JSDS: JOURNAL OF STATISTICS AND DATA SCIENCE
VOLUME 1, No 2, October 2022
e-ISSN: 2828-9986
https://ejournal.unib.ac.id/index.php/jsds/index

Measles Disease Analysis in Bengkulu Province Using Zero Inflated Poisson Regression and Zero Inflated Negative Binomial Regression

Ilham Alifa Azagi [1*]
[1] Departement of Statistics, IPB University, Indonesia

* Corresponding Author: ilhamalifa98@gmail.com

Article Info

Article History:
Received: September 30, 2022
Revised: October 27, 2022
Accepted: October 28, 2022
Available Online: October 31, 2022

Key Words:
Campak
Regresi ZIP
Regresi ZINB
AIC

Abstract

Zero Inflated Poisson regression (ZIP) and Zero Inflated Negative Binomial (ZINB) regression were used if there was overdispersion and no multicollinearity in the data. This study aims to analyze measles in Bengkulu Province using the ZIP and ZINB regression models. Among them are selecting the best model, seeing the influential variables from the best model, and predicting the results of the best model. The data used is one dependent variable, namely the number of measles cases (Y) in each puskesmas and six independent variables (X) namely the percentage of measles immunization, the amount of malnutrition, the percentage of exclusive breastfeeding, the percentage of vitamin A, the percentage of proper sanitation, and the percentage of healthy house. The results of this study, the ZIP regression model formed is a discrete model for $\hat{\mu}$, namely $\ln(\hat{\mu}_i)$=-5.042-0.007X_1-0.014X_3+0.094X_4 and a zero inflation model for $\hat{\omega}$, namely logit($\hat{\omega}_i$)= -3.656+0.101X_4-0.054X_6, while the ZINB regression model formed is a discrete model for $\hat{\mu}$, namely $\ln(\hat{\mu}_i)$=-9,289+0.120X_4 and a zero inflation model for $\hat{\pi}$, namely logit($\hat{\pi}_i$)=- 17.841+0.205X_4. The AIC value of the ZINB regression model is 255.249, which is smaller than the AIC value of the ZIP regression model of 331.467, so the ZINB regression model is better to use. The influential variable in this study is the percentage of vitamin A administered. There is not much difference between predicted results and the actual data.

1. INTRODUCTION

According to the World Health Organization (WHO), health is a state of complete mental, physical and social well-being not only free from disease but functioning normally. According to the Ministry of Health, health is a normal and prosperous state in which a person's physical, social and mental activities can be carried out without obstacles, meaning that there is continuity between physical health. One of the health problems in Indonesia is measles, which mostly affects children under the age of five. According to WHO data in 2015, Indonesia is one of the 10 countries with the highest number of measles cases in the world. The Indonesian Ministry of Health found that the number of measles cases in Indonesia is very high and tends to increase from 2015 to 2017. The Zero Inflated Poisson (ZIP) and Zero Inflated Negative Binomial (ZINB) regression models can be used to solve this problem. A previous study entitled "Comparison of Zero-Inflated Poisson Regression (ZIP) and Zero-Inflated Negative Binomial Regression (ZINB) on Overdispersed Data" was conducted by Dewanti, Susilawati, and Srinadi in 2016. In 2017, the study entitled Best Regression Model using Zero Inflated Poisson (ZIP) and Zero Inflated Negative Binomial (ZINB) conducted by Kurniawan. The case study in this study is the under-five mortality rate at Tirto Health Center, Pekalongan City in 2016, with the conclusion that the best Poisson model is between Zero Inflated Poisson (ZIP) and Zero Inflated Negative Binomial (ZINB) is a Zero Inflated Poisson (ZIP) model with an AIC value (101.1387) which is smaller than Zero Inflated Negative Binomial (ZINB) with an AIC value (103.1392). Based on the background, the authors conducted an analysis of measles in Bengkulu Province using Zero Inflated Poisson (ZIP) regression and Zero Inflated Negative Binomial (ZINB) regression. This is to get the best model among the ZIP and ZINB regressions, the influencing factors from the best model, and predict from the best model.

