

Article

Parameter Estimation and Hypothesis Testing of Multivariate Poisson Inverse Gaussian Regression

Selvi Mardalena, Purhadi Purhadi * , Jerry Dwi Trijoyo Purnomo and Dedy Dwi Prastyo 

Department of Statistics, Faculty of Science and Data Analytics, Institut Teknologi Sepuluh Nopember (ITS), Jawa Timur 60111, Indonesia; selvi.mardalena95@gmail.com (S.M.); jerry@statistika.its.ac.id (J.D.T.P.); dedy-dp@statistika.its.ac.id (D.D.P.)

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

Received: 3 September 2020; Accepted: 16 October 2020; Published: 20 October 2020



Abstract: Multivariate Poisson regression is used in order to model two or more count response variables. The Poisson regression has a strict assumption, that is the mean and the variance of response variables are equal (equidispersion). Practically, the variance can be larger than the mean (overdispersion). Thus, a suitable method for modelling these kind of data needs to be developed. One alternative model to overcome the overdispersion issue in the multi-count response variables is the Multivariate Poisson Inverse Gaussian Regression (MPIGR) model, which is extended with an exposure variable. Additionally, a modification of Bessel function that contain factorial functions is proposed in this work to make it computable. The objective of this study is to develop the parameter estimation and hypothesis testing of the MPIGR model. The parameter estimation uses the Maximum Likelihood Estimation (MLE) method, followed by the Newton–Raphson iteration. The hypothesis testing is constructed using the Maximum Likelihood Ratio Test (MLRT) method. The MPIGR model that has been developed is then applied to regress three response variables, i.e., the number of infant mortality, the number of under-five children mortality, and the number of maternal mortality on eight predictors. The unit observation is the cities and municipalities in Java Island, Indonesia. The empirical results show that three response variables that are previously mentioned are significantly affected by all predictors.

Keywords: overdispersion; mixed Poisson; multivariate inverse gaussian regression poisson (MPIGR); exposure; number of mortality

1. Introduction

The relationship between predictor variables and Poisson distributed response variables can be analyzed while using Poisson regression. However, some cases cannot deal with the assumption in Poisson regression, namely equidispersion, which indicates that the mean is equal to the variance of the response variable. In real data, the variance can be larger than the mean (overdispersion). Thus, a proper model for analyzing such a kind of data needs to be developed. Multivariate data consists of two or more correlated response variables. The researchers have conducted several studies in the univariate Poisson regression, but the development of multivariate Poisson regression is still lacking [1].

The Poisson Inverse Gaussian Regression (PIGR) model is one of the alternative models to overcome the overdispersion issue. This model is built from mixed Poisson distribution, which is a combination of the Poisson distribution and the inverse Gaussian distribution. The PIG distribution was firstly introduced by Holla in 1996, with the characteristics of count data having a high initial value and considerably skewed to the right curve [2]. The univariate IGPR models have been widely utilized in previous studies [3–8]. Otherwise, there are few studies of the Multivariate Poisson Inverse Gaussian

Regression (MPIGR) model. Ghitany et al. (2012) previously discussed research on the MPIGR model. That study examined several multivariate mixed Poisson regression models and their application. Parameter estimation was done using the MLE method by the Expectation and Maximization (EM) algorithm [9].

This study extends the MPIGR model by adding an exposure variable. Furthermore, the third modification of Bessel function contain factorial functions which results very large numbers that is incomputable by a computer. The results that cause the programming cannot be completed and the parameter estimator is not obtained. Therefore, the factorial function must be modified into a simpler form so that it does not produce incomputable numbers. Once the model specification has been determined, this research discusses the parameter estimation and hypothesis testing. The MLE method, followed by the Newton–Raphson iteration, is employed for the parameter estimation in this work. The next aim is to develop the hypothesis testing using the MLRT approach that has been done in this study. The proposed MPIGR model, along with its parameter estimation and hypothesis testing, is applied in order to model the relationship of three correlated response variables and eight predictors. The three response variables are the number of infant mortality, the number of under-five children mortality, and the number of maternal mortality. The observation unit is the cities and municipalities in Java Island, Indonesia. Furthermore, the exposure variable is used as a weight for each unit of observation. The exposure variable in this research is the number of live births.

