199018.