1.1 Poisson Distribusi

Poisson is a discrete distribution used to estimate the probability that a given outcome will occur exactly y times in standardized units with the average occurrence of events per unit constant. A discrete random variable Y is said to have a Poisson distribution with parameter $\mu > 0$ if it has the probability density function as follows:

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

1.2 Negative Binomial Distribution

An experiment that has the same nature as a binomial experiment, but on a negative binomial distribution the action is repeated until a certain number of successes are reached. In the Negative Binomial distribution, it is possible to know the probability of Y success in n actions that have been determined. It is also known the probability of r-th success in the y-th action. So in a negative binomial experiment, the number of successes is certain while the number of trials is random [1]. A random variable Y that has a negative binomial distribution has the probability density function as follows:

$$f(y;r,\pi) = C_{r-1}^{y-1}\pi^{r}(1-\pi)^{y-r} \qquad y = r, r+1, \ldots \quad (2)$$

1.3 Regression

Regression is a measurement of the relationship between two or more variables which is expressed in the form of a relationship/function. Regression requires a firm separation between the independent variable and the dependent variable, usually sequentially with X and Y. Regression analysis aims to explain or model the relationship between variables. Regression analysis is an advanced analysis that is used to predict how far the change in the value of a variable is. Thus, regression analysis can assist in making decisions whether the rise and fall of a dependent variable can be done by increasing or decreasing the independent variable [2].

1.4 Poisson Regression

Poisson regression is a regression model in which the dependent variable (Y) is discretely distributed or the standard model for discrete data [3]. Let Y be the number of events that occur in one period with the parameter value of the Poisson distribution (μ). The expected value and variance of the Poisson distribution is μ [4]. Poisson regression model [5] can be described as:

$$Y_i = \mu_i + \varepsilon_i$$
$$ln(\mu_i) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_k X_{ik}$$
$$\mu_i = e^{X_i^T \beta}$$

1.5 Multicollinearity test

The multicollinearity test aims to see whether the regression model found a correlation between the independent variables, where every change in an independent variable will cause the other independent variables to change. The method for testing the existence of this multicollinearity can be seen from the Tolerance Value or Variance Inflation Factor (VIF). If VIF > 10 or if tolerance value < 0.1, multicollinearity occurs. If VIF < 10 or if tolerance value > 0.1, there is no multicollinearity [6]. VIF generates an index of the amount by which the variance of each regression coefficient increases relatively. An increase occurs when the independent variable situation is correlated. The VIF formula is:

$$VIF = \frac{1}{1 - R_j^2}$$

1.6 Overdispersion

Overdispersion is a deviation of assumptions that often occurs in Poisson regression caused by the number of observations being zero. Overdispersion in Poisson regression results in the standard deviation of the estimated parameter being much smaller than the actual value (underestimate) and the significance test of the independent variable is much larger than the actual value (overestimate), so that the resulting conclusion becomes invalid. Homogeneous variance shows that the dispersion value in the data is 1. To detect overdispersion in Poisson regression, Pearson's Chi-Square is divided into degrees of freedom. The formula for the Pearson's Chi-Square statistical test is as follows:

$$\chi^2 = \sum_{i=1}^{n} \frac{(y_i - E(Y))^2}{Var(Y)}$$

1.7 Excess of Zeros

One of the problems in Poisson regression is the excess zeros (Excess Zeros). Excess zeros can be seen in the proportion of response variables with a value of zero which is greater than other discrete data. In addition, Poisson regression also becomes no longer accurate in describing the actual data. Excess zeros is one of the causes of overdispersion [7].

1.8 Zero Inflated Poisson (ZIP) Regression

The ZIP regression model was used to analyze more zero observations than estimated. In the ZIP regression, the case with the dependent variable contains a zero value, which is greater than 50% [8]. If Y is a dependent random variable that has a ZIP distribution, the value of zero is assumed to appear in two ways corresponding to different circumstances. The first state occurs with probability _i and produces only zero values, while the other state occurs with probability (1-ω_i) and has a Poisson distribution called sampling zeros. This two-state process gives a simple two-component mixed distribution with the following probability function [9]:

$$p(Y = y_i) = \begin{cases} \omega_i + (1-\omega_i)e^{-\mu} & , y_i = 0 \\ (1-\omega_i)\dfrac{e^{-\mu}\mu^{y_i}}{y_i!} & , y_i > 0 \end{cases} \tag{3}$$

