Bivariate zero inflated generalized Poisson regression model in the number of pregnant maternal mortality and the number of postpartum maternal mortality in the Central Java Province in 2017

Q Aini¹, Purhadi², Irhamah³

¹,²,³ Statistics Department, Institut Teknologi Sepuluh Nopember, Jalan Arif Rahman Hakim, Surabaya, 60111

Email: ¹qurotulaini.2707@gmail.com, ²purhadi@statistika.its.ac.id, ³irhamah@statistika.its.ac.id

Abstract. Excess zero is one of the problems in Generalized Poisson regression where the number of responses is contained zero exceeds 60 percent. One of the statistical methods have been developed is Zero Inflated Generalized Poisson Regression (ZIGPR). If there are two response variables, the appropriate regression analysis is Bivariate ZIGPR (BZIGPR). This study aims to determine the factors that influence the number of pregnant maternal mortality and the number of postpartum maternal mortality in 91 sub districts in Pekalongan Residency, Central Java Province 2017 through by the BZIGPR. The variables are percentage of K1 pregnancy examinations, percentage of K4 pregnancy examinations, percentage of deliveries assisted by health workers, percentage of pregnant women who received Fe3, percentage of TT2 + immunization in pregnant women, ratio of midwives per 100,000 population and percentage of obstetric complications management. The estimation of BZIGPR parameters using the Maximum Likelihood Estimation (MLE) method which results in a non-linear form so that, it is solved by the Berndt Hall-Hall Hausman (BHHH) iteration method. The result hypothesis testing using the Maximum Likelihood Ratio Test (MLRT) is reject null hypothesis. BZIGPR produces 2 regression models \( \hat{\ln \mu} \) and \( \logit \hat{p} \). The predictor variables that affect response \( Y_1 \) and \( Y_2 \) on model \( \ln \hat{\mu} \) are all of predictor variables. The predictor variables that affect response \( Y_1 \) on model \( \logit \hat{p} \) are all of predictor variables while the predictor variables that affect response \( Y_2 \) are all of predictor variables except percentage of pregnant TT2 + immunization women and percentage of midwives per 100,000 population.

1. Introduction

One of regression analysis whose response variable follows the Poisson distribution is Poisson regression which is used to analyze data counts that indicate the frequency of an event where the number of events is a non-negative integer. Poisson regression has a weakness where the mean distribution must be equal to the variance or equidispersion so that with the case of under/overdispersion the Poisson regression model cannot be used. To overcome this the Generalized Poisson Distribution (GPD) model which has one additional parameter that makes the distribution more flexible in dealing with under/overdispersion [1]. Another problem is excess zero where the number of observed zero responses exceeds the number of zero responses estimated by the model.
Zero-Inflated Poisson (ZIP) regression model developed to overcome excess zero in response variables [2]. The parameter estimation results in ZIP model will be biased when overdispersion occurs on data that has excess zero.

Zero Inflated Generalized Poisson (ZIGP) regression model is a combined model of the ZIP regression model with the Generalized Poisson regression model in which the ZIP regression model is a suitable model used in conditions where many responses are zero, while the GP model is suitable for use when violations of the equidispersion assumption occur in the Poisson distribution[3]. The development of univariate ZIGP into bivariate ZIGP (BZIGP) was developed and has two types of ZIGP, type I which has a common parameter of zero inflation and type II which has its own parameter of zero can be applied in a wider range [4]. One of the data that follows the Poisson distribution is the Maternal Mortality Rate (MMR) where the frequency of an event from the number of events has no upper limit and is a non-negative integer.

The suitable distribution for the data is the Poisson distribution [5]. Reducing MMR is one of the 17 global goals of the Sustainable Development Goals (SDGs) with the first target of reducing the ratio of maternal deaths to <70 deaths per 100,000 live births by 2030 [6]. Central Java Province had the highest maternal mortality rate in Indonesia in 2017 and Pekalongan Residency had more maternal mortality rate than other residencies in Central Java Province which was 23.38 percent [6]. Data on the number of pregnant and postpartum maternal mortality in the Pekalongan Residency in 2017 has a positive correlation of 0.27 and is suspected to be under / overdispersion and contains excess zero so that BZIGP regression can be applied to this data. The purpose of this study is to assess the parameter estimation using MLE and hypothesis testing using MLRT. The results of this study are expected to determine the factors that have a significant influence on the number of maternal mortality and the number of postpartum maternal mortality in Pekalongan residency in Central Java Province in 2017.

