

**ZERO-INFLATED POISSON REGRESSION TO DETERMINE
THE FACTORS THAT INFLUENCE
THE MATERNAL MORTALITY RATE IN BANDUNG 2016**

Sri Rahayu

Faculty of Mathematics and Science, Mathematics Departement
Universitas Padjadjaran
sri rahayu14.45@gmail.com

Nurul Gusriani

Faculty of Mathematics and Science, Mathematics Departement
Universitas Padjadjaran

Iin Irianingsih

Faculty of Mathematics and Science, Mathematics Departement
Universitas Padjadjaran

ABSTRACT. *In Indonesia, especially Bandung, there are still many cases of maternal deaths during pregnancy, birth, and postpartum which can increase the Maternal Mortality Rate (MMR). Cases of maternal deaths are mostly due to exsanguination, infection, anemia, and the other causes that related to pregnancy. To reduce the MMR of Bandung, we can analyze to determine the factors that influence the MMR in order to maximize the maternal health care programs so as to prevent the possibility of death. The analysis is using Zero-Inflated Poisson (ZIP) regression because maternal mortality data is the result of counting that allows overdispersion due to excess zeros. Regression parameter estimation use expectation maximization algorithm followed by the Newton-Raphson iteration method. The analysis result showed that of the five suspected factors to affect the MMR -such as the first tetanus toxoid immunization (TT1), the provision of 90 Fe tablets (Fe3), postpartum care, pregnancy complications, and the first antenatal care (K1)- only TT1 and postpartum care are significantly affect the MMR.*

Keywords: *maternal mortality rate, overdispersion, excess zeros, expectation maximization, Newton-Raphson, Zero-Inflated Poisson.*

ABSTRAK. *Di Indonesia, khususnya kota Bandung, masih banyak terjadi kasus kematian ibu pada masa kehamilan, persalinan, dan nifas yang dapat meningkatkan Angka Kematian Ibu (AKI). Kematian ibu banyak disebabkan karena pendarahan, infeksi, anemia, dan penyebab lainnya yang berkaitan dengan kehamilan. Salah satu usaha untuk menurunkan AKI di kota Bandung yaitu dengan melakukan analisis guna mengetahui faktor-faktor yang memengaruhi AKI dengan tujuan untuk memaksimalkan program pelayanan kesehatan ibu, sehingga dapat mencegah kemungkinan terjadinya kematian. Analisis dilakukan menggunakan regresi *Zero Inflated Poisson* (ZIP). Hal ini disebabkan data kematian ibu merupakan data hasil *counting* yang memungkinkan terjadinya *overdispersion* karena *excess zeros*. Estimasi parameter regresi dilakukan*

menggunakan algoritma *expectation maximization* yang dilanjutkan dengan metode iterasi *Newton-Raphson*. Hasil analisis menunjukkan bahwa dari lima faktor yang diduga memengaruhi AKI yaitu faktor imunisasi *tetanus toxoid* yang pertama (TT1), konsumsi tablet Fe sebanyak 90 tablet (Fe3), pelayanan ibu nifas, komplikasi kehamilan, dan kunjungan kehamilan yang pertama (K1), hanya dua faktor yang secara signifikan memengaruhi AKI yaitu faktor TT1 dan pelayanan ibu nifas.

Kata kunci: angka kematian ibu, *overdispersion*, *excess zeros*, *expectation maximization*, *Newton-Raphson*, *Zero Inflated Poisson*.

1. INTRODUCTION

The Maternal Mortality Rate (MMR) is the number of maternal deaths due to pregnancy, childbirth, and within 42 days postpartum every 100,000 live births. The MMR is a measurement of the woman's death risk that related to pregnancy, as well as being an indicator of the quality of health care and health status of people in a region. Other than that, the MMR useful to describe the level awareness of healthy living behavior, nutritional status, maternal health, and level of health service. The MMR of Bandung in 2016 reached 62 every 100,000 life of birth, with 27 cases of maternal deaths (Dinas Kesehatan Kota Bandung).

