Available online at http://scik.org

Commun. Math. Biol. Neurosci. 2023, 2023:23

https://doi.org/10.28919/cmbn/7881

ISSN: 2052-2541

LENGTH OF HOSPITAL STAY MODEL OF COVID-19 PATIENTS WITH QUANTILE BAYESIAN WITH PENALTY LASSO

FERRA YANUAR[1,*], ATHIFA SALSABILA DEVA[1], AIDINIL ZETRA[2], MAIYASTRI[1]

[1]Department of Mathematics and Data Science, Andalas University, Padang, Indonesia

[2]Department of Political Science, Andalas University, Padang, Indonesia

Abstract: This study aims to identify the best model for the length of hospital stay of COVID-19 patients in West Sumatra Province using Bayesian LASSO quantile regression method and Bayesian Adaptive LASSO quantile regression method. The quantile analysis is employed in Bayesian concept to produce more effective and natural estimated values, especially for data with non-normal distribution. The combination of the LASSO method and Adaptive LASSO as a variable selection method was applied to obtain the best model and produce estimated values that are close to the actual values. A comparison of the estimated values generated from the two methods was conducted using data from 1737 COVID-19 patients at M. Djamil General Hospital in Padang from March to December 2020. The result obtained is that the Bayesian Adaptive LASSO quantile regression method generally yields a shorter 95% confidence interval, with MAD (Median Absolute Deviation), MSE (Mean Squared Error), RMSE (Root of Mean Squared Error) values smaller than those produced by the Bayesian LASSO quantile regression method. The length of hospital stay of COVID-19 patients in West Sumatra is significantly influenced by age, the diagnosis of COVID-19 patients in the positive category, the patient's discharge status in the cured and death categories, and the

*Corresponding author

E-mail address: ferrayanuar@sci.unand.ac.id

Received January 09, 2023

number of comorbidities. Below the 0.50 quantile, the length of hospital stay for patients diagnosed with positive COVID-19 who were then declared cured is around three days and 4 hours longer than the length of stay for patients diagnosed with Person Under Supervision (PerUS). It is approximately 9 hours and 50 minutes longer than the length of stay of COVID-19 patients forced to go home. The length of stay of COVID-19 patients who died was around 16 hours 31 minutes shorter than the length of stay of COVID-19 patients who were forced to discharge from the hospital.

Keywords: length of hospitalization; COVID-19; Bayesian quantile regression; LASSO; adaptive LASSO.

2020 AMS Subject Classification: 62F15.

1. INTRODUCTION

COVID-19 infects the human respiratory system and causes several symptoms, such as fever, cough, and shortness of breath. The COVID-19 pandemic that emerged in early 2020 significantly impacted various countries, including Indonesia. The COVID-19 shock has adversely affected Indonesia's long-term income mobilization capacity [1]. COVID-19 will also impact the effectiveness of health services because many patients need immediate treatment. Therefore, significant action is required to address this pressing health problem while strengthening the primary health care system [2]. In Indonesia, especially West Sumatra Province, until early June 2021, 43,281 COVID-19 patients were treated at the hospital. The length of stay of COVID-19 patients in the hospital is influenced by certain factors [3]. Identifying which factors significantly affect the length of stay of COVID-19 patients is necessary. The estimated length of hospital stay model obtained can be used for the benefit of health service activities, the need for health facilities, and preparation for making decisions related to preparedness for COVID-19 to reduce the negative impacts on the health, economic and social sectors [4]–[6].

West Sumatra is one of ten provinces with the most cases of COVID-19 in Indonesia [7]. The limited facilities and infrastructure in West Sumatra caused the medical teams to experience difficulties in treating the overflowing COVID-19 patients. This study examines the length of hospital stay model for COVID-19 patients in West Sumatra using the Bayesian LASSO quantile regression and Adaptive LASSO methods. This method was used because the preliminary analysis

found that the data distribution was not normally distributed. The use of quantile analysis in the Bayesian concept aims to produce more effective and natural model parameter estimates, especially for data that is not normally distributed and violates other classical assumptions [8]–[10]. The combination of the LASSO method and Adaptive LASSO as a method of selecting independent variables as well as a method of regularization in Bayesian quantile analysis is carried out to obtain the best model and produce estimated values that are close to the estimated actual values [11]–[14].

Studies related to the Bayesian quantile regression method using the LASSO method and Adaptive LASSO began with the concept of quantile regression by Koenker and Basset [15] and the concept of regularization of the LASSO method from Thisbirani's research[16]. The Bayesian quantile regression method using the Asymmetric Laplace distribution (ALD) for its likelihood function was first introduced by Yu & Mooyed [10]. The combination of the Bayesian quantile regression method and LASSO was carried out by Li et al. [17] by adding regularization parameters to the parameter estimation process. Kozumi and Kobayashi investigated numerical simulations of the Gibbs Sampling algorithm in the Bayesian quantile regression method [18]. Research on the Bayesian quantile regression method using the LASSO and Adaptive LASSO methods was developed by Alhamzawi et al. [12] and Tang et al. [13]. Hamid and Al-Husseini examined the Bayesian LASSO concept in composite quantile regression [19]. Yanuar et al. modeled the low birth weight of newborns using Bayesian quantile regression [20] and by combining the Bootstrap method [21]. Alhamzawi and Mallick developed the reciprocal LASSO method on the Bayesian quantile regression method [22]. Algamal et al discussed DNA selection using Bayesian quantile regression [23]. Several Bayesian studies discussing COVID-19, including Yanuar et al., modeled the length of hospital stay of COVID-19 patients using Bayesian quantile regression [24] and examined the performance of the Bayesian concept in Structural Equation Modeling (SEM) analysis of health behavior during the pandemic of COVID-19 in West Sumatra [20]. Therefore, this study aims to construct the length of hospital stay of COVID-19 patients in West Sumatra using modification of Quantile regression methods.