2. Literature

2.1. Bivariate zero-inflated generalized Poisson distribution

ZIGP distribution is a combination of the Bernoulli distribution and the BGP distribution [4]. This study uses ZIGP type II models with multiplicative factors. 
\( (Y, Y_2)^T \sim ZIGP \eta (p_1, p_2, \lambda_1, \lambda_2, \varphi_1, \varphi_2, \eta) \) the joint probability mass function will be derived as:

1. \( P(Y_i = 0, Y_2 = 0) = (p_1 + p_2 + (1 - p_1)e^{-\lambda_1} + (1 - p_2)e^{-\lambda_2} + (1 - p_1)(1 - p_2)e^{(\lambda_1 + \lambda_2)} G_i \)
2. \( P(Y_i = 0, Y_2 > 0) = (1 - p_2) \frac{\lambda_2^{y_2}}{y_2!} \exp[-\lambda_2(1 + \varphi_2 y_2)](p_1 + (1 - p_1) e^{-\lambda_1}) G_2 \)
3. \( P(Y_i > 0, Y_2 = 0) = (1 - p_1) \frac{\lambda_1^{y_1}}{y_1!} \exp[-\lambda_1(1 + \varphi_1 y_1)](p_2 + (1 - p_2) e^{-\lambda_2}) G_3 \)
4. \( P(Y_i > 0, Y_2 > 0) = (1 - p_1)(1 - p_2) \prod_{i=1}^{2} \frac{\lambda_i^{y_i}}{y_i!} \exp[-\lambda_i(1 + \varphi_i y_i)] G_4 \)

where \( G_i = 1 + \eta(1 - g_i)(1 - g_2), G_2 = [1 + \eta(1 - g_1)(e^{-\gamma_2} - g_2), G_2 = [1 + \eta (1 - g_2)(e^{-\gamma_1} - g_1)] \),
\( G_1 = [1 + \eta (e^{-\gamma_1} - g_1)(e^{-\gamma_2} - g_2)] \), \( g_i = E(e^{Y_i}) = \exp[\lambda_i(t_i - 1)] \) and \( \ln t_i - \varphi_i \lambda_i(t_i - 1) + 1 = 0; t = 1, 2 \)
\( \text{Mean } E(Y_i) = (1 - p_1) \frac{\lambda_1}{1 - \varphi_1 \lambda_1} \) and \( \text{Var}(Y_i) = (1 - p_1) \left[ \frac{\lambda_1}{(1 - \varphi_1 \lambda_1)} + p_1 \mu^2 \right] \)

2.2. Overdispersion

Overdispersion of Poisson regression occurs when the variance of the response variable is greater than mean. To detect overdispersion used variance test (VT). If variance test > 1 called overdispersion, if variance test < 1 called underdispersion and if variance test = 1 called equidispersion.
\[ VT = \sum_{i=1}^{n} \frac{(y_i - \hat{y})^2}{\hat{y}} \]

### 2.3. Bivariate zero-inflated generalized Poisson regression

Marginal mean from BZIGP distribution is \( \mu_i(x_i) = \frac{\lambda_i}{1 - \phi_i \lambda_i} \) or \( \lambda_i = \frac{\mu_i}{1 + \varphi_i \mu_i} \); \( i = 1, 2, ..., n; l = 1, 2 \)

Joint model \( \ln \mu_i = x_i^T \beta_i \) and logit \( p_{i0} = x_i^T \gamma_i \) [2] are:

| \( Y_1 \) | \( Y_2 \) |
|---|---|
| 0 | \( >0 \) | 0 | \( >0 \) |

\[ p_1 = \frac{e^{x_i^T \gamma_i}}{1 + e^{x_i^T \gamma_i}} \quad 1 - p_1 = \frac{1}{1 + e^{x_i^T \gamma_i}} \quad p_2 = \frac{e^{x_i^T \gamma_i}}{1 + e^{x_i^T \gamma_i}} \quad 1 - p_2 = \frac{1}{1 + e^{x_i^T \gamma_i}} \]

Where \( x_i = [x_{i1}, x_{i2}, \ldots, x_{i7}]^T \), \( \beta_i = [\beta_{i0}, \beta_{i1}, \beta_{i2} \ldots, \beta_{i7}]^T \), \( \gamma_i = [\gamma_{i0}, \gamma_{i1}, \gamma_{i2} \ldots, \gamma_{i7}]^T \) by substituting the above equation into equation (1) we get the probability density function of BZIGP regression as follows:

1. \( P(Y_1 = 0, Y_2 = 0) \)

\[ = \frac{e^{x_i^T \gamma_i}}{1 + e^{x_i^T \gamma_i}} \quad \frac{e^{x_i^T \gamma_i}}{1 + e^{x_i^T \gamma_i}} \quad \frac{1}{1 + e^{x_i^T \gamma_i}} \quad \frac{1}{1 + e^{x_i^T \gamma_i}} \quad \frac{1}{1 + e^{x_i^T \gamma_i}} \quad \frac{1}{1 + e^{x_i^T \gamma_i}} \quad \frac{1}{1 + e^{x_i^T \gamma_i}} \quad \frac{1}{1 + e^{x_i^T \gamma_i}} \]

2. \( P(Y_1 = 0, Y_2 > 0) \)

\[ = \frac{1}{1 + e^{x_i^T \gamma_i}} \quad GP(\gamma_2 | \beta_2, \varphi_2) \left( \frac{e^{x_i^T \gamma_i}}{1 + e^{x_i^T \gamma_i}} + \frac{1}{1 + e^{x_i^T \gamma_i}} \right) \quad \left( 1 + \eta \left( 1 - g_1 \right) \left( e^{y_2} - g_2 \right) \right) \]