In this paper, analysis of the factors that influence MMR is done through the data of number of maternal deaths in Bandung 2016 obtained from the official website of Bandung's health service. The data of number of maternal deaths in Bandung 2016 shown that the data are counting data that classified into discrete data. The results of this analysis are supposed to help the government programs in resolving the MMR problems that are concerned to increase.

Analysis for discrete data can use Poisson regression which assumes that the average and the variance of the dependent variable with Poisson distribution has the same value (equidispersion) (Winkelmann, 2003). The data about death is a rare event data that contains many zero values (excess zeros) which has important meaning for the research. This can cause one of the assumption violations called overdispersion. Data which has overdispersion due to excess zeros can be analyzed by Zero-inflated Poisson (ZIP) regression.

ZIP regression model formed to determine the factors that influence the MMR in Bandung 2016. Regression parameter estimation in this paper using the

Expectation Maximization (EM) algorithm and then followed by the Newton-Raphson iteration method.

2. METHODS

2.1 Poisson Regression

Poisson regression assumes that the dependent variable (Y) which has a Poisson distribution with the following probability functions (Kutner *et al*, 2005):

$$P(Y_i = y_i) = \frac{e^{-\mu_i} \mu_i^{y_i}}{y_i!}, \quad y_i = 0, 1, 2, \dots \quad (1)$$

$E(Y_i) = \mu_i$ and $Var(Y_i) = \mu_i$ are respectively the average value and the variance value of Poisson distributions. The estimated model for Poisson regression is formed using a generalized linear model with an estimate of the regression parameter β is given by $\hat{\beta}$, therefore the following model form is obtained by (Jong-Min & Sunghae, 2016):

$$\begin{aligned} \ln(\hat{\mu}_i) &= \hat{\beta}_0 + \hat{\beta}_1 x_{i_1} + \hat{\beta}_2 x_{i_2} + \dots + \hat{\beta}_m x_{i_m} = \mathbf{x}_i^T \hat{\boldsymbol{\beta}} \\ \hat{y} &= \hat{\mu}_i = e^{(\mathbf{x}_i^T \hat{\boldsymbol{\beta}})} \end{aligned} \quad (2)$$

2.2 Parameters Estimation of Poisson Regression

Poisson regression parameter estimation using the Maximum Likelihood Estimation (MLE) method, followed by the Newton-Raphson iteration with these steps:

1. Forming the Poisson regression likelihood function using the equation (1) then transformed it into a log-likelihood function (Winkelmann, 2003):

$$\ln[L(\hat{\beta})] = \sum_{i=1}^n [-e^{\mathbf{x}_i^T \hat{\boldsymbol{\beta}}} + y_i \mathbf{x}_i^T \hat{\boldsymbol{\beta}} - \ln(y_i!)] \quad (3)$$

2. Determining the value of $\hat{\boldsymbol{\beta}}^{(0)}$.
3. Multiplying \mathbf{X} by $\hat{\boldsymbol{\beta}}^{(k)}$ to form gradient vector which contains the first partial derivative value of equation (3) againsts $\hat{\boldsymbol{\beta}}$. Gradient vector can be obtained with this formula:

$$\mathbf{g}(\hat{\boldsymbol{\beta}}^{(k)}) = \mathbf{X}^T \mathbf{p} + \mathbf{X}^T \mathbf{y} \quad (4)$$

where \mathbf{X}^T is the transpose matrix of \mathbf{X} , \mathbf{y} is the gradient vector of dependent variable, and \mathbf{p} is a vector which contains $p_i = -\exp(\mathbf{x}_i^T \hat{\boldsymbol{\beta}}^{(k)})$ with $i = 1, 2, \dots, n$.