1.9 Zero Inflated Poisson (ZIP) Regression Parameters Estimation

The estimation of ZIP regression parameters uses the maximum likelihood method. This method is the most frequently used method for estimating parameters. The combined model for μ and ω is as follows [8]:

$$ln\ (\mu_i\ ) = \boldsymbol{X_i^T \beta} \text{ and } logit(\omega_i) = ln\ \left(\frac{\omega_i}{1-\omega_i}\right) = \boldsymbol{X_i^T \gamma} \tag{4}$$

1.10 Zero Inflated Negative Binomial (ZINB) Regression

The ZINB model can be used to model count data or discrete data with many zero values in the dependent variable (zero inflation) and overdispersion occurs. If Y_i is a random variable with i= 1,2,...,n then the value of the dependent variable occurs in two circumstances. The first state is called a zero state and produces only zero observations, while the second state is called a negative binomial state which has a negative binomial distribution with probability (1-π_i) so it is called sampling zeroes [10]. The probability function of the Zero Inflated Negative Binomial (ZINB) regression model can be expressed as follows:

$$P(Y_i = y_i) = \begin{cases} \pi_i + (1-\pi_i)\left(\dfrac{1}{1+\kappa\mu_i}\right)^{\frac{1}{\kappa}}, & y_i = 0 \\ (1-\pi_i)\dfrac{\Gamma\left(y_i + \frac{1}{\kappa}\right)}{\Gamma\left(\frac{1}{\kappa}\right) y_i!}\left(\dfrac{1}{1+\kappa\mu_i}\right)^{\frac{1}{\kappa}}\left(\dfrac{\kappa\mu_i}{1+\kappa\mu_i}\right)^{y_i}, & y_i > 0 \end{cases} \tag{5}$$

1.11 Zero Inflated Negative Binomial (ZINB) Regression Parameter Estimation

The estimation of ZINB regression parameters using the maximum likelihood method. This method is the most frequently used method for estimating parameters. The combined model for μ and ω is as follows [8]:

$$ln\ (\mu_i\ ) = \boldsymbol{X_i^T}\ \boldsymbol{\beta} \text{ and } logit(\pi_i) = ln\ \left(\frac{\pi_i}{1-\pi_i}\right) = \boldsymbol{X_i^T}\boldsymbol{\gamma} \quad (6)$$

1.12 Model Fit Test

Testing the suitability of the ZIP and ZINB regression models can use the Likelihood Ratio (LR) value [11]. The hypothesis of the likelihood ratio test is as follows:

$H_0 : \beta_1 = \beta_2 = \cdots = \beta_p = \gamma_1 = \gamma_2 = \cdots = \gamma_p = 0$ (All independent variables have no effect on the dependent variable)

$H_1 :$ At least one of $\beta_j \neq 0$ or $\gamma_j \neq 0$ (There are independent variables that have an influence on the dependent variable)

The statistics of the likelihood ratio test [4] are as follows:

$$G = -2\ ln\left[\frac{L_0}{L_1}\right] = -2(ln\ L_0 - ln\ L_1)$$

1.13 Parameter Testing

The ZIP and ZINB regression parameters were tested partially, namely with the and parameters. The test is carried out in two ways, namely testing the log model parameters and testing the logit model parameters. The test statistics used are by using the Wald test [12] are as follows:

1. Parameter β

Test the significance of the parameter ln(μ)=$\boldsymbol{X\beta}$ for the ZIP regression model and the significant test of the parameter ln(μi)=X_i^T for the ZINB regression model.

$H_0 : \beta_j = 0$

$H_1 : \beta_j \neq 0$

Test statistic:

$$W_j = \left(\frac{\hat{\beta}_j}{SE(\hat{\beta}_j)}\right)^2$$

Rejection criteria:

Reject H_0 if $W_j > \chi^2_{(\alpha,1)}$ at significant level (α). This means that the independent variable (X_j) has a significant influence on the dependent variable (Y).

