

Article

Estimation and Hypothesis Testing for the Parameters of Multivariate Zero Inflated Generalized Poisson Regression Model

Dewi Novita Sari ^{1,2} , Purhadi Purhadi ^{1,*} , Santi Puteri Rahayu ¹ and Irhamah Irhamah ¹

¹ Department of Statistics, Faculty of Science and Data Analytics, Institut Teknologi Sepuluh Nopember, Surabaya 60111, Indonesia; dnovita@bps.go.id (D.N.S.); santi_pr@statistika.its.ac.id (S.P.R.); irhamah@statistika.its.ac.id (I.I.)

² BPS—Statistics of Banyumas Regency, Purwokerto 53114, Indonesia

* Correspondence: purhadi@statistika.its.ac.id

Abstract: We propose a multivariate regression model called Multivariate Zero Inflated Generalized Poisson Regression (MZIGPR) type II. This model further develops the Bivariate Zero Inflated Generalized Poisson Regression (BZIGPR) type II. This study aims to develop parameter estimation, test statistics, and hypothesis testing, both simultaneously and partially, for significant parameters of the MZIGPR model. The steps of the EM algorithm for obtaining the parameter estimator are also described in this article. We use Berndt–Hall–Hall–Hausman (BHHH) numerical iteration to optimize the EM algorithm. Simultaneous testing is carried out using the maximum likelihood ratio test (MLRT) and the Wald test to partially assess the hypothesis. The proposed MZIGPR model is then used to model the three response variables: the number of maternal childbirth deaths, the number of postpartum maternal deaths, and the number of stillbirths with four predictors. The units of observation are the sub-districts of the Pekalongan Residency, Indonesia. The indicate overdispersion in the data on the number of maternal childbirth deaths and stillbirths, and underdispersion in the data on the number of postpartum maternal deaths. The empirical studies show that the three response variables are significantly affected by all the predictor variables.

Keywords: MZIGPR type II; EM algorithm; BHHH algorithm; likelihood ratio test



Citation: Sari, D.N.; Purhadi, P.; Rahayu, S.P.; Irhamah, I. Estimation and Hypothesis Testing for the Parameters of Multivariate Zero Inflated Generalized Poisson Regression Model. *Symmetry* **2021**, *13*, 1876. <https://doi.org/10.3390/sym13101876>

Academic Editor:
Alessandro Sarracino

Received: 13 August 2021
Accepted: 28 September 2021
Published: 5 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The Poisson regression model is commonly used to analyze data in which the response variable follows a Poisson distribution. Even though the Poisson regression model has been widely applied in various disciplines, it involves an assumption that, in some cases, is difficult to realize—namely, equidispersion. Equidispersion is the condition wherein sample variance is equal to the mean. When the sample variance is greater (less) than the sample mean, it is called overdispersion (underdispersion). The use of Poisson regression on over- or underdispersed data tends to make the standard error and test statistics derived from the model inaccurate, resulting in invalid conclusions [1,2].

As alternatives, several distribution models have been developed, such as negative binomial distribution (NBD) and log-normal Poisson distribution, which can be used to resolve overdispersion. Alternative models for count data that can overcome both over- and underdispersion include double Poisson distribution, gamma count distribution, and generalized Poisson distribution (GPD). The first two distribution models are weak in their probability function, which is complex, and the variance and mean have no explicit form [3]. Research on GPD has developed it into generalized Poisson regression (GPR) known as the GP-1 regression model (or the classic GPR model [4]), GP-2 regression (or the restricted GPR model [5]), the bivariate GPR (BGPR) model [6], and the multivariate GPR (MGPR) model [7].

Another problem with Poisson regression is the quantity of response data with a value of zero (excess zero). Lambert [8] developed zero inflated Poisson distribution (ZIPD) to

resolve this problem. The ZIP model assumes that the population, or observations, consist of two groups. A single observation is included in a group that is always zero (zero states) with a probability of p , and is included in a group with a zero value and a positive value in the count data generated by Poisson distribution (Poisson state) with a probability of $1-p$. The development of the ZIP model has led to the emergence of several studies on bivariate ZIP (BZIP) and multivariate ZIP (MZIP) models [9–11].

The ZIPR model is the proper way to handle excess zero in the Poisson model, but is not suitable for controlling overdispersion. A study by Ridout et al. in 1998 showed that the results of ZIP parameter estimation are biased when there is an overdispersion of excess zero data. In such cases, the Zero Inflated Negative Binomial (ZINB) model performs better than the ZIP model with smaller AIC and BIC values [12]. However, the iteration technique used for parameter estimation in the ZINB model often fails to converge. Therefore, Famoye and Singh (2006) introduced the Zero Inflated Generalized Poisson (ZIGP) model as an alternative to the ZINB model [13]. Czado et al. [14] defined ZIGP distribution as an analog of ZIP distribution with additional zero inflation parameters.

The ZIGP regression model was formed based on ZIGP distribution. Thus, the ZIGP regression model can model count data with over/underdispersion and excess zero. The ZIGP model for univariate response has been applied in many aspects, such as domestic violence, entrepreneurship, biology, health, and zoology [1,13–16].

Zhang and Huang (2015) introduced the bivariate ZIGP (BZIGP), which is an extension of the univariate ZIGP. Zhang and Huang proposed type I and type II BZIGP distribution with a flexible correlation structure that can be used both when the correlation is positive or negative, and when there is over or underdispersion. The difference between the two marginal distributions of BZIGP gives the zero inflation parameters. BZIGP type I has the same zero inflation parameter on both marginals, whereas the two marginals of BZIGP type II have different zero inflation parameters [17].

BZIGP type I distribution was later developed into MZIGP type I distribution by Huang et al. in 2017. MZIGP type I was applied to data with two responses: the number of days children were inactive in the previous four weeks due to sickness, and the number of days children were bedridden in the previous four weeks due to sickness [18]. MZIGP type I distribution only considers zero and non-zero response pairs. This study proposes a multivariate regression model that can be used on count data with excess zeros and several types of dispersion, and allows for several combinations of response pairs. The proposed model is called a multivariate Zero Inflated Generalized Poisson Regression (MZIGPR) type II (henceforth MZIGPR^(II)).

Based on the previously mentioned background, the aims of this study are: (i) to construct an MZIGPR^(II) model, (ii) to obtain a parameters estimator, (iii) to test the significance of the model as well as the significance of the individual parameter, and (iv) to apply the MZIGPR^(II) model to real data. The case study used here includes factors that influence the number of maternal childbirth deaths (Y_1), the number of postpartum maternal deaths (Y_2), and the number of stillbirths (Y_3). The units of observation are 91 sub-districts in Pekalongan Residency, Indonesia. The predictor variables include the percentage of child-births assisted by health workers, the percentage of TT2+ vaccination in pregnant women, the percentage of obstetric complications that were handled, and the ratio of midwives in the population (per 10,000).