4. Forming the inverse of the Hessian matrix $[\mathbf{H}(\hat{\boldsymbol{\beta}}^{(k)})]^{-1}$ which contains the second partial derivative value of equation (3) against $\hat{\boldsymbol{\beta}}$ or by calculating the following equation:

$$[\mathbf{H}(\hat{\boldsymbol{\beta}}^{(k)})]^{-1} = (\mathbf{X}^T \mathbf{Q} \mathbf{X})^{-1} \quad (5)$$

where $\mathbf{Q} = \text{diag}(\mathbf{p})$.

5. Calculating $\hat{\boldsymbol{\beta}}^{(k+1)}$ by the following equation (Winkelmann, 2003):

$$\hat{\boldsymbol{\beta}}^{(k+1)} = \hat{\boldsymbol{\beta}}^{(k)} - [\mathbf{H}(\hat{\boldsymbol{\beta}}^{(k)})]^{-1} \mathbf{g}(\hat{\boldsymbol{\beta}}^{(k)}) \quad (6)$$

6. Iterating step 1 to 5 until $\|\hat{\boldsymbol{\beta}}^{(k+1)} - \hat{\boldsymbol{\beta}}^{(k)}\| \leq \tau$ are fulfilled, where τ is a very small number, in most cases $\tau = 10^{-5}$ (Winkelmann, 2003).

2.3 Score Test

Excess zeros is a situation of many zeros in a research data. Excess zeros can cause violation of equidispersion assumption, for example condition where the value of the dependent variable variance exceeds the average value or called overdispersion (McCullagh & Nelder, 1989). Overdispersion in excess zeros data can be detected using a score test with the following test statistics (Rahayu, 2014):

$$s_\omega = \frac{(n_0 - e^{\hat{\mu}_0})^2}{ne^{\hat{\mu}_0}(1 - e^{\hat{\mu}_0}) - n(e^{\hat{\mu}_0})^2} \quad (7)$$

The test hypothesis used is $H_0: \omega = 0$ against the alternative $H_1: \omega \neq 0$. If the value of $|s_\omega| > \chi_{(\alpha;1)}^2$ then reject H_0 , where the value of $\chi_{(\alpha;1)}^2$ is obtained from the chi-square table. It means that excess zeros data contains overdispersion. So, Poisson regression cannot be used to analyze and to model the data.

2.4 Zero-Inflated Poisson Regression

Zero-Inflated Poisson (ZIP) regression assume the dependent variable (Y) has a Poisson distribution with the probability density function as follows (Jong-Min & Sunghae, 2016):

$$P(Y_i = y_i) = \begin{cases} \omega_i + (1 - \omega_i)e^{-\mu_{iZIP}}, & y_i = 0 \\ (1 - \omega_i) \frac{e^{-\mu_{iZIP}} \mu_{iZIP}^{y_i}}{y_i!}, & y_i > 0 \end{cases} \quad (8)$$

$E(Y_i) = (1 - \omega_i)\mu_{iZIP}$ and $Var(Y_i) = (1 - \omega_i)(\mu_{iZIP} + \omega_i\mu_{iZIP}^2)$ are respectively the average value and the variance value. ZIP regression has two model which formed based on link function in the Generalized Linear Model (GLM). The Poisson model (μ_{iZIP}) has a regression parameter β^* with an estimated value $\widehat{\beta}^*$ and the independent variable matrix \mathbf{X} has the following estimation model form (Mouatassim & Ezzahid, 2012):

$$\hat{\mu}_{iZIP} = e^{(\mathbf{x}_i^T \widehat{\beta}^*)} \quad (9)$$

The logistic model (ω_i) has a regression parameter γ with an estimated value $\hat{\gamma}$ and the independent variable matrix \mathbf{X}^* . By using the logit link function, the estimation of the logistic model is (Mouatassim & Ezzahid, 2012):

$$\text{logit}(\widehat{\omega}_i) = \ln\left(\frac{\widehat{\omega}_i}{1 - \widehat{\omega}_i}\right) = \mathbf{x}_i^{*T} \widehat{\gamma}$$

so,

$$\widehat{\omega}_i = \frac{e^{(\mathbf{x}_i^{*T} \widehat{\gamma})}}{1 + e^{(\mathbf{x}_i^{*T} \widehat{\gamma})}} \quad (10)$$

and

$$(1 - \widehat{\omega}_i) = \frac{1}{1 + e^{(\mathbf{x}_i^{*T} \widehat{\gamma})}} \quad (11)$$

The independent variables matrix of Poisson model can be different from logistic model with different number of regression parameters.