3. \( P(Y_1 > 0, Y_2 = 0) \)

\[ = \frac{1}{1 + e^{x_i^T \gamma_i}} \quad GP(\gamma_1 | \beta_1, \varphi_1) \left( \frac{e^{x_i^T \gamma_i}}{1 + e^{x_i^T \gamma_i}} + \frac{1}{1 + e^{x_i^T \gamma_i}} \right) \quad \left( 1 + \eta \left( e^{y_1} - g_1 \right) \left( 1 - g_2 \right) \right) \]

4. \( P(Y_1 > 0, Y_2 > 0) \)

\[ = \frac{1}{1 + e^{x_i^T \gamma_i}} \quad \prod_{l=1}^{2} \left( \frac{e^{x_i^T \gamma_i}}{1 + e^{x_i^T \gamma_i}} \right)^{\gamma_{il}} \quad \left( 1 + \varphi_i \gamma_i \right) ^{\gamma_{il} - 1} \quad \left( \frac{1 + \varphi_i \gamma_i}{1 + \varphi_i e^{x_i^T \gamma_i}} \right) \quad \left( 1 + \eta \prod_{l=1}^{2} \left( e^{y_{il}} - g_{il} \right) \right) \]

The BZIGP regression models:

\[ \ln \mu_i = \hat{\beta}_{i0} + \hat{\beta}_{i1} x_{i1} + \hat{\beta}_{i2} x_{i2} + \hat{\beta}_{i3} x_{i3} + \hat{\beta}_{i4} x_{i4} + \hat{\beta}_{i5} x_{i5} + \hat{\beta}_{i6} x_{i6} + \hat{\beta}_{i7} x_{i7} \cdot i = 1, 2, \ldots, n \]

\[ \text{logit } \hat{p}_{i0} = \gamma_{i0} + \hat{\gamma}_{i1} x_{i1} + \hat{\gamma}_{i2} x_{i2} + \hat{\gamma}_{i3} x_{i3} + \hat{\gamma}_{i4} x_{i4} + \hat{\gamma}_{i5} x_{i5} + \hat{\gamma}_{i6} x_{i6} + \hat{\gamma}_{i7} x_{i7} \]

### 2.4. Estimation parameters BZIGPR model

Steps of parameter estimation used MLE:

1. Make the likelihood function

\[ (Y_{i1}, Y_{i2})^T \sim \text{ZIGP}_y(\theta) \text{ where } \theta = [\gamma_1, \gamma_2, \beta_1, \beta_2, \varphi_1, \varphi_2, \eta]^T \]
\[ A_i = f_i(y_{i1}, y_{i2}) = P(Y_{i1} = 0, Y_{i2} = 0), i = 1, 2, ..., n_i \]
\[ B_i = f_i(y_{i1}, y_{i2}) = P(Y_{i1} = 0, Y_{i2} > 0), i = 1, 2, ..., n_i \]
\[ C_i = f_i(y_{i1}, y_{i2}) = P(Y_{i1} > 0, Y_{i2} = 0), i = 1, 2, ..., n_i \]
\[ D_i = f_i(y_{i1}, y_{i2}) = P(Y_{i1} > 0, Y_{i2} > 0), i = 1, 2, ..., n_i \]

\[ L(\gamma_1, \gamma_2, \beta_1, \beta_2, \varphi_1, \varphi_2, \eta) = \prod_{i=1}^{n} (A_i)^{\lambda_i - c_i - d_i} (B_i)^{n_i} (C_i)^{c_i} (D_i)^{d_i} \]  

(3)

2. Based on (2) the log likelihood function for BZIGPR can be written as follows:

\[ l = \sum_{i=1}^{n} (1 - b_i - c_i - d_i) \ln A_i + \sum_{i=1}^{n} (b_i) \ln B_i + \sum_{i=1}^{n} (c_i) \ln C_i + \sum_{i=1}^{n} (d_i) \ln D_i \]  

(4)

\[ \ln A_i = \sum_{i=1}^{n_i} \ln \left( \frac{1}{1 + \exp(x_i^T \gamma_i)} + \frac{1}{1 + \exp(x_i^T \gamma_2)} \right) + \ln(A_i + A_2 + A_3) \]  

where

\[ A_1 = \exp(x_i^T \gamma_1) \left( \exp(x_i^T \gamma_2) + \frac{\exp(\gamma_i \beta_2)}{1 + \varphi_2 \exp(x_i^T \beta_2)} \right) \]
\[ A_2 = \exp(x_i^T \gamma_1) \exp(\gamma_i \beta_2) \]
\[ A_3 = \exp\left( \frac{\exp(x_i^T \beta_1)}{1 + \varphi_1 \exp(x_i^T \beta_1)} - \frac{\exp(x_i^T \beta_2)}{1 + \varphi_2 \exp(x_i^T \beta_2)} \right) \left( 1 + \eta(1 - g_1)(1 - g_2) \right) \]

\[ \ln B_i = \sum_{i=1}^{n_i} \ln \left( \frac{1}{1 + e^{x_i^T \gamma_i}} + y_{i2} \right) \ln \left( 1 + \varphi_2 e^{x_i^T \beta_2} \right) + \ln \left( 1 + \varphi_1 e^{x_i^T \beta_1} \right) - \ln y_{i2} \]  