The decline in maternal mortality and neonatal mortality has become a global issue because it arises in all countries worldwide, and not only in developing countries. The first and second targets of the third goal of the Sustainable Development Goals (SDGs) are to reduce the maternal mortality ratio to less than 70 deaths per 100,000 live births, and to reduce neonatal deaths to at least 12 deaths per 1000 live births, by 2030 [19]. Despite showing a decline, Indonesia's maternal and neonatal mortality rates are still relatively high, at 305 maternal deaths per 100,000 live births and 15 neonatal deaths per 1000 live births [20,21]. Pekalongan Residency has more maternal deaths and stillbirths than other

residencies in Central Java Province, Indonesia—24.0% and 20.79% [22]. To achieve the SDGs by 2030, greater efforts are required.

The discussion in this study is divided into several sections. In Section 2, we discuss the MZIGPR^(II) model. Section 3 discusses the estimation of the parameters of the MZIGPR^(II) model using the EM algorithm method, and the testing of the hypothesis of the MZIGPR^(II) model using maximum likelihood ratio test (MLRT). The proposed approach is used to model data on the numbers of maternal childbirth deaths, postpartum maternal deaths, and stillbirths. The discussion and conclusions can be found in the last two sections.

2. Multivariate Zero Inflated Generalized Poisson Regression Type II (MZIGPR^(II)) Model

A k -dimensional multivariate ZIGP (MZIGP^(II)) has been constructed via a mixture of multivariate GP, univariate GP, and a distribution degenerate at the original point $(0, 0, \dots, 0)$. Suppose Y_1 is a case of maternal childbirth death and Y_2 is a case of stillbirth. The sub-districts can then be categorized as follows, based on death cases: (i) sub-districts that have no cases of childbirth maternal deaths and stillbirths ($Y_1 = 0, Y_2 = 0$), (ii) sub-districts with cases of childbirth maternal deaths, but no cases of stillbirths ($Y_1 > 0, Y_2 = 0$), (iii) sub-districts with no cases of childbirth maternal deaths, but stillbirths ($Y_1 = 0, Y_2 > 0$), and (iv) sub-districts with cases of childbirth maternal deaths and stillbirths ($Y_1 > 0, Y_2 > 0$). To obtain a k -dimensional ZIGP multivariate model, we use a mixture of a degenerate on the zero points $(0, 0, \dots, 0)$, k distributions with univariate GP for one type of death and $k - 1$ zeros (for example, $(GP(\lambda_1, \varphi_1), 0, \dots, 0), \dots, (0, \dots, 0, GP(\lambda_k, \varphi_k))$), $\binom{k(k-1)}{2} = C_2^k$ distributions with bivariate GP for two types of death and $k - 2$ zeros, \dots , k distributions each with $(k - 1)$ -dimensional GP for $k - 1$ types of death and 1 zero, and finally, a k -dimensional GP for all k types of death. Since the amount of mixture formed depends on the number of responses used, the following discussion will focus on a model with three responses.

This MZIGP^(II) distribution is used mainly for situations in which the observed events are mostly zero. Therefore, if the observed events are mostly non-zero, then the k -dimensional $Y > 0$ will represent all mixtures of l -dimensional $Y > 0$ and $(k - l)$ dimensional $Y = 0$, where $k \geq l \geq 2$ [10]. Suppose $(Y_1, Y_2, Y_3) \sim MZIGP^{(II)}(p_1, p_2, p_3; \lambda_1, \lambda_2, \lambda_3, \varphi_1, \varphi_2, \varphi_3)$. Thus, the pmf of MZIGPR^(II) can be written as follows:

$$P(Y_{1i} = y_{1i}, Y_{2i} = y_{2i}, Y_{3i} = y_{3i}) = \begin{cases} A_i, & \text{if } y_{1i} = 0, y_{2i} = 0, y_{3i} = 0 \\ B_i, & \text{if } y_{1i} > 0, y_{2i} = 0, y_{3i} = 0 \\ C_i, & \text{if } y_{1i} = 0, y_{2i} > 0, y_{3i} = 0 \\ D_i, & \text{if } y_{1i} = 0, y_{2i} = 0, y_{3i} > 0 \\ E_i, & \text{if } y_{1i} > 0, y_{2i} > 0, y_{3i} > 0 \end{cases} \quad (1)$$

The joint pmf $P(Y_{1i} = y_{1i}, Y_{2i} = y_{2i}, Y_{3i} = y_{3i})$ in Equation (1) (which can be seen in detail in Appendix A), with functions μ_{ki} and p_{ki} , satisfies

$$\begin{aligned} \mu_{ki} &= \exp(\mathbf{x}_i^T \boldsymbol{\beta}_k) \\ p_{ki} &= \frac{\exp(\mathbf{x}_i^T \boldsymbol{\gamma}_k)}{1 + \exp(\mathbf{x}_i^T \boldsymbol{\gamma}_k)} \text{ and } (1 - p_{ki}) = \frac{1}{1 + \exp(\mathbf{x}_i^T \boldsymbol{\gamma}_k)}, \quad k = 1, 2, 3; i = 1, 2, \dots, n \end{aligned} \quad (2)$$

where φ is the dispersion parameter, γ is the zero inflation parameter, \mathbf{x}_i is the i -th row of the covariate matrix \mathbf{X} , $\boldsymbol{\beta}_k$ is the q -dimensional regression parameter column vector, and $\boldsymbol{\gamma}_k$ is the q -dimensional zero inflation parameter column vector.

The MZIGPR^(II) model reduces to the MGPR model when the zero inflation parameter $\gamma = 0$, and it reduces to the MZIPR model when the dispersion parameter $\varphi = 0$.

