M A T H L I N E
JURNAL MATEMATIKA DAN PENDIDIKAN MATEMATIKA

ISSN 2502-5872 (Print)
ISSN 2622-3627 (Elektronik)

Volume 10 Number 2, May 2025, 373-385

MODELING CLASSIFICATION OF STUNTING TODDLER HEIGHT USING BAYESIAN BINARY QUANTILE REGRESSION WITH PENALIZED LASSO

Lilis Harianti Hasibuan [1], Ferra Yanuar[2*], Dodi Devianto[3], Maiyastri[4]
[1]Departement of Mathematics, Universitas Islam Negeri Imam Bonjol Padang, West Sumatera, Indonesia
[2,3,4]Departement of Mathematics and Data Science, Universitas Andalas, West Sumatera, Indonesia
*Correspondence: ferrayanuar@sci.unand.ac.id

ABSTRACT

Stunting is a child who has a height that is shorter than the age standard. One of the main indicators of stunting is a height that is lower than the standard for toddlers. Stunting in Indonesia is of great concern due to the high prevalence of stunting. Stunting children are at risk of impaired cognitive development, which will result in the development of human resources. This study aims to develop a classification model to detect stunted toddlers based on height using the Bayesian binary quantile regression method with LASSO (Least Absolute Shrinkage and Selection Operator). This method was chosen because of its ability to handle multicollinearity and variable selection problems automatically, as well as provide better estimates on non-normally distributed data. The data used in this study includes five independent variables such as age, weight at birth, gender, how to measure height and nutritional status. The results showed that independent variables that significantly affect the height of stunting toddlers can be a concern to reduce the problem of stunting in Indonesia. The results of model show that variable age, weight at birth, and nutritional status have a significant influence to classification of stunting toddler height. Indicator of model goodness is seen from the quantile that has the smallest MSE value. The model that has the smallest MSE is in quantile 0.25 with an MSE value of 0.1622.
Keywords: Stunting; Binary Quantile Regression, Bayesian, LASSO.

How to Cite: Hasibuan, L. H., Yanuar, F., Devianto, D., & Maiyastri. (2025). Modeling Classification Of Stunting Toddler Height Using Bayesian Binary Quantile Regression With Penalized Lasso. Mathline: Jurnal Matematika dan Pendidikan Matematika, *10*(2), 373-385. http://doi.org/10.31943/mathline.v10i2.928

PRELIMINARY

Stunting is a condition of impaired growth in children characterized by shorter height than their age standard (Leroy & Frongillo, 2019). Based on data from the World Health Organization (WHO), stunting is a serious nutritional problem in many countries, including Indonesia. This condition can affect children's physical and cognitive development and contribute to the increased risk of infant and child mortality (Mustakim et al., 2022). According to data from Indonesia's Central Bureau of Statistics (BPS) Sumatra Barat, the prevalence of stunting in Indonesia is still quite high despite various efforts to tackle it. Statistical modelling is an effective tool for analyzing the factors that influence the height growth of stunted children. Based on the data that has been obtained, West Sumatra Province still has a stunting proportion of 27.47%; this is still close to the value of Indonesia's stunting proportion, which is 27.67% (Yasril & Sari, 2022).

In the cases of stunting of toddlers in West Sumatra, the factors that affect stunting include the age of the baby when measured after the age of 2 years (Vitaloka et al., 2019). A part from that, the height gain of stunted toddlers is strongly influenced by the nutritional intake obtained by children (Kang et al., 2018)(Ariati et al., 2018). Stunting toddlers are measured standing or recumbent (Wilson et al., 2011). Not only that, one of the studies conducted said that gender and weight also affect the classification of stunting toddlers (Huriah & Nurjannah, 2020). The factors assumed to influence height classification can be modelled with regression analysis.

Regression analysis aims to see the influence between the dependent variable and one or more the independent variable (Harianti Hasibuan et al., 2022)(Sholih et al., 2024)(Wizsa & Rahmi, 2025). However, data in the field often violates the assumption of normality; one of the models for data that violates the assumption of normality can be overcome with quantile regression (Yanuar et al., 2016)(Hasibuan et al., 2025).

Quantile regression is regression analysis used to estimate the relationship between independent variables and specific quantiles (percentiles) of the dependent variable's distribution, rather than just the mean (as in ordinary least squares regression) (Yanuar et al., 2020),(Hasibuan, et al., 2024). A quantile is a value that divides a probability distribution or a set of data into intervals with equal probabilities. In other words, it's a cutoff point below which a certain proportion of the data falls.

One approach that is gaining popularity in this research is binary quantile regression, which can handle non-normal data distribution and provide more robust estimates compared to conventional linear regression. Binary quantile regression allows us to model the relationship between the independent and response variables in different quantiles with response binary (dichotomous) (Benoit & Van den Poel, 2012).

In addition, the binary quantile regression approach can be strengthened by using Bayesian methods, which allow for handling uncertainty in the model. The Bayesian method is an approach in statistics used to update or adjust our beliefs (or estimates) of an event based on new data or information obtained (Hasibuan et al., 2025). Bayesian method is considered good for estimating parameters because the parameters obtained are considered as random variables that have a certain distribution pattern. This led to the development of a study combining binary quantile regression with the Bayes method, which is considered good at modeling both small and large samples (Benoit et al., 2013).

