Determinants of Stunting Among Under-Five Years Children Using the Ordinal Logistic Regression Model

Marisa Rifada¹,³*, Nur Chamidah¹,³, Putri Nuraini¹, Faris Dwiki Gunawan¹, Lailatul Muniroh²,³

¹ Department of Mathematics, Universitas Airlangga, Surabaya, Indonesia
² Department of Health Nutrition, Universitas Airlangga, Surabaya, Indonesia
³ Research Group of Statistical Modeling in Life Science, Faculty of Science and Technology, Universitas Airlangga, Surabaya, Indonesia
*Corresponding author: marisa.rifada@fst.unair.ac.id

ABSTRACT

Stunting is a kind of chronic nutritional problem among under-five years children that characterized by a shorter height compared to other children of their age. The ordinal logistic regression is one statistical method that used to analyze the ordinal response variable with three or more categories and predictor variables that are categorical or continuous. This method has been used in various fields, one of them is in the health field. In this study, ordinal logistic regression is used to describe the correlation between stunting status of children and birth weight, birth length, mother's age, mother's height, mother's knowledge about nutrition, parenting, health services, hygiene and sanitation, number of children and number of family members. The stunting status of children in this study is divided into three kind which are severely stunted, moderately stunted, and normal stunted. The aim of this study is to describe the characteristics of children and to determine the factors that influence the stunting status of children in priority villages for stunting prevention i.e. Bucorwetan and Alaspadan village in the Probolinggo district, East Java, Indonesia. The prevalence of stunting was 30% (22.5% moderately, 7.5% severely). The mother's height and health services were the important determinants of stunting.

Keywords: Children, Ordinal Logistic Regression, Stunting.

1. INTRODUCTION

Phenomenon of stunting is a major nutritional problem that occurs in Indonesia. Stunting will be one of the strategic issues that are priorities for health development in Indonesia for the next five years (2020-2024). Stunting is a kind of chronic nutritional problem in toddlers which is characterized by a shorter height than other children of their age.

In 2017, 22.2%, or around 150.8 million toddlers in the world were stunted. According to stunting prevalence data that collected by the World Health Organization (WHO), among the third countries, Indonesia is the highest prevalence in the South-East Asia Regional (SEAR). The average prevalence of stunting among under-five years children in Indonesia in 2005-2017 is 36.4% [1]. The results of Basic Health Research in 2018 showed that as many as 30.8% or around 7 million toddlers were stunted. This number is still far from the minimum stunting rate set by WHO, which is 20%. In East Java Province, the prevalence of stunting among under-five years children is 32.81%. One of the areas in East Java that has a fairly high prevalence rate of stunting is Probolinggo Regency, which is 39.9% [2].

It is important to reduce the incidence of stunting in toddlers as early as possible to avoid long-term adverse effects such as stunted child development. Stunting can affect brain development so that the child's intelligence level is not optimal. This is at risk of reducing productivity as an adult. Stunting also makes children more susceptible to disease. Children who are stunted have a higher risk of suffering from chronic diseases in adulthood [1].

Several studies examining the factors that are at risk of influencing the incidence of stunting in toddlers have been carried out. Ni'mah and Nadhiroh...
[3] using the Chi-Square and Fisher Exact tests and concluded that there is a correlation between the birth length, a history of exclusive breastfeeding, family income, mother’s education, and mother’s knowledge about nutrition against the incidence of stunting among toddlers. Another factor related to the incidence of stunting is birth weight. Research in Nepal shows that babies with low birth weight have a higher risk of becoming stunted [4]. The birth length of the baby is also associated with the incidence of stunting. Research in Kendall shows that babies with short birth lengths have a high risk of stunting in toddlers [5]. This study aims to develop a risk model for the incidence of stunting among under-five years children using an ordinal logistic regression approach. If the form of the model is known, it is expected that the incidence of stunting among under-five years children can be anticipated from the start and preventive steps can be taken against the factors that have a significant effect.

2. LITERATURE REVIEW

2.1. Multicollinearity Detection

Multicollinearity is one of the problems in regression that occurs because of the high correlation between the predictor variables. To detect cases of multicollinearity, it can be done by calculating the VIF (Variance Inflation Factor) value which is defined as the following equation:

\[ VIF_j = \frac{1}{1-R_j^2} \]

where \( R_j^2 \) is the coefficient of determination between one predictor variable \( X_j \) and another predictor variables. The predictor variable is said to have multicollinearity if the VIF value > 10.