3. Inference

3.1. Parameter Estimation of MZIGPR^(III) Model

Let the random variable $(Y_{1i}, Y_{2i}, Y_{3i})^T \stackrel{iid}{\sim} ZIGP(\gamma_1, \gamma_2, \gamma_3, \beta_1, \beta_2, \beta_3, \varphi_1, \varphi_2, \varphi_3)$ for $i = 1, 2, 3, \dots, n$. Define

$$\begin{aligned} I_0 &= \{i : y_{1i} = 0, y_{2i} = 0, y_{3i} = 0; i = 1, 2, 3, \dots, n\} \\ I_1 &= \{i : y_{1i} > 0, y_{2i} = 0, y_{3i} = 0; i = 1, 2, 3, \dots, n\} \\ I_2 &= \{i : y_{1i} = 0, y_{2i} > 0, y_{3i} = 0; i = 1, 2, 3, \dots, n\} \\ I_3 &= \{i : y_{1i} = 0, y_{2i} = 0, y_{3i} > 0; i = 1, 2, 3, \dots, n\} \\ I_4 &= \{i : y_{1i} > 0, y_{2i} > 0, y_{3i} > 0; i = 1, 2, 3, \dots, n\} \end{aligned}$$

The number of elements of I_0, I_1, I_2, I_3, I_4 is n_0, n_1, n_2, n_3 and n_4 , where $n_4 = n - n_0 - n_1 - n_2 - n_3$. There is no limit to the sample size used, regardless of whether the sample sizes are the same or different. In this study, we used different sample sizes of n_0, n_1, n_2, n_3 and n_4 .

Suppose $\theta = \beta_1, \beta_2, \beta_3, \gamma_1, \gamma_2, \gamma_3, \varphi_1, \varphi_2, \varphi_3, \eta_{12}, \eta_{13}, \eta_{23}$; then, the likelihood function of the observation data can be written as follows:

$$L(\theta|Y_{obs}) = \prod_{i=1}^n f(y_{1i}, y_{2i}, y_{3i}) = \prod_{i \in I_0} A_i \prod_{i \in I_1} B_i \prod_{i \in I_2} C_i \prod_{i \in I_3} D_i \prod_{i \in I_4} E_i \tag{3}$$

Parameter estimation is carried out using the EM algorithm. We employ **a**, **b**, **c**, and **d** as unobserved variables or latent variables.

The latent variable **a** divides n_0 into $A_0 + A_1 + A_2 + A_3 + A_4 + A_5 + A_6 + A_7$, where $A_7 = n_0 - A_0 - A_1 - A_2 - A_3 - A_4 - A_5 - A_6$; the latent variable **b** divides n_1 into $B_0 + B_1 + B_2 + B_3$ where $B_3 = n_1 - B_0 - B_1 - B_2$; the latent variable **c** divides n_2 into $C_0 + C_1 + C_2 + C_3$ where $C_3 = n_2 - C_0 - C_1 - C_2$, and the latent variable **d** divides n_3 into $D_0 + D_1 + D_2 + D_3$, where $D_3 = n_3 - D_0 - D_1 - D_2$. Thus, the distribution of **a**, **b**, **c** and **d** is

$$\begin{aligned} \mathbf{a}|(Y_{obs}, \theta) &\sim \text{Multinomial} \left(n_0; \frac{f_0}{f}, \frac{f_1}{f}, \frac{f_2}{f}, \frac{f_3}{f}, \frac{f_4}{f}, \frac{f_5}{f}, \frac{f_6}{f}, \frac{f_7}{f} \right) \\ \mathbf{b}|(Y_{obs}, \theta) &\sim \text{Multinomial} \left(n_1; \frac{h_0}{h}, \frac{h_1}{h}, \frac{h_2}{h}, \frac{h_3}{h} \right) \\ \mathbf{c}|(Y_{obs}, \theta) &\sim \text{Multinomial} \left(n_2; \frac{j_0}{j}, \frac{j_1}{j}, \frac{j_2}{j}, \frac{j_3}{j} \right) \\ \mathbf{d}|(Y_{obs}, \theta) &\sim \text{Multinomial} \left(n_3; \frac{l_0}{l}, \frac{l_1}{l}, \frac{l_2}{l}, \frac{l_3}{l} \right) \end{aligned} \tag{4}$$

where $f = \sum_{s=0}^7 f_s, h = \sum_{s=0}^3 h_s, j = \sum_{s=0}^3 j_s, l = \sum_{s=0}^3 l_s$ and f_s, h_s, j_s, l_s refer to Appendix A (i)–(iv).

The likelihood function for complete data denoted by $Y_{com} = \{Y_{obs}, f_0, f_1, f_2, f_3, f_4, f_5, f_6, f_7, h_0, h_1, h_2, h_3, j_0, j_1, j_2, j_3, l_0, l_1, l_2, l_3\}$ is a function that will be maximized using the EM algorithm. This likelihood function is proportional to

$$\begin{aligned} L(\theta|Y_{com}) &\propto (f_0)^{a_0} (f_1)^{a_1} (f_2)^{a_2} (f_3)^{a_3} (f_4)^{a_4} (f_5)^{a_5} (f_6)^{a_6} (f_7)^{a_7} \\ &\times \prod_{i \in I_1} (1 - p_1) \frac{\lambda_1^{y_1} (1 + \varphi_1 y_1)^{y_1 - 1}}{y_1!} \exp(-\lambda_1 (1 + \varphi_1 y_1)) (h_0)^{b_0} (h_1)^{b_1} (h_2)^{b_2} (h_3)^{b_3} \\ &\times \prod_{i \in I_2} (1 - p_2) \frac{\lambda_2^{y_2} (1 + \varphi_2 y_2)^{y_2 - 1}}{y_2!} \exp(-\lambda_2 (1 + \varphi_2 y_2)) (j_0)^{c_0} (j_1)^{c_1} (j_2)^{c_2} (j_3)^{c_3} \\ &\times \prod_{i \in I_3} (1 - p_3) \frac{\lambda_3^{y_3} (1 + \varphi_3 y_3)^{y_3 - 1}}{y_3!} \exp(-\lambda_3 (1 + \varphi_3 y_3)) (l_0)^{d_0} (l_1)^{d_1} (l_2)^{d_2} (l_3)^{d_3} \\ &\times \prod_{i \in I_4} (1 - p_1)(1 - p_2)(1 - p_3) \prod_{k=1}^3 \frac{\lambda_k^{y_k} (1 + \varphi_k y_k)^{y_k - 1}}{y_k!} \exp(-\lambda_k (1 + \varphi_k y_k)) Q \end{aligned}$$