As the third objective of the Sustainable Development Goals (SDGs), infant, under-five children, and maternal mortality are estimated to decrease every year. However, in 2017 alone, the under-five children accounted for 5.4 million death, with 2.5 million death occurring in the first month of life (neonatal), 1.6 million occurring at age 1–11 months (infant), and 1.3 million at age 1–4 years. There are 118 countries that already had an under-five mortality rate below the SDG target of a mortality rate at least as low as 25 deaths per 1000 births. At the same time, approximately 50 countries need to be accelerated to achieve the SDG target by 2030. As one of low-middle income country, the under-five mortality rate in Indonesia in 2017 is 28 deaths per 1000 births for male and 22 deaths per 1000 births for female. Nevertheless, efforts to reduce mortality inequity within country should be intensified [10].

The maternal mortality ratio (MMR) is one of statistical measures of maternal mortality. MMR is the number of maternal deaths during a given time period per 100,000 live births during the same time period. MMR is considered to be low if it is less than 100, moderate if it is 100–299, high if it is 300–499, very high if it is 500–999, and extremely high if it is greater than or equal to 1000 maternal deaths per 100,000 live births. SDGs target for maternal mortality by 2030 is reducing the global MMR to less than 70 per 100,000 births, with no country having a maternal mortality rate of more than twice the global average. The MMR of Indonesia in 2017 is 177, which is categorized as moderate and must be reduced in order to achieve the SDGs target [11]. There is a relationship among maternal mortality, infant mortality, and under-five mortality. In previous study, it stated that the maternal mortality cause the infant to be more likely to die than to survive and the survival trajectory of these children is far worse than those of mothers who not die [12].

Based on the issue above, analysis of the MPIGR model is performed in order to determine the factors that influence the mortality number of infants, under-five children, and maternal in Java, Indonesia. The remaining part of this paper is constructed, as follows. Section 2 provides the material and method in more detail. Results and conclusion are addressed in the last section.

2. Materials and Methods

2.1. Multivariate Poisson Inverse Gaussian Distribution (MPIGD)

The MPIGD is a mixed Poisson distribution that consists of two or more correlated response variables. Let Y_1, Y_2, \dots, Y_m as the response variables with the assumption $Y_j \sim \text{Poisson}(\mu_j), j = 1, 2, \dots, m$, with its mean and variance are the same or they can be written in the following regression equation:

$$\mu_j = E(Y_j | \mu_j) = \text{Var}(Y_j | \mu_j) = f(\mathbf{X}; \boldsymbol{\beta}) = \text{Exp}(\mathbf{X}\boldsymbol{\beta}). \quad (1)$$

which is called as equidispersion. For some cases, equidispersion is too restrictive and rarely satisfied. Oftentimes, the conditional variance will exceed the conditional mean (overdispersion), which is likely to result from positive contagion and unobserved heterogeneity. An error term ε_j is introduced to μ_j ,

$$\text{Exp}(\mathbf{X}^T \boldsymbol{\beta} + \varepsilon_j) = \mu_j \text{Exp}(\varepsilon_j) = \mu_j v_j \quad (2)$$

Thus, $v\mu_1, \dots, v\mu_m$ are the mean for each response variable now, $Y_j \sim \text{Mixed Poisson}(\mu_j v_j)$ [8,13]. The mixed Poisson distribution depends on the specific distribution of random variable V . In this study, variable V is an inverse Gaussian distributed. Hence, Y_1, Y_2, \dots, Y_m are Poisson inverse Gaussian distribution (PIGD), with the probability mass function (pmf) being given by [3],

$$P(y_1, \dots, y_m; \mu, \tau) = \int_0^\infty \prod_{j=1}^m \frac{e^{(-v\mu_j)} (v\mu_j)^{y_j}}{y_j!} g(v; \tau) dv \quad (3)$$

the $g(v; \tau)$ implies the probability density function of Inverse Gaussian of random variable V [9].

$$g(v; \tau) = (2\pi\tau v^3)^{-0.5} e^{-(v-1)^2/2\tau v}, v > 0 \quad (4)$$

where $E(v) = 1$ and $\text{Var}(V) = \tau$.

Based on Equation (1), the marginal of PIGD is as follows:

$$P(Y_1 = y_1, \dots, Y_m = y_m | \mu, \tau) = (2z\pi^{-1})^{\frac{1}{2}} e^{\frac{1}{\tau} z} K_s(z) (z\tau)^{-\sum_{j=1}^m y_j} \prod_{j=1}^m \frac{(\mu_j)^{y_j}}{y_j!} \quad (5)$$

with $s = \sum_{j=1}^m y_j - \frac{1}{2}$, $z = \frac{1}{\tau} \sqrt{1 + 2\tau \sum_{j=1}^m \mu_j}$ and $K_s(z) = K_{(\sum_{j=1}^m y_j) - \frac{1}{2}}(\frac{1}{\tau} \sqrt{1 + 2\tau \sum_{j=1}^m \mu_j})$ is the third modification of Bessel function [14]. The property of the PIGD can be found in [15,16].