2.2. Ordinal Logistic Regression

The ordinal logistic regression is one of statistical method that used for analyzing response variables which have an ordinal scales, it should consist of three or more categories. Predictor variables that can be included in the model are categorical or continuous data that consist of two or more variables [6]. Models that can be used for ordinal logistic regression are cumulative logit models. For example, \( Y_i \) is an ordinal scale response variable that has \( q \) categories and \( x_i = [x_{i1}, x_{i2}, \ldots, x_{ip}]^T \) is a vector of the predictor variable in the \( i \)-th observation where \( i = 1, 2, \ldots, n \), then the cumulative logit model can be defined as follows:

\[
\text{logit}[P(Y_i \leq j \mid x_i) = \theta_j + x_i^T \beta]
\]

where \( j = 1, 2, \ldots, q-1 \). \( P(Y_i \leq j \mid x_i) \) is the cumulative probability of the \( j \)-th category response variable with terms \( x_i \), \( \theta_j \) is the intercept parameter and \( \beta = (\beta_1, \beta_2, \ldots, \beta_p)^T \) as usual denotes the vector of the regression coefficients. The logit link function can be defined as follows:

\[
\text{logit}[P(Y_i \leq j \mid x_i)] = \ln \left( \frac{P(Y_i \leq j \mid x_i)}{1-P(Y_i \leq j \mid x_i)} \right)
\]

based on Equation (2) and Equation (3), the ordinal logistic regression model can be stated as follows:

\[
\ln \left( \frac{P(Y_i \leq j \mid x_i)}{1-P(Y_i \leq j \mid x_i)} \right) = \theta_j + x_i^T \beta
\]

Therefore, from Equation (4) we can get

\[
P(Y_i \leq j \mid x_i) = \frac{\exp(\theta_j + x_i^T \beta)}{1 + \exp(\theta_j + x_i^T \beta)}
\]

If there are three response categories, then the cumulative probability of the \( j \)-th category response variable is as follows:

\[
P(Y_i \leq 1 \mid x_i) = \frac{\exp(\theta_1 + x_i^T \beta)}{1 + \exp(\theta_1 + x_i^T \beta)}
\]

\[
P(Y_i \leq 2 \mid x_i) = \frac{\exp(\theta_2 + x_i^T \beta)}{1 + \exp(\theta_2 + x_i^T \beta)}
\]

Suppose \( \pi_j(x_i) = P(Y_i = j \mid x_i) \); \( j = 1, 2, \ldots, q \) is the probability that the response variable in the \( i \)-th observation has \( j \)-th category, then \( P(Y_i \leq j \mid x_i) = \pi_1(x_i) + \pi_2(x_i) + \ldots + \pi_j(x_i) = 1 \), so that

\[
\pi_1(x_i) = \frac{\exp(\theta_1 + x_i^T \beta)}{1 + \exp(\theta_1 + x_i^T \beta)}
\]

\[
\pi_2(x_i) = \frac{\exp(\theta_2 + x_i^T \beta)}{1 + \exp(\theta_2 + x_i^T \beta)} - \frac{\exp(\theta_1 + x_i^T \beta)}{1 + \exp(\theta_1 + x_i^T \beta)}
\]

\[
\pi_3(x_i) = 1 - \frac{\exp(\theta_3 + x_i^T \beta)}{1 + \exp(\theta_3 + x_i^T \beta)}
\]

The probability value for each response category is used as a guide for classification. An observation will be classified as the response variable of the \( j \)-th category based on the greatest probability value.
2.3. Estimating Parameters

Estimating parameters of ordinal logistic regression models can be obtained using the Maximum Likelihood Estimation (MLE) method. The first step in the MLE method is to make the likelihood function of the sample with \( n \) independent observations \((Y_i, x_i)\) as follows:

\[
L(\varphi) = \prod_{i=1}^{n} \pi_{j}^{y_{ij}} \pi_{2}^{y_{2j}} \ldots \pi_{q}^{y_{qj}} (x_i)_{\varphi}^{y_{mj}}
\]

where \( n_{ji} \) will be 1 if \( Y_j = j \) and will be 0 if \( Y_j \neq j \), and \( \varphi = (\beta_1, \beta_2, \ldots, \beta_{q-1}, \beta_2, \ldots, \beta_p)^T \) is a parameter which is estimated by maximizing the likelihood function and \( \gamma_j(x_i) = P(Y_i \leq j | x_i) \). Next, the In-transform function is performed as follows:

\[
\ell = \ln (L(\varphi)) = \sum_{i=1}^{n} \sum_{j=1}^{q} \left[ n_{ji} \ln \left( \frac{\exp(\beta_j + x_i^T \beta)}{1 + \exp(\beta_j + x_i^T \beta)} \right) - \frac{\exp(\beta_{j-1} + x_i^T \beta)}{1 + \exp(\beta_{j-1} + x_i^T \beta)} \right]
\]