The ln-likelihood function $L(\theta|Y_{com}) = \ell(\theta|Y_{com})$ is stated as follows:

$$\begin{aligned} \ell(\theta|Y_{com}) = & a_0 \ln(f_0) + a_1 \ln(f_1) + a_2 \ln(f_2) + a_3 \ln(f_3) + a_4 \ln(f_4) + a_5 \ln(f_5) + a_6 \ln(f_6) + a_7 \ln(f_7) \\ & + \sum_{i \in I_1} \ln \left((1 - p_1) \frac{\lambda_1^{y_1} (1 + \varphi_1 y_1)^{y_1 - 1}}{y_1!} \exp(-\lambda_1 (1 + \varphi_1 y_1)) (h_0)^{b_0} (h_1)^{b_1} (h_2)^{b_2} (h_3)^{b_3} \right) \\ & + \sum_{i \in I_2} \ln \left((1 - p_2) \frac{\lambda_2^{y_2} (1 + \varphi_2 y_2)^{y_2 - 1}}{y_2!} \exp(-\lambda_2 (1 + \varphi_2 y_2)) (j_0)^{c_0} (j_1)^{c_1} (j_2)^{c_2} (j_3)^{c_3} \right) \\ & + \sum_{i \in I_3} \ln \left((1 - p_3) \frac{\lambda_3^{y_3} (1 + \varphi_3 y_3)^{y_3 - 1}}{y_3!} \exp(-\lambda_3 (1 + \varphi_3 y_3)) (l_0)^{d_0} (l_1)^{d_1} (l_2)^{d_2} (l_3)^{d_3} \right) \\ & + \sum_{i \in I_4} \ln \left[(1 - p_1)(1 - p_2)(1 - p_3) \sum_{k=1}^3 \frac{\lambda_k^{y_k} (1 + \varphi_k y_k)^{y_k - 1}}{y_k!} \exp(-\lambda_k (1 + \varphi_k y_k)) Q \right] \end{aligned}$$

where $\lambda_{ki} = \frac{\mu_{ki}}{1 + \varphi_k \mu_{ki}}$, in which μ_{ki} and p_{ki} refer to Equation (2).

The solution of the first derivative of $\ell(\theta|Y_{com})$ concerning the parameters γ and θ does not have an explicit form. Therefore, the maximum likelihood estimator for the MZIGPR(II) model's parameters is obtained through the EM algorithm with the Berndt–Hall–Hall–Hausman (BHHH) iteration. The E-step in the EM algorithm involves replacing a_s, b_s, c_s and d_s with $a_s^{(c)}, b_s^{(c)}, c_s^{(c)}$ and $d_s^{(c)}$, where

$$\begin{aligned} a_s^{(c)} &= E \left(a_s \mid y_i, \theta^{(c)} \right) = \frac{n_0 f_s}{f}, \quad s = 0, 1, 2, 3, 4, 5, 6, 7 \\ b_s^{(c)} &= E \left(b_s \mid y_i, \theta^{(c)} \right) = \frac{n_1 h_s}{h}, \quad s = 0, 1, 2, 3 \\ c_s^{(c)} &= E \left(c_s \mid y_i, \theta^{(c)} \right) = \frac{n_2 j_s}{j}, \quad s = 0, 1, 2, 3 \\ d_s^{(c)} &= E \left(d_s \mid y_i, \theta^{(c)} \right) = \frac{n_3 k_s}{k}, \quad s = 0, 1, 2, 3 \end{aligned}$$

The M-step in the EM algorithm is the step to update $\hat{\theta}^{(c+1)} = \hat{\theta}^{(c)} - \mathbf{H}^{-1}(\hat{\theta}^{(c)}) \mathbf{g}(\hat{\theta}^{(c)})$, where $\hat{\theta}^{(c)}$ is the vector estimator at the c -th iteration and $\mathbf{g}(\hat{\theta})$ is the gradient vector. $\mathbf{H}(\theta)$ is defined as $\mathbf{H}(\theta) = -\sum_{i=1}^n \mathbf{g}_i(\theta) \mathbf{g}_i^T(\theta)$, where $\mathbf{g}_i(\theta)$ is the gradient vector at the i -th observation. The iteration will stop when $\hat{\theta}^{(c)}$ converges or $|\hat{\theta}^{(c+1)} - \hat{\theta}^{(c)}| < \varepsilon$, where ε is a very small positive real number.

3.2. Hypothesis Testing of the MZIGPR(II) Model

Hypothesis testing of the MZIGPR(II) model was conducted via both simultaneous and partial parameter testing. Simultaneous parameter testing is used to determine the significance of parameters simultaneously in the model. The hypothesis for the simultaneous testing of β and γ is as follows:

$$\begin{aligned} H_0 : \beta_{k1} = \dots = \beta_{kq} = \gamma_{k1} = \dots = \gamma_{kq} = 0 \\ H_1 : \text{at least one } \beta_{kr} \neq 0 \text{ and } \gamma_{kr} \neq 0; \quad k = 1, 2, \dots, m; \quad r = 1, 2, \dots, q \end{aligned} \tag{5}$$

The statistic test is $G^2 = -2 \ln \left(\frac{L(\hat{\omega})}{L(\hat{\Omega})} \right) = 2 [\ln(L(\hat{\Omega})) - \ln(L(\hat{\omega}))]$. Reject the null hypothesis if $G^2 > \chi_{\alpha, v}^2$ with significance level α , and v is the difference $n(\Omega) - n(\omega)$.

A partial test is carried out to determine which parameters have a significant effect on the model. The hypothesis for the partial testing of β and γ is

a. Partial testing of β

$$\begin{aligned} H_0 : \beta_{kr} = 0 \\ H_1 : \beta_{kr} \neq 0; \quad k = 1, 2, \dots, m; \quad r = 1, 2, \dots, q \end{aligned}$$

The statistic test is $Z_b = \frac{\hat{\beta}_{kr}}{se(\hat{\beta}_{kr})}$. Reject the null hypothesis if $|Z_b| > Z_{\alpha/2}$.

b. Partial testing of γ

$$H_0 : \gamma_{kr} = 0$$

$$H_1 : \gamma_{kr} \neq 0; k = 1, 2, \dots, m; r = 1, 2, \dots, q$$

The statistic test is $Z_g = \frac{\hat{\gamma}_{kr}}{se(\hat{\gamma}_{kr})}$. Reject the null hypothesis if $|Z_g| > Z_{\alpha/2}$.

4. Application of Data

The proposed MZIGPR^(III) model is applied to model the numbers of maternal childbirth deaths, postpartum maternal deaths, and stillbirths. The data used in this study are secondary data collected by the Central Java Provincial Health Office in 2017. There are seven regencies and municipalities with 91 sub-districts in the Karesidenan Pekalongan, Central Java.

This study employs three response variables (Y) and four predictor variables (X). The three response variables are the number of maternal childbirth deaths, represented as Y_1 , the number of postpartum maternal deaths, Y_2 , and the number of stillbirths, Y_3 . The four predictor variables are the percentage of childbirths assisted by medical personnel in a sub-district (X_1), the percentage of TT2 + immunization in pregnant women in a sub-district (X_2), the percentage of obstetric complications handled in a sub-district (X_3), and the ratio of midwives in the population (per 10,000) in a sub-district (X_4).