2. PRELIMINARIES

2.1. Factors of Length of Hospital Stay for Patients with COVID-19

The length of stay of COVID-19 patients in the hospital is assumed to be influenced by age [25]–[27]. The age group of 18 – 25 years has a severe level of vulnerability [25]. Elderly patients have a higher death risk [28] and are more helpless, and are treated longer [5], [27], [29], [30]. The length of stay of COVID-19 patients is also influenced by gender. COVID-19 is more dangerous for men[28], with a susceptibility rate of 33% [25]. Female patients with critical illnesses have more extended stays [4]. Diagnostic factors related to COVID-19 also affect the patient's length of stay in the hospital [3]–[5], [24]. Patients with a positive diagnosis and Patient Under Supervision (PaUS) undergo hospitalization for at least 14 days [3]. Diagnosis and administration of D-Dimer [31] and ceftriaxone in female patients reduce the length of stay [4]. Another factor that is assumed to affect the length of stay of COVID-19 patients is the illnesses (comorbidities) that the patients experienced before. Patients with hypertension, diabetes [4], [30], [32], coronary arteries [32], kidney complications [33], organ failure and decreased leukocytes[34], fever [35], obesity [30] or patients who have had more than 2 comorbidities [6], will tend to undergo more extended hospitalization than other patients who do not have comorbidities. The discharge status of COVID-19 patients is also a determining factor in the length of stay of COVID-19 patients [3], [5], [24], [33]. COVID-19 patients are declared cured after being treated in a hospital for at least 14 days [3]. Patients who are declared dead have a shorter stay, on average 5 or 6 days [33].

2.2. Real Data

The data used in this study were COVID-19 patients data obtained directly from the M. Djamil Central General Hospital (RSUP) Padang City as a referral hospital in West Sumatra Province. The data used was 1737 patients hospitalized at the hospital from March to December 2020. Table 1 below presents a description of the data used.

Table 1. Statistics Descriptive of COVID-19 Patients in West Sumatra

Variable	Category	Frequency	Percentage (%)
Gender	Male	888	51,12 %
	Female	849	48,88 %
Diagnose COVID-19	Person Under Supervision (PerUS)	7	0,4 %
	Patients Under Supervision (PaUS)	1273	73,29 %
	Positive	457	26,31 %
Discharge Status	Cured	1374	79,10 %
	Die	280	16,12 %
	Outpatient	13	0,75 %
	Forced discharge	70	4,03 %

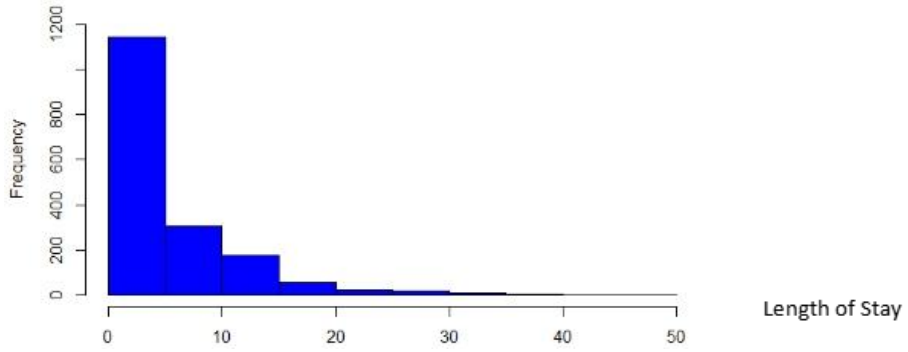


Figure 1. Histogram of Length of Hospitalization for COVID-19 Patients in West Sumatra.

Table 1 informs that of the 1737 COVID-19 patients treated, 51.12% were male patients, 73.29% were patients diagnosed with PaUS (Patients under Supervision), and 79.10% were declared cured. In Figure 1, the histogram of the length of stay of COVID-19 patients is more skewed to the left and asymmetrical as a normal distribution curve should be. The data are not normally distributed, and the quantile regression method is appropriate for data modeling like this. In this study, variables such as gender, COVID-19 diagnosis, and discharge status were independent variables with categorical data types. Therefore, these variables were changed to dummy variables for the estimation process in the regression analysis. Table 2 below presents the formation of dummy variables.

Table 2. Dummy Variables for Gender, COVID-19 diagnosis, and Discharge Status.

Gender (X_2)			
		X_{2D1}	
Male		1	
Female		0	
COVID-19 Diagnosis (X_3)			
	X_{3D1}	X_{3D2}	
PerUS	0	0	
PaUS	1	0	
Positive	0	1	
Discharge Status (X_4)			
	X_{4D1}	X_{4D2}	X_{4D3}
Cured	1	0	0
Die	0	1	0
Outpatient	0	0	1
Forced discharge	0	0	0

2.3. Quantile Regression Method

If a vector $\boldsymbol{y} = (y_1, y_2, \cdots y_n)'$ is declared as the response variable and $\boldsymbol{x} = (x_1, x_2, \cdots x_k)'$ is defined as the predictor variable, then the linear regression equation model for the τ^{th} quantile, where $0 < \tau < 1$ with n samples and k predictors for $i = 1,2, \ldots, n$ is written as:

$$y_i = \beta_{0\tau} + \beta_{1\tau}x_{i1} + \beta_{2\tau}x_{i2} + \cdots + \beta_{k\tau}x_{ik} + \varepsilon_i, \quad (1)$$

with $\boldsymbol{\beta}(\tau)$ as the parameter vector and $\boldsymbol{\varepsilon}$ as the residual vector. The estimated value of the parameters in the quantile regression equation $\widehat{\boldsymbol{\beta}}(\tau)$ is obtained by minimizing the following equation [15]:

$$\sum_{i=1}^{n} \rho_\tau(y_i - x_i'\boldsymbol{\beta}), \quad (2)$$

with $\rho_\tau(u) = u\big(\tau - I(u < 0)\big)$is the equivalent *loss function* with the equation:

$$\rho_\tau(\varepsilon) = \varepsilon\big(\tau I(\varepsilon > 0) - (1 - \tau)I(\varepsilon < 0)\big), \quad (3)$$

$I(.)$ is an indicator function, which has a value of 1 when $I(.)$ is true and 0 otherwise.