From the results of the derivation in Equation (11), a nonlinear function is obtained so that the Newton Raphson iteration method is used in solving as follows:

\[
\varphi^{(i+1)} = \varphi^{(i)} - (H^{(i)})^{-1} r^{(i)}
\]

where \( r^{(i)} = \frac{d\ell}{d\varphi^{(i)}}; i = 0, 1, \ldots \) and Hessian matrix \( H^{(i)} = \frac{d^2\ell}{d\varphi d\varphi} \). The parameter vector estimate will be obtained when the process has converged, expressed as \( \max |\varphi^{(i+1)} - \varphi^{(i)}| < \delta \), with \( \delta \) is a small positive number.

2.4. Simultaneous Test

Simultaneous testing is carried out to determine the effect of predictor variables simultaneously on the response variable with the following hypothesis:

\( H_0: \beta_1 = \beta_2 = \ldots = \beta_k; k = 1, 2, \ldots, p \)

\( H_1: \) there is at least one \( \beta_k \neq 0; k = 1, 2, \ldots, p \)

The test statistic used to test the hypothesis is the likelihood ratio test which is defined as follows [7]:

\[
G^2 = -2 \ln \left[ \prod_{i=1}^{n} \pi_{1}(x_i)_{\varphi}^{n_{1i}} \pi_{2}(x_i)_{\varphi}^{n_{2i}} \ldots \pi_{q}(x_i)_{\varphi}^{n_{qi}} \right]
\]

where \( n_1 = \sum_{i=1}^{n} n_{1i} \), \( n_2 = \sum_{i=1}^{n} n_{2i} \), \( n_q = \sum_{i=1}^{n} n_{qi} \) and \( n = n_1 + n_2 + \ldots + n_q \). The area of rejection for \( H_0 \) is \( G^2 > \chi^2(\alpha, p) \) or p-value less than \( \alpha \).

2.5. Partial Test

This partial test is used to determine the effect of each predictor variable on the response variable with the following hypothesis:

\( H_0: \beta_k = 0; k = 1, 2, \ldots, p \)

\( H_1: \beta_k \neq 0; k = 1, 2, \ldots, p \)

The test statistic used to test this hypothesis is the Wald test statistic which is defined as follows [7]:

\[
W_k = \frac{\beta_k}{s(\beta_k)}; k = 1, 2, \ldots, p
\]

The area of rejection for \( H_0 \) is \( |W_k| > Z_{(\alpha/2)} \) or p-value less than \( \alpha \).

2.6. Goodness-of-Fit Test

The goodness-of-fit test is used to determine whether the model is appropriate or not, this test is done by comparing the actual model with the fit model with the following hypothesis:

\( H_0: \) Ordinal logistic regression models are appropriate

\( H_1: \) Ordinal logistic regression model are not appropriate

The test statistic used is the deviance test which is defined as follows:

\[
D = 2 \sum_{i=1}^{n} \left[ n_{qi} \ln \left( \frac{\pi_{1}(x_i)}{\pi_{1}(x_i)} \right) + n_{2i} \ln \left( \frac{\pi_{2}(x_i)}{\pi_{2}(x_i)} \right) + \ldots + n_{qi} \ln \left( \frac{\pi_{q}(x_i)}{\pi_{q}(x_i)} \right) \right]
\]

where \( n_{qi} \) is the number of observations in the ith category of the ordinal response variable and the jth predictor variable.
The area of rejection for $H_0$ if $D > \chi^2_{\alpha,(q-1)(J-p-1)}$, where $J$ is the number of level combinations of different predictor variables [7].

2.6. Classification Accuracy

Apparent Error Rate (APER) is a value used to see the probability of error in classifying objects. Suppose the response variable $Y$ has $q$ categories, then the APER value is

$$\text{APER} = \frac{\sum_{i,j} n_{ij}}{\sum_{i,j} n_{ij}} \times 100\%$$ (16)

where $n_{ij}$ is the number of $Y$ events in category $i$ from the observation results classified into $Y$ events in category $j$ from the prediction results [8]. Based on the APER value in Equation (16), classification accuracy is obtained $100\% - \text{APER}$.