2.5 Parameter Estimation of Zero-inflated Poisson Regression

The Expectation Maximization (EM) algorithm is used to estimate the Zero-inflated Poisson (ZIP) regression parameter with the following steps (Mouatassim & Ezzahid, 2012):

- Determining the distribution of indicator variable (Z)

Assume the dependent variable (Y) is related to the indicator variable (Z), such as:

$$Z_i = \begin{cases} 1, & y_i = 0 \\ 0, & y_i > 0 \end{cases} \quad (12)$$

so Z_i has Bernoulli distribution.

- Determining the joint probability distribution function of Y_i and Z_i :

$$f(y_i, z_i) = \left(\frac{1}{1 + e^{(\mathbf{x}_i^* T \hat{\gamma})}} \right)^{1-z_i} \left(\frac{e^{(\mathbf{x}_i^* T \hat{\gamma})}}{1 + e^{(\mathbf{x}_i^* T \hat{\gamma})}} \right)^{z_i} \left(\frac{e^{-e^{(\mathbf{x}_i^* T \hat{\beta}^*)}} e^{y_i (\mathbf{x}_i^* T \hat{\beta}^*)}}{y_i!} \right)^{1-z_i} \quad (13)$$

- Forming the likelihood function of equation (13) and then transforming it into log-likelihood form, therefore obtained:

$$\ln[L(\hat{\beta}^*, \hat{\gamma})] = \ln[L(\hat{\beta}^*)] + \ln[L(\hat{\gamma})] - \sum_{i=1}^n \{(1 - z_i) \ln[y_i!]\}$$

where

$$\ln[L(\hat{\beta}^*)] = \sum_{i=1}^n \{(1 - z_i)(-e^{(\mathbf{x}_i^* T \hat{\beta}^*)} + y_i \mathbf{x}_i^* T \hat{\beta}^*)\} \quad (14)$$

and

$$\ln[L(\hat{\gamma})] = \sum_{i=1}^n \left\{ z_i \mathbf{x}_i^* T \hat{\gamma} - \ln \left[1 + e^{(\mathbf{x}_i^* T \hat{\gamma})} \right] \right\} \quad (15)$$

- The expectation stage aims to find conditional expectations of indicator variables ($E(Z_i|Y_i)$) which will be used to determine the value of $z_i^{(k)}$:

$$z_i^{(k)} = \begin{cases} \frac{1}{1 + e^{(-\mathbf{x}_i^* T \hat{\gamma}^{(k)} - e^{(\mathbf{x}_i^* T \hat{\beta}^{*(k)})})}}, & y_i = 0 \\ 0, & y_i > 0 \end{cases} \quad (16)$$

- The maximization stage aims to find the convergent values of $\widehat{\boldsymbol{\beta}}^*$ and $\widehat{\boldsymbol{\gamma}}$ from the log-likelihood function in equations (14) and (15) using the Newton-Raphson iteration method.