2.4. Bayesian Quantile Regression Method

Yu and Moyeed [10] suggested that minimizing the loss function of the quantile regression is equivalent to maximizing the likelihood function of the Asymmetric Laplace Distribution (ALD)

because the loss function in quantile regression is identical to the ALD likelihood function. ALD is used to form the likelihood function so that the estimator becomes more effective and natural or close to the true value so that the correct estimation process can be produced. The ALD distribution is a continuous probability distribution. Random variable ε with ALD distribution with probability density function $f(\varepsilon)$, i.e.:

$$f_\tau(\varepsilon) = \tau(1-\tau)exp\big(-\rho_\tau(\varepsilon)\big), \tag{4}$$

with $0 < \tau < 1$ and $\rho_\tau(\varepsilon)$as defined in equation (3). ALD has a combined representation of several distributions based on the exponential and the normal distribution, which are used to form the likelihood function. Let Z be a random variable with an exponential distribution $Z \sim \exp(1)$, and U is a random variable with a standard normal distribution $U \sim N(0,1)$. If $\varepsilon\, is$ the random variable that has an ALD distribution, it can be expressed as [15]:

$$\varepsilon = \theta z + pu\sqrt{z}, \tag{5}$$

where $\theta = \frac{1-2\tau}{(1-\tau)\tau}$ and $p^2 = \frac{2}{(1-\tau)\tau}$. Based on equation (5), the likelihood function used in parameter $\boldsymbol{\beta}$ estimation for the τ^{th} quantile in the Bayesian quantile regression analysis is formulated in equation (6) as follows[14]:

$$L(\boldsymbol{\beta}_\tau, \boldsymbol{v}, \sigma) \propto \left(\prod_{i=1}^n (\sigma v_i)^{-\frac{1}{2}}\right)\left(exp\left(-\sum_{i=1}^n \frac{\left(y_i-(x_i'\boldsymbol{\beta}_\tau+\theta v_i)\right)^2}{2p^2\sigma v_i}\right)\right), \tag{6}$$

with $\sigma > 0$ as the scale parameter and $v_i = \sigma z_i$ with $\exp(\sigma)$ distribution. Prior distribution was chosen for parameters $\boldsymbol{\beta}_\tau \sim N(\boldsymbol{b}_0, \boldsymbol{B}_0)$, $v_i \sim \exp(\sigma)$, and $\sigma \sim IG(a,b)$. The corresponding posterior distribution is obtained as follows:

$$(\boldsymbol{\beta}_\tau|\boldsymbol{v},\sigma,\boldsymbol{y}) \sim N\left[\left(\boldsymbol{B}_0{}^{-1} + \sum_{i=1}^n \frac{\boldsymbol{x}_i\boldsymbol{x}_i'}{p^2\sigma v_i}\right)^{-1}\left(\boldsymbol{B}_0{}^{-1}\boldsymbol{b_0} + \sum_{i=1}^n \frac{\boldsymbol{x}_i(y_i-\theta v_i)}{p^2\sigma v_i}\right), \left(\boldsymbol{B}_0{}^{-1} + \sum_{i=1}^n \frac{\boldsymbol{x}_i\boldsymbol{x}_i'}{p^2\sigma v_i}\right)^{-1}\right];, \tag{7}$$

$$(v_i|\boldsymbol{\beta}_\tau,\sigma,\boldsymbol{y}) \sim GIG\left(\frac{1}{2}, \left(\frac{\left(y_i-x_i'\boldsymbol{\beta}_\tau\right)^2}{p^2\sigma}\right), \left(\frac{2}{\sigma} + \frac{\theta^2}{p^2\sigma}\right)\right);, \tag{8}$$

$$(\sigma|\boldsymbol{\beta}_\tau,\boldsymbol{v},\boldsymbol{y}) \sim IG\left(\left(a+\frac{3n}{2}\right), \left(b + \sum_{i=1}^n v_i + \sum_{i=1}^n \frac{\left(y_i-(x_i'\boldsymbol{\beta}_\tau+\theta v_i)\right)^2}{2p^2 v_i}\right)\right). \tag{9}$$

2.5. Bayesian Quantile Regression with Penalty LASSO and Adaptive LASSO.

The prior distribution of the regression model for the τ^{th} quantile for n samples with k independent variables using the Bayesian LASSO quantile regression method is:

$$\begin{aligned} f(\beta_j) &\propto \frac{1}{\sqrt{2\pi s_j}} \exp\left(-\frac{\beta_j^2}{2s_j}\right), \\ f(v_i) &\propto \frac{1}{\sigma} \exp\left(-\frac{v_i}{\sigma}\right), \\ f(\sigma) &\propto \sigma^{-a-1} \exp\left(-\frac{b}{\sigma}\right), \\ f(s_j) &\propto \exp\left(-\frac{\eta_j^2}{2} s_j\right), \\ f(\eta^2) &\propto \frac{1}{d^{-c}\Gamma(c)} (\eta^2)^{c-1} \exp(-d\eta^2), \end{aligned} \tag{10}$$

with $\boldsymbol{s} = (s_0, s_1, \cdots, s_k),\ i = 1,2,\cdots,n,\ \sigma > 0,\ v_i > 0,\ a > 0,\ b > 0,\ s_j > 0,\ \eta^2 > 0,\ c > 0,$ $d > 0$. Based on equation (10), the posterior distribution of the quantile Bayesian LASSO regression model is obtained as follows:

$$\begin{aligned} f(\beta_j|\sigma, \boldsymbol{v}, \boldsymbol{s}, \eta^2, \boldsymbol{y}, \boldsymbol{\beta}_{-j}) &\sim N\left(\frac{\frac{1}{p^2\sigma}\sum_{i=1}^{n}\frac{\hat{y}_{ij}x_{ij}}{v_i}}{\frac{1}{s_j}+\frac{1}{p^2\sigma}\sum_{i=1}^{n}\frac{x_{ij}^2}{v_i}}, \frac{1}{\frac{1}{s_j}+\frac{1}{p^2\sigma}\sum_{i=1}^{n}\frac{x_{ij}^2}{v_i}}\right), \\ f(v_i|\sigma, \boldsymbol{v}_{-i}, \boldsymbol{s}, \eta^2, \boldsymbol{y}, \boldsymbol{\beta}) &\sim GIG\left(\frac{1}{2}, \frac{(y_i - \boldsymbol{x}_i'\boldsymbol{\beta})^2}{p^2\sigma}, \frac{2}{\sigma} + \frac{\theta^2}{p^2\sigma}\right), \\ f(\sigma|\boldsymbol{v}, \boldsymbol{s}, \eta^2, \boldsymbol{y}, \boldsymbol{\beta}) &\sim IG\left(a + \frac{3n}{2}, b + \sum_{i=1}^{n}\left(\frac{\left(y_i - (\boldsymbol{x}_i'\boldsymbol{\beta} + \boldsymbol{\theta}v_i)\right)^2}{2p^2v_i} + v_i\right)\right), \\ f(s_j|\sigma, \boldsymbol{v}, \eta^2, \boldsymbol{y}, \boldsymbol{\beta}, \boldsymbol{s}_{-j}) &\sim GIG\left(\frac{1}{2}, \beta_j, \eta^2\right), \\ f(\eta^2|\sigma, \boldsymbol{v}, \boldsymbol{s}, \boldsymbol{y}, \boldsymbol{\beta}) &\sim Gamma\left(c + k, d + \sum_{j=0}^{k}\frac{s_j}{2}\right). \end{aligned} \tag{11}$$