3. RESEARCH METHODOLOGY

3.1. Data Source

The data used in this study are sample data of toddlers recorded in the One Posyandu Application and survey sampling in priority villages for stunting prevention i.e. Bucorwetan and Alaspandan village in Probolinggo district, East Java, Indonesia. The sample used was 80 toddlers.

3.2. Research Variable

The response variable $(Y)$ used in this study is the nutritional status of toddlers based on the height/age index which is categorized into 3, i.e. severely stunted $(Y = 1)$, moderately stunted $(Y = 2)$ and normal $(Y = 3)$. Meanwhile, the predictor variable $(X)$ consists of 11 continuous variables, namely birth weight $(X_1)$, birth length $(X_2)$, parity $(X_3)$, mother's age $(X_4)$, mother's height $(X_5)$, mother's knowledge of nutrition $(X_6)$, parenting $(X_7)$, health services $(X_8)$, hygiene and sanitation $(X_9)$, number of children $(X_{10})$ and number of family members $(X_{11})$.

3.3. Methods

The steps of data analysis in this research are as follows:

- Detection of multicollinearity
- Ordinal logistic regression analysis
  - Estimated logistic regression model building
- Test the significance of parameters simultaneously and partially
- Ordinal logistic regression model building
- Interpretation of factors affecting the nutritional status of toddlers based on Height/Age index
- Goodness-of-Fit Test using the deviance test statistic
- Calculating the classification accuracy value of the model formed

4. RESULTS AND DISCUSSION

4.1. Characteristics of Research Variables

Analysis of the characteristics of the variables in this study describes the condition of toddlers in the sample. The first characteristic analysis was carried out on the response variable, namely the nutritional status of toddlers based on height/age index. The results of the analysis can be shown in the following Figure 1:

![Figure 1. Percentage of nutritional status of toddlers based on height/age index](image)

Based on Figure 1, it can be seen that the prevalence of stunting cases in the sample is 30% with moderately stunted is 22.5% and severely stunted is 7.5%. Furthermore, the characteristic analysis of the predictor variables is as follows:

- Descriptive analysis to determine the characteristics of each variable
Table 1. Descriptive statistics on predictor variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Median</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>X₁</td>
<td>2.978</td>
<td>3</td>
<td>1.820</td>
<td>4.8</td>
<td>0.468</td>
</tr>
<tr>
<td>X₂</td>
<td>49.7</td>
<td>49</td>
<td>37</td>
<td>57</td>
<td>2.523</td>
</tr>
<tr>
<td>X₃</td>
<td>1.789</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>0.741</td>
</tr>
<tr>
<td>X₄</td>
<td>29.913</td>
<td>30</td>
<td>20</td>
<td>43</td>
<td>5.408</td>
</tr>
<tr>
<td>X₅</td>
<td>149.959</td>
<td>150</td>
<td>139</td>
<td>165</td>
<td>4.846</td>
</tr>
<tr>
<td>X₆</td>
<td>63.25</td>
<td>70</td>
<td>30</td>
<td>100</td>
<td>15.326</td>
</tr>
<tr>
<td>X₇</td>
<td>74.5</td>
<td>80</td>
<td>20</td>
<td>100</td>
<td>21.103</td>
</tr>
<tr>
<td>X₈</td>
<td>84.688</td>
<td>100</td>
<td>25</td>
<td>100</td>
<td>18.007</td>
</tr>
<tr>
<td>X₉</td>
<td>62</td>
<td>60</td>
<td>0</td>
<td>100</td>
<td>20.277</td>
</tr>
<tr>
<td>X₁₀</td>
<td>1.825</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>0.759</td>
</tr>
<tr>
<td>X₁₁</td>
<td>4.813</td>
<td>5</td>
<td>2</td>
<td>9</td>
<td>1.274</td>
</tr>
</tbody>
</table>

Based on Table 1, it can be seen that the variable $X₁$ (birth weight) has a mean of 2.978 kg but in the minimum value there is still a birth weight that is less than 2 kg which is 1.82 kg. This shows that there are still toddlers who are born prematurely. Furthermore, the variable $X₄$ (mother’s age) has a the median value of 30 years but in the minimum and maximum ages of 20 years and 43 years it is a period beyond the age of the mother to produce healthy babies. So there is a possibility that toddlers are born less healthy. In other variables, the variable $X₇$ (parenting) has a median value of 80 and a minimum value of 20. This proves that the implementation of good parenting is not optimal for toddlers. Meanwhile, the variable $X₉$ (hygiene and sanitation) has a median value of 60 but it has the minimum value of 0. This indicates that there is a lack of hygiene and sanitation in toddlers.