Bayesian binary quantile regression is a method to model the probability of a binary event, at a given quantile level, while incorporating prior information and updating it based on new data with a Bayesian approach. The Bayesian binary quantile regression (BBQR) method provides advantages in terms of flexibility and robustness to outliers and allows the use of data with more complex distributions (Benoit & Van den Poel, 2012). Further research was developed by Rahim (Alhamzawi et al., 2012) who shrank the parameter value so that there was no overfitting that was too large with Least Absolute Shrinkage and Selection Operator (LASSO). This approach can also provide more reliable results in the context of limited or imperfect data. LASSO is used as a regulation parameter in order to robust the obtained parameters used in Bayesian Quantile Regression (Yanuar et al., 2023)(Benoit & Van den Poel, 2012).

Children's height is one of the main indicators in monitoring the nutritional status and health of children, especially in toddlers. Optimal growth is strongly influenced by various factors, including adequate nutritional intake, environmental factors, and good health care. One of the main challenges in addressing stunting is the inability to accurately predict and measure the factors that influence growth. The application of Bayesian binary quantile regression with LASSO in modelling the classification of height gain in stunted toddlers can provide new insights into determining more targeted intervention strategies (Ma et al., 2023). This study aims to develop a model that can predict the factors that affect the height growth of stunted toddlers, with a focus on the use of Bayesian binary quantile Regression (BBQR) and Bayesian binary quantile regression (BBLQR) techniques are expected to provide a more comprehensive understanding of the growth dynamics of stunted toddlers. This study aims to compare and build a classification model of height gain in stunting toddlers using Bayesian binary quantile regression (BBQR) and (BBLQR). Hopefully, the results of this study can provide a deeper understanding of the factors that influence height growth in stunted toddlers and provide more effective policy recommendations in efforts to overcome stunting in Indonesia.

MODELING PROCEDURE

Quantile Regression

Quantile regression is a statistical technique used to analyze the relationship between the independent variable (predictor) and the dependent variable (response) at various quantiles (percentiles) of the dependent variable distribution rather than just at the mean (as is done in ordinary linear regression) (Hasibuan et al., 2024). This method provides a complete picture of the relationship between variables than ordinary linear regression, which only focuses on the mean or centre of the data distribution. Quantile regression is appropriate when data is not normally distributed ha, has a highly skewed distribution, or has significant outliers. The quantile regression equation can be written as (Koenker & Bassett Jr, 1978):

$$Q_\tau(Y|X) = X\beta(\tau) \tag{1}$$

Where:

$Q_\tau(Y|X)$ is quantile τ from dependent variable Y given X.

$\beta(\tau)$ is a vector of parameter quantile regression in quantile τ, which is the quantile used. The estimated value of the parameters in the quantile regression equation $\widehat{\boldsymbol{\beta}_\tau}$ is obtained by minimizing the following equation (Davino et al., 2013):

$$\sum_{i=1}^{n} \rho_\tau(y_i - \boldsymbol{x}_i'\boldsymbol{\beta}_\tau). \tag{2}$$

With $\rho_\tau(u) = u(\tau - I(u < 0))$is the *loss function* with the equation [10]:

$$\rho_\tau(\epsilon) = \varepsilon\big(\tau I(\varepsilon \geq 0) - (1-\tau)I(\varepsilon < 0)\big). \tag{3}$$

It is an indicator function, which has a value of 1 when $I\,(\,.\,)$it is true and 0 otherwise. This aims to reduce the prediction error in the higher quantiles (penalizing the error in the upper quantiles more) and vice versa in the lower quantiles

Bayesian Binary Quantile Regression

Binary quantile regression (BQR) was introduced by Benoit et al. (Benoit & Van den Poel, 2012). BQR model for τ^{th} quantile and *n* samples and *k* predictor for $i = 1,2,\dots,n$ is written as (Yu et al., 2003):

$$y_i^* = \beta_{0\tau} + \beta_{1\tau}x_{i1} + \beta_{2\tau}x_{i2} + \cdots + \beta_{k\tau}x_{ik} + \varepsilon_i. \tag{4}$$

$$y_i^* = \boldsymbol{x}_i'\boldsymbol{\beta}_\tau + \varepsilon_i. \tag{5}$$

Where $\boldsymbol{x} = (x_{i1}, x_{i2}, \dots, x_{ik})'$is the independent variable for the sample$i = 1,2,\dots,n$, $\boldsymbol{\beta}(\tau)$ the parameter vector, and $\boldsymbol{\varepsilon}$ as the residual vector andy_i is the observed response of $i^{th}$the subject determined by the latent unobserved responsey_i^*(Benoit et al., 2013).

$$y_i = \begin{cases} 1, \text{if } y_i^* \geq 0 \\ 0, \text{otherwise.} \end{cases} \tag{6}$$