2. Parameter γ

The significant test of the logit(ω)=$\boldsymbol{X\gamma}$ parameter for the ZIP regression model and the significant test of the $logit(\pi_i) = ln\left(\frac{\pi_i}{1-\pi_i}\right) = \boldsymbol{X_i'}\boldsymbol{\gamma}$ for the ZINB regression model.

$H_0 : \gamma_j = 0$

$H_1 : \gamma_j \neq 0$

Test statistic:

$$W_j = \left(\frac{\hat{\gamma}_j}{SE(\hat{\gamma}_j)}\right)^2$$

Rejection criteria:

Reject H_0 if $W_j > \chi^2_{(\alpha,1)}$ at significant level (α). This means that the independent variable (X_j) has a significant influence on the dependent variable (Y).

1.14 Model Selection

The better model selection for ZIP and ZINB regression can be seen from the AIC (Akaike Information Criterion) value that can be used for selecting the better model. The AIC value is calculated based on the maximum likelihood value and the number of parameters in the formed regression model. The equation for calculating the AIC value is as follows:

$$AIC = -2 \ln L(\hat{\theta}) + 2k$$

2. METHOD

This study uses a quantitative approach using discrete data. The research carried out includes the type of applied research, namely the application of basic science in certain fields. The basic science in this research is the Zero Inflated Poisson (ZIP) regression model and the Zero Inflated Negative Binomial (ZINB) regression model which will be applied to measles cases. The data used is secondary data. Based on the number of measles cases in every Puskesmas in Bengkulu Province, the number of observational data is 179. The data is obtained from the publication of the Regency/City Health Office in 2018, namely the health profile of each Regency/City in Bengkulu Province in 2017.

The stages of the research carried out are as follows:

1. Describe the data using descriptive statistics.
2. Perform multicollinearity testing.
3. Test assumptions in Poisson regression analysis by using dispersion estimates in Poisson regression.
4. Perform ZIP regression analysis by finding the regression model and assessing the parameters of the ZIP regression model.
5. Conduct a ZIP regression model suitability test.
6. Performing a significant parameter test on the ZIP regression model partially.
7. Interpret the ZIP regression model.
8. Perform ZINB regression analysis by searching for regression models and assessing the parameters of the ZINB regression model.
9. Performing the ZINB regression model suitability test.
10. Performing a significant test of the parameters on the ZINB regression model partially.
11. Interpret the ZINB regression model.
12. Comparing the AIC value of the ZIP regression model and the ZINB regression to see the best model used to model the number of measles.
13. Predict the best regression model results.

3. RESULTS AND DISCUSSION

3.1 Descriptive statistics

Measles case modeling with ZIP and ZINB regression uses one dependent variable (Y), namely the number of measles cases and six independent variables, namely the percentage of measles immunization (X_1), the amount of malnutrition (X_2), the percentage of exclusive breastfeeding (X_3), the percentage of vitamin A administration. (X_4), the percentage of proper sanitation (X_5), and the percentage of healthy houses (X_6).

The descriptive statistics for each variable are shown in Table 1. Based on Table 1, the number of measles cases (Y) has a variance of 18.045, a standard deviation of 4.248, and an average of 1.011. The number of measles cases also has a minimum value of 0 in 158 Puskesmas from 5 districts/cities in Bengkulu Province and a maximum value of 41 occurring in Bentiring Health Center, Central Bengkulu Regency.

Table 1. Descriptive Statistics

Variable	Minimum	Maximum	Variance	Standard Deviation	Average
Y	0	41	18,045	4,248	1,011
X_1	1,350	181,300	1243,36	35,261	84,170
X_2	0	5	0,670	0,818	0,408
X_3	0	100	709,093	26,629	58,430
X_4	24,120	144,600	314,241	17,727	88,710
X_5	0	165,300	1447,213	38,042	47,440
X_6	0	187,690	950,867	30,836	64,750

3.2 Multicollinearity Test

The results of the VIF calculation for each independent variable are as follows:

Table 2 Multicollinearity Test

Independent Variable	VIF
X_1	1,060782
X_2	1,020944
X_3	1,221189
X_4	1,321654
X_5	1,209956
X_6	1,221912

Based on Table 2, the VIF value for each independent variable is less than 10. This means that there is no multicollinearity between independent variables or no correlation between independent variables..