The prior distribution of the regression model for the τ^{th} quantile from n samples with k independent variables using the Bayesian Adaptive LASSO quantile regression method as follows:

$$\begin{aligned} f(\beta_j) &\propto \frac{1}{\sqrt{2\pi s_j}} \exp\left(-\frac{\beta_j^2}{2s_j}\right), \\ f(v_i) &\propto \sigma \exp(-\sigma v_i), \end{aligned} \tag{12}$$

$$f(\sigma) \propto \frac{b^a}{\Gamma(a)}\sigma^{a-1}\exp(-b\sigma),$$

$$f(s_j) \propto \frac{\sigma}{2\lambda_j^2}\exp\left(-\frac{\sigma}{2\lambda_j^2}s_j\right),$$

$$f(\lambda_j^2) \propto \frac{\delta^\alpha}{\Gamma(\alpha)}(\lambda_j^2)^{-\alpha-1}\exp\left(-\frac{\delta}{\lambda_j^2}\right),$$

with $\boldsymbol{s} = (s_0, s_1, \cdots, s_k),\ i = 1,2,\cdots,n,\ \sigma > 0,\ v_i > 0,\ a > 0,\ b > 0,\ s_j > 0,\ \lambda_j^2 \geq 0,\ \alpha > 0,$ $\delta > 0$. Based on equation (12) posterior distribution of the regression model Bayesian Adaptive LASSO quantiles is obtained as:

$$f(\beta_j|\sigma, \boldsymbol{v}, \boldsymbol{s}, \boldsymbol{\lambda}^2, \boldsymbol{y}, \boldsymbol{\beta}_{-j}) \sim N\left(\frac{\frac{\sigma}{p^2}\sum_{i=1}^n \frac{\hat{y}_{ij}x_{ij}}{v_i}}{\frac{1}{s_j}+\frac{\sigma}{p^2}\sum_{i=1}^n \frac{x_{ij}^2}{v_i}}, \frac{1}{\frac{1}{s_j}+\frac{\sigma}{p^2}\sum_{i=1}^n \frac{x_{ij}^2}{v_i}}\right),$$

$$f(v_i|\sigma, \boldsymbol{v}_{-i}, \boldsymbol{s}, \boldsymbol{\lambda}^2, \boldsymbol{y}, \boldsymbol{\beta}) \sim GIG\left(\frac{1}{2}, \frac{\sigma(y_i - \boldsymbol{x}_i'\boldsymbol{\beta})^2}{p^2}, 2\sigma + \frac{\sigma\theta^2}{p^2}\right),$$

$$f(\sigma|\boldsymbol{v}, \boldsymbol{s}, \boldsymbol{\lambda}^2, \boldsymbol{y}, \boldsymbol{\beta}) \sim Gamma\left(a + k + \frac{3n}{2}, b + \sum_{i=1}^n\left(\frac{\left(y_i - (\boldsymbol{x}_i'\boldsymbol{\beta} + \boldsymbol{\theta}v_i)\right)^2}{2p^2v_i} + v_i\right) + \sum_{j=0}^k \frac{s_j}{2\lambda_j^2}\right), \quad (13)$$

$$f(s_j|\sigma, \boldsymbol{v}, \boldsymbol{\lambda}^2, \boldsymbol{y}, \boldsymbol{\beta}, \boldsymbol{s}_{-j}) \sim GIG\left(\frac{1}{2}, \beta_j, \frac{\sigma}{\lambda_j^2}\right),$$

$$f(\lambda_j^2|\sigma, \boldsymbol{v}, \boldsymbol{s}, \boldsymbol{y}, \boldsymbol{\beta}, \boldsymbol{\lambda}_{-j}^2) \sim IG\left(1 + \alpha, \delta + \frac{\sigma s_j}{2}\right).$$

3. MAIN RESULTS

The parameter estimation process was carried out by determining the *mean* and the variance of each parameter formulated in the posterior distribution obtained from both methods. Furthermore, these results were applied to data on hospital length of stay of COVID-19 patients in West Sumatra Province to formulate a regression model using the R software. COVID-19 patients' length of stay model was estimated using the Bayesian LASSO quantiles regression method (BLQRM) and Bayesian Adaptive LASSO quantile regression method (BALQRM). Table 3 presents the comparison of parameter estimation results, width confidence interval width of 95% for quantiles 0.10; 0.25; 0.50; 0.75, and 0.90.