2.2. Multivariate Poisson Inverse Gaussian Regression (MPIGR)

The MPIGR model is a development model from the univariate Inverse Gaussian Poisson Regression (PIGR) with two or more correlated response variables. Let $(Y_{i1}, Y_{i2}, \dots, Y_{im}) \sim \text{MPIGR}(v\mu_{ij})$, where $i = 1, 2, \dots, n$ and $j = 1, 2, \dots, m$ then the MPIGR model can be stated as follows:

$$\log\left(\frac{E(Y_{ij})}{q_i}\right) = \mathbf{x}_i^T \boldsymbol{\beta}_j \quad (6)$$

$$E(Y_{ij}) = q_i e^{\mathbf{x}_i^T \boldsymbol{\beta}_j} \quad (7)$$

with $E(Y_{ij}) = \mu_{ij} = q_i \exp(\mathbf{x}_i^T \boldsymbol{\beta}_j)$, the q_i is an exposure variable which is defined as the weight of the i -th unit observation, $\mathbf{x}_i^T = [1 \quad x_{1i} \quad x_{2i} \quad \dots \quad x_{ki} \quad \dots \quad x_{pi}]$ is the vector of predictor variables with $(p+1)$ dimension on the i -th observation ($i = 1, 2, \dots, n$), $\boldsymbol{\beta}_j^T = [\beta_{j0} \quad \beta_{j1} \quad \beta_{j2} \quad \dots \quad \beta_{jk} \quad \dots \quad \beta_{jp}]$ is a $(p+1) \times 1$ vector of regression coefficient associated with the j -th response variable ($j = 1, 2, \dots, m$).

3. Results

3.1. Parameter Estimation of MPIGR Model

Estimation of the MPIGR model parameters is obtained using the Maximum Likelihood Estimation (MLE) method by maximizing the likelihood function. Two parameters will be estimated, namely $\boldsymbol{\beta}_j$ and τ . The first step of the MLE method is by taking n random sample,

$(Y_{i1}, Y_{i2}, \dots, Y_{ij}, \dots, Y_{im}, X_{1i}, X_{2i}, \dots, X_{ki}, \dots, X_{pi})$ with $j = 1, 2, \dots, m, k = 1, 2, \dots, p$ and $i = 1, 2, \dots, n$. The joint probability density function of $Y_{i1}, Y_{i2}, \dots, Y_{im}$ is:

$$f(y_{ij}|\beta_j, \tau; j = 1, 2, 3) = \left(\frac{2}{\pi\tau}\right)^{\frac{1}{2}} e^{\frac{1}{\tau}} K_{S_i}(z_i) \left(1 + 2\tau \sum_{j=1}^3 \mu_{ij}\right)^{-\frac{(2 \sum_{j=1}^3 y_{ij}-1)}{4}} \prod_{j=1}^3 \frac{\mu_{ij}^{y_{ij}}}{y_{ij}!} \tag{8}$$

where $\mu_{ij} = q_i \exp(\mathbf{x}_i^T \beta_j)$. The likelihood function of Equation (8) is as below:

$$L(\beta_j, \tau; j = 1, 2, \dots, m) = \prod_{i=1}^n f(y_{ij}|\beta_j, \tau; j = 1, 2, \dots, m) \\ = e^{\frac{n}{\tau}} \left(\frac{2}{\pi}\right)^{\frac{n}{2}} \left(\frac{1}{\tau}\right)^{\frac{n}{2}} \prod_{i=1}^n \left(K_{S_i}(z_i) \left(1 + 2\tau \sum_{j=1}^m q_i e^{\mathbf{x}_i^T \beta_j}\right)^{-\frac{(2 \sum_{j=1}^m y_{ij}-1)}{4}} \prod_{j=1}^m \frac{(q_i e^{\mathbf{x}_i^T \beta_j})^{y_{ij}}}{y_{ij}!} \right) \tag{9}$$

The log-likelihood function from Equation (9) is:

$$l(\beta_j, \tau; j = 1, 2, \dots, m) = \log L(\beta_j, \tau, j = 1, 2, \dots, m) \\ = \frac{n}{\tau} - \frac{n}{2} \log \tau + \frac{n}{2} \log\left(\frac{2}{\pi}\right) + \sum_{i=1}^n \log(K_{S_i}(z_i)) - \sum_{i=1}^n \left(\frac{2 \sum_{j=1}^m y_{ij}-1}{4}\right) \log\left(1 + 2\tau \sum_{j=1}^m q_i e^{\mathbf{x}_i^T \beta_j}\right) \\ + \log\left(\sum_{i=1}^n \sum_{j=1}^m q_i \log(y_{ij})\right) + \sum_{i=1}^n \sum_{j=1}^m y_{ij} \mathbf{x}_i^T \beta_j - \log\left(\sum_{i=1}^n \sum_{j=1}^m y_{ij}!\right) \tag{10}$$