Combining the quantile regression technique with the binary selection regression model, the following binary quantile regression models can be obtained (Benoit et al., 2013):

$$Q_\tau(y_i^*|x_i) = \boldsymbol{x}_i'\boldsymbol{\beta}_\tau + \varepsilon_i, \tag{7}$$

Where$Q_\tau(y_i^*|x_i) = \inf\{(y_i|F(y_i|x_i) \geq \tau)\}$ conditional quantileβ_τ is a parameter of theτ^{th} quantile. We use transformations that can be performed:

$$Q_\tau(y_i|x_i) = Q_\tau(h(y_i^*|x_i) = h(Q_\tau(y_i^*|x_i) = h(x_i'\beta_\tau). \quad (8)$$

With $h(x_i'\beta_\tau) = I(x_i'\beta_\tau > 0)$, suppose $\rho_\tau(u) = \frac{|u|-(2p-1)u}{2}$ is a test function. The parameter β_τ is determined by the following formula (Yanuar, Yozza, et al., 2023):

$$\sum_{i=1}^{n} \rho_\tau(y_i - h(\boldsymbol{x}_i'\boldsymbol{\beta}_\tau)). \quad (9)$$

Where $h(\boldsymbol{x}_i'\boldsymbol{\beta}_\tau) = I(\boldsymbol{x}_i'\boldsymbol{\beta}_\tau > 0)$ is the indicator variable equal to 1 if $I(\boldsymbol{x}_i'\boldsymbol{\beta}_\tau)$ it is true and zero otherwise. In Bayesian Quantile Regression, the Asymmetric Laplace Distribution is used to model the error term in the regression model. The key idea is that the ALD can model the Asymmetry of the data is important when predicting different quantiles of the dependent variable (Yu & Moyeed, 2001). When performing quantile regression, we are trying to find the regression coefficients that minimize a weighted sum of absolute deviations. In the case of the ALD, this corresponds to the quantile loss function, where the loss depends on the quantile of interest. ALD is used in the process of forming a random variable ε is ALD distributed with a likelihood density function $f(\varepsilon)$:

$$f_\tau(\varepsilon) = \tau(1-\tau)\exp(-\rho_\tau(\varepsilon)). \quad (10)$$

with $0 < \tau < 1$ and $\rho_\tau(\varepsilon)$ being the loss function ε, the error of the estimation, and $I(\varepsilon)$ the indicator function.

Suppose $(Z \sim exp\ (1))$ and $U \sim N\ (0,1)$. With ε is an ALD distributed random variable then ε can be expressed in the following equation:

$$\varepsilon = \theta z + pu\sqrt{z}. \quad (11)$$

Where $\theta = \frac{1-2\tau}{(1-\tau)\tau}$ and $p^2 = \frac{2}{(1-\tau)\tau}$ (Yu & Moyeed, 2001). Parameter $\boldsymbol{\beta}$ estimation for the τ^{th} quantile in the Bayesian quantile regression is formulated in equation (12) is: (Benoit et al., 2013):

$$L(\boldsymbol{y}_i^*|\boldsymbol{\beta},\boldsymbol{\sigma},\boldsymbol{v}) = (\prod_{i=1}^{n}(\sigma v_i)^{-\frac{1}{2}})\left(\exp\left(-\frac{(y_i^* - (\boldsymbol{x}_i'\boldsymbol{\beta}_\tau + \theta v_i))^2}{2p^2\sigma v_i}\right)\right). \quad (12)$$

With $\sigma > 0$ the scale parameter and dan $v_i = \sigma z_i$ spreading exp (σ) distribution. Based on equation (14), the full conditional distribution y_i^* is truncated normal distribution:

$$\boldsymbol{y}_i^*|\boldsymbol{\beta},\boldsymbol{\sigma},\boldsymbol{v} = \begin{cases} N(\boldsymbol{x}_i'\boldsymbol{\beta}_\tau + \theta v_i, p^2\sigma v_i)I(y_i^* > 0), \text{if } \boldsymbol{y}_i = 1 \\ N(\boldsymbol{x}_i'\boldsymbol{\beta}_\tau + \theta v_i, p^2\sigma v_i)I(y_i^* > 0), \text{if } \boldsymbol{y}_i = 0 \end{cases}. \quad (13)$$

The prior distribution used in this study are $\boldsymbol{\beta}_\tau \sim N(b_0, B_0)$,

$v_i \sim \exp(\sigma)$ and $\sigma \sim IG(a,b)$. While, posterior distribution for each prior are as follows:

$(\boldsymbol{\beta}|, \sigma, v, \boldsymbol{y}_i^*) \sim N[(\boldsymbol{B}_0^{-1} + \boldsymbol{x}_i(p^2\sigma v)^{-1}\boldsymbol{x}_i')^{-1}(\boldsymbol{B}_0^{-1}\boldsymbol{b}_0 + \boldsymbol{x}_i(p^2\sigma v)^{-1}x_i')^{-1}y_i^* - x_i(p^2\sigma v)^{-1}\theta v_i), (\boldsymbol{B}_0^{-1} + \boldsymbol{x}_i(p^2\sigma v)^{-1}\boldsymbol{x}_i')^{-1}]$

$$(v_i|\boldsymbol{\beta}, \sigma, \boldsymbol{y}_i^*) \sim GIG(\frac{1}{2}, \left(\frac{(y_i^* - x_i'^{\prime}\beta_\tau)^2}{p^2\sigma}\right), (\frac{2}{\sigma} + \frac{\theta^2}{p^2\sigma})), \quad (14)$$

$$(\sigma|\boldsymbol{\beta}, \boldsymbol{v}, \boldsymbol{y}_i^*) \sim IG(a + \frac{3n}{2}, (b + \sum_{i=1}^n v_i + \sum_{i=1}^n + \left(\frac{(y_i^* - (x_i'\beta_\tau + \theta v_i))^2}{2p^2\sigma}\right))..$$