Table 1 shows that the mean values of the response variables Y_1 , Y_2 , and Y_3 are 0.18, 0.67, and 2.76, respectively. As regards the data spread, measured by the coefficient of variation (CoV), the number of maternal childbirth deaths (Y_1) showed the largest CoV, 300.88. This indicates that the variable Y_1 is more heterogeneous than Y_2 and Y_3 . The heterogeneity of Y_1 is shown in Figure 1. Even though the histograms of Y_1 , Y_2 , and Y_3 show asymmetrical curves, Y_1 's curve looks more skewed to the right compared to the other two response variables.

In the assumed MZIGP^(II) distribution, the response data with a value of zero consist of two states, namely, the zero state and the generalized Poisson (GP) state. Thus, the response data Y_1 , Y_2 , and Y_3 , which are zero ($Y_{ki} = 0; k = 1, 2, 3; i = 1, \dots, n$), also consist of two states. The first $Y_{ki} = 0$ (zero states) means no maternal childbirth or postpartum deaths and no stillbirths in a sub-district. The second $Y_{ki} = 0$ (GP state) means that there are maternal childbirth and postpartum deaths, and stillbirths, in a sub-district. However, during the last year when the data were collected, there were no maternal childbirth or postpartum deaths, or stillbirths. The zero value in the GP state is assumed to derive from a certain distribution; in this case, the MGP distribution.

Table 1. Variable description.

Variables ($n = 91$)	Mean	Standard Deviation	Coefficient of Variation	Min.	Max.
The number of maternal childbirth deaths (Y_1)	0.18	0.53	300.88	0	4
The number of postpartum maternal deaths (Y_2)	0.67	0.96	142.49	0	3
The number of stillbirths (Y_3)	2.76	2.60	94.13	0	10
The percentage of medically assisted births (X_1)	97.75	4.01	4.11	79.59	100
The percentage of TT2+ vaccination in pregnant women (X_2)	78.22	22.51	28.78	0.64	100
The percentage of obstetric complications handling (X_3)	30.17	9.6	31.82	10.57	61.61
The ratio of midwives per 10,000 population (X_4)	4.87	1.79	36.68	1.81	13.32

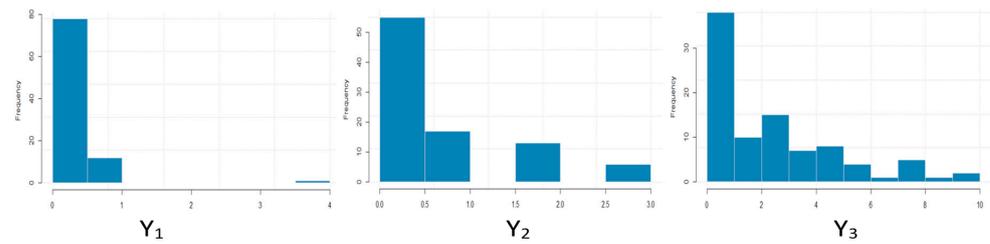


Figure 1. Histogram of Y_1 , Y_2 , and Y_3 .

MZIGPR^(II) modeling including all predictor variables yields parameter estimates, as shown in Table 2. Hypothesis testing simultaneously on these parameters determines whether each predictor affects the response variable differently. Simultaneous testing gives the result that the statistic $G^2 = 89092.513$ is greater than $\chi^2_{0.05,24} = 36.415$, so H_0 is rejected. The decision to reject H_0 indicates that at least one predictor variable affects the number of maternal deaths during childbirth and postpartum, as well as stillbirths.

Table 2. The results of the estimation of the MZIGPR^(II) model’s parameters.

Parameter	Estimate	Standard Error	Z	p-Value
The number of maternal childbirth deaths (Y_1)				
β_{01}	-4.496	8.04×10^{-8}	-5.59×10^7	$p < 0.001$
β_{11}	0.062	8.05×10^{-6}	7.73×10^3	$p < 0.001$
β_{21}	-0.020	3.97×10^{-6}	-5.07×10^3	$p < 0.001$
β_{31}	-0.029	1.80×10^{-6}	-1.63×10^4	$p < 0.001$
β_{41}	-0.908	2.35×10^{-7}	-3.87×10^6	$p < 0.001$
γ_{01}	2.091	2.13×10^{-7}	9.84×10^6	$p < 0.001$
γ_{11}	-0.012	2.08×10^{-5}	-5.64×10^2	$p < 0.001$
γ_{21}	0.010	1.63×10^{-5}	6.06×10^2	$p < 0.001$
γ_{31}	0.008	7.94×10^{-6}	1.01×10^3	$p < 0.001$
γ_{41}	0.362	1.10×10^{-6}	3.30×10^5	$p < 0.001$
The number of postpartum deaths (Y_2)				
β_{02}	-5.448	8.36×10^{-8}	-6.51×10^7	$p < 0.001$
β_{12}	-0.020	5.61×10^{-7}	-3.57×10^4	$p < 0.001$
β_{22}	-0.026	1.36×10^{-6}	-1.94×10^4	$p < 0.001$
β_{32}	0.062	9.51×10^{-6}	6.53×10^3	$p < 0.001$
β_{42}	-1.424	4.87×10^{-13}	-2.93×10^{12}	$p < 0.001$
γ_{02}	-1.424	4.87×10^{-13}	-2.93×10^{12}	$p < 0.001$
γ_{12}	-0.124	4.58×10^{-11}	-2.71×10^9	$p < 0.001$
γ_{22}	-0.051	2.10×10^{-11}	-2.41×10^9	$p < 0.001$
γ_{32}	0.011	1.89×10^{-11}	5.81×10^8	$p < 0.001$
γ_{42}	0.197	2.35×10^{-12}	8.37×10^{10}	$p < 0.001$
The number of stillbirths (Y_3)				
β_{03}	-0.880	2.71×10^{-6}	-3.25×10^5	$p < 0.001$
β_{13}	0.029	2.64×10^{-4}	1.10×10^2	$p < 0.001$
β_{23}	0.014	1.91×10^{-4}	7.24×10^1	$p < 0.001$
β_{33}	0.013	8.70×10^{-5}	1.44×10^2	$p < 0.001$
β_{43}	-0.723	1.38×10^{-5}	-5.24×10^4	$p < 0.001$
γ_{03}	1.207	5.33×10^{-11}	2.27×10^{10}	$p < 0.001$
γ_{13}	-0.087	5.17×10^{-9}	-1.68×10^7	$p < 0.001$
γ_{23}	-0.058	2.44×10^{-9}	-2.37×10^7	$p < 0.001$
γ_{33}	-0.016	1.75×10^{-9}	-8.96×10^6	$p < 0.001$
γ_{43}	-0.064	2.67×10^{-10}	-2.38×10^8	$p < 0.001$

Partial tests were performed to determine which variables had a significant influence on the model. Table 2 shows that all predictors significantly affect the three responses at a significance level of 0.05.

