Negative Binomial Regression Analysis of Factors Influencing Stunting Cases in Central Lombok Regency

Elma Yulia Putri Ananda[1], Suwardi Annas [2*], Hisyam Ihsan [1], Sukarna [1] and Aswi Aswi [2]

[1] Mathematics Department, Universitas Negeri Makassar, Makassar
[2] Statistics Department, Universitas Negeri Makassar, Makassar
*Corresponding author: suwardi_annas@unm.ac.id

Received: 6 August 2024
Revised: 2 September 2024
Accepted: 19 September

ABSTRACT – Poisson regression is commonly used to model count data, relying on the crucial assumption of equidispersion, where the mean and variance are equal. However, this assumption is often violated in real-world data, which can exhibit overdispersion or underdispersion. When this occurs, the standard Poisson model becomes unsuitable, leading to biased and inaccurate parameter estimates. To address overdispersion in count data, Negative Binomial Regression (NBR) is a viable alternative, as it incorporates an additional parameter to account for variability greater than the mean. Stunting, a condition characterized by significantly impaired growth in infants, has been a primary concern for the Indonesian government during the 2019-2024 period, particularly in Central Lombok district. Reducing stunting rates is critical to ensuring an optimal quality of life for future generations. Despite extensive research on stunting, the application of NBR to analyze factors influencing stunting cases in Central Lombok Regency has not yet been explored. This study aims to implement the NBR model to identify the determinants of stunting in Central Lombok. Data were collected from 29 community health centers (PUSKESMAS) in Central Lombok. The findings indicate that an increase in the number of malnourished toddlers is associated with a corresponding rise in stunting cases. Similarly, a higher prevalence of low-birth-weight infants is linked to an elevated incidence of stunting.

Keywords– Poisson Regression, Negative Binomial Regression, Overdispersion, Stunting, Central Lombok Regency

I. INTRODUCTION

Regression analysis is a statistical technique employed to examine the relationships between predictor variables and response variables [1], [2]. Typically, the data utilized in this analysis is continuous [3]. However, there are instances where discrete data is present, particularly in the response variable, where it takes the form of count data—non-negative values that represent the number of events occurring within a specified interval of time, space, or volume [4]. Such cases can be effectively addressed through Poisson regression analysis [5]. A key assumption of Poisson regression is that the variance of the data is equal to its mean, a condition known as equidispersion [6]. In practice, however, equidispersion is seldom observed; count data often exhibit either overdispersion (where variance exceeds the mean) or underdispersion (where variance is less than the mean) [2], [7].

Negative Binomial Regression (NBR) extends Poisson regression by incorporating a Gamma distribution for the mean, making it suitable for addressing overdispersion [7], [8]. Consequently, NBR offers greater flexibility than Poisson regression. In terms of parameter estimation, the NBR model employs the Maximum Likelihood Estimation (MLE) method [6]. Therefore, NBR is anticipated to provide a robust modeling approach for managing instances of overdispersion.

Stunting is a condition that impairs a child's physical growth and development due to chronic undernutrition, particularly from pregnancy through the age of two [9], [10]. Children who are stunted typically exhibit shorter stature compared to their peers and are often underweight for their age [11]. The long-term consequences of stunting include chronic health issues, cognitive impairments, and reduced productivity in adulthood. It serves as a critical indicator of nutritional deficiencies and limited access to adequate nutrition, especially during vital periods of growth and development [12].

According to United Nations (UN) statistics from 2020, over 149 million children under five globally are affected by stunting, representing 22% of this age group; among them, 6.3 million are stunted toddlers in Indonesia [13]. The United Nations Children's Fund (UNICEF) identifies several factors contributing to stunting, including malnutrition in children under two years old, inadequate maternal nutrition during pregnancy, and poor sanitation practices [9], [10]. Currently, the prevalence of stunting in Indonesia is 21.6%, with a target of reducing this figure to 14% by 2024 [14]. To achieve this goal, concerted efforts are essential, beginning at the family level, which is the smallest unit of society. Notably, in 2022, West Nusa Tenggara reported a particularly high stunting prevalence of 32.7%.

Several studies related to child and maternal health have been conducted, including analyses of factors affecting stunting using Geographically Weighted Regression (GWR) [14], investigations into overdispersion with NBR in maternal mortality cases in Bandung City [7], and applications of GWR models to infant mortality data [15]. Other notable studies include NBR modeling of infant mortality data in Jombang Regency [16], analyses of human mobility and COVID-19 through negative binomial regression [17], and the use of NBR for modeling infant mortality in East Kalimantan [18].

However, none of these studies have applied NBR to stunting cases or analyzed the factors influencing stunting. The objectives of this study were twofold: first, to assess the implementation of the NBR model in analyzing stunting cases in Central Lombok, and second, to identify the factors that influence stunting in this region. The urgency of this research lies in its potential to enhance the NBR model for analyzing factors related to stunting in Central Lombok Regency, West Nusa Tenggara. This approach addresses the limitations associated with the equidispersion assumption, thereby providing a more flexible and robust modeling framework for data exhibiting overdispersion.