Table 3. The parameter Estimated and Width of 95% Confidence Interval using BLQRM and BALQRM

Indicator Variables	BLQRM		BALQRM	
	Estimated Mean	Width 95% CI	Estimated Mean	Width 95% CI
$\tau = 0{,}10$				
Intercept	3.0480*	0.1564	3.0535*	**0.1235**
Age (X_1)	0.0008	0.0496	0.0009	**0.0489**
Man (X_{2D1})	0.0068	0.0621	0.0065	**0.0603**
PaUS (X_{3D1})	0.1858	**0.8703**	0.1838	0.8791
Positive (X_{3D2})	1.0326*	0.8950	1.0379*	**0.8946**
Cured (X_{4D1})	0.8171*	0.1437	0.8179*	**0.1422**
Die (X_{4D2})	-0.0141	0.1591	-0.0104	**0.1485**
Outpatient (X_{4D3})	-0.0153	0.2371	-0.0118	**0.2193**
Comorbidities (X_5)	-0.0005	0.0556	0.0001	**0.0552**
$\tau = 0{,}25$				
Intercept	3.6863*	0.1461	3.6872*	**0.1430**
Age (X_1)	-0.0048	0.0513	-0.0049	**0.0503**
Man (X_{2D1})	0.0036	0.0480	0.0038	**0.0478**
PaUS (X_{3D1})	0.0007	0.9828	0.0352	**0.9519**
Positive (X_{3D2})	1.7964*	1.0160	1.8334*	**0.9726**
Cured (X_{4D1})	0.5718*	0.4722	0.5812*	**0.4631**
Die (X_{4D2})	-0.2417*	0.4351	-0.2313*	**0.4231**
Outpatient (X_{4D3})	0.0489	**0.2881**	0.0483	0.2934
Comorbidities (X_5)	0.0165	**0.0694**	0.0162	0.0694
$\tau = 0{,}50$				
Intercept	5.3256*	0.2611	5.3322*	**0.2474**
Age (X_1)	-0.0877*	0.2270	-0.0843*	**0.2239**
Man (X_{2D1})	0.0288	0.1823	0.0280	**0.1742**
PaUS (X_{3D1})	-0.0693	1.7368	0.0140	**1.4243**
Positive (X_{3D2})	3.0768*	1.7648	3.1673*	**1.4652**
Cured (X_{4D1})	0.4080*	0.4806	0.4099*	**0.4385**
Die (X_{4D2})	-0.6908*	0.4578	-0.6882*	**0.4043**
Outpatient (X_{4D3})	0.0884	**0.3120**	0.0863	0.3152
Comorbidities (X_5)	0.1577*	0.2969	0.1517*	**0.2858**

$\tau = 0{,}75$				
Intercept	7.1231*	**0.3388**	7.1224*	0.3448
Age (X_1)	-0.3371*	0.3129	-0.3362*	**0.3083**
Man (X_{2D1})	-0.0598	0.3123	-0.0521	**0.2994**
PaUS (X_{3D1})	-0.6599	2.7494	-0.4024	**2.4624**
Positive (X_{3D2})	3.6690*	2.8012	3.9353*	**2.5463**
Cured (X_{4D1})	0.3883*	**0.5181**	0.3779*	0.5502
Die (X_{4D2})	-0.7418*	**0.5436**	-0.7424*	0.5489
Outpatient (X_{4D3})	0.0978	**0.3454**	0.1047	0.3493
Comorbidities (X_5)	0.6024*	0.3262	0.5979*	**0.3160**
$\tau = 0{,}90$				
Intercept	10.3410*	0.5242	10.3403*	**0.5066**
Age (X_1)	-0.8044*	**0.5181**	-0.7936*	0.5466
Man (X_{2D1})	-0.0908	0.4815	-0.0755	**0.4493**
PaUS (X_{3D1})	-1.1160	3.5413	-0.7760	**2.8633**
Positive (X_{3D2})	4.8764*	3.5749	5.2213*	**2.9534**
Cured (X_{4D1})	0.7436*	**0.8348**	0.7153*	0.8474
Die (X_{4D2})	-0.7030*	**0.8583**	-0.7129*	0.8669
Outpatient (X_{4D3})	0.0985	0.4763	0.1000	**0.4363**
Comorbidities (X_5)	0.8218*	**0.5837**	0.8101*	0.5919

***** Significant at $\alpha = 0{,}05$

Variables that are statistically significant in influencing the length of stay of COVID-19 patients for the two methods used are the age variable in the quantile 0.50; 0.75; 0.90, positive diagnosis variable in each quantile, death variable in the quantile other than 0.10, variable number of comorbidities in quantiles 0.50; 0.75; 0.90. Overall, it can be concluded that the length of stay model for COVID-19 patients using the Bayesian Adaptive LASSO quantile regression method produces an estimated 95% confidence interval width which tends to be smaller than the estimated 95% confidence interval width resulting from the Bayesian LASSO quantile regression method. Next, the comparison of the error values from the application of the two methods is presented in Table 4 below.

Table 4. MAD, MSE, and RMSE for each Selected Quantile Model.

Quantile	Model	MAD	MSE	RMSE
$\tau = 0{,}10$	BLQRM	3.339820	10.05405	3.17082
	BALQRM	**3.335335**	**10.01959**	**3.16537**
$\tau = 0{,}25$	BLQRM	2.924132	6.411903	2.53217
	BALQRM	**2.923741**	**6.407665**	**2.53134**
$\tau = 0{,}50$	BLQRM	2.539746	0.796494	0.89246
	BALQRM	**2.539451**	**0.784802**	**0.88589**
$\tau = 0{,}75$	BLQRM	3.014133	0.819771	0.90541
	BALQRM	**3.013949**	**0.818452**	**0.90468**
$\tau = 0{,}90$	BLQRM	5.199166	17.00303	4.12347
	BALQRM	**5.198545**	**16.99684**	**4.12272**

The Bayesian Adaptive LASSO quantile regression method generally produces the smallest estimated values of MAD, MSE, and RMSE than those produced by the Bayesian LASSO quantile regression method. Based on Table 3 and Table 4, the best model obtained for the length of stay of COVID-19 patients in West Sumatra Province is a model using the Bayesian Adaptive LASSO quantile regression method at quantile 0.50 which is formulated below:

$$\hat{y} = 5.3322 - 0.0843X_1 + 0.0280X_{2D1} + 0.0140X_{3D1} + 3.1673X_{3D2} + 0.4099X_{4D1} - 0.6882X_{4D2} + 0.0863X_{4D3} + 0.1517X_5 \quad (16)$$