where

\[ B_i = \left( \frac{-e^{x_i^T \beta_1} (1 + \varphi_1 y_{i1})}{1 + \varphi_1 e^{x_i^T \beta_1}} \right) + \ln \left( \frac{1}{1 + e^{x_i^T \gamma_i}} + \ln \left( e^{x_i^T \gamma_i} + \exp \left( \frac{-e^{x_i^T \beta_1}}{1 + \varphi_1 e^{x_i^T \beta_1}} \right) \right) \right) G_2 \]

\[ \ln C_i = \sum_{i=1}^{n_i} \ln \left( \frac{1}{1 + e^{x_i^T \gamma_2}} + y_{i1} \right) \ln \left( 1 + \varphi_1 e^{x_i^T \beta_1} \right) + \ln \left( 1 + \varphi_2 e^{x_i^T \beta_2} \right) - \ln y_{i1} \]  

where

\[ C_i = \left( \frac{-e^{x_i^T \beta_1} (1 + \varphi_1 y_{i1})}{1 + \varphi_1 e^{x_i^T \beta_1}} \right) + \ln \left( \frac{1}{1 + e^{x_i^T \gamma_2}} + \ln \left( e^{x_i^T \gamma_2} + \exp \left( \frac{-e^{x_i^T \beta_2}}{1 + \varphi_2 e^{x_i^T \beta_2}} \right) \right) \right) G_3 \]

\[ \ln D_i = \sum_{i=1}^{n_i} \ln \left( \frac{1}{1 + e^{x_i^T \gamma_2}} + \frac{1}{1 + e^{x_i^T \gamma_2}} + y_{i1} \right) \ln \left( 1 + \varphi_1 e^{x_i^T \beta_1} \right) + \ln \left( 1 + \varphi_2 e^{x_i^T \beta_2} \right) - \ln y_{i1} \]  

where

\[ D_i = \left( \frac{-e^{x_i^T \beta_1} (1 + \varphi_1 y_{i1})}{1 + \varphi_1 e^{x_i^T \beta_1}} \right) + \ln \left( 1 + \eta \left( e^{x_i^T \gamma_2} \right) \right) G_4 \]

3. Look for the first derivative of function \( l \) with respect to parameters \( \gamma_1, \gamma_2, \beta_1, \beta_2, \varphi_1, \varphi_2, \eta \) then matched with zero.

4. If MLE of parameters are not available in closed form then we adopt the BHHH algorithm to calculate the complete data MLE of parameters with steps [7]:

a. Determine the initial value \( \theta \) and \( m=0 \) with \( \varepsilon > 0 \) for the convergence tolerance limit.

\[ \theta = \left[ \beta_{1(0)}, \beta_{2(0)}, \gamma_{1(0)}, \gamma_{2(0)}, \varphi_1, \varphi_2, \eta \right]^T \]

where \( \beta_{1(0)}, \beta_{2(0)}, \gamma_{1(0)} \) dan \( \gamma_{2(0)} \) obtain from ZIGPR univariat estimation parameters.
b. Calculate gradient vector $g(\hat{\theta}_m) = \left[ \frac{\partial L(\bullet)}{\partial \hat{\beta}_1^{(0)}}, \frac{\partial L(\bullet)}{\partial \hat{\beta}_2^{(0)}}, \ldots, \frac{\partial L(\bullet)}{\partial \hat{\eta}} \right]^T$

de

where $L(\bullet) = \ln L(\gamma_1, \gamma_2, \beta_1, \beta_2, \phi_1, \phi_2, \eta)$.

c. Get Hessian matrix $H(\hat{\theta}_m) = -\sum_{i=1}^{n} g(\hat{\theta}_m) (\hat{\theta}_m)^T$

d. Substituting $\hat{\theta}_m$ into $g(\hat{\theta}_m)$ and Hessian $H(\hat{\theta}_m)$.

e. Do iteration starting form m=0 with function $\hat{\theta}_{m+1} = \hat{\theta}_m - H^{-1}(\hat{\theta}_m) g(\hat{\theta}_m)$, iteration will be stop if $\left\| \hat{\theta}_{m+1} - \hat{\theta}_m \right\| \leq \varepsilon$ where $\varepsilon$ is very small positive numbers close to 0.

f. Repeat step b and other with m=m+1.

2.5. Hypothesis testing of BZIGPR

Hypothesis testing of BZIGPR models used Maximum Likelihood Ratio Test (MLRT) with hypothesis:

1. Simultaneous testing of $\beta$ parameters

$H_0: \beta_1 = \beta_2 = \ldots = \beta_\nu = 0, l = 1,2$

$H_1: \text{minimal } \beta_\nu \neq 0, r = 1,2,\ldots,7$

Test statistics $G^2 = -2 \ln \left[ \frac{L(\hat{\Omega})}{L(\hat{\omega})} \right]$ (5)

where $L(\hat{\Omega}) = \text{parameter's set under population}, \quad \hat{\Omega} = \{\hat{\beta}_1, \hat{\beta}_2, \hat{\gamma}_1, \hat{\gamma}_2, \hat{\phi}_1, \hat{\phi}_2, \hat{\eta}\}$

$L(\hat{\omega}) = \text{parameter's set under } H_0, \quad \hat{\omega} = \{\hat{\beta}_{10}, \hat{\beta}_{20}, \hat{\gamma}_1, \hat{\gamma}_2, \hat{\phi}_1, \hat{\phi}_2, \hat{\eta}\}$.