II. LITERATURE REVIEW

A. Multicollinearity Test

Multicollinearity refers to a situation in regression modeling where independent variables are highly correlated, leading to violations of the regression assumptions. This can complicate the evaluation and interpretation of the model [19]. To assess multicollinearity, the Variance Inflation Factor (VIF) method can be employed. The formula for calculating the test statistic is as follows:

$$VIF_h = \frac{1}{1-R_h^2}, h = 1,2,\ldots,q \qquad (1)$$

VIF_h : Variance Inflation Factor for the h-th predictor variable. It quantifies the extent to which multicollinearity increases the variance of a regression coefficient.

R_h^2 : The coefficient of determination obtained when the h-th predictor variable is regressed against all the other predictor variables in the model. It represents the proportion of variance in the h-th predictor that can be explained by the other predictors.

h: An index representing each of the predictor variables in the regression model, where h =1, 2 ,..., q

q: The total number of predictor variables in the regression model

A regression model is considered free of multicollinearity if the VIF values for the predictor variables are below 10.

B. Poisson Regression

Count data and response variables that follow a Poisson distribution can be analyzed using Poisson regression [2]. This model is a type of generalized linear model that describes the relationship between the dependent variable and the independent variables. The Poisson regression model is defined as follows:

$$f(y,\mu) = \frac{e^{-\mu}\mu^y}{y!}; y = 0,1,2,3,4\ldots \qquad (2)$$

$f(y,\mu)$: The probability mass function of the Poisson distribution. It gives the probability of observing a count y given the mean rate μ

y: the count data, or response variable, represents the number of occurrences of an event

μ: the mean parameter of the Poisson distribution, representing the expected value of the count data and the rate at which events occur in a fixed interval of time or space

e: approximately equal to 2.71828, is the base of the natural logarithm used in the exponent to model the probability distribution

Under the assumptions of Poisson regression, the mean μ is equal to both the expected value and the variance of the response variable Y

$E(Y) = \mu$ and $Var(Y) = \mu$

C. Overdispersion

Equidispersion is an assumption in statistical modeling indicating that the mean and variance of the data are equal. Specifically, equidispersion implies that the variance of the response variable equals its mean: $Var(Y) = E(Y)$. Overdispersion occurs when the variance of a response variable exceeds its mean, expressed as $Var(Y) > E(Y)$. Conversely, underdispersion occurs when the variance of the response variable is less than its mean, expressed as $Var(Y) < E(Y)$. To detect overdispersion, we can examine the values of the deviance and Pearson chi-squared statistic relative to the degrees of freedom. If these values exceed 1, it suggests that overdispersion is present. The following formulas are used to compute these statistics [20]:

1) Deviance

The deviance can be calculated using the following formula:

$$\phi_1 = \frac{D^2}{db};\ D^2 = 2\sum_{i=1}^{n}\{y_i \ln\left(\frac{y_i}{\hat{\mu}_i}\right) - (y_i - \mu_i)\} \qquad (3)$$

The symbol ϕ_1 represents the scaled deviance statistic. It is calculated by dividing the deviance D^2 by the degrees of freedom db

D^2 is the deviance, which measures the difference between the fitted model and a saturated model (a model with perfect fit).

db: the degrees of freedom, defined as $n - k$, where n is the number of observations and k is the number of parameters in the model

$\hat{\mu}_i$: represents the fitted mean of the response variable for the i-th observation

n : denotes the number of observations.

k : the number of parameters, including the constant term

If ϕ_1 is greater than 1, it suggests that there is overdispersion in the model.

2) *Pearson chi-squared*

The Pearson chi-squared statistic can be calculated using the following formula:

$$\phi_2 = \frac{\chi^2}{db};\ \chi^2 = \sum_{i=1}^{n} \frac{(y_i - \mu_i)^2}{var(y_i)}$$

In this equation, db is defined as $n - k$ where k represents the number of parameters, including the constant term, and n denotes the number of observations. The value χ^2 indicates the Pearson chi-squared statistic.
ϕ_2: represents the normalized Pearson chi-squared statistic, which measures the discrepancy between observed and predicted variances. It indicates whether the observed variability exceeds expectations under the model assumptions.
χ^2: is the Pearson chi-squared statistic, which measures the deviation of observed values from expected values.

- If $\phi_2 > 1$, it suggests overdispersion, meaning that the variance in the data is greater than what the model assumes
- If $\phi_2 \approx 1$, it suggests that the model fits the data reasonably well under the assumption of equidispersion.
- If $\phi_2 < 1$, it suggests underdispersion, meaning the variance in the data is less than what the model assumes.

D. *Negative Binomial Regression (NBR)*

NBR is a statistical model used to analyze the relationship between predictor variables and a count-based response variable. It is particularly useful for addressing overdispersion issues that can arise in Poisson regression, which may lead to inaccurate conclusions [4], [16], [20], [21].