Bayesian Binary LASSO Quantile Regression

In Binary Bayesian Quantile Regression, we are interested in modelling the conditional quantiles of a binary response variable dichotomous with a Bayesian approach. The method combines quantile regression, which focuses on estimating specific quantiles of the response distribution, with the Bayesian framework to incorporate uncertainty in the model's parameters. The key components and steps involved in estimating the parameters in this setting (Benoit & Van den Poel, 2012),(Benoit et al., 2013):

$$\beta_{LASSO} = \min_{\beta \epsilon \mathbb{R}} \sum_{i=1}^n \rho_\tau(\boldsymbol{y}_i^* - \boldsymbol{x}_i'\boldsymbol{\beta}) + \lambda \sum_{j=1}^k |\beta_j|. \quad (15)$$

Where λ is a non-negative variable penalty coefficient. The prior distribution $\boldsymbol{\beta}_\tau, \eta^2, \zeta, \sigma, \boldsymbol{s}, \boldsymbol{v}, \delta$ used for n th sample with k predictor according to for used in Bayesian LASSO binary quantile regression is:

$$f(\beta|\eta^2, s_j) = \prod_{j=1}^k \int_0^\infty \frac{1}{\sqrt{2\pi s_j}} exp\,(-\frac{{\beta_j}^2}{2s_j})\frac{\eta^2}{2} exp\,(\frac{-\eta^2}{2}s_j)ds_j$$

, (16)

$$f(\eta^2|\delta, \zeta) = \frac{\zeta^\delta}{\Gamma(\delta)}\eta^{2(\delta-1)}\exp(-\zeta\eta^2),$$

$$f(\zeta|\delta) = 1,$$

$$f(\sigma) = \sigma^{a_1^{-1}}\exp(-a_2\sigma),$$

$$f(s_j|\eta^2) = \frac{\eta^2}{2}\exp\left(-\frac{\eta^2}{2}s_j\right),$$

$$f(v_i|\sigma) = \sigma\exp(-v_i\sigma),$$

$$f(\delta|\zeta, \eta^2) = \frac{(\zeta\eta^2)^\delta}{\Gamma(\delta)}.$$

with $\eta = \sigma\lambda$, $\eta^2 \sim Gamma(\eta^2, \zeta^{-1})$, $\boldsymbol{s} = (s_1, \dots, s_k), i = 1,2 \dots k$, $\boldsymbol{v} = (v_1, \dots, v_n)$, $\sigma > 0$, $a_1 > 0$, $a_2 > 0$, $\eta^2 > 0$, $\zeta > 0, \delta > 0$. Based on equation (18), the joint posterior distribution Bayesian LASSO binary quantile regression is obtained as follows:

$$f(\boldsymbol{\beta_\tau}|\eta^2,\zeta,\sigma,\boldsymbol{s},\boldsymbol{v},\delta,\boldsymbol{y}^*)\sim N\left(\frac{\frac{\sigma\sum_{i=1}^{n}\hat{y}_{ij}x_{ij}}{2v_i}}{\frac{1}{s_j}+\sigma\sum_{i=1}^{n}\frac{x_{ij}^2}{2v_i}},\frac{1}{\frac{1}{s_j}+\sigma\sum_{i=1}^{n}\frac{x_{ij}^2}{2v_i}}\right)$$

$$f(\eta^2|\boldsymbol{\beta_\tau},\zeta,\sigma,\boldsymbol{s},\boldsymbol{v},\delta,\boldsymbol{y}^*)\sim Gamma\left(\zeta+k,v+\sum_{j=0}^{k}\frac{s_j}{2}\right),$$

$$f(v|\boldsymbol{\beta_\tau},\eta^2,\zeta,\sigma,\boldsymbol{s},\boldsymbol{v},\delta,\boldsymbol{y}^*)\sim Gamma(\zeta,\eta^2),$$

$$f(\zeta|\boldsymbol{\beta_\tau},\eta^2,\sigma,\boldsymbol{s},\boldsymbol{v},\delta,\boldsymbol{y}^*)\sim Gamma(\zeta,\eta^2),$$

$$f(v_i|\boldsymbol{\beta_\tau},\eta^2,v,\zeta,\sigma,\boldsymbol{s},\boldsymbol{v},\delta,\boldsymbol{y}^*)\sim GIG\left(\frac{1}{2},\left(\frac{(y_i^*-x_i'\beta_\tau)^2}{p^2\sigma}\right),\left(\frac{2}{\sigma}+\frac{\theta^2}{p^2\sigma}\right)\right),$$