Newton-Raphson iteration steps to estimate ZIP regression parameters are as follows:

1. Determining the value of $\widehat{\boldsymbol{\beta}}^{*(0)}$ and $\widehat{\boldsymbol{\gamma}}^{(0)}$.
2. Forming a **z**ik vector with an entry $z_i^{(k)}$ where $i = 1, 2, \dots, n$.
3. Multiplying the matrix \mathbf{X} by $\widehat{\boldsymbol{\beta}}^{*(k)}$ and the matrix \mathbf{X}^* by $\widehat{\boldsymbol{\gamma}}^{(k)}$ to form the gradient vector as follows:
 - Gradient vector for Poisson model containing the first partial derivative value of equation (14) against $\widehat{\boldsymbol{\beta}}^*$.

$$\mathbf{g}(\widehat{\boldsymbol{\beta}}^{*(k)}) = \mathbf{X}^T(\mathbf{M}\mathbf{u}) + \mathbf{X}^T(\mathbf{M}\mathbf{y}) \quad (17)$$

where \mathbf{M} is the diagonal matrix with the entry $m_{i,i} = 1 - z_i^{(k)}$ and \mathbf{u} is the vector with the entry $u_i = -\exp(\mathbf{x}_i^T \widehat{\boldsymbol{\beta}}^{*(k)})$.

- Gradient vector for logistik model containing the first partial derivative value of equation (15) against $\widehat{\boldsymbol{\gamma}}$.

$$\mathbf{g}(\widehat{\boldsymbol{\gamma}}^{(k)}) = \mathbf{X}^{*T} \mathbf{z} \mathbf{i} \mathbf{k} - \mathbf{X}^{*T} (\mathbf{R}^{-1} \mathbf{v}) \quad (18)$$

where \mathbf{R}^{-1} is the diagonal matrix with the entry $r_{i,i} = \frac{1}{1 + \exp(\mathbf{x}_i^{*T} \widehat{\boldsymbol{\gamma}})}$ and

\mathbf{v} is the vector with the entry $v_i = \exp(\mathbf{x}_i^{*T} \widehat{\boldsymbol{\gamma}})$.

4. Forming the inverse of the Hessian matrix as follows:
 - $[\mathbf{H}(\widehat{\boldsymbol{\beta}}^{*(k)})]^{-1}$ containing the second partial derivative value of equation (14) against $\widehat{\boldsymbol{\beta}}^*$ which is inverted.

$$[\mathbf{H}(\widehat{\boldsymbol{\beta}}^{*(k)})]^{-1} = (\mathbf{X}^T \mathbf{S} \mathbf{X})^{-1} \quad (19)$$

where $\mathbf{S} = \mathbf{M}[\text{diag}(\mathbf{u})]$.

- $[\mathbf{H}(\widehat{\boldsymbol{\gamma}}^{(k)})]^{-1}$ containing the second partial derivative value of equation (15) against $\widehat{\boldsymbol{\gamma}}$ which is inverted.

$$[\mathbf{H}(\hat{\boldsymbol{\gamma}}^{(k)})]^{-1} = (\mathbf{X}^{*T} \mathbf{T} \mathbf{X}^*)^{-1} \quad (20)$$

where $\mathbf{T} = (\mathbf{R}^{-1})^2 \text{diag}(-\mathbf{v})$.

5. Calculating $\hat{\boldsymbol{\beta}}^{*(k+1)}$ and $\hat{\boldsymbol{\gamma}}^{(k+1)}$ using the following equation [9]:

$$\hat{\boldsymbol{\beta}}^{*(k+1)} = \hat{\boldsymbol{\beta}}^{*(k)} - [\mathbf{H}(\hat{\boldsymbol{\beta}}^{*(k)})]^{-1} \mathbf{g}(\hat{\boldsymbol{\beta}}^{*(k)}) \quad (21)$$

and

$$\hat{\boldsymbol{\gamma}}^{(k+1)} = \hat{\boldsymbol{\gamma}}^{(k)} - [\mathbf{H}(\hat{\boldsymbol{\gamma}}^{(k)})]^{-1} \mathbf{g}(\hat{\boldsymbol{\gamma}}^{(k)}) \quad (22)$$

6. Iterating until it $\|\hat{\boldsymbol{\beta}}^{*(k+1)} - \hat{\boldsymbol{\beta}}^{*(k)}\| \leq \tau$ and $\|\hat{\boldsymbol{\gamma}}^{(k+1)} - \hat{\boldsymbol{\gamma}}^{(k)}\| \leq \tau$ are fulfilled, where τ is a very small number, in most case $\tau = 10^{-5}$.