The Negative Binomial distribution is characterized by the following mean and variance:

Mean: $E(Y) = \mu$
Variance: $Var(Y) = \mu + \theta\mu^2$
The probability mass function of the Negative Binomial distribution is given by:

$$f(y; \mu, \theta) = \frac{\Gamma(y + \frac{1}{\theta})}{\Gamma(\frac{1}{\theta}) y!} \left(\frac{1}{1+\theta\mu}\right)^{\frac{1}{\theta}} \left(\frac{\theta\mu}{1+\theta\mu}\right)^{y}, y = 0, 1, 2, \dots \quad (4)$$

In the special case where $\theta = 0$, the Negative Binomial distribution reduces to the Poisson distribution, where the variance equals the mean:
$Var(Y) = \mu$
In this scenario, both the expected value and variance are identical:
$\boldsymbol{E(Y) = Var(Y) = \mu}$
The exponential distribution function of the negative binomial distribution can be expressed as:

$$f(y; \mu, \theta) = exp\left\{ y \ln\left(\frac{\theta\mu}{1+\theta\mu}\right) + \frac{1}{\theta} \ln\left(\frac{1}{1+\theta\mu}\right) + \ln\left(\frac{\Gamma(y + \frac{1}{\theta})}{\Gamma(\frac{1}{\theta}) y!}\right) \right. \quad (5)$$

$\boldsymbol{\theta}$: This dispersion parameter controls overdispersion in the Negative Binomial distribution. Larger values of $\boldsymbol{\theta}$ indicate greater overdispersion.
$\boldsymbol{\Gamma}$: the Gamma function, which generalizes the factorial function to continuous values. For integer arguments, $\boldsymbol{\Gamma(n) = (n-1)!}$

E. *Negative Binomial Regression Parameter Test*

The objective of this test is to assess the effect of predictor variables (X) on the response variable (Y), both partially and simultaneously [7], [18].
Simultaneous testing is performed using the likelihood ratio test with the following hypotheses:

- Null Hypothesis H_0 : $\beta_1 = \beta_2 = \cdots = \beta_n = 0$
- Alternative Hypothesis H_1 : $\beta_j \neq 0$; for at least $j = 1, 2, 3, \dots, n$

where:
$L(\hat{\Omega})$: the likelihood value when all predictor variables (X) are included.
$L(\hat{\omega})$: represents the likelihood value when the predictor variables (X) are excluded.
Partial testing evaluates the individual effects of each predictor variable on the response variable. The hypotheses for this test are:
H_0: $\beta_j = 0$

H_1: $\beta_j \neq 0$, $j = 1, 2, 3, \dots, n$
The test statistic is calculated as:

$$z = \frac{\hat{\beta}_j}{se(\hat{\beta}_i)}$$

$\hat{\beta}_j$: the estimated coefficient for the j-th predictor variable.
$se(\hat{\beta}_i)$: the standard error of $\hat{\beta}_j$
Decision rule:
Reject H_0 if $|Z_{count}| > Z_{(n-k-1,\frac{\alpha}{2})}$ which $Z_{(n-k-1,\frac{\alpha}{2})}$ is the critical value from the standart normal distribution.
Alternatively, if the *p-value* < ∝, H_0 should be rejected.

F. *Selection of the Best Model*

To identify the most suitable model, the Akaike Information Criterion (AIC) is commonly used. The AIC is based on the log-likelihood function and is calculated as follows [4], [15]:

$$AIC = -2 \ln L + 2p$$

where
L : the likelihood of the model, and
p : the number of parameters in the model. A lower AIC indicates a better model
In addition to AIC, the Bayesian Information Criterion (BIC) and deviance are also used for model comparison. The Log Likelihood value is another important criterion; the model with the highest Log Likelihood value is generally preferred.

III. METHODOLOGY

Data and Research Variables

The data utilized in this study includes the number of stunted children in 2023 (29 community health center), which serves as the response variable and was obtained from the Central Lombok District Health Office. The predictor variables include the number of malnourished toddlers, the number of toddlers receiving complete basic immunization, the number of deliveries attended by health workers, the incidence of LBW, and the number of newborns receiving immediate newborn care (IMD). The selection of this predictor variable is informed by previous research on factors influencing stunting cases [12], [22]–[24]. The variables used are presented in Table 1.

Table 1 Variables and Operational definition of variables

	Variables	Operational definition of variables
Y	Toddler stunting	Children with short stature (< -2 SD) [25].
X_1	Malnourished Toddlers	Underweight toddlers are defined as those with a Z-score between -2.0 and -3 or and or arm circumference of 11.5-12.5 cm in Children aged 6-59 months [26].
X_2	Toddlers receive complete basic immunization	Toddlers who have received a complete schedule of basic immunizations, which includes several types of vaccines such as polio, BCG, DPT, and others [27].
X_3	Delivery of pregnant women assisted by health workers	Pregnant women deliver with the assistance of a qualified health worker at a healthcare facility [28].
X_4	Low birth weight	Babies with a birth weight of less than 2,500 grams are weighed one hour after birth [29].
X_5	Newborns receive IMD	Babies begin self-feeding shortly after birth [30].