Equation (16) can be interpreted as following: if there is an increase in age for 1 year, then the length of stay of COVID-19 patients below a quantile of 0.50 decreases by 0.0843 days or for 2 hours, assuming other variables are constant. The length of stay of a COVID-19 patient diagnosed positive under quantile 0.50 is 3.1673 days or about 3 days 4 hours longer than the length of stay of a COVID-19 patient diagnosed with PerUS under quantile 0.50 assuming other variables are constant. The length of stay of a COVID-19 patient who is declared cured under the 0.50 quantile is 0.4099 days or about 9 hours 50 minutes longer than the length of stay of the patients who were forced to go home under the 0.50 quantile, assuming the variable else constant. The length of stay of a COVID-19 patient who died under a quantile of 0.50 was 0.6882 days or around 16 hours 31 minutes shorter than the length of stay of the patients who were forced to go home under a quantile

of 0.50, assuming other variables constant. If there is an increase in the number of comorbid by 1, then the length of stay of COVID-19 patients below 0.50 quantile increases by 0.1517 days or for 3 hours 38 minutes, assuming other variables are constant. Furthermore, it is necessary to perform a convergence test for each model parameter resulting from applying the Bayesian Adaptive LASSO quantile regression method. The convergence test is identified by looking at the results of the trace plot, density plot, and ACF (Autocorrelation) plot for each model parameter. The results of convergence test are presented in Figures 3 and 4 below.

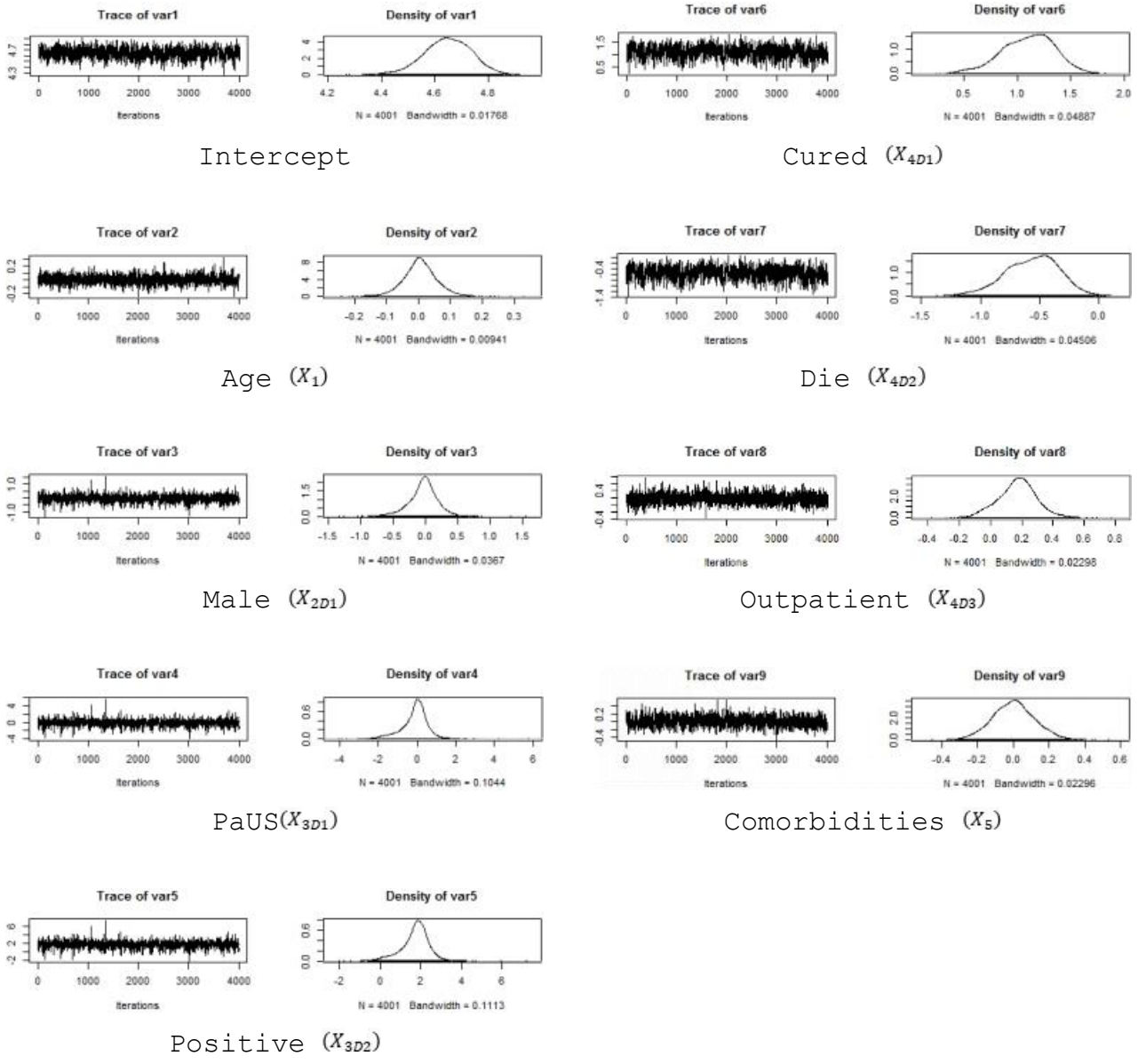
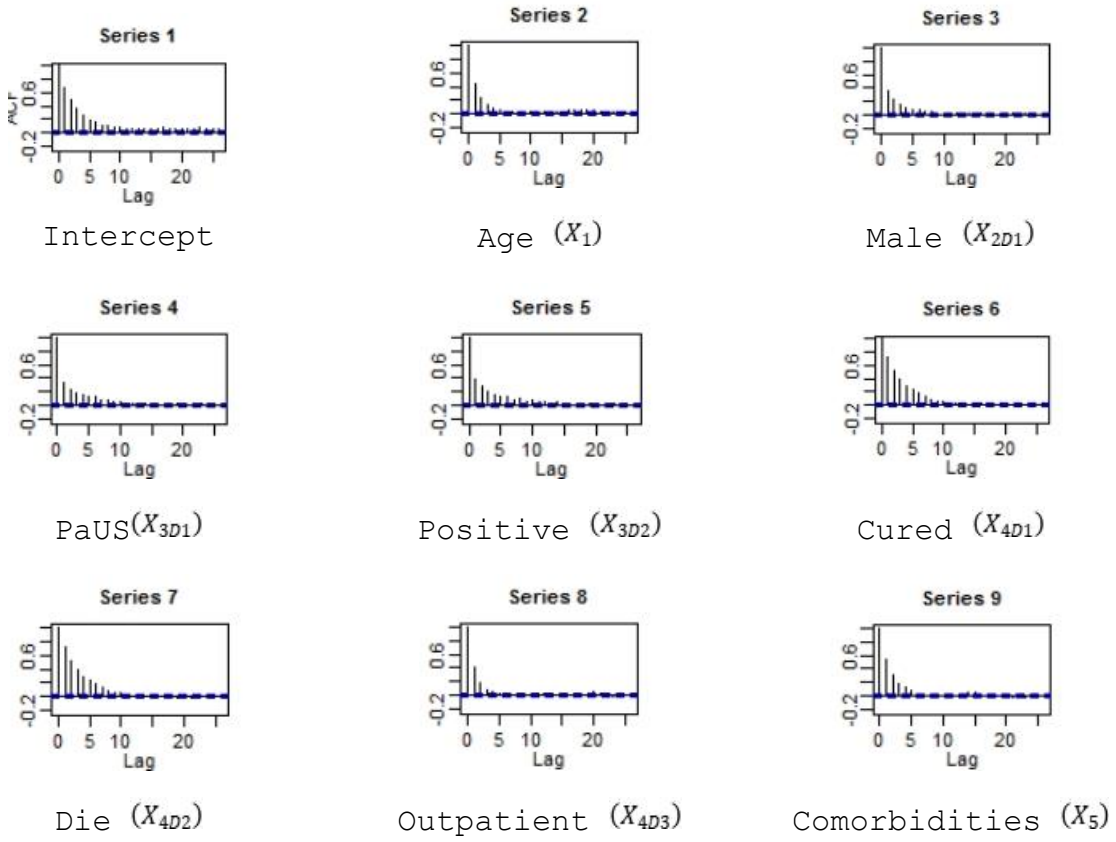