2. Simultaneous testing of $\gamma$ parameters

$H_0: \gamma_{10} = \gamma_{11} = \ldots = \gamma_{17} = 0, l = 1,2$

$H_1: \text{minimal } \gamma_{\nu} \neq 0, r = 1,2,\ldots,7$

Statistics test $G^2 = -2 \ln \left[ \frac{L(\hat{\Omega})}{L(\hat{\omega})} \right]$ where

$L(\hat{\Omega}) = \text{parameter's set under population}, \quad \hat{\Omega} = \{\hat{\beta}_1, \hat{\beta}_2, \hat{\gamma}_1, \hat{\gamma}_2, \hat{\phi}_1, \hat{\phi}_2, \hat{\eta}\}$

$L(\hat{\omega}) = \text{parameter's set under } H_0, \quad \hat{\omega} = \{\hat{\beta}_{10}, \hat{\beta}_{20}, \hat{\gamma}_{10}, \hat{\gamma}_{20}, \hat{\phi}_1, \hat{\phi}_2, \hat{\eta}\}$.

The distribution approach of chi square distribution with degrees of freedom $v$ (number of parameters under population - number of parameters under $H_0$). If $D > \chi^2_{v, \alpha}$, so that reject $H_0$ then carry out a partial testing to determine which parameters affect the response with hypothesis:

$H_0: \beta_\nu = 0, l = 1,2; r = 1,2,\ldots,7$  \quad $H_0: \gamma_\nu = 0, l = 1,2; r = 1,2,\ldots,7$

$H_1: \beta_\nu \neq 0$  \quad $H_1: \gamma_\nu \neq 0$

$Z = \frac{\hat{\beta}_\nu}{se(\hat{\beta}_\nu)}$ reject $H_0$ if $|Z| > t_{\alpha/2}$  \quad $Z = \frac{\hat{\beta}_\nu}{se(\hat{\beta}_\nu)}$ reject $H_0$ if $|Z| > t_{\alpha/2}$
2.6. Correlation test and multicollinearity

Correlation is an indicator used in a linear relationship between two variables with hypothesis testing [8]:

H0: There are no relationship between \( Y_1 \) and \( Y_2 \)
H1: There are a relationship between \( Y_1 \) and \( Y_2 \)

\[
t = \frac{r_{Y_1Y_2}\sqrt{n-2}}{\sqrt{1-(r_{Y_1Y_2})^2}}, \quad -1 < r < 1
\]

where \( r_{Y_1Y_2} = \frac{\sum_{i=1}^{n}(y_{1i} - \bar{Y}_1)(y_{2i} - \bar{Y}_2)}{\sqrt{\sum_{i=1}^{n}(y_{1i} - \bar{Y}_1)^2 \sum_{i=1}^{n}(y_{2i} - \bar{Y}_2)^2}} \)

reject \( H_0 \) if \( |t| > t_{\alpha/2(n-2)} \)

While to detect multicollinearity used Variance Inflation Factor (VIF).

\[
VIF = \frac{1}{1-R_j^2}, \quad j = 1, 2, ..., k
\]

\( R_j^2 \) is determination coefficient between predictor variable. VIF>10 indicate that multicollinearity occur [9].

2.7. Maternal mortality

Maternal mortality is maternal mortality rate due to pregnancy and childbirth per 100,000 live births in a certain region and time. The maternal death is the result of all causes related to pregnancy or treatment and is not caused by accident, injury or other incidental reasons. These causes are included in the dimensions of maternal death which are formulated as 4Too 3Too late means too young (<20 years), too old (>35 years), too many children (having more than three children), too close the birth distance (<2 years), being late to reaching a health facility, late getting help / handling, and being late to recognize danger signs of pregnancy and childbirth [10].

3. Methods

3.1. Research variable

Data sourced from health office of Central Java Province in 2017 with research unit spread in 91 sub districts in Pekalongan Residency.

| Table 1. Research variables |
|-----------------------------|
| Notation | Definitions |
| **Response variables** | |
| \( Y_1 \) | Number of pregnant mortality in a sub district |
| \( Y_2 \) | Number of postpartum mortality in a sub district |
| **Predictor variables** | |
| \( X_1 \) | Percentage of K1 pregnancy examinations in a sub district |
| \( X_2 \) | Percentage of K4 pregnancy examinations in a sub district |
| \( X_3 \) | Percentage of deliveries assisted by health workers in a sub district |
| \( X_4 \) | Percentage of pregnant TT2 + immunization women in a sub district |
| \( X_5 \) | Percentage of pregnant women who received Fe3 a sub district |
| \( X_6 \) | Percentage of obstetric complications management a sub district |
| \( X_7 \) | Ratio of midwives per 100,000 population a sub district |