The steps involved in the negative binomial regression analysis are as follows:

1. Data were collected from 29 community health centers, including Kopang, Muncan, Waja Geseng, Teratak, Tanak Beak, Aik Darek, Mantang, Janapria, Langko, Pringgarata, Bagu, Bonjeruk, Ubung, Puyung, Darek, Batujangkih, Praya, Aik Mual, Pengadang, Batunyala, Batujai, Penujak, Mangkung, Ganti, Mujur, Teruwai, Sengkol, Kuta, and Awang community health centers.
2. The data were cleaned to ensure accuracy and consistency.
3. A multicollinearity test of the independent variables was conducted using the *car* library in R Software version 4.4.0.
4. Poisson regression analysis was performed using the *AER* library in R Software version 4.4.0.
5. Conduct a significance test for the parameters of the Poisson regression model.
6. An overdispersion test was conducted by performing the Deviance test using the *DHARMa* library.
7. A Negative Binomial analysis was executed using the MASS library.
8. Parameter significance was assessed through both simultaneous and partial tests.
9. To determine the best model among the three—Poisson regression model, Negative Binomial regression model, and Negative Binomial regression model with significant variables—based on the criteria of the best model according to the AIC value.

INFERENSI, Vol. 7(3), November. 2024. ISSN: 0216-308X (Print) 2721-3862 (Online)

10. Finally, the results were interpreted to draw meaningful conclusions.

IV. RESULTS AND DISCUSSIONS

A. Descriptive Statistics

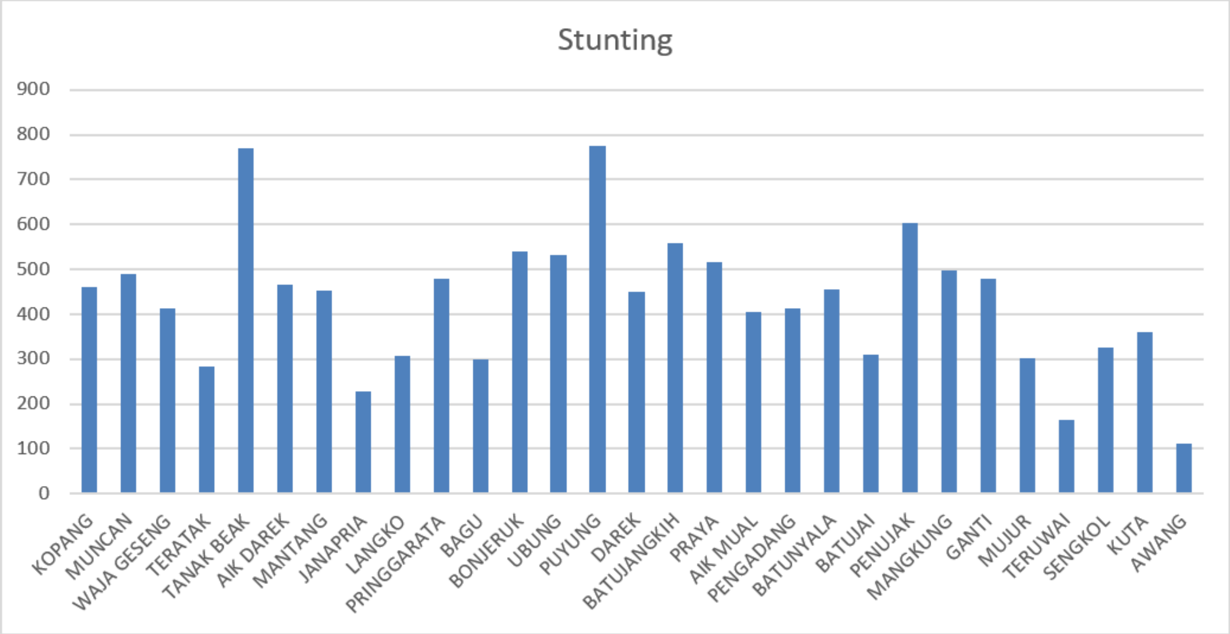

Figure 1 Number of stunting cases in 2023

Table 2 The number of stunting cases in each subdistrics and Puskesmas

No	Subdistricts	Puskesmas	Number of Stunting Cases	Total
1	Batukliang	Mantang	453	918
		Aik Darek	465	
2	Batukliang Utara	Teratak	283	1054
		Tanak Beak	771	
3	Janapria	Janapria	229	537
		Langko	308	
4	Jonggat	Ubung	533	1848
		Bonjeruk	540	
		Puyung	775	
5	Kopang	Muncan	490	903
		Wajageseng	413	
6	Praya	Praya	517	1382
		Aik Mual	405	
		Kopang	460	
7	Praya Barat	Penujak	604	1411
		Mangkung	497	
		Batujai	310	
8	Praya Barat Daya	Darek	451	1008
		Batu Jangkih	557	
9	Praya Tengah	Pengadang	412	866
		Batunyala	454	
10	Praya Timur	Mujur	301	781
		Ganti	480	
11	Pringgarata	Pringgarata	478	776
		Bagu	298	
12	Pujut	Sengkol	326	962
		Kuta	361	
		Teruwai	164	
		Awang	111	