2.6 Likelihood Ratio Test

The method used to test the overall of ZIP regression parameter is the Likelihood Ratio Test (LRT) with the following test statistics (Agresti, 2002; Hosmer & Lemeshow, 2000):

$$\begin{aligned} G_{ZIP} = & 2n \ln(1 + e^{\hat{\gamma}_0}) \\ & - 2 \sum_{i=1}^n \left[\ln(1 + e^{\mathbf{x}_i^{*T} \hat{\boldsymbol{\gamma}}}) + z_i^{(k)} \hat{\gamma}_0 - z_i^{(k)} \mathbf{x}_i^{*T} \hat{\boldsymbol{\gamma}} \right. \\ & - (1 - z_i^{(k)}) e^{\hat{\beta}_0^*} + (1 - z_i^{(k)}) e^{\mathbf{x}_i^{*T} \hat{\boldsymbol{\beta}}^*} + (1 - z_i^{(k)}) y_i \hat{\beta}_0^* \\ & \left. - (1 - z_i^{(k)}) y_i \mathbf{x}_i^{*T} \hat{\boldsymbol{\beta}}^* \right] \end{aligned} \quad (23)$$

Test hypothesis used is $H_0: \beta_1^* = \beta_2^* = \dots = \beta_m^* = \gamma_1 = \gamma_2 = \dots = \gamma_r = 0$ against alternative H_1 : there is at least one $\beta_j^* \neq 0$ or $\gamma_l \neq 0$ ($j = 1, 2, \dots, m$ and $l = 1, 2, \dots, r$). If the value of $G_{ZIP} > \chi_{(\alpha, db)}^2$ (db : the number of independent variables in the ZIP regression model include Poisson model and logistic model) then reject H_0 , where the value of $\chi_{(\alpha, db)}^2$ obtained from chi-square table. If H_0 is accepted then the data cannot be modeled using ZIP regression.

2.7 Wald Test

Wald test method is used to partially test the parameters of the ZIP model and has the following testing procedure:

- For Poisson model

Hypothesis: $H_0: \beta_j^* = 0$ against alternative $H_1: \beta_j^* \neq 0$, where $j = 1, 2, \dots, m$.

Test statistics (Winkelmann, 2003):

$$W_{\beta_j^*} = \frac{\widehat{\beta}_j^{*2}}{Var(\widehat{\beta}_j^*)} \quad (24)$$

The value of $Var(\widehat{\beta}_j^*)$ is obtained from the main diagonal entry (j, j) of the inverse Hessian matrix in equation (19). Test criteria: reject H_0 if test statistics fulfill $W_{\beta_j^*} > \chi_{(\alpha; 1)}^2$. If H_0 is accepted then the independent variable of $\widehat{\beta}_j^*$ must be removed from the model.

- For logistic model

Hypothesis: $H_0: \gamma_l = 0$ against alternative $H_1: \gamma_l \neq 0$, where $l = 1, 2, \dots, r$.

Test statistics (Winkelmann, 2003):

$$W_{\gamma_l} = \frac{\widehat{\gamma}_l^2}{Var(\widehat{\gamma}_l)} \quad (25)$$

The value of $Var(\widehat{\gamma}_l)$ is obtained from the main diagonal entry (l, l) of the inverse Hessian matrix in equation (20). Test criteria: reject H_0 if test statistics fulfill $W_{\gamma_l} > \chi_{(\alpha; 1)}^2$. If H_0 is accepted then the independent variable of $\widehat{\gamma}_l$ must be removed from the model.

3. RESULT AND DISCUSSION

The research is performed by taking the dependent variable (Y), which is the number of maternal death cases in Bandung 2016 that is 30 observations and the independent variable (X), which is the suspected factors to affect MMR, including TT1 or the first tetanus toxoid immunization (X_1), Fe3 or provision 90 Fe tablets for pregnant women (X_2), postpartum care (X_3), pregnancy complication (X_4), and K1 or first antenatal care (X_5).