5. Discussion

There are two types of MZIGPR^(II) models—the log model and the logit model. The log model states that the probability Y_i of the GP state is affected by significant variables. In contrast, the logit model states that the probability Y_i of the zero states is affected by significant variables. The GP state yields zero observations and positive integers generated by the GP distribution, whereas the zero states produce zero observations only.

The MZIGPR^(II) model for the number of maternal deaths during childbirth or postpartum, and the number of stillbirths, based on the data of Table 2, is constructed as follows:

a. Regression model of maternal childbirth deaths:

$$\begin{aligned}\log(\hat{\mu}_{1i}) &= -4.496 + 0.062X_{1i} - 0.020X_{2i} - 0.029X_{3i} - 0.908X_{4i} \\ \text{logit}(\hat{p}_{1i}) &= 2.091 - 0.012X_{1i} + 0.010X_{2i} + 0.008X_{3i} + 0.362X_{4i}\end{aligned}\quad (6)$$

b. Regression model for postpartum maternal deaths:

$$\begin{aligned}\log(\hat{\mu}_{2i}) &= -5.448 - 0.020X_{1i} - 0.026X_{2i} + 0.062X_{3i} + 0.819X_{4i} \\ \text{logit}(\hat{p}_{2i}) &= -1.424 - 0.124X_{1i} - 0.051X_{2i} + 0.011X_{3i} + 0.197X_{4i}\end{aligned}\quad (7)$$

c. Regression model for stillbirths:

$$\begin{aligned}\log(\hat{\mu}_{3i}) &= -0.880 + 0.029X_{1i} + 0.014X_{2i} + 0.013X_{3i} - 0.723X_{4i} \\ \text{logit}(\hat{p}_{3i}) &= 1.207 - 0.087X_{1i} - 0.058X_{2i} - 0.016X_{3i} - 0.064X_{4i}\end{aligned}\quad (8)$$

The estimated dispersion parameters for φ_1 , φ_2 , and φ_3 are 1.091, -0.148 and 0.516, respectively, which means that the dispersion is not zero. The values of φ_1 and φ_3 indicate the existence of overdispersion in the data on the number of maternal childbirth deaths and stillbirths. In contrast, the values of φ_2 indicate underdispersion in the data on the number of postpartum maternal deaths.

The log and logit models in Equations (6)–(8) show that the variables that significantly affect the GP state also affect the zero state. Based on the regression coefficient of the MZIGPR model, there is a pattern of relationships between several predictor variables and responses that contradict the existing theory. For example, the log model shows that the variable coefficient X_1 (the percentage of childbirths assisted by medical personnel) is positive. This means that every 1% increase in the number of childbirths assisted by medical personnel will increase the average number of maternal childbirth deaths (Y_1) by 1.06 people and the average number of stillbirths (Y_3) by 1.03 people, but decrease the average number of postpartum maternal deaths (Y_2) by 0.98 people when the other predictor variables were held constant. Thus, X_1 shows an inappropriate relationship pattern with Y_1 and Y_3 .

The coefficient of variable X_2 (the percentage of TT2+ vaccination in pregnant women) is negative. If all other variables are constant, then every 1% increase in pregnant women vaccinated with TT2+ will reduce the average number of maternal childbirth deaths (Y_1) by 0.98 people and the average number of postpartum maternal deaths (Y_2) by 0.97 people. Still, the number of stillbirths (Y_3) increased by 1.01 people. In this case, X_2 has an inappropriate relationship with Y_3 . The remaining predictor variables are interpreted in the same way as in the logit model for X_1 and X_2 .

In the logit model, the X_1 variable has a negative value. This means that every one percent increase in the number of childbirths assisted by medical personnel will reduce the average number of maternal survives during childbirth (Y_1) by 0.99 people, the average number of maternal survives during postpartum (Y_2) by 0.88 people, and the average number of live births (Y_3) by 0.92 people, when the other predictor variables are held constant.

In the logit model, the variable X_1 has a negative value. This means that for every one percent increase in medically assisted births, the average number of mothers who survive childbirth (Y_1) decreases by 0.99 people. A one percent increase in X_1 will also reduce the mean number of mothers who survive after birth (Y_2) and the number of live births (Y_3) by 0.88 and 0.92 people, respectively, if the other predictor variables are constant.

The relationship pattern between X_1 and the three responses is inappropriate. The remaining variables in the logit model are interpreted in the same way as the X_1 variable in the logit model.

6. Conclusions

In this study, we developed an MZIGPR^(II) model, along with its parameter estimation and hypothesis testing. The parameter estimation is performed using the EM algorithm, followed by the BHHH iteration. The maximum likelihood ratio test (MLRT) and Wald tests have been used to test the significance of the model and the individual parameters, respectively. The proposed MZIGPR^(II) model is applied to model the number of maternal childbirth deaths, the number of postpartum maternal deaths, and the number of stillbirths.

MZIGPR^(II) can be used in cases of both overdispersion and underdispersion. We found underdispersion in the data on the number of postpartum maternal deaths and overdispersion in the data on the number of maternal childbirth deaths and stillbirths. The empirical results show that all four predictors affect the three response variables.

The main limitation of this study is that the data used are under-representational, because they were only collected at public health centers (puskesmas). Several variables thought to affect the numbers of maternal deaths during childbirth and postpartum, as well as stillbirths, were not used due to limited data availability at the sub-district level. In further research, the predictor variables used in this study can be replaced or added to other more relevant variables, according to the existing theory.

The MZIGPR^(II) model's global parameters can be used in all observation locations. Differences in regional characteristics and geographical conditions pertaining in each observation location make this global model less accurate. MZIGPR^(II) modeling with spatial aspects can be undertaken in further research.

Author Contributions: Conceptualization, D.N.S., P.P., S.P.R. and I.I.; methodology, P.P. and S.P.R.; software, D.N.S. and I.I.; validation, P.P., S.P.R. and I.I.; formal analysis, D.N.S. and P.P.; investigation, D.N.S.; data curation, D.N.S.; writing—original draft preparation, D.N.S.; writing—review and editing, P.P., S.P.R. and I.I.; supervision, P.P., S.P.R. and I.I.; project administration, P.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Deputy Bidang Penguatan Riset dan Pengembangan, Ministry of Research and Technology/National Research and Innovation Agency (Kemenristek or RISTEK-BRIN), the Republic of Indonesia, via grant Penelitian Disertasi Doktor (PDD) in 2021 with grant number 3/E1/KP.PTNBH/2021.