$$f(s_i|\boldsymbol{\beta_\tau},\eta^2,v,\zeta,\sigma,\boldsymbol{s},\boldsymbol{v},\delta,\boldsymbol{y}^*)\sim GIG\left(\frac{1}{2},\beta_j^2,\eta^2\right),$$

$$f(\sigma|\boldsymbol{\beta_\tau},\eta^2,v,\zeta,\boldsymbol{s},\boldsymbol{v},\delta,\boldsymbol{y}^*)\sim GIG\left(a+\frac{3n}{2},(b+\sum_{i=1}^{n}\left(\frac{(y_i^*-(\boldsymbol{x}_i'\boldsymbol{\beta_\tau}+\sigma v_i)^2}{2p^2v_i}\right)+v_i\right)$$

Data

The data used in this study are secondary data. The data was obtained directly from the Air Dingin Health Center, Padang City. The amount of data amounted to 148 data from 2022 to 2023. The dependent variable (Y) in this study is the classification of height of stunting toddlers based on Height-for-age Z-score (HAZ) is an indicator used to assess children's growth status by comparing a child's height at a given age against the WHO healthy child growth standards. Based on the literature, the independent variables that are assumed to affect the dependent variable are age, weight at birth, gender, how to measure height, nutritional status. The description of the research variables can be seen in the following Table below:

Table1. Descriptive Statistics of data category

Variable	Category	Frequency	Percentage
Height Gain	Short	122	82.43%
	Very Short	26	17.57%
Gender	Male	91	61.48%
	Female	57	38.52%
How to measure height	Stand up	106	71.62%
	Recumbent	42	28.38%
Nutrition Status	Malnutrition	5	3.37%
	Undernutrition	16	10.81%
	Normal Nutrition	109	73.64%
	Overnutrition	5	3.37%
	At risk of overnutrition	13	8.78%

Data Analysis

a. Creating statistic descriptive data.

b. Analysis Bayesian Binary LASSO Quantile Regression (BBLQR)

   1. Prior Distribution Determination, likelihood function and posterior distribution parameter.
   2. Determine the posterior mean of each parameter.
   3. Estimating model parameters and significance tests in each quantile.
   4. Estimating the value of the width of the 95% confidence interval in each quantile.
   5. Calculating the Mean Square Error (MSE) value in each quantile.

c. Comparing the estimation results, 95% confidence interval width, MSE value of each quantile between the two methods.

d. Selecting the best model and checking the trace plot and density plot.

e. Conclusions and interpretations on the selected model.

RESULT AND DISCUSSION

This stage, parameter estimation will be carried out using Software R with the Gibbs Sampling approach. The quantile chosen in this study is the quantile $\tau = 0.05; 0.25; 0{,}55; 0.75; 0.95$. The selection of quantiles in research is usually based on the purpose of the data analysis and the characteristics of the data distribution. Quantiles divide the data into equal parts and are used to understand the distribution of the data. The results of parameter estimation using the BBQR and BBLQR methods can be presented in the Table below:

Tabel 2. Estimated value of β for each quantile – τ using BBQR and BBLQR

Independent Variables	BBQR		BBLQR	
	Parameter $(\widehat{\beta})$	Width CI 95%	Parameter $(\widehat{\beta})$	Width CI 95%
$\tau = 0.05$				
Intercept	-5.0319	5.7095	-0.0220	0.1480
X_1(Age)	1.2234	1.3475	0.0107	0.0516
X_2(weight at birth)	-0.0004	0.0042	-0.0007	0.0003
X_3(Gender)	-0.4602	2.4955	-0.0021	0.0745
X_4 (how to measure height)	-0.7775	4.0454	-0.0015	0.1110
X_5(nutritional status)	0.0736*	1.4374	-0.0037*	0.3540
$\tau = 0.25$				
Intercept	-0.3043	2.7222	0.5090	1.0500
X_1(Age)	0.5017*	0.6679	0.0388*	0.1140
X_2(weight at birth)	-0.0001*	0.0017	-0.0001*	0.0004
X_3(Gender)	-0.1994	1.1414	-0.0111	0.1760
X_4 (how to measure				

height)	-0.3795	1.7478	0.3805	0.8860
X_5(nutritional status)	-0.1194*	0.6613	-0.0310*	0.1670
$\tau = 0.55$				
Intercept	1.0900	2.9033	0.9990	0.1180
X_1(Age)	0.2180	0.6189	0.0013	0.0166
X_2(weight at birth)	0.0004*	0.0017	-0.0041*	0.0002
X_3(Gender)	-0.1070	1.2064	0.0003	0.3250
X_4 (how to measure height)	-0.1380	1.8105	0.0007	0.0562
X_5(nutritional status)	-0.1440*	0.7782	-0.0017*	0.0378
$\tau = 0.75$				
Intercept	1.7913	3.6718	1.0000	0.0696
X_1(Age)	0.1772	0.7479	0.0070	0.0107
X_2(weight at birth)	0.0001*	0.0023	0.00006*	0.0001
X_3(Gender)	-0.0566	1.4677	-0.0003	0.0216
X_4 (how to measure height)	-0.0882	2.2695	-0.0018	0.3590
X_5(nutritional status)	-0.1166*	1.0092	-0.00002*	0.0222
$\tau = 0.95$				
Intercept	4.1742	6.6331	0.0101	0.0035
X_1(Age)	0.17806	1.4182	0.0003	0.0060
X_2(weight at birth)	0.0004*	0.0047	0.0001*	0.0008
X_3(Gender)	0.0117	2.8245	-0.0004	0.0170
X_4(how to measure height)	-0.080	4.2424	-0.0026	0.0024
X_5 (nutritional status)	0.0114*	1.8904	0.0003*	0.0090