3.1 Result of Parameter Estimation and Test Score for Poisson Regression Model

Estimated parameter model of Poisson regression that convergent using Newton-Raphson is obtained in the 6th iteration where $\tau = 10^{-5}$ and $\widehat{\boldsymbol{\beta}}^{(0)} = \mathbf{0} = [0 \ 0 \ \dots \ 0]^T$, so that the Poisson regression model is as follows:

$$\hat{\mu} = \exp(-0,9968 - 0,9128X_1 - 3,7063X_2 - 5,1820X_3 - 1,2523X_4 + 5,5745X_5) \quad (26)$$

Detection of overdispersion in excess zeros data use the score test method that calculated from equation (7). Based on the hypothesis $H_0: \omega = 0$ and $H_1: \omega \neq 0$, the value of the test statistic is $|s_\omega| = 9,8911 > \chi_{(0,05;1)}^2 = 3,84$ so H_0 is rejected at a significance level of 5%. The result shows that the dependent variable in the data has overdispersion caused by excess zeros. Poisson regression cannot be used to analyze or model the data.

3.2 Analysis Result Using the Zero-inflated Poisson Regression

Estimated parameter of Zero-Inflated Poisson (ZIP) regression model which is convergent using the Newton-Raphson iteration method is obtained in the 51st iteration where $\tau = 10^{-5}$, $\widehat{\boldsymbol{\beta}}^{*(0)} = \widehat{\boldsymbol{\gamma}} = \mathbf{0} = [0 \ 0 \ \dots \ 0]^T$, the matrix of \mathbf{X} equals to the matrix of \mathbf{X}^* which include X_1, X_2, X_3, X_4 , and X_5 , so the Poisson regression model is as follows:

Table 1. Results of parameter estimation and significance test of ZIP regression

Regression Parameters	Estimated Value	W	Regression Parameters	Estimated Value	W
$\widehat{\beta}_0^*$	8.3560		$\widehat{\gamma}_0$	12.6506	
$\widehat{\beta}_1^*$	0.6596	0.5140	$\widehat{\gamma}_1$	9.1412	4.8126
$\widehat{\beta}_2^*$	-6.3461	1.9459	$\widehat{\gamma}_2$	5.2294	0.0853
$\widehat{\beta}_3^*$	-16.8231	6.3957	$\widehat{\gamma}_3$	-44.6521	3.8128
$\widehat{\beta}_4^*$	-1.3969	1.1778	$\widehat{\gamma}_4$	-3.5313	0.8008
$\widehat{\beta}_5^*$	-0.7801	0.0243	$\widehat{\gamma}_5$	-20.8007	1.4002

The overall significance test of the ZIP parameters regression model uses the Likelihood Ratio Test (LRT) with the test statistic values calculated using equation (23). Based on the hypothesis $H_0 : \beta_1^* = \beta_2^* = \beta_3^* = \beta_4^* = \beta_5^* = \gamma_1 =$

$\gamma_2 = \gamma_3 = \gamma_4 = \gamma_5 = 0$ and H_1 : there is at least one " \neq " sign, the test statistic value is $G_{ZIP} = 30.7313 > \chi^2_{(0.05;10)} = 18.31$, which rejects H_0 at a significance level of 5%. So, there is at least one suspected factor that affect the MMR which contributes to change the number of maternal deaths.