4.2. Multicollinearity Detection

Multicollinearity detection is performed to determine whether there is a linear relationship between predictor variables in the model. Multicollinearity detection correlates between predictor variables, where the estimation of multicollinearity cases occurs if there is a high correlation between the predictor variables. The existence of multicollinearity cases can also be seen based on the VIF value with a limit of 10. If the VIF value of a variable is more than 10 then multicollinearity occurs. The following table shows the multicollinearity detection of the variables in this study:

Table 2. Multicollinearity test of predictor variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Birth weight</td>
<td>2.516</td>
</tr>
<tr>
<td>Birth length</td>
<td>2.421</td>
</tr>
<tr>
<td>Parity</td>
<td>18.773</td>
</tr>
<tr>
<td>Mother’s age</td>
<td>1.869</td>
</tr>
<tr>
<td>Mother’s height</td>
<td>1.190</td>
</tr>
<tr>
<td>Mother's knowledge of nutrition</td>
<td>1.204</td>
</tr>
<tr>
<td>Parenting</td>
<td>1.183</td>
</tr>
<tr>
<td>Health services</td>
<td>1.109</td>
</tr>
<tr>
<td>Hygiene and sanitation</td>
<td>1.303</td>
</tr>
<tr>
<td>Number of children</td>
<td>17.241</td>
</tr>
<tr>
<td>Number of family members</td>
<td>1.120</td>
</tr>
</tbody>
</table>

Based on Table 2, it was found that the variable $X₃$ (parity) and $X₁₀$ (number of children) had a VIF value more than 10 so that they had multicollinearity cases and were excluded from the model.

4.3. Ordinal Logistic Regression

Ordinal logistic regression modeling was done by including all predictor variables that did not have cases of multicollinearity. To perform ordinal logistic regression testing, a simultaneous test was conducted first. The results of the simultaneous test for these variables can be seen in the following table:

Table 3. Concurrent Test with Likelihood Ratio

<table>
<thead>
<tr>
<th>Model</th>
<th>G²</th>
<th>Chi-Square</th>
<th>df</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept Only</td>
<td>124.730</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final</td>
<td>101.282</td>
<td>23.448</td>
<td>9</td>
<td>0.005</td>
</tr>
</tbody>
</table>
Based on Table 3, shows that the value of $G^2$ is 23.448, which means that $G^2 > \chi^2_{(0.05;9)} = 16.919$ or the significance value of 0.005 is less than the alpha value of 5% so that the decision taken is to reject $H_0$.

**Table 4. Partial Parameter Estimation**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>Wald</th>
<th>Sig.</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>([Y = 1])</td>
<td>33.038</td>
<td>12.069</td>
<td>7.494</td>
<td>0.006</td>
<td>9.384 - 56.693</td>
</tr>
<tr>
<td>([Y = 2])</td>
<td>35.198</td>
<td>12.167</td>
<td>8.369</td>
<td>0.004</td>
<td>11.351 - 59.044</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.131</td>
<td>0.977</td>
<td>0.018</td>
<td>0.893</td>
<td>-1.783 - 2.045</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.221</td>
<td>0.166</td>
<td>1.772</td>
<td>0.183</td>
<td>-0.104 - 0.545</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>-0.037</td>
<td>0.055</td>
<td>0.455</td>
<td>0.500</td>
<td>-0.144 - 0.070</td>
</tr>
<tr>
<td>$\beta_4$</td>
<td>0.174</td>
<td>0.062</td>
<td>7.788</td>
<td>0.005*</td>
<td>0.052 - 0.297</td>
</tr>
<tr>
<td>$\beta_5$</td>
<td>-0.033</td>
<td>0.020</td>
<td>2.614</td>
<td>0.106</td>
<td>-0.072 - 0.007</td>
</tr>
<tr>
<td>$\beta_6$</td>
<td>0.015</td>
<td>0.014</td>
<td>1.184</td>
<td>0.277</td>
<td>-0.012 - 0.043</td>
</tr>
<tr>
<td>$\beta_7$</td>
<td>0.032</td>
<td>0.015</td>
<td>4.505</td>
<td>0.034*</td>
<td>0.002 - 0.062</td>
</tr>
<tr>
<td>$\beta_8$</td>
<td>-0.018</td>
<td>0.015</td>
<td>1.407</td>
<td>0.236</td>
<td>-0.046 - 0.011</td>
</tr>
<tr>
<td>$\beta_9$</td>
<td>-0.149</td>
<td>0.217</td>
<td>0.473</td>
<td>0.492</td>
<td>-0.575 - 0.276</td>
</tr>
</tbody>
</table>