Data Availability Statement: Dataset not publicly available.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Let $Z \sim \text{Bernoulli}(1 - p)$, and $(U_1, U_2, U_3)^T \sim \text{MGP}(\lambda_1, \lambda_2, \lambda_3, \varphi_1, \varphi_2, \varphi_3)$. The joint pmf of $(Y_1, Y_2, Y_3)^T$ is given by $P(Y_1 = y_1, Y_2 = y_2, Y_3 = y_3) = P(Z_1 U_1 = y_1, Z_2 U_2 = y_2, Z_3 U_3 = y_3)$, so that,

- i. If $Y_1 = 0$, $Y_2 = 0$, and $Y_3 = 0$, we have

$$\begin{aligned}
 P(Y_1 = 0, Y_2 = 0, Y_3 = 0) &= P(Z_1 = 0, Z_2 = 0, Z_3 = 0) + P(Z_1 = 0, Z_2 = 0, Z_3 = 1, U_3 = 0) + P(Z_1 = 0, Z_2 = 1, U_2 = 0, Z_3 = 0) \\
 &+ P(Z_1 = 1, U_1 = 0, Z_2 = 0, Z_3 = 0) + P(Z_1 = 0, Z_2 = 1, U_2 = 0, Z_3 = 1, U_3 = 0) \\
 &+ P(Z_1 = 1, U_1 = 0, Z_2 = 0, Z_3 = 1, U_3 = 0) + P(Z_1 = 1, U_1 = 0, Z_2 = 1, U_2 = 0, Z_3 = 0) \\
 &+ P(Z_1 = 1, U_1 = 0, Z_2 = 1, U_2 = 0, Z_3 = 1, U_3 = 0) \\
 &= f_0 + f_1 + f_2 + f_3 + f_4 + f_5 + f_6 + f_7
 \end{aligned}$$

where $f_0 = (p_1 p_2 p_3)$ $f_1 = (p_1 p_2 (1 - p_3) e^{-\lambda_3})$ $f_2 = (p_1 p_3 (1 - p_2) e^{-\lambda_2})$ $f_3 = (p_2 p_3 (1 - p_1) e^{-\lambda_1})$
 $f_4 = (p_1 (1 - p_2) (1 - p_3) e^{-\lambda_2 - \lambda_3} Q)$ $f_5 = (p_2 (1 - p_1) (1 - p_3) e^{-\lambda_1 - \lambda_3} Q)$
 $f_6 = (p_3 (1 - p_1) (1 - p_2) e^{-\lambda_1 - \lambda_2} Q)$ $f_7 = ((1 - p_1) (1 - p_2) (1 - p_3) e^{-\lambda_1 - \lambda_2 - \lambda_3} Q)$

ii. If $Y_1 > 0, Y_2 = 0,$ and $Y_3 = 0,$ we have

$$\begin{aligned}
 P(Y_1 = y_1, Y_2 = 0, Y_3 = 0) &= P(Z_1 = 1, U_1 = y_1, Z_2 = 0, Z_3 = 0) + P(Z_1 = 1, U_1 = y_1, Z_2 = 1, U_2 = 0, Z_3 = 0) \\
 &+ P(Z_1 = 0, U_1 = y_1, Z_2 = 1, Z_3 = 0, U_3 = 0) + P(Z_1 = 0, U_1 = y_1, Z_2 = 0, U_2 = 0, Z_3 = 0, U_3 = 0) \\
 &= \left[(1 - p_1) \frac{\lambda_1^{y_1} (1 + \varphi_1 y_1)^{y_1 - 1}}{y_1!} \exp(-\lambda_1 (1 + \varphi_1 y_1)) \right] (h_0 + h_1 + h_2 + h_3)
 \end{aligned}$$

where $h_0 = p_2 p_3$ $h_1 = p_3 (1 - p_2) \exp(-\lambda_2) Q$ $h_2 = p_2 (1 - p_3) \exp(-\lambda_3) Q$ $h_3 = (1 - p_2) (1 - p_3) \exp(-\lambda_2 - \lambda_3) Q$

iii. If $Y_1 = 0, Y_2 > 0,$ and $Y_3 = 0,$ we have

$$\begin{aligned}
 P(Y_1 = 0, Y_2 = y_2, Y_3 = 0) &= P(Z_1 = 0, Z_2 = 1, U_2 = y_2, Z_3 = 0) + P(Z_1 = 1, U_1 = 0, Z_2 = 1, U_2 = y_2, Z_3 = 0) \\
 &+ P(Z_1 = 0, Z_2 = 1, U_2 = y_2, Z_3 = 1, U_3 = 0) + P(Z_1 = 1, U_1 = 0, Z_2 = 1, U_2 = y_2, Z_3 = 1, U_3 = 0) \\
 &= \left[(1 - p_2) \frac{\lambda_2^{y_2} (1 + \varphi_2 y_2)^{y_2 - 1}}{y_2!} \exp(-\lambda_2 (1 + \varphi_2 y_2)) \right] (j_0 + j_1 + j_2 + j_3)
 \end{aligned}$$

where $j_0 = p_1 p_3$ $j_1 = p_3 (1 - p_1) \exp(-\lambda_1) Q$ $j_2 = p_1 (1 - p_3) \exp(-\lambda_3) Q$ $j_3 = (1 - p_1) (1 - p_3) \exp(-\lambda_1 - \lambda_3) Q$

iv. If $Y_1 = 0, Y_2 = 0,$ and $Y_3 > 0,$ we have

$$\begin{aligned}
 P(Y_1 = 0, Y_2 = 0, Y_3 = y_3) &= P(Z_1 = 0, Z_2 = 0, Z_3 = 1, U_3 = y_3) + P(Z_1 = 1, U_1 = 0, Z_2 = 0, Z_3 = 1, U_3 = y_3) \\
 &+ P(Z_1 = 0, Z_2 = 1, U_2 = 0, Z_3 = 1, U_3 = y_3) + P(Z_1 = 1, U_1 = 0, Z_2 = 1, U_2 = 0, Z_3 = 1, U_3 = y_3) \\
 &= \left[(1 - p_3) \frac{\lambda_3^{y_3} (1 + \varphi_3 y_3)^{y_3 - 1}}{y_3!} \exp(-\lambda_3 (1 + \varphi_3 y_3)) \right] (l_0 + l_1 + l_2 + l_3)
 \end{aligned}$$

where $l_0 = p_1 p_2$ $l_1 = p_2 (1 - p_1) \exp(-\lambda_1) Q$ $l_2 = p_1 (1 - p_2) \exp(-\lambda_2) Q$ $l_3 = (1 - p_1) (1 - p_2) \exp(-\lambda_1 - \lambda_2) Q$