3.2. Data analysis steps

The data analysis steps are:

a. Make descriptive analysis of response and predictor variables
b. Detect Overdispersion using variance test (VT), underdispersion if VT < 1, overdispersion if VT > 1 and equidispersion if VT = 1.
c. Correlation testing of response variable using Pearson correlation test.
d. Detection of multicollinearity between predictor variables use VIF. VIF > 10 indicate multicollinearity.
e. Parameter estimation use MLE, if it is not produced close form so use Berndt Hall-Hall Hausman (BHHH) iteration.
f. Hypothesis testing use MLRT test.
g. Interpret BZIGPR model

4. Result

4.1. Descriptive statistics

Table 2. Descriptive statistics of variables

| Variable | Minimum | Maximum | Mean | Variance |
|----------|---------|---------|------|----------|
| \( Y_1 \) | 0.00 | 3.00 | 0.41 | 0.49 |
| \( Y_2 \) | 0.00 | 3.00 | 0.67 | 0.91 |
| \( X_1 \) | 46.60 | 100.00 | 97.89 | 42.63 |
| \( X_2 \) | 49.51 | 100.00 | 92.74 | 50.63 |
| \( X_3 \) | 79.59 | 100.00 | 97.75 | 16.11 |
| \( X_4 \) | 0.64 | 100.00 | 78.22 | 506.92 |
| \( X_5 \) | 61.17 | 100.00 | 93.18 | 50.80 |
| \( X_6 \) | 10.57 | 61.61 | 30.17 | 92.14 |
| \( X_7 \) | 18.14 | 100.00 | 48.15 | 246.45 |

Based on the above table, it can saw that mean from number of the pregnant maternal mortality is 0.41 and mean from postpartum maternal mortality is 0.67 with maximum case is 3.

4.2. Correlation and multicollinearity

Table 3. Correlation

| Variable | \( Y_1 \) | P-value |
|----------|---------|---------|
| \( Y_2 \) | 0.27 | 0.01 |

Checking the relationship between response variables is done before modelling [11]. Examination of correlation using the Pearson correlation test that produces a correlation value between the number of pregnant maternal mortality and number of postpartum maternal mortality is 0.27 and with a p-value of 0.01. it can be concluded between the number of pregnant maternal mortality and number of postpartum maternal mortality correlated and modelling can be continued

Table 4. VIF of predictor variables

| VIF | X_1 | X_2 | X_3 | X_4 | X_5 | X_6 | X_7 |
|------|-----|-----|-----|-----|-----|-----|-----|
| 2.78 | 2.77 | 1.78 | 1.12 | 1.57 | 1.33 | 1.32 |

The VIF value of all predictor variables is less than 10 so that there is no multicollinearity and all predictor variables can be modelled. The VIF value of all predictor variables is less than 10 so that there is no multicollinearity and all predictor variables can be modelled.

4.3. Overdispersion

Disperse parameter using VT and the result are \( VT_1 = 108.108 \) and \( VT_2 = 122.492 \) so it can be concluded that overdispersion occurred in the number of maternal mortality and in the number of postpartum maternal mortality at Pekalongan Residency in 2017.
4.4. Parameter estimation
Since the complete data MLEs of $\theta$ are not available in close form, we adopt the BHNN algorithm to calculate the complete data MLEs of $\theta$. The result of parameters estimation of $\theta$ are:

Table 5. Parameter estimation

| Par     | Pregnant maternal mortality estimation | Par     | Postpartum maternal mortality estimation |
|---------|----------------------------------------|---------|----------------------------------------|
| $\gamma_{10}$ | 0.0100 | $\beta_{10}$ | -10.4856 |
| $\gamma_{11}$ | 0.0143 | $\beta_{11}$ | 0.0317 |
| $\gamma_{12}$ | 0.0139 | $\beta_{12}$ | -0.0412 |
| $\gamma_{14}$ | 0.0124 | $\beta_{14}$ | 0.0097 |
| $\gamma_{15}$ | 0.0139 | $\beta_{15}$ | 0.0069 |
| $\gamma_{16}$ | 0.0120 | $\beta_{16}$ | 0.0440 |
| $\gamma_{17}$ | 0.0121 | $\beta_{17}$ | 0.0060 |
| $\varphi_1$  | 1.0640 |                    |         |
| $\varphi_2$  | 1.2108 |                    |         |
| $\eta$      | 0.2690 |                    |         |

4.5. Hypothesis testing
1. Simultaneous testing of $\beta$ parameters
   $H_0: \beta_{1} = \beta_{2} = \ldots = \beta_{23} = 0$
   $H_1$: minimal $\beta_{ol} \neq 0, l = 1, 2; r = 1, 2, \ldots, 7$
   Test statistics $G^2 = 2 \ln \left( L(\hat{\theta}) - L(\hat{\omega}) \right) = 3576.365, G^2 > \chi^2_{0.05;14} = 23.685 \Rightarrow$ reject $H_0$

2. Simultaneous testing of $\gamma$ parameters
   $H_0: \gamma_{11} = \gamma_{12} = \ldots = \gamma_{27} = 0$
   $H_1$: minimal $\gamma_{ol} \neq 0, l = 1, 2; r = 1, 2, \ldots, 7$
   Test statistics $G^2 = 2 \ln \left( L(\hat{\theta}) - L(\hat{\omega}) \right) = 3531.279, D > \chi^2_{0.05;14} = 23.685 \Rightarrow$ reject $H_0$
It can be concluded that at least one parameter $\beta$ and $\gamma$ affects the predictor variables so that to determine the influential parameters a partial test is performed. Partial testing of parameters $\beta$