***significantly $\alpha = 0,05$.**

Table 2 shows the variables of age, weight at birth and nutritional status are significant at the quantile in both methods, and variable age and nutritional status significant in all quantile in both methods because the confidence interval does not contain zero. Variable of gender, how to measure height not significantly in all quantile. MSE is used for the indicator of model goodness from the comparison of the BBQR and BBLQR methods. MSE values can be presented in Table 3 below:

Table 3. MSE Value BBQR and BBLQR

Quantile	MSE	
	BBQR	BBLQR
0,05	10.5670	0.8353
0,25	0.60516	**0.1622**
0,55	0.26533	0.1730
0,75	1.24016	0.1758
0,95	15.2410	0.1785

Based on the Table 3 above, it can be seen that the quantile that has the smallest MSE value is quantile 0.25. With these results, the model for the classification of stunting toddler height is as follows:

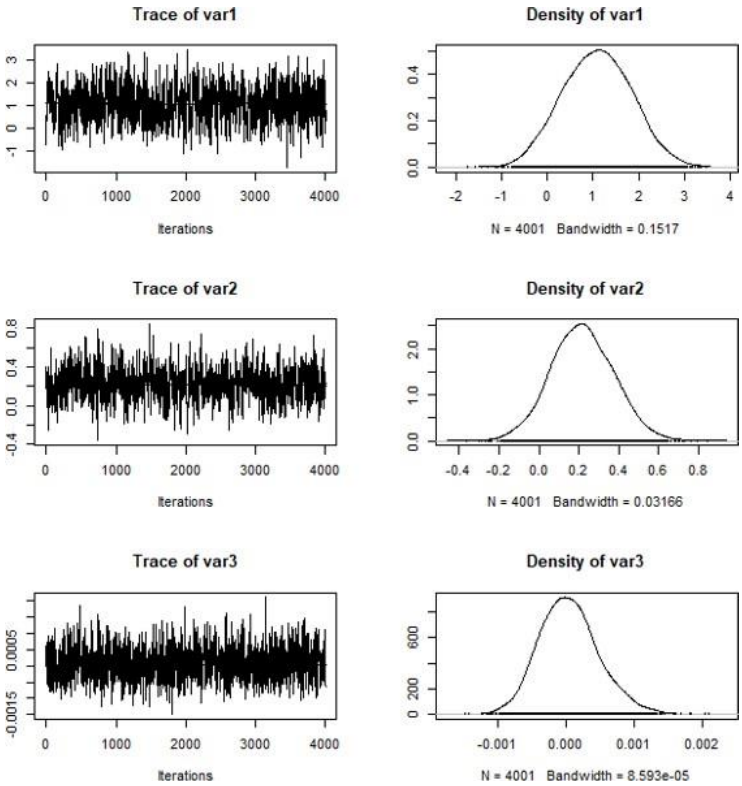
$$Logit(p_i) = 0.5090 + 0.0388X_1 - 0.0001X_2 - 0.0111X_3 + 0.3805X_4 - 0.0310X_5$$

Based equation above can be interpreted by looking at odds ratio values in Table 4.

Table 4. Odd Ratio at Quantile $\tau = 0.25$

Independent Variables	Estimated Mean $(\widehat{\beta})$	Odd Ratio
X_1 (Age)	0.5090	1.6637
X_2 (weight at birth)	0.0388	1.0396
X_3 (Gender)	-0.0001	0.9999
X_4 (how to measure height)	-0.0111	0.9889
X_5 (nutritional status)	0.3805	1.4631

Based on Table 4 above, the stunting height classification model can be interpreted. For every additional year of child age, the chance of stunting is 50%. If the toddler's weight increases by 1 kg, then the chance of experiencing stunting is 3%. Furthermore, to see the convergence of each parameter in the 0.25 quantile, trace plots and density plots were used. It can be seen that the parameters have converged and stabilized around the mean. More clearly related to this



convergence can be seen in the figure below:

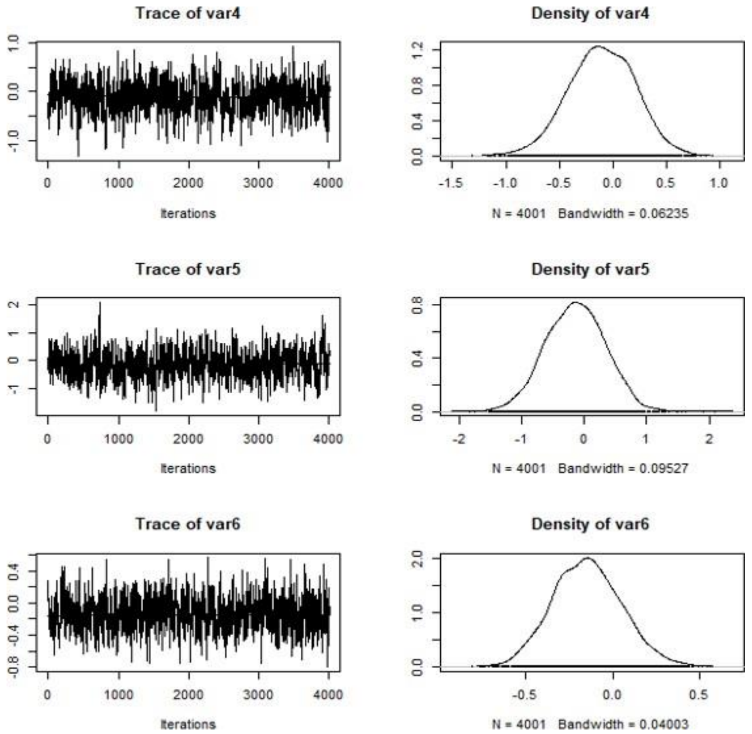


Figure 1. Trace plot and Density Plot for parameters in quantile 0.25

In Figure 1, the trace plot shows that the parameter values generated by the Markov Chain Monte Carlo (MCMC) algorithm at each iteration have converged or stabilized. It can be seen that the parameter value is already at the limit of the 95% confidence interval, and the trace plot shows a flat random movement. Density plot shows the posterior distribution of the parameters after MCMC iteration. It is like a smoothed histogram to estimate the probability distribution. Density plot shows the posterior distribution of the regression parameters in each quantile is very important because in quantile regression the errot distribution is not assumed to be normal like ordinary regression. From the results, the parameters are close to the normal distribution or have stabilized.

CONCLUSION

Based on the results of data analysis, this study can be concluded that the classification model of stunting toddler height obtained the best model in quantile 0.25 seen from the smallest MSE value of all quantiles. Independent variables that have a significant influence are age, nutritional status. These results can be used as input for related agencies and add insight related to the methods used for research needs. This research has limitations related to the fact that it is still vulnerable to overfitting, so that parameter regulation is needed, as well as testing the accuracy of model parameters with more robust methods.

ACKNOWLEDGMENT

Thank you for the support from the Mathematics Study Program, Faculty of Science and Technology, Universitas Islam Negeri Imam Bonjol Padang. We would also like to thank the Department of Mathematics and Data Science, Andalas University.

REFERENCES

Alhamzawi, R., Yu, K., & Benoit, D. F. (2012). Bayesian adaptive Lasso quantile regression. *Statistical Modelling*, *12*(3), 279–297. https://doi.org/10.1177/1471082X110120030

Ariati, N. N., Padmiari, I. A. E., Sugiani, P. P. S., & Suarni, N. N. (2018). Description of nutritional status and the incidence of stunting children in early childhood education programs in Bali-Indonesia. *Bali Medical Journal (Bali Med J) 2018, Volume 7, Number 3*, *7*(3), 723–726.

Benoit, D. F., Alhamzawi, R., & Yu, K. (2013). Bayesian lasso binary quantile regression. *Computational Statistics*, *28*, 2861–2873. https://doi.org/10.1007/s00180-013-0439-0

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. https://doi.org/10.1002/jae.1216

Davino, C., Furno, M., & Vistocco, D. (2013). *Quantile regression: theory and applications* (Vol. 988). John Wiley & Sons.

Harianti Hasibuan, L., Musthofa, S., Studi Matematika, P., & Imam Bonjol Padang, U. (2022). Journal of Science and Technology Penerapan Metode Regresi Linear Sederhana Untuk Prediksi Harga Beras di Kota Padang. *Journal of Science and Technology*, *2*(1), 85–95. https://doi.org/10.15548/jostech.v2i1.3802

Hasibuan, L. H., Putri, D. M., & Jannah, M. (2022). Simple Linear Regression Method to Predict Cooking Oil Prices in the Time of Covid-19. *Logaritma: Jurnal Ilmu-Ilmu Pendidikan Dan Sains*, *10*(01), 81–94.