v. If $Y_1 > 0, Y_2 > 0,$ and $Y_3 > 0,$ we have

$$\begin{aligned}
 P(Y_1 = y_1, Y_2 = y_2, Y_3 = y_3) &= P(Z_1 = 1, U_1 = y_1, Z_2 = 1, U_2 = y_2, Z_3 = 1, U_3 = y_3) \\
 &= (1 - p_1) (1 - p_2) (1 - p_3) \prod_{k=1}^3 \frac{\lambda_k^{y_k} (1 + \varphi_k y_k)^{y_k - 1}}{y_k!} \exp(-\lambda_k (1 + \varphi_k y_k)) Q
 \end{aligned}$$

for the case in which at least two of the y_k 's are greater than zero.
 where $Q = 1 + \eta_{12}(\exp^{-y_1} - g_1)(\exp^{-y_2} - g_2) + \eta_{13}(\exp^{-y_1} - g_1)(\exp^{-y_3} - g_3) + \eta_{23}(\exp^{-y_2} - g_2)(\exp^{-y_3} - g_3)$, $g_k = E(\exp(-Y_k)) = \exp(\lambda_k(t_k - 1))$ with $\ln t_k - \varphi_k \lambda_k(t_k - 1) + 1 = 0$; $k = 1, 2, 3$, $\mu_{ki} = \frac{\lambda_{ki}}{1 - \varphi_k \lambda_{ki}}$ or $\lambda_{ki} = \frac{\mu_{ki}}{1 + \varphi_k \mu_{ki}}$.

References

1. Ozmen, I.; Famoye, F. Count Regression Models with an Application to Zoological Data Containing Structural Zeros. *Data Sci. J.* **2007**, *5*, 491–502. [[CrossRef](#)]
2. McCullagh, P.; Nelder, J.A. *Generalized Linear Models*, 2nd ed.; Chapman and Hall: London, UK, 1989.
3. Famoye, F. A New Bivariate Generalized Poisson Distribution. *Stat. Neerl.* **2010**, *64*, 112–124. [[CrossRef](#)]
4. Consul, P.C.; Famoye, F. Generalized Poisson Regression Model. *Commun. Stat. Theory Methods.* **1992**, *21*, 89–109. [[CrossRef](#)]
5. Wang, W.; Famoye, F. Modeling Household Fertility Decisions with Generalized Poisson Regression. *J. Popul. Econ.* **1997**, *10*, 273–283. [[CrossRef](#)] [[PubMed](#)]
6. Famoye, F. Comparisons of Some Bivariate Regression Models. *J. Stat. Comput. Simul.* **2012**, *82*, 937–949. [[CrossRef](#)]
7. Famoye, F. A Multivariate Generalized Poisson Regression Model. *Commun. Stat. Theory Methods.* **2015**, *44*, 497–511. [[CrossRef](#)]
8. Jiang, X.; Paul, S.R. A Zero-Inflated Bivariate Poisson Regression Model and Application to Some Dental Epidemiological Data. *Calcutta Stat. Assoc. Bull.* **2009**, *61*, 241–244. [[CrossRef](#)]
9. Dong, C.; Richards, S.H.; Clarke, D.B.; Zhou, X.; Ma, Z. Examining Signalized Intersection Crash Frequency Using Multivariate Zero-Inflated Poisson Regression. *Saf. Sci.* **2014**, *70*, 63–69. [[CrossRef](#)]
10. Li, C.; Lu, J.; Park, J.; Kim, K.; Brinkley, P.A.; Peterson, J.P. Multivariate Zero-Inflated Poisson Models and Their Application. *Technometrics* **1999**, *41*, 29–38. [[CrossRef](#)]
11. Liu, Y.; Tian, G.L. Type I Multivariate Zero-Inflated Poisson Distribution with Applications. *Comput. Stat. Data Anal.* **2015**, *83*, 200–222. [[CrossRef](#)]
12. Ridout, M.; Demetrio, C.G.B.; Hinde, J. Models for Count Data with Many Zeros. In Proceedings of the XIXth International Biometric Conference, Cape Town, South Africa, 13–18 December 1998; pp. 179–192.
13. Famoye, F.; Singh, K.P. Zero-Inflated Generalized Poisson Regression Model with an Application to Domestic Violence Data. *Data Sci. J.* **2006**, *4*, 117–130. [[CrossRef](#)]
14. Czado, C.; Erhardt, V.; Min, A.; Wagner, S. Zero-inflated Generalized Poisson Models with Regression Effects on the Mean, Dispersion and Zero-inflation Level Applied to Patent Outsourcing Rates. *Stat. Methodol.* **2007**, *7*, 125–153. [[CrossRef](#)]
15. Cui, Y.; Yang, W. Zero-inflated Generalized Poisson Regression Mixture Model for Mapping Quantitative Trait Loci Underlying Count Trait with Many Zeros. *J. Theor. Biol.* **2009**, *256*, 276–285. [[CrossRef](#)] [[PubMed](#)]
16. Almasi, A.; Eshragian, M.R.; Moghimbeigi, A.; Rahimi, A.; Mohammad, K.; Fallahigilan, S. Multilevel Zero-inflated Generalized Poisson Regression Modelling for Dispersed Correlated Count Data. *Stat. Methodol.* **2016**, *30*, 1–14. [[CrossRef](#)]
17. Zhang, C.; Tian, G.; Huang, X. Two New Bivariate Zero-Inflated Generalized Poisson Distribution with a Flexible Correlation Structure. *Stat. Optim. Inf. Comput.* **2015**, *3*, 105–137. [[CrossRef](#)]
18. Huang, X.F.; Tian, G.L.; Zhang, C.; Jiang, X. Type I Multivariate Zero-Inflated Generalized Poisson Distribution with Applications. *Stat. Its Interface* **2017**, *10*, 291–311. [[CrossRef](#)]
19. United Nations Development Programme. *Indicators and Data Mapping to Measure Sustainable Development Goals (SDGs) Targets: Case of Indonesia 2015*; UNDP: Jakarta, Indonesia, 2016.
20. BPS-Statistics Indonesia. *Profil Penduduk Indonesia Hasil SUPAS 2015*; BPS-Statistics Indonesia: Jakarta, Indonesia, 2016.
21. BKKBN. *BPS-Statistics Indonesia, Kementerian Kesehatan and USAID. Survei Kesehatan dan Demografi Indonesia 2017*; BKKBN: Jakarta, Indonesia, 2018.
22. Dinas Kesehatan. *Profil Kesehatan Provinsi Jawa Tengah Tahun 2017*; Dinas Kesehatan: Semarang, Indonesia, 2018.