3.3 Zero Inflated Poisson (ZIP) Regression Modeling

The first step in modeling the ZIP regression is to estimate the parameters. The ZIP regression model was carried out three times to obtain all independent variables (X) that affect the dependent variable (Y). The parameter estimation results are as follows:

Table 3. Results of ZIP Regression Parameter Estimation

Parameter	Estimate	Standard Error	z	Pr $(>\|z\|)$	W_j
$\hat{\beta}_0$	-5,042	1,039	-4,855	$1{,}2 \times e^{-6}$	23,549
$\hat{\beta}_1$	-0,007	0,003	-2,588	0,010	5,444
$\hat{\beta}_3$	-0,014	0,003	-4,393	$1{,}12 \times e^{-10}$	21,778
$\hat{\beta}_4$	0,094	0,012	7,954	$1{,}81 \times e^{-15}$	61,361
$\hat{\gamma}_0$	-3,656	4,241	-0,862	0,389	0,743
$\hat{\gamma}_4$	0,101	0,045	2,244	0,025	5,038
$\hat{\gamma}_6$	-0,054	0,016	-3,370	0,001	11,391

Based on Table 3, the ZIP regression model has two estimated parameters, namely β and γ. The β parameter is used when forming a discrete model for $\hat{\mu}$ while the γ parameter is used when forming a zero inflation model for $\widehat{\omega}$. The ZIP regression model formed is as follows:

1. Discrete model for $\hat{\mu}$

$$ln(\hat{\mu}_i) = -5{,}042 - 0{,}007X_{1i} - 0{,}014X_{3i} + 0{,}094X_{4i}$$

2. Zero inflation model for $\widehat{\omega}$

$$logit(\widehat{\omega}_i) = -3{,}656 + 0{,}101X_{4i} - 0{,}054X_{6i}$$

3.4 Zero Inflated Poisson (ZIP) Regression Model Interpretation

The ZIP regression model was carried out until all the variables tested had an effect on the dependent variable (*Y*). The equation of the ZIP regression model formed from the parameter significance test for the and parameters is as follows:

1. Discrete model for $\hat{\mu}$

$$ln(\hat{\mu}_i) = -5{,}042 - 0{,}007X_{1i} - 0{,}014X_{3i} + 0{,}094X_{4i}$$

The interpretation of the model is:

a. The X_{1i} coefficient is -0.007. It means that every 1% addition of measles immunization will cause a decrease in the probability of the number of measles cases by exp(-0.007) = 0.993 times the original number of measles cases, if other variables are constant.
b. The X_{3i} coefficient is -0.014. It means that every 1% addition of exclusive breastfeeding will cause a decrease in the probability of the number of measles cases by exp(-0.014) = 0.986 times the original number of measles cases, if other variables are constant.
c. The X_{4i} coefficient is 0.094. It means that every 1% addition of vitamin A will cause an increase in the probability of the number of measles cases by exp (0.094) = 1.098 times the original number of measles cases, if other variables are constant.

2. Zero inflation model for $\hat{\omega}$

$$logit(\hat{\omega}_i) = -3{,}656 + 0{,}101X_{4i} - 0{,}054X_{6i}$$

The interpretation of the model is:

a. The X_{4i} coefficient is 0.101. It means that for every 1% addition of vitamin A, it will cause an increase in the chance of the number of measles cases by exp(0,1005)=1,1057 times.
b. The X_{6i} coefficient is -0.054. It means that every 1% addition of healthy houses, it will cause a decrease in the chance of the number of measles cases by exp (-0.05361) = 0.9478 times.