Then partial significance test of the ZIP parameters regression model use the wald test with test statistic values calculated using equation (24) for the Poisson model and equation (25) for the logistic model. Based on Table 1, the results show that in the Poisson model, significant independent variable is only X_3 or postpartum maternal service factor, whereas in the logistic model significant independent variable is only X_1 or TT1 factor. Variables that do not contribute significantly must be removed from each model and parameters with the new independent variable matrix are re-estimated. Matrix of \mathbf{X} includes X_3 and matrix of \mathbf{X}^* includes X_1 , with the following results:

Table 2. Results of re-estimation of parameter values and significant test of the ZIP regression

Regression Parameters	Estimated Value	G_{ZIP}	W	Wald Test Criteria
$\hat{\beta}_0^*$	0.9742	12.2805 (rejects H_0)		
$\hat{\beta}_3^*$	-7.6225		3.9306	rejects H_0
$\hat{\gamma}_0$	-3.8031			
$\hat{\gamma}_1$	4.8529		4.9813	rejects H_0

Based on Table 2 the Poisson model on ZIP regression obtained as follows:

$$\hat{\mu}_{ZIP} = \exp(0.9742 - 7.6225X_3)$$

this means that if the average number of postpartum mothers who get health services (X_3) increases by one unit, then the number of maternal deaths will decrease by $\exp(-7.6225)$ or by 0.0005 times. Meanwhile the obtained logistic model on ZIP regression is as follows:

$$\hat{\omega} = \frac{\exp(-3.8031 + 4.8529X_1)}{1 + \exp(-3.8031 + 4.8529X_1)}$$

this means that if the number of pregnant women who get TT1 immunization (X_1) increases by one unit then the probability maternal deaths not happening will

increase by $\exp(4.8529)$ or by 128.1114 times. The estimated value of the dependent variable, $E(Y) = \hat{y}$ can be calculated using the following equation:

$$\hat{y} = (1 - \hat{\omega})\hat{\mu}_{ZIP} = \frac{\exp(0.9742 - 7.6225X_3)}{1 + \exp(-3.8031 + 4.8529X_1)}$$

4. CONCLUSION

After analyzing the data of maternal deaths in Bandung 2016 using Zero-Inflated Poisson (ZIP) regression it can be concluded that the factors that significantly affect the MMR in Bandung 2016 are the first tetanus toxoid immunization (TT1) and postpartum maternal care. The government of Bandung is expected to pay more attention to TT1 immunization for pregnant women which is useful to prevent maternal deaths due to infection. In addition, the postpartum care program needs to be improved and make sure that all postpartum women get health care.

In this paper overdispersion was caused only because of excess zeros. For the next research is expected to use a method that can handle all cause of overdispersion.

REFERENCES

- Agresti, A., *Categorical Data Analysis Second Edition*, John Wiley & Sons Inc., New York (USA), 2002.
- Dinas Kesehatan Kota Bandung, *Profil Kesehatan Kota Bandung Tahun 2016.*, <https://124.81.106.202/dashboard.php?page=profil dinas>, accessed on April 6, 2018.
- Hosmer, D. J. dan Lemeshow, S., *Applied Logistic Regression Second Edition*, John Wiley & Sons Inc., New York (USA), 2000.
- Jong-Min, K. dan Sunghae, J., *Zero-Inflated Poisson and Negative Binomial Regressions for Technology Analysis*, *International Journal of Software Engineering and Its Applications*, **10**(12) (2016), 431-448.
- Kutner, MH., *et al.*, *Applied Linear Statistical Models Fifth Edition*, McGraw-Hill, New York (USA), 2005.

- McCullagh, P., & Nelder, J.A. 1989. *Generalized Linear Models Second Edition*. London & New York: Chapman and Hall.
- Mouatassim, Y., & Ezzahid, EH. 2012. Poisson Regression and Zero-inflated Poisson Regression: Application to Private Health Insurance Data. *European Actuarial Journal ISSN: 2190-9733*. Germany: Springer.
- Rahayu, LP. 2014. *Kajian Overdispersi pada Regresi Poisson dan Zero-Inflated Poisson Untuk Beberapa Karakteristik Data*. Thesis. Institut Pertanian Bogor.
- Winkelmann, Rainer. 2003. *Econometric Analysis of Count Data Fourth Edition*. Germany: Springer.