Then, Equation (10) derive to the parameter $\theta = [\beta_1^T \quad \beta_2^T \quad \dots \quad \beta_j^T \quad \dots \quad \beta_m^T \quad \tau]^T$ and we get:

$$\frac{\partial l(\theta)}{\partial \beta_1} = \sum_{i=1}^n \frac{\partial \log K_{S_i}(z_i)}{\partial \beta_1} - \sum_{i=1}^n \left(\frac{2 \sum_{j=1}^3 y_{ij}-1}{2(1+2\tau \sum_{j=1}^3 q_i e^{\mathbf{x}_i^T \beta_j})}\right) \tau q_i e^{\mathbf{x}_i^T \beta_1} \mathbf{x}_i^T + \sum_{i=1}^n y_{i1} \mathbf{x}_i^T \\ = \sum_{i=1}^n \frac{\partial \log K_{S_i}(z_i)}{\partial \beta_1} - \sum_{i=1}^n \left(\frac{2 \sum_{j=1}^3 y_{ij}-1}{2(1+2\tau \sum_{j=1}^3 \mu_{ij})}\right) \tau \mu_{i1} \mathbf{x}_i^T + \sum_{i=1}^n y_{i1} \mathbf{x}_i^T$$

where $\frac{\partial \log K_{\sum_{j=1}^3 y_{ij}-\frac{1}{2}}(z_i)}{\partial \beta_1} = \left(-\left(\sqrt{1 + 2\tau \sum_{j=1}^3 \mu_{ij}}\right) M(\vec{\mathbf{y}}_i) + \frac{\tau \left(\sum_{j=1}^3 y_{ij}-\frac{1}{2}\right)}{\sqrt{1+2\tau \sum_{j=1}^3 \mu_{ij}}}\right) \left(\frac{\mu_{i1} \mathbf{x}_i^T}{\sqrt{1+2\tau \sum_{j=1}^3 \mu_{ij}}}\right)$, then

$$\frac{\partial l(\theta)}{\partial \beta_1} = \sum_{i=1}^n \left\{ \left(-\left(\sqrt{1 + 2\tau \sum_{j=1}^3 \mu_{ij}}\right) M(\mathbf{y}_i) + \frac{\tau \left(\sum_{j=1}^3 y_{ij}-\frac{1}{2}\right)}{\sqrt{1+2\tau \sum_{j=1}^3 \mu_{ij}}}\right) \frac{\mu_{i1} \mathbf{x}_i^T}{\sqrt{1+2\tau \sum_{j=1}^3 \mu_{ij}}} - \frac{\left(2 \sum_{j=1}^3 y_{ij}-1\right) \tau \mu_{i1} \mathbf{x}_i^T}{2 \left(1+2\tau \sum_{j=1}^3 \mu_{ij}\right)} + \sum_{j=1}^3 y_{ij} \mathbf{x}_i^T \right\} \\ = \sum_{i=1}^n \left\{ \left(-\frac{2 \left(1+2\tau \sum_{j=1}^3 \mu_{ij}\right) M(\mathbf{y}_i) \mu_{i1} \mathbf{x}_i^T}{2 \left(1+2\tau \sum_{j=1}^3 \mu_{ij}\right)} + \frac{2 \left(1+2\tau \sum_{j=1}^3 \mu_{ij}\right) y_{i1} \mathbf{x}_i^T}{2 \left(1+2\tau \sum_{j=1}^3 \mu_{ij}\right)}\right) \right\} \\ \frac{\partial l(\theta)}{\partial \beta_1} = \sum_{i=1}^n [y_{i1} - M(\mathbf{y}_i) \mu_{i1}] \mathbf{x}_i^T \tag{11}$$

with the same procedure, generally we get:

$$\frac{\partial l(\theta)}{\partial \beta_j} = \sum_{i=1}^n [y_{ij} - M(y_i)\mu_{ij}] \mathbf{x}_i^T \tag{12}$$