3.5 Zero Inflated Negative Binomial (ZINB) Regression Modeling

The first step in modeling the ZINB regression is to estimate the parameters. The ZINB regression model was carried out twice until all the independent variables (*X*) tested affected the dependent variable (*Y*). The parameter estimation results are as follows:

Table 4. Result of ZINB . Regression Parameter Estimation

Parameter	Estimate	Standard Error	z	Pr $(>\|z\|)$	W_j
$\hat{\beta}_0$	-9,289	3,146	-2,952	0,003	8.718
$\hat{\beta}_4$	0,120	0,037	3,241	0,001	10.519
$\hat{\gamma}_0$	-17,841	8,165	-2,185	0,029	4.774
$\hat{\gamma}_4$	0,205	0,083	2,470	0,014	6.100

Based on Table 4, the ZINB regression model has two estimated parameters, namely β and γ. The β parameter is used when forming a discrete model for $\hat{\mu}$, while the γ parameter is used when forming a zero inflation model for $\hat{\pi}$. The ZINB regression model formed is as follows:

1. Discrete Model for $\hat{\mu}$

$$ln(\hat{\mu}_i) = -9{,}289 + 0{,}120X_{4i}$$

2. Zero Inflation Model for $\hat{\pi}$

$$logit(\hat{\pi}_i) = -17{,}841 + 0{,}205X_{4i}$$

3.6 Zero Inflated Negative Binomial (ZINB) Regression Model Interpretation

The ZINB regression model was carried out until all the variables tested had an effect on the dependent variable (*Y*). The equation of the ZINB regression model formed from the parameter significance test for the β and γ parameters is as follows:

1. Discrete Model for $\hat{\mu}$

$$ln(\hat{\mu}_i) = -9{,}289 + 0{,}120X_{4i}$$

The interpretation of the model is:

The X_{4i} coefficient of 0.120. It means that every 1% addition of vitamin A will cause an increase in the probability of the number of measles cases by exp(0.120) = 1.128 times the original number of measles cases, if other variables are constant.

2. Zero Inflation Model for $\hat{\pi}$

$$logit(\hat{\pi}_i) = -17{,}841 + 0{,}205X_{4i}$$

The interpretation of the model is:

The X_{4i} coefficient of 0.205. It means that for every 1% addition of vitamin A will cause a decrease in the chance of the number of measles cases by exp(0,205) = 1,228 times.

3.7 Goodness of the Model

The comparison of the goodness of the ZIP and ZINB regression models on the number of measles cases in Bengkulu Province in 2017 can be seen from the AIC (Akaike Information Criterion) value. The calculation of the AIC value of each model is calculated based on the maximum likelihood value and the number of parameters in the formed regression model. The results of the calculation of the value are as follows:

Table 5. Akaike Information Criterion

Regression Model	AIC
ZIP	331,467
ZINB	255,249

Based on Table 5, the ZINB regression model is better than the ZIP regression model. This is because the AIC value of the ZINB regression model is smaller than the AIC value of the ZIP regression model.

3.7 Best Model Result Prediction

The best model obtained in this study is the ZINB regression model. Therefore, predictions will be made using this model. Y prediction results are depicted in the graph as follows:

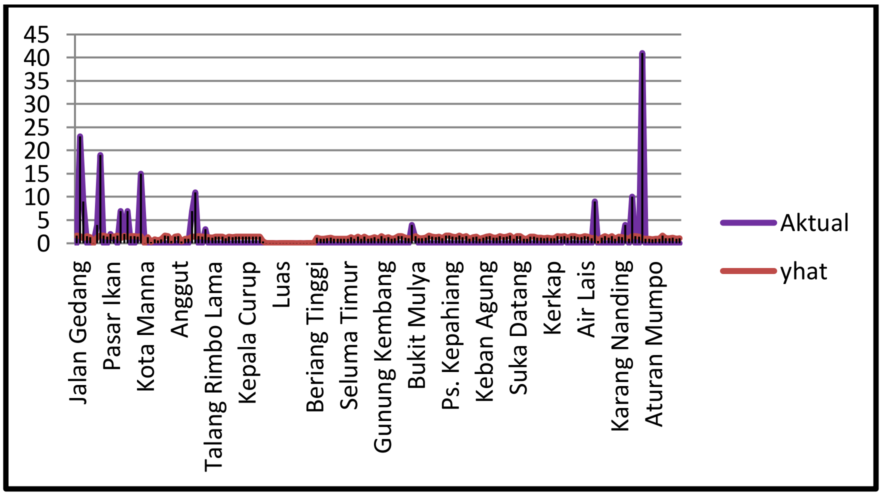

Figure 1. Graph of Actual Data and $\hat{Y}$.