The distribution of reported stunting cases across the 29 community health centers (Puskesmas) is presented in Figure 1 and Table 2. Based on these figures, Jonggat Subdistrict records the highest number of stunting cases, with a total of 1,848 cases. Of the three health centers in Jonggat, the Puyung Health Center has the highest number of stunting cases among all 29 centers. The high prevalence of stunting in Jonggat is largely attributed to inadequate nutritional intake. In contrast, Janapria Subdistrict has the fewest reported cases, with 537 cases distributed across two health centers, Janapria and Langko. However, the lowest number of cases overall is found at the Awang Health Center in Pujut Subdistrict, with only 111 cases.

Descriptive statistics, including the minimum, maximum, mean, variance, and standard deviation for both the dependent and independent variables, are presented in Table 3.

Table 3 Descriptive analysis of the variables

Variables	Min	Max	Mean	Variance	Standard Deviation
Y	111	775	429	22972.29	151.5661
X_1	122	918	401	29342.06	171.2953
X_2	107	516	305	9681.15	98.39284
X_3	24	927	355	24092.89	155.2189
X_4	0	81	18.97	296.4631	17.2181
X_5	104	961	347.4	44245.1	210.3452

B. Multicollinearity Test

When a correlation exists among independent variables, regression analysis may not be applicable. Tolerance values and the Variance Inflation Factor (VIF) can be utilized to assess the presence of multicollinearity among independent variables [18], [23]. The results of the multicollinearity test, conducted using the *car* package in R software, are presented in Table 4.

Tabel 4 The values of VIF for each independent variable

Variables	VIF
X_1	1.89
X_2	3.61
X_3	3.75
X_4	1.14
X_5	1.61

The multicollinearity test utilizes the *vif* function. According to Table 4, the five variables have a VIF value of less than 10, indicating that there is no multicollinearity among the variables. Therefore, these variables can be used in the construction of Poisson regression models.

C. Poisson Regression Modeling

The results of the parameter estimation tests for the Poisson regression model are given in Table 5:

Tabel 5 Poisson Regression Test Value

Parameters	Value	Std. Error	Z value	P value
β_0	5.595000	0.0381900	146.523	<2e-16 ***
β_1	0.0012050	0.0000755	15.9560	<2e-16 ***
β_2	-0.0004084	0.0001847	-2.2110	0.027 *
β_3	-0.0000654	0.0001132	-0.5780	0.563
β_4	0.0064960	0.0005486	11.842	<2e-16 ***
β_5	-0.0000657	0.0000549	-1.1970	0.231

Based on Table 5, the following model is obtained after estimating the Poisson Regression:

$$(\hat{\mu}_i) = \exp\,(5.595{+}0.001205X_1{-}0.0004084X_2{-}0.0000654X_3{+}0.006496X_4{-}0.00006568X_5)$$

Among the five explanatory variables, three are statistically significant and have an impact on the dependent variable at the 5% significance level ($p < 0.05$): malnourished toddlers (X_1), toddlers receiving complete basic immunization (X_2), and those with low birth weight (X_4).

D. Overdispersion Test

Overdispersion refers to a situation where the variance of a dataset exceeds its mean. Based on the results (Table 1) of the descriptive analysis of the research data, it is evident that the variance of the dependent variable, stunting cases (Y), is greater than its mean, with a variance of 22972.29 compared to a mean of 429. Additionally, the presence of overdispersion in the stunting data from Central Lombok Regency can be assessed by examining the Deviance and Chi-Square values (Table 6). When these values are divided by their respective degrees of freedom, the resulting ratios exceed 1, indicating overdispersion.

Table 6 Results of Overdispersion Test

Criteria	df	Value	Ratio
Null Deviance	28	1570.92	56.10
Residual Deviance	23	785.91	34.17

The residual deviance value is 785.91 with 23 degrees of freedom. When the residual deviance (785.91) is divided by the degrees of freedom (23), the resulting value is 34.17 which exceeds 1, indicating the presence of overdispersion in the Poisson regression model. Overdispersion can compromise the quality of the Poisson regression model by increasing the error rate. One effective approach to mitigate the effects of overdispersion is to replace the Poisson distribution assumption with a Negative Binomial distribution.

E. Negative Binomial Regression Model

The results of the parameter estimation using the Negative Binomial approach with R software using "mass" are presented in Table 7.