Figure 3. Trace Plot and Density Plot Model Parameters at Quantile 0.50

In Figure 3 the resulting trace plot has formed two horizontal linear lines for each parameter, thus it is said that the estimated values of the model parameters have converged to a certain value. Figure 3 also shows that the density plot produced for each parameter estimate already resembles a normal distribution curve, that is, it is symmetrical, meaning that the estimated model parameter values are normally distributed. As for Figure 4, the ACF plot generated for each parameter shows an autocorrelation value that slowly goes to zero with increasing lag. It means that the estimated value is generated towards stability and then reaches convergence so that it is concluded that the resulting estimated value is acceptable.



Figures 4. The plot of ACF Model Parameters at Quantile 0.50.

4. CONCLUSIONS

In this study, it was proven that the Bayesian Adaptive LASSO quantile regression method is better for modeling the length of stay of COVID-19 patients. This method can produce estimates of the width of the 95% confidence interval, MAD, MSE, and RMSE values that are smaller than the

Bayesian LASSO quantile regression method. The results of implementing the LASSO Adaptive Bayesian quantile regression method are that the length of stay of COVID-19 patients in West Sumatra Province is influenced by age, diagnoses related to COVID-19 (with PDP and positive categories), number of comorbidities, and patient discharge status (with cured categories, died, and outpatient). Thus, it can be concluded that to reduce the length of stay of COVID-19 patients in West Sumatra, individuals who are elderly and have comorbidities should be careful of their surroundings so as not to be infected with the Corona virus. The duration of hospital stay for individuals with these characteristics will take longer (if the person concerned is diagnosed with the COVID-19 virus) than for individuals with other conditions.

ACKNOWLEDGMENTS

DRPM, the Deputy for Strengthening Research and Development of the Ministry of Research and Technology, Indonesia, funded this research, with Contract Number 086/E5/PG.02.00.PT/2022 and by Andalas University, Indonesia, under Contract Number T/72/UN.16.17/PT.01.03/IS-RPT/2022.

CONFLICT OF INTERESTS

The author(s) declare that there is no conflict of interests.

REFERENCES

[1] H. Iswahyudi, The persistent effects of covid-19 on the economy and fiscal capacity of Indonesia, Jurnal Ekonomi dan Pembangunan 29 (2021), 113–130. https://doi.org/10.14203/jep.29.2.2021.113-130.

[2] UNICEF, Impact of COVID-19 on health in Indonesia, Report, (2022). [Online]. Available: https://www.unicef.org/indonesia/reports/impact-covid-19-maternal-newborn-child-and-adolescent-health-indonesia

[3] Pedoman Pencegahan dan Pengendalian Coronavirus Disease (COVID-19) Revisi ke-5, (2020). [Online]. Available: https://covid19.go.id/p/protokol/pedoman-pencegahan-dan-pengendalian-coronavirus-disease-covid-19-revisi-ke-5.

[4] A.K. Alahmari, Z.S. Almalki, A.A. Albassam, et al. Factors associated with length of hospital stay among COVID-19 patients in Saudi Arabia: A retrospective study during the first pandemic wave, Healthcare. 10 (2022), 1201. https://doi.org/10.3390/healthcare10071201.

[5] S. Çetin, A. Ulgen, H. Sivgin, et al. A study on factors impacting length of hospital stay of COVID-19 inpatients, J. Contemp. Med. 11 (2021), 396–404. https://doi.org/10.16899/jcm.911185.

[6] G. Thiruvengadam, M. Lakshmi, R. Ramanujam, A Study of factors affecting the length of hospital stay of COVID-19 patients by cox-proportional hazard model in a south Indian tertiary care hospital, J. Primary Care Commun. Health. 12 (2021), 215013272110002. https://doi.org/10.1177/21501327211000231.

[7] COVID-19 Distribution Map. https://covid19.go.id/peta-sebaran-covid19.

[8] C. Muharisa, F. Yanuar, D. Devianto, Simulation study the using of bayesian quantile regression in nonnormal error, CAUCHY. 5 (2018), 121–126. https://doi.org/10.18860/ca.v5i3.5633.

[9] F. Yanuar, The simulation study to test the performance of quantile regression method with heteroscedastic error variance, CAUCHY. 5 (2017), 36–41. https://doi.org/10.18860/ca.v5i1.4209.

[10] K. Yu, R.A. Moyeed, Bayesian quantile regression, Stat. Probab. Lett. 54 (2001), 437–447. https://doi.org/10.1016/S0167-7152(01)00124-9.