Figure 1 illustrates the relationship between actual and predicted data. Actual data is data on the number of measles cases in Bengkulu Province in 2017. Predicted data are calculated using the ZINB regression. The visual result obtained is a graph between the actual and predicted data. ZINB regression is good enough to be used to model the data since there is no significant different between the predicted and actual data. In addition, there is no negative value on the predicted data.

4. CONCLUSION

The conclusions obtained from this research are as follows:

1. ZIP Regression Model in Measles Cases in Bengkulu Province in 2017
   a. Discrete Model for $\hat{\mu}$
   $$ln(\hat{\mu}_i) = -5{,}042 - 0{,}007X_{1i} - 0{,}014X_{3i} + 0{,}094X_{4i}$$
   b. Zero Inflation Model for $\hat{\omega}$
   $$logit(\hat{\omega}_i) = -3{,}656 + 0{,}101X_{4i} - 0{,}054X_{6i}$$
2. ZINB Regression Model in Measles Cases in Bengkulu Province in 2017
   a. Discrete Model for $\hat{\mu}$
   $$ln(\hat{\mu}_i) = -9{,}289 + 0{,}120X_{4i}$$
   b. Zero Inflation Model for $\hat{\pi}$
   $$logit(\hat{\pi}_i) = -17{,}841 + 0{,}205X_{4i}$$
3. The AIC value of the ZINB and ZIP regression model consecutively are 255.2492 and 331.4668. The ZINB regression is better than ZIP regression since it has smaller AIC value.
4. Based on the ZINB regression model, a significant factor influencing the number of measles cases in Bengkulu Province is the percentage of vitamin A administration.
5. Based on the difference between actual and predicted data and the value of the predicted data, the ZINB regression is good enough to be used to model the data.

REFERENCES

[1] Novianti, P. & I. Sriliana. 2016. *Handout Pengantar Teori Peluang*. Bengkulu : Universitas Bengkulu.

[2] Kurniawan, R. & B. Yuniarto. 2016. *Analisis Regresi Dasar dan Penerapannya dengan R*. Jakarta : Kencana.Kusuma, W., D. Komalasari, & M. Hadijati. 2013. Model Regresi *Zero Inflated Poisson* pada Data Overdispersion. *Jurnal Matematika*. No.2, Vol.3, 71-85.

[3] Hilbe, J.M. 2011. *Negative Binomial Regression (2th ed.)*. New York : Cambridge University Press.

[4] Nurmaleni dan N.R. Rahayu. 2017. Faktor-Faktor yang Mempengaruhi Angka Penderita Gizi Buruk pada Balita di Papua Tahun 2015 dengan Metode Regresi *Zero Inflated Poisson* (ZIP). *Jurnal LOG!K@*. No.1, 1-14.

[5] Myers, R.H. 1990. *Classical and Modern Regression with Appilications. 2nd ed*. Boston: PW-KENT Publishing Company Boston.

[6] Priyatno, D. 2014. *SPSS Pengolahan Data Terpraktis*. Yogyakarta : Andi Offset.

[7] Kurniawan, R. & B. Yuniarto. 2016. *Analisis Regresi Dasar dan Penerapannya dengan R*. Jakarta : Kencana.

[8] Lambert, D. 1992. *Zero-Inflated Poisson Regression, with an Application to Defects in Manufacturing*. *Technometrics*, No.1, Vol.34, 1-14.

[9] Jansakul, N. & J.P. Hinde. 2002. *Score Tests for Zero-Inflated Poisson Models*. *Computational Statistics & Data Analysis*. 40, 75-96.

[10] Garay, A.M., E.M. Hashimoto, E.M.M. Ortega., & V.H. Lachos. 2011. *On Estimation and Influence Diagnostics for Zero-Inflated Negative Binomial Regression Models*. *Computational Statistics and Data Analysis*. 55, 1304-1318.

[11] Hilbe, J.M. 2011. *Negative Binomial Regression (2th ed.)*. New York : Cambridge University Press.

[12] Myers, R.H., D.C. Montgomery, G.G. Vining, & T.J. Robinson. 2010. *Generalized Linear Models with Application in Engineering and The Sciences (2th ed.)*. New Jersey : John Wiley and Sons.