* significant at alpha = 0.05

Based on Table 4, it can be seen that two variables have a significant value less than the alpha value of 5%, i.e. 0.005 and 0.034. They are the variables $X_5$ (mother’s height) and $X_8$ (health services). The height of the parents is related to the physical growth of the child. Genetic factors in mothers, namely height affect the incidence of stunting in toddlers. Likewise regarding health services for toddlers will affect the fulfillment of toddler nutrition.

Thus an ordinal logistic regression model can be built based on variables of mother’s height and health services. The results can be seen in the following table:

**Table 5. Parameter Estimates for Logit Model**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>Wald</th>
<th>Sig.</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>([Y = 1])</td>
<td>25.797</td>
<td>8.613</td>
<td>8.971</td>
<td>0.003</td>
<td>8.916 - 42.678</td>
</tr>
<tr>
<td>([Y = 2])</td>
<td>27.722</td>
<td>8.697</td>
<td>10.160</td>
<td>0.001</td>
<td>10.675 - 44.768</td>
</tr>
<tr>
<td>$\beta_5$</td>
<td>0.174</td>
<td>0.057</td>
<td>9.357</td>
<td>0.002</td>
<td>0.062 - 0.285</td>
</tr>
<tr>
<td>$\beta_8$</td>
<td>0.032</td>
<td>0.014</td>
<td>5.204</td>
<td>0.023</td>
<td>0.005 - 0.060</td>
</tr>
</tbody>
</table>

* significant at alpha = 0.05

The logit model is structured based on significant parameters in Table 5 with the following results:

$$\logit P(Y_i \leq 1|x_i) = 25.797 + 0.174X_5 + 0.032X_8$$

$$\logit P(Y_i \leq 2|x_i) = 27.722 + 0.174X_5 + 0.032X_8$$

### 4.4. Model Fit Test

This test aims to determine the suitability of the regression model. The model is considered suitable if the p-value at Deviance is more than 5%. The results of the model suitability test can be seen in the following table:

**Table 6. Model Suitability Test**

<table>
<thead>
<tr>
<th>Chi-Square</th>
<th>df</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson</td>
<td>140.625</td>
<td>82</td>
</tr>
<tr>
<td>Deviance</td>
<td>83.029</td>
<td>82</td>
</tr>
</tbody>
</table>

Based on Table 6, it can be seen that the p-value at Deviance of 0.447 indicates more than 5% so that a decision to accept $H_0$ can be taken and it can be concluded that the model obtained is appropriate.
4.5. Accuracy Of Model Classification

The accuracy of model classification is done by calculating the value of the classification accuracy between the actual value and the predicted value obtained from the model that has been formed.

Table 7. The accuracy of model classification

<table>
<thead>
<tr>
<th>Observation</th>
<th>Predicted Response Category</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Severely stunteds</td>
</tr>
<tr>
<td>Severe stunteds</td>
<td>0</td>
</tr>
<tr>
<td>Moderately stunteds</td>
<td>0</td>
</tr>
<tr>
<td>Normal</td>
<td>0</td>
</tr>
</tbody>
</table>

Based on Table 7, it is found that the model classification accuracy value is 70%.

5. CONCLUSION

Analysis of the characteristics of the variables in this study describes that the prevalence of stunting cases is 30% with moderately stunted is 22.5% and severely stunted is 7.5%, there are still toddlers who are born prematurely, mothers have period beyond the age of the mother to produce healthy babies and the implementation of good parenting is not optimal for toddlers. Based on the ordinal logistic regression model, two variables are significant determinants of stunting among under-five years children, i.e mother’s height and health services. The model classification accuracy value is 70%.

ACKNOWLEDGMENTS

The authors specially appreciate to Universitas Airlangga for funding this research with a contract number 824/UN3.14/PT/2020. We also thank to editors and peer reviewers who have given useful suggestions and criticisms for the improvement of this paper.

REFERENCES