Table 7 Parameter Value of Negative Binomial regression

Parameters	Value	Std. Error	Z value	P value
β_0	5.4894116	0.2111848	25.993	< 2e-16
β_1	0.0012424	0.0004209	2.952	0.00316
β_2	-0.0003329	0.0010131	-0.329	0.74244
β_3	0.0000299	0.0006540	0.046	0.96353
β_4	0.0079273	0.0032524	2.437	0.01480
β_5	-0.0000534	0.0003158	-0.169	0.86573

The significant predictor variables at a 5% significance level are X_1 and X_4. Consequently, the Negative Binomial regression model for stunting cases in Central Lombok Regency can be expressed as follows:

$$(\hat{\mu}_i) = \exp\,(5.4894116{+}0.0012424X_1{+}0.0079273X_4)$$

Based on this equation, the variable representing malnourished toddlers (X_1), with a parameter estimate of $\hat{\beta}$= 0.0012424, yields an exponential value of 1.001243. This indicates that an increase in the number of malnourished toddlers is associated with a corresponding increase in the number of stunted toddlers. Similarly, the Low-Birth-Weight variable (X_4), with a parameter estimate of $\hat{\beta}$ = 0.0079273, results in an exponential value of 1.007959. This suggests that an increase in the number of low-birth-weight infants is associated with a higher prevalence of stunted toddlers. The results of this study are consistent with research conducted by [31]–[33], which indicates that the prevalence of malnutrition among infants and the incidence of low birth weight positively contribute to the increase of stunting.

F. Selection of the Best Model

The best model is selected from the various models generated by considering the smallest AIC value. A lower AIC value indicates a superior model fit, suggesting that it better explains the data.

Table 8 Table of AIC values

Model	AIC
Poisson	1024.9
Negative Binomial	370.39
Negative Binomial with variables (X_1 and X_4)	364.69

Table 8 demonstrates that the most effective model for predicting the stunting rate in Central Lombok Regency is the Negative Binomial model, which incorporates the variables X_1 and X_4 and has an AIC value of 364.69.

This study has several limitations. First, the NBR model assumes constant overdispersion across all observations. If the level of overdispersion varies across different subgroups, the model may not fully capture the complexity of the data. Additionally, NBR does not account for spatial correlation. In spatial data, such as stunting cases, nearby regions may exhibit similar stunting rates due to shared environmental factors or socioeconomic conditions, and ignoring spatial autocorrelation can lead to biased estimates. Furthermore, if key covariates influencing stunting are omitted from the model, the results may be biased, leading to incorrect inferences. This could be due to unmeasured confounders or incorrect model specification. While NBR is more flexible than Poisson regression, it may still oversimplify complex relationships between variables, such as interactions between covariates or nonlinear relationships.

To address these limitations, future work could implement spatial regression models like Spatial Lag or Spatial Error models to account for spatial dependency. Additionally, exploring Bayesian Spatial Models, such as Bayesian Spatial Conditional Autoregressive (CAR) models, could provide a more nuanced understanding of the spatial distribution of stunting cases. If longitudinal data are available, applying models such as Random Effects or Fixed Effects Negative Binomial models could help in understanding the temporal dynamics of stunting cases and how the factors influencing stunting evolve over time. Furthermore, conducting cross-validation or sensitivity analyses to assess the robustness of the model's predictions can enhance the reliability of the findings. Comparing NBR with alternative models, such as Zero-Inflated or Hurdle models, might also yield valuable insights.

V. CONCLUSIONS AND SUGGESTIONS

The NBR model is an effective approach for addressing overdispersion in count data, as it includes an additional parameter to account for variability that exceeds the mean. In this study, the NBR model was employed to examine the determinants of stunting in Central Lombok.

The results of the Negative Binomial regression model for stunting cases across 29 community health centers in Central Lombok Regency, NTB Province, in 2023 are expressed as follows:

$$(\hat{\mu}_i) = \exp\,(5.4894116 + 0.0012424X_1 + 0.0079273X_4)$$

The significant factors influencing the incidence of stunting in Central Lombok Regency, as identified by the Negative Binomial regression, are malnourished toddlers (X_1) and low birth weight (X_4) with an AIC value of 364.69.

The results reveal that an increase in the number of malnourished toddlers is associated with a corresponding rise in stunting cases. Additionally, a higher prevalence of low-birth-weight infants is linked to an elevated incidence of stunting.

The author suggests that future research should explore comparisons of models using various criteria, including spatial aspects. Additionally, it is recommended that the Central Lombok Regency government focus more attention on community health centers with high stunting rates and implement strategies to control the factors contributing to the rise in stunting cases.

ACKNOWLEDGEMET

The authors wish to express their gratitude to the Kemenristekdikti, Indonesia for its support. This work was funded by Kemenristekdikti under the Postgraduate Program-Master's Thesis Research scheme for 2024, contract number 2773/UN36.11/LP2M/2024.

REFERENCES

[1] A. Khairunnisa, Ihsan, and M. Patandianan, "Risk Burden of Heat-Related Morbility Due to Urban Heat Island Effect in Tamalanrea District, Makassar," *J. Masy. Kesehat.*, vol. 19, no. 4, pp. 149–158, 2023, doi: 10.30597/mkmi.v19i4.26691.