   $H_0: \beta_{ol} = 0, l = 1, 2; r = 1, 2, \ldots, 7$
   $H_1$: $\beta_{ol} \neq 0$

Table 6. Parameter’s estimation of $\beta$

| Parameter | Pregnant maternal mortality estimation | Z-table | P-value | Parameter | Postpartum maternal mortality estimation | Z-table | P-value |
|-----------|---------------------------------------|---------|---------|-----------|----------------------------------------|---------|---------|
| $\beta_{10}$ | -10.4856 | -436572.26 | 0.0000 | $\beta_{20}$ | -30.5023 | -1629897.96 | 0.0000 |
| $\beta_{11}$ | 0.0317 | 26.0414 | 0.0000 | $\beta_{21}$ | 0.3514 | 415.7019 | 0.0000 |
| $\beta_{12}$ | -0.0412 | -39.6578 | 0.0000 | $\beta_{22}$ | 0.0173 | 25.7143 | 0.0000 |
| $\beta_{14}$ | 0.0097 | 5.2435 | 0.0000 | $\beta_{23}$ | -0.0375 | -43.0567 | 0.0000 |
| $\beta_{15}$ | 0.0069 | 5.9201 | 0.0000 | $\beta_{24}$ | 0.0089 | 3.1957 | 0.0014 |
| $\beta_{16}$ | 0.0440 | 78.8105 | 0.0000 | $\beta_{25}$ | -0.0268 | -47.4829 | 0.0000 |
| $\beta_{17}$ | 0.0060 | 5.3661 | 0.0000 | $\beta_{26}$ | 0.0254 | 34.7496 | 0.0000 |
| $\beta_{27}$ |                    |         |        | $\beta_{27}$ | -0.0159 | -19.7507 | 0.0000 |
Based on the results of partial test stated with a significance level of 0.05. the predictor variables that affects to the response the number of pregnant maternal mortality and the number of maternal mortality are all predictor variables.

2. Partial testing of parameters $\gamma$  
   
   \[ H_0 : \gamma = 0; r = 1, 2, \ldots, 7 \]  
   \[ H_1 : \gamma \neq 0 \]

| Table 7. Parameter’s estimation of parameters $\gamma$ |
|-----------------------------------------------|
| Parameter | Pregnant mortality | Postpartum mortality |
|-----------|--------------------|---------------------|
|           | Estimation | Z- table | P-value | Estimation | Z- table | P-value |
| $\gamma_{10}$ | 0.0100 | 683.1607 | 0.0000 | $\gamma_{20}$ | 0.0120 | 1495.414 | 0.0000 |
| $\gamma_{11}$ | 0.0143 | 11.1830 | 0.0000 | $\gamma_{21}$ | 0.0109 | 7.3192 | 0.0000 |
| $\gamma_{12}$ | 0.0139 | 13.4849 | 0.0000 | $\gamma_{22}$ | 0.0109 | 7.7084 | 0.0000 |
| $\gamma_{14}$ | 0.0124 | 1.9886 | 0.0467 | $\gamma_{23}$ | 0.0109 | 11.5491 | 0.0000 |
| $\gamma_{15}$ | 0.0139 | 12.1278 | 0.0000 | $\gamma_{24}$ | 0.0102 | 1.8438 | 0.0652 |
| $\gamma_{16}$ | 0.0120 | 14.7161 | 0.0000 | $\gamma_{25}$ | 0.0105 | 9.0523 | 0.0000 |
| $\gamma_{17}$ | 0.0121 | 8.8958 | 0.0000 | $\gamma_{26}$ | 0.0125 | 5.8631 | 0.0000 |

Based on that table with a significance level of 0.05. the predictor variables that affects to the response $Y_1$ are all predictor variables. percentage of pregnant TT2 + immunization women ($X_4$) and ratio of midwives per 100,000 population ($X_7$), while on the response $Y_2$ are percentage K1 pregnancy examinations ($X_1$), percentage of K4 pregnancy examinations ($X_2$), percentage of deliveries assisted by health workers ($X_3$), percentage of pregnant women who received Fe3 ($X_5$) and percentage of obstetric complications management ($X_6$).

4.5.1. Regression model of the number of pregnant maternal mortality for Poisson state  
   \[ \ln \hat{\mu}_i = -10.4856 + 0.0317x_{i1} - 0.0412x_{i2} + 0.0097x_{i4} + 0.0069x_{i5} + 0.0440x_{i6} + 0.0060x_{i7} \]  
   Every additional 1% the percentage of K1 pregnancy examination it will increase the number of pregnant maternal mortality by exp (0.0317) = 1.032 times from the previous number if the other variables are constant. Every additional 1% the percentage of K4 pregnancy examination it will reduce the number of pregnant maternal mortality by exp (-0.0412) = 0.9596 times from the previous number if the other variables are constant. Every additional 1% the percentage of pregnant TT2 + immunization, percentage of pregnant women who received Fe, percentage of obstetric complications management in pregnant women and ratio of midwives per 100,000 population it will increase the number of pregnant maternal mortality each by 1.0097, 1.0069, 1.045, 1.006 times from the previous number if other variables are constant.