Hasibuan, L. H., Putri, D. M., Jannah, M., & Musthofa, S. (2022). Analisis Metode Single Exponential Smoothing Dan Metode Regresi Linear Untuk Prediksi Harga Daging Ayam Ras. *Math Educa Journal*, *6*(2), 120–130. https://doi.org/10.15548/mej.v6i2.3872

Hasibuan, L. H., Yanuar, F., Devianto, D., & Maiyastri, M. (2024). Quantile Regression Analysis; Simulation Study With Violation of Normality Assumption. *JOSTECH Journal of Science and Technology*, *4*(2), 133–142. https://doi.org/10.15548/jostech.v4i2.9643

Hasibuan, L. H., Yanuar, F., Devianto, D., Maiyastri, M., & Apriona, A. (2025). Bayesian Method for Linear Regression Modeling; Simulation Study with Normality Assumption. *JOSTECH Journal of Science and Technology*, *5*(1), 81–92. https://doi.org/10.15548/jostech.v5i1.10994

Hasibuan, L. H., Yanuar, F., Harahap, V. P., & Qalbi, L. (2024). Performance Quantile Regression and Bayesian Quantile Regression in Dealing with Non-normal Errors (Case Study on Simulated Data). *Numerical: Jurnal Matematikadan Pendidikan Matematika*, *8*. https://doi.org/10.25217/numerical.v8i2.4922

Huriah, T., & Nurjannah, N. (2020). Risk factors of stunting in developing countries: A scoping review. *Open Access Macedonian Journal of Medical Sciences*, *8*(F), 155–160. https://doi.org/10.3889/oamjms.2020.4466

Kang, Y., Aguayo, V. M., Campbell, R. K., Dzed, L., Joshi, V., Waid, J. L., Gupta, S. D., Haselow, N. J., & West Jr, K. P. (2018). Nutritional status and risk factors for stunting in preschool children in Bhutan. *Maternal & Child Nutrition*, *14*, e12653. https://doi.org/10.1111/mcn.12653

Koenker, R., & Bassett Jr, G. (1978). Regression quantiles. *Econometrica: Journal of the Econometric Society*, 33–50. https://doi.org/10.2307/1913643

Leroy, J. L., & Frongillo, E. A. (2019). Perspective: what does stunting really mean? A critical review of the evidence. *Advances in Nutrition*, *10*(2), 196–204. https://doi.org/10.1093/advances/nmy101

Ma, Z., Han, Z., & Wang, M. (2023). Bayesian variable selection and estimation in binary quantile regression using global-local shrinkage priors. *Communications in Statistics-Simulation and Computation*, 1–15. https://doi.org/10.1080/03610918.2023.2293638

Mustakim, M. R. D., Irawan, R., Irmawati, M., & Setyoboedi, B. (2022). Impact of Stunting on Development of Children between 1-3 Years of Age. *Ethiopian Journal of Health Sciences*, *32*(3). 10.4314/ejhs.v32i3.13

Sholih, A. S., Hasibuan, L. H., & Rianjaya, I. D. (2024). Analisis Regresi Logistik Ordinal Terhadap Faktor-Faktor yang Mempengaruhi Predikat Kelulusan Mahasiswa Sarjana UIN Imam Bonjol Padang. *MAp (Mathematics and Applications) Journal*, *6*(2), 99–110. https://doi.org/10.15548/map.v6i2.10075

Vitaloka, F. S. W., Setya, D. N., & Widyastuti, Y. (2019). *Hubungan Status Anemia Ibu Hamil dengan Kejadian Stunting Balita Usia 24-59 Bulan di Wilayah Kerja Puskesmas Gedangsari II Gunung Kidul*. Poltekkes Kemenkes Yogyakarta.

Wilson, H. J., Dickinson, F., Griffiths, P. L., Azcorra, H., Bogin, B., & InÊS Varela-Silva, M. (2011). How useful is BMI in predicting adiposity indicators in a sample of Maya children and women with high levels of stunting? *American Journal of Human Biology*, *23*(6), 780–789. https://doi.org/10.1002/ajhb.21215

Wizsa, U. A., & Rahmi, A. (2025). Classification Of KIP-K Scholarship Using Logistic Regression, Classification Trees, and Boosting Based On Decision Support System. *Mathline: Jurnal Matematika Dan Pendidikan Matematika*, *10*(1), 221–235. https://doi.org/10.31943/mathline.v10i1.837

Yanuar, F., Deva, A. S., Maiyastri, M., Yozza, H., & Zetra, A. (2021). Modeling Length of Hospital Stay for Patients With COVID-19 in West Sumatra Using Quantile Regression Approach. *CAUCHY: Jurnal Matematika Murni Dan Aplikasi*, *7*(1), 118–128. https://doi.org/10.18860/ca.v7i1.12995

Yanuar, F., Deva, A. S., & Zetra, A. (2023). Length of hospital stay model of COVID-19 patients with quantile Bayesian with penalty LASSO. *Commun. Math. Biol. Neurosci.*, *2023*, Article-ID. https://doi.org/10.28919/cmbn/7881

Yanuar, F., Zetra, A., Putri, A. R., & Asdi, Y. (2020). *Bayesian LASSO Quantile Regression: An Application to the Modeling of Low Birth Weight*.

Yasril, A. I., & Sari, M. (2022). Aplikasi Structural Equation Modeling Untuk Mengkaji Faktor Penyebab Stunting Di Provinsi Sumatera Barat. *Jurnal Endurance*, *7*(1), 1–12. http://doi.org/10.22216/jen.v7i1.656

Yu, K., Lu, Z., & Stander, J. (2003). Quantile regression: applications and current research areas. *Journal of the Royal Statistical Society Series D: The Statistician*, *52*(3), 331–350. https://doi.org/10.1111/1467-9884.00363