[2] A. Aswi, S. A. Astuti, and S. Sudarmin, "Evaluating the Performance of Zero-Inflated and Hurdle Poisson Models for Modeling Overdispersion in Count Data," *Inferensi*, vol. 5, no. 1, p. 17, 2022, doi: 10.12962/j27213862.v5i1.12422.

[3] A. Rahayu, "Model-model Regresi untuk Mengatasi Masalah Overdispersi pada Regresi Poisson," *J. peqguruan Conf. Ser.*, vol. 1, no. April, pp. 1–5, 2020, doi: http://dx.doi.org/10.35329/jp.v2i1.1866.

[4] Y. A. Ulfa, A. M. Soleh, and B. Sartono, "Handling of Overdispersion in the Poisson Regression Model with Negative Binomial for the Number of New Cases of Leprosy in Java," *Indones. J. Stat. Its Appl.*, vol. 5, no. 1, pp. 1–13, 2021, doi: 10.29244/ijsa.v5i1p1-13.

[5] N. H. Fitrial and A. Fatikhurrizqi, "Pemodelan Jumlah Kasus Covid-19 Di Indonesia Dengan Pendekatan Regresi Poisson Dan Regresi Binomial Negatif," *Semin. Nas. Off. Stat.*, vol. 2020, no. 1, pp. 65–72, 2021, doi: 10.34123/semnasoffstat.v2020i1.465.

[6] P. C. Ambarwati, I. Indahwati, and M. N. Aidi, "Kajian Simulasi Overdispersi Pada Regresi Poisson Dan Binomial Negatif Terboboti Geografis Untuk Data Balita Gizi Buruk," *Indones. J. Stat. Its Appl.*, vol. 4, no. 3, pp. 484–497, 2020, doi: 10.29244/ijsa.v4i3.684.

[7] H. M. Winata, "Mengatasi Overdispersi dengan Regresi Binomial Negatif pada Angka Kematian Ibu di Kota Bandung," *J. Gaussian*, vol. 11, no. 4, pp. 616–622, 2023, doi: 10.14710/j.gauss.11.4.616-622.

[8] A. Sauddin, N. I. Auliah, and W. Alwi, "Pemodelan Jumlah Kematian Ibu di Provinsi Sulawesi Selatan Menggunakan Regresi Binomial Negatif," *J. MSA (Mat. dan Stat. serta Apl.)*, vol. 8, no. 2, p. 42, 2020, doi: 10.24252/msa.v8i2.17409.

[9] A. Aswi, S. Sukarna, and N. Nurhilaliyah, "Pemetaan Risiko Relatif Kasus Stunting di Provinsi Sulawesi Selatan," *Sainsmat J.*

Ilm. Ilmu Pengetah. Alam, vol. 11, no. 1, p. 11, 2022, doi: 10.35580/sainsmat111325202022.

[10] A. A. Azis and A. Aswi, "Spatial Clustering of Stunting Cases in Indonesia: a Bayesian Approach," *Commun. Math. Biol. Neurosci.*, vol. 2023, pp. 1–11, 2023, doi: 10.28919/cmbn/7898.

[11] A. Aswi, B. Poerwanto, Sudarmin, and Nurwan, *Bayesian Spatial Modelling of Stunting Cases in South Sulawesi Province: Influential Factors and Relative Risk*, vol. 2023, no. Icsmtr. Atlantis Press International BV, 2023. doi: 10.2991/978-94-6463-332-0_11.

[12] A. Aswi and S. Sukarna, "Pemodelan Spasial Bayesian dalam Menentukan Faktor yang Mempengaruhi Kejadian Stunting di Provinsi Sulawesi Selatan," *J. Math. Comput. Stat.*, vol. 5, no. 1, pp. 1–11, 2021.

[13] U. Stunting, "Cegah Stunting, Wapres Minta Keluarga Indonesia Prioritaskan Kebutuhan Gizi Anak dan Sanitasi," *RI, Kementerian Sekretariat Negara Presiden, Sekretariat Wakil*, 2023.

[14] Rokom, "Prevalensi Stunting di Indonesia Turun ke 21,6% dari 24,4%," *Sehat Negriku*, 2023.

[15] L. E. Afri, Aunuddin, and A. Djuraidah, "Model regresi binomial negatif terboboti geografis untuk data kematian bayi," *Forum Stat. dan Komputasi Indones. J. Stat.*, vol. 17, no. 2, pp. 33–39, 2012.

[16] D. K. Wardani and A. Wulandari, "Pemodelan Negative Binomial Regression Pada Data Jumlah Kematian Bayi Di Kabupaten Jombang," *Transform. J. Pendidik. Mat. dan Mat.*, vol. 4, no. 2, pp. 311–320, 2020, doi: 10.36526/tr.v4i2.968.