4.5.2. Regression model of the number of pregnant maternal mortality for zero state  
   \[ \text{logit} \hat{p}_i = -0.01 + 0.0143x_{i1} + 0.0139x_{i2} + 0.0124x_{i4} + 0.0139x_{i5} + 0.0120x_{i6} + 0.0121x_{i7} \]  
   Every additional 1% the percentage of K1 pregnancy examinations it will increase risk of pregnant maternal mortality by 1.014 times. Every additional 1% percentage of K4 pregnancy examinations, percentage of pregnant TT2 + immunization, percentage of pregnant women who received Fe, percentage of obstetric complications management in pregnant women and ratio of midwives per 100,000 population it will increase risk of pregnant maternal mortality each by 1.014, 1.014, 1.012, 1.014, 1.012, 1.012 times.

a. Regression model of postpartum maternal mortality for Poisson state  
   \[ \ln \hat{\mu}_x = -30.5023 + 0.3514x_{i1} + 0.0173x_{i2} - 0.0375x_{i4} + 0.0089x_{i5} - 0.0268x_{i5} + 0.0254x_{i6} - 0.0159x_{i7} \]
d. Regression model of postpartum maternal mortality for zero state
\[ \logit \hat{\mu}_2 = 0.0120 + 0.0109x_{1i} + 0.0109x_{2i} + 0.0102x_{4i} + 0.0105x_{5i} + 0.0125x_{6i} + 0.0106x_{7i} \]

5. Conclusion
Based on the results of the study it can be concluded:
1. Parameter estimation using MLE are not closed formed so we use the BHHH algorithm to calculate the complete data MLEs parameters.
2. Hypothesis testing using the MLRT test and the simultaneous test results in at least one parameter that influences the response so that it is continued by a partial test to determine the influential predictor variables.
3. Regression models on the number of pregnant maternal mortality:
   \[ \ln \hat{\mu}_1 = -10.4856 + 0.0317x_{1i} - 0.0412x_{2i} + 0.0097x_{4i} + 0.0069x_{5i} + 0.0440x_{6i} + 0.0060x_{7i} \]
   \[ \logit \hat{\mu}_1 = -0.01 + 0.0143x_{1i} + 0.0139x_{2i} + 0.0124x_{4i} + 0.0139x_{5i} + 0.0120x_{6i} + 0.0121x_{7i} \]
   Regression models on the number of postpartum maternal mortality:
   \[ \ln \hat{\mu}_2 = -30.5023 + 0.3514x_{1i} + 0.0173x_{2i} - 0.0375x_{3i} + 0.0089x_{4i} - 0.0268x_{5i} + 0.0254x_{6i} - 0.0159x_{7i} \]
   \[ \logit \hat{\mu}_2 = 0.0120 + 0.0109x_{1i} + 0.0109x_{2i} + 0.0109x_{3i} + 0.0102x_{4i} + 0.0105x_{5i} + 0.0125x_{6i} + 0.0106x_{7i} \]
4. The linear interpolation method, scant methods, etc. can be alternative methods for iteration for the next studies also add other predictor variables.

Acknowledgment
The authors thank to central java Provincial Health Department for providing access to the data and thank to Institute Technology Sepuluh Nopember. We are grateful to the referees for careful checking of the details and for helpful comment which improve this paper.

References
[1] Consul P C and Jain G C 1973 A generalization of the poisson distribution (London:Technometrics) vol 15 p 791-799
[2] Lambert D 1992 Zero-Inflated Poisson Regression With an Application to Defects in Manufacturing (London:Technometrics) vol 34 p 1-14
[3] Famoye F and Singh K P 2006 Zero-Inflated Generalized Poisson Regression Model with an Application to Domestic Violence Data Journal of Data Science 4 p 117-130
[4] Zhang C, Tian G and Huang X 2015 Two New Bivariate Zero-Inflated Generalized Poisson Distribution With a Flexible Correlation Structure (Perth: International Academic Press) vol 3 p 105-137
[5] Agresti A 2002 Categorical Data Analysis (Second Edition) John Wiley & Sons, Inc New Jersey
[6] ASEAN Secretariat 2017 ASEAN Statistical Report on Millenium Development Goals 2017 ASEAN Secretariat Jakarta
[7] Cameron, C and Trivedi, P K 2005 Microeconometric Methods and applications UK Cambridge University Press
[8] Kawamura K 1973 The Structure of Bivariate Poisson Distribution (Tokyo:Kodai Mathematical Journal) vol 25 p 246-256
[9] Gujarati D N 2004 Basic Econometrics 4th edition McGraw Hill New York
[10] Dinas Kesehatan 2018 Profil Kesehatan Provinsi Jawa Tengah Tahun 2018 Dinkes Jawa Tengah Semarang
[11] Purhadi, Dewi Y S and Amaliana L 2015 Zero Inflated Poisson and Geographically Weighted Zero-Inflated Poisson Regression Model: Application to Elephantiasis (Filariasis) Counts Data (Dubai:Journal of Mathematics and Statistics) vol 11 p 52-60