[11] R. Alhamzawi, Model selection in quantile regression models, J. Appl. Stat. 42 (2015), 445–458. https://doi.org/10.1080/02664763.2014.959905.

[12] R. Alhamzawi, K. Yu, Variable selection in quantile regression via Gibbs sampling, J. Appl. Stat. 39 (2012), 799–813. https://doi.org/10.1080/02664763.2011.620082.

[13] Q. Tang, H. Zhang, S. Gong, Bayesian regularized quantile regression analysis based on asymmetric Laplace distribution, J. Appl. Math. Phys. 08 (2020), 70–84. https://doi.org/10.4236/jamp.2020.81006.

[14] F. Yanuar, A. Zetra, Length-of-stay of hospitalized COVID-19 patients using bootstrap quantile regression, IAENG Int. J. Appl. Math. 51 (2021), 799-810.

[15] R. Koenker, G. Bassett, Regression quantiles, Econometrica. 46 (1978), 33. https://doi.org/10.2307/1913643.

[16] R. Tibshirani, Regression shrinkage and selection via the lasso, J. R. Stat. Soc.: Ser. B (Methodol.) 58 (1996), 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x.

[17] Q. Li, N. Lin, R. Xi, Bayesian regularized quantile regression, Bayesian Anal. 5 (2010), 533–556. https://doi.org/10.1214/10-BA521.

[18] H. Kozumi, G. Kobayashi, Gibbs sampling methods for Bayesian quantile regression, J. Stat. Comput. Simul. 81 (2011), 1565–1578. https://doi.org/10.1080/00949655.2010.496117.

[19] F.H.H. Alhusseini, New Bayesian lasso composite quantile regression, Amer. Sci. Res. J. Eng. Technol. Sci. 29 (2017), 138-152.

[20] F. Yanuar, H. Yozza, A. Zetra, Modified quantile regression for modeling the low birth weight, Front. Appl. Math. Stat. 8 (2022), 890028. https://doi.org/10.3389/fams.2022.890028.

[21] F. Yanuar, H. Yozza, F. Firdawati, et al. Applying bootstrap quantile regression for the construction of a low birth weight model, Makara J. Health Res. 23 (2019), 5. https://doi.org/10.7454/msk.v23i2.9886.

[22] R. Alhamzawi, H. Mallick, Bayesian reciprocal LASSO quantile regression, Commun. Stat. – Simul. Comput. 51 (2022), 6479–6494. https://doi.org/10.1080/03610918.2020.1804585.

[23] Z.Y. Algamal, R. Alhamzawi, H.T. Mohammad Ali, Gene selection for microarray gene expression classification using Bayesian Lasso quantile regression, Computers Biol. Med. 97 (2018) 145–152. https://doi.org/10.1016/j.compbiomed.2018.04.018.

[24] F. Yanuar, A.S. Deva, M. Maiyastri, et al. Modeling length of hospital stay for patients with COVID-19 in West Sumatra using quantile regression approach, CAUCHY. 7 (2021), 118–128. https://doi.org/10.18860/ca.v7i1.12995.

[25] S.H. Adams, M.J. Park, J.P. Schaub, et al. Medical vulnerability of young adults to severe COVID-19 illness—data from the national health interview survey, J. Adolescent Health. 67 (2020), 362–368. https://doi.org/10.1016/j.jadohealth.2020.06.025.

[26] B.G. Kaufman, C.A. Sueta, C. Chen, et al. Are trends in hospitalization prior to hospice use associated with hospice episode characteristics? Amer J. Hosp. Palliat. Care. 34 (2017), 860–868. https://doi.org/10.1177/1049909116659049.

[27] K. Yuki, M. Fujiogi, S. Koutsogiannaki, COVID-19 pathophysiology: A review, Clinic. Immunol. 215 (2020), 108427. https://doi.org/10.1016/j.clim.2020.108427.

[28] C. Gebhard, V. Regitz-Zagrosek, H.K. Neuhauser, et al. Impact of sex and gender on COVID-19 outcomes in Europe, Biol. Sex Differ. 11 (2020), 29. https://doi.org/10.1186/s13293-020-00304-9.

[29] P.Q. Thai, D.T.T. Toan, D.T. Son, et al. Factors associated with the duration of hospitalisation among COVID-19 patients in Vietnam: A survival analysis, Epidemiol. Infect. 148 (2020), e114. https://doi.org/10.1017/S0950268820001259.

[30] R. Zaboli, S.M. Mousavi, M. Bahadori, et al. Influential factors on the hospitalization length of COVID-19 patients: A systematic review, Trauma Mon. 27 (2022), 82-99. https://doi.org/10.30491/tm.2022.312499.1395.

[31] F.J. Chen, F.R. Li, J.Z. Zheng, et al. Factors associated with duration of hospital stay and complications in patients with COVID-19, J. Public Health Emerg. 5 (2021), 6. https://doi.org/10.21037/jphe-20-74.

[32] Y. Du, L. Tu, P. Zhu, et al. Clinical features of 85 fatal cases of COVID-19 from Wuhan. A retrospective observational study, Amer. J. Respir. Crit. Care Med. 201 (2020), 1372–1379. https://doi.org/10.1164/rccm.202003-0543OC.

[33] A.A. Alharbi, K.I. Alqumaizi, I. Bin Hussain, et al. Hospital length of stay and related factors for COVID-19 inpatients among the four southern regions under the proposed southern business unit of Saudi Arabia, J. Multidiscip. Healthcare. 15 (2022), 825–836. https://doi.org/10.2147/JMDH.S362625.

[34] A. Birhanu, B.T. Merga, G.M. Ayana, et al. Factors associated with prolonged length of hospital stay among COVID-19 cases admitted to the largest treatment center in Eastern Ethiopia, SAGE Open Med. 10 (2022), 205031212110703. https://doi.org/10.1177/20503121211070366.

[35] A. Guo, J. Lu, H. Tan, et al. Risk factors on admission associated with hospital length of stay in patients with COVID-19: a retrospective cohort study, Sci. Rep. 11 (2021), 7310. https://doi.org/10.1038/s41598-021-86853-4.