[17] L. I. Oztig and O. E. Askin, "Human mobility and coronavirus disease 2019 (COVID-19): a negative binomial regression analysis," *Public Health*, vol. 185, pp. 364–367, 2020, doi: 10.1016/j.puhe.2020.07.002.

[18] M. Fathurahman, "Regresi Binomial Negatif untuk Memodelkan Kematian Bayi di Kalimantan Timur," *Eksponensial*, vol. 13, pp. 79–86, 2022.

[19] T. W. Utami, "Analisis regresi binomial negatif untuk mengatasi overdispersion regresi poisson pada kasus demam berdarah dengue," *J. Stat. Univ. Muhammadiyah Semarang*, vol. 1, no. 2, pp. 0–6, 2013.

[20] F. Fitri, F. M. Sari, N. F. Gamayanti, and I. T. Utami, "Infant Mortality Case: An Application of Negative Binomial Regression in order to Overcome Overdispersion in Poisson Regression," *Eksata Berk. Ilm. Bid. MIPA*, vol. 22, no. 03, pp. 200–210, 2021, doi: https://doi.org/10.24036//eksakta/vol22-iss2/272 Eksakta BerkalaiIlmiah Bidang MIPA http://www.eksakta.ppj.unp.ac.id/index.php/eksakta Article.

[21] R. A. Salenussa, M. S. N. Van Delsen, and G. Haumahu, "Modelling Negative Binomial Regression To Resolve Overdisperssion(Case Studi : The Number of Families at Risk of Stunting in Maluku Province in 2021)," vol. 4, no. 2, 2023.

[22] F. Zubedi, M. A. Aliu, Y. Rahim, and F. A. Oroh, "Analisis Faktor-Faktor Yang Mempengaruhi Stunting Pada Balita Di Kota Gorontalo Menggunakan Regresi Binomial Negatif," *Jambura J. Probab. Stat.*, vol. 2, no. 1, pp. 48–55, 2021, doi: 10.34312/jjps.v2i1.10284.

[23] F. Cholid, D. Trishnanti, and H. Al Azies, "Pemetaan Faktor-Faktor yang Mempengaruhi Stunting pada Balita dengan Geographically Weighted Regression (GWR) Mapping of The Factors Affecting Stunting on Toddlers with Geographically Weighted Regression (GWR)," *SEMNAkes 2019 "Improving Qual. Heal. Tharough Adv. Res. HHelatH Sci.*, no. March 2020, pp. 156–165, 2019, doi: 10.17605/OSF.IO/9MZU7.

[24] K. Widayati, I. kadek A. D. Putra, and N. L. M. A. Dewi, "Determinant Factor for Stunting in Toddler," *J. Aisyah J. Ilmu Kesehat.*, vol. 6, no. 1, pp. 9–16, 2021, doi: 10.30604/jika.v6i1.381.

[25] Menteri Kesehatan Republik Indoesia, "Peraturan Menteri Kesehatan Repunlik Indonesia Nomor 2 Tahun 2020," vol. 21, no. 1, pp. 1–9, 2020.

[26] H. SUYUTI, "Peraturan Menteri Kesehatan Republik Indonesia Nomor 29 Tahun 2019," pp. 5–10, 2019.

[27] Menteri Kesehatan Republik Indonesia, "Peraturan Menteri Kesehatan Republik Indonesia Nomor 12 Tahun 2017," pp. 2–4, 2017.

[28] B. LINGGA and P. K. RIAU, "Persalinan Melalui Tenaga Kesehatan dan Fasilitas Kesehatan," pp. 1–29, 2022.

[29] M. Rivki, A. M. Bachtiar, T. Informatika, F. Teknik, and U. K. Indonesia, *Pendek (Stunting) di Indonesia, Masalah dan Solusinya*, no. 112. 2015.

[30] C. Segovia, "Inisiasi Menyusu Dini dan Air Susu Ibu Ekslusif," *Tesis Dr.*, vol. 2014, no. June, pp. 1–2, 2014.

[31] I. P. Sari, Y. Ardillah, and A. Rahmiwati, "Berat bayi lahir dan kejadian stunting pada anak usia 6-59 bulan di Kecamatan Seberang Ulu I Palembang," *J. Gizi Indones. (The Indones. J. Nutr.*, vol. 8, no. 2, pp. 110–118, 2020, doi: 10.14710/jgi.8.2.110-118.

[32] M. Ihwal, "Arus Jurnal Sains dan Teknologi (AJST) Analisis Faktor-faktor yang Mempengaruhi Banyaknya Kasus Stunting," vol. 2, no. 1, 2024.

[33] A. Adityaningrum *et al.*, "Faktor Penyebab Stunting Di Indonesia: Analisis Data Sekunder Data Ssgi Tahun 2021 Factors Causing Stunting in Indonesia: 2021 Ssgi Secondary Data Analysis," *Jambura J. Epidemiol.*, vol. 3, no. 1, pp. 1–10, 2021.

© 2024 by the authors. This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (http://creativecommons.org/licenses/by-sa/4.0/).