Furthermore, the first derivative of the MPIGR log-likelihood function to parameter τ is formulated as below:

$$\begin{aligned} \frac{\partial l(\theta)}{\partial \tau} &= \frac{\partial}{\partial \tau} \left[\frac{n}{\tau} - \frac{n}{2} \log \tau + \frac{n}{2} \log\left(\frac{2}{\pi}\right) + \sum_{i=1}^n \log(K_{s_i}(z_i)) - \sum_{i=1}^n \frac{\left(2 \sum_{j=1}^3 y_{ij} - 1\right) \log\left(1 + 2\tau \sum_{j=1}^3 \mu_{ij}\right)}{4} \right. \\ &\quad \left. + \log\left(\sum_{i=1}^n \sum_{j=1}^3 q_i \log(y_{ij})\right) + \sum_{i=1}^n \sum_{j=1}^3 y_{ij} \mathbf{x}_i^T \beta_j - \log\left(\sum_{i=1}^n \sum_{j=1}^3 y_{ij}!\right) \right] \\ &= -\frac{n}{\tau^2} - \frac{n}{2\tau} + \sum_{i=1}^n \frac{\partial \log K_{s_i}(z_i)}{\partial \tau} - \sum_{i=1}^n \frac{\left(2 \sum_{j=1}^3 y_{ij} - 1\right) \sum_{j=1}^3 \mu_{ij}}{1 + 2\tau \sum_{j=1}^3 \mu_{ij}} \end{aligned}$$

where $\frac{\partial \log K_{s_i}(z_i)}{\partial \tau} = \frac{\left(1 + \tau \sum_{j=1}^3 \mu_{ij}\right) \left(\sum_{j=1}^3 y_{ij} - \frac{1}{2}\right)}{\tau \left(1 + 2\tau \sum_{j=1}^3 \mu_{ij}\right)} + \frac{M(y_i) \left(1 + \tau \sum_{j=1}^3 \mu_{ij}\right)}{\tau^2}$ and we got:

$$\begin{aligned} \frac{\partial l(\theta)}{\partial \tau} &= -\frac{n}{\tau^2} - \frac{n}{2\tau} - \sum_{i=1}^n \frac{\left(1 + \tau \sum_{j=1}^3 \mu_{ij}\right) \left(\sum_{j=1}^3 y_{ij} - \frac{1}{2}\right)}{\tau \left(1 + 2\tau \sum_{j=1}^3 \mu_{ij}\right)} + \sum_{i=1}^n \frac{M(y_i) \left(1 + \tau \sum_{j=1}^3 \mu_{ij}\right)}{\tau^2} \\ &\quad + \sum_{i=1}^n \frac{\left(2 \sum_{j=1}^3 y_{ij} - 1\right) \sum_{j=1}^3 \mu_{ij}}{1 + 2\tau \sum_{j=1}^3 \mu_{ij}} \\ &= -\frac{n}{2\tau} + \frac{n}{2\tau} + \sum_{i=1}^n \frac{M(y_i) \left(1 + \tau \sum_{j=1}^3 \mu_{ij}\right) - 1}{\tau^2} - \sum_{i=1}^n \sum_{j=1}^3 \frac{y_{ij}}{\tau} \\ \frac{\partial l(\theta)}{\partial \tau} &= \sum_{i=1}^n \frac{M(y_i) \left(1 + \tau \sum_{j=1}^m \mu_{ij}\right) - 1}{\tau^2} - \sum_{i=1}^n \sum_{j=1}^m \frac{y_{ij}}{\tau} \end{aligned} \tag{13}$$

where

$$M(y_i) = \frac{1}{\sqrt{1 + 2\tau \sum_{j=1}^m \mu_{ij}}} \frac{K_{\left(\sum_{j=1}^m y_{ij}\right) + \frac{1}{2}}(z_i)}{K_{\left(\sum_{j=1}^m y_{ij}\right) - \frac{1}{2}}(z_i)} \tag{14}$$

According to [14], the third modification of Bessel function in Equation (12) can be written, as follows:

$$K_{\left(\sum_{j=1}^m y_{ij}\right) + \frac{1}{2}}(z_i) = \left(\frac{\pi}{2z_i}\right)^{1/2} e^{-z_i} \sum_{m=0}^{\sum_{j=1}^m y_{ij}} \frac{\left\{\left(\sum_{j=1}^m y_{ij}\right) + m\right\}!}{\left\{\left(\sum_{j=1}^m y_{ij}\right) - m\right\}! m!} (2z_i)^{-m} \tag{15}$$

$$K_{\left(\sum_{j=1}^m y_{ij}\right)-\frac{1}{2}}(z_i) = K_{\left[\left(\sum_{j=1}^m y_{ij}\right)-1\right]+\frac{1}{2}}(z_i) = \left(\frac{\pi}{2z_i}\right)^{1/2} e^{-z_i} \sum_{m=0}^{\left(\sum_{j=1}^m y_{ij}\right)-1} \frac{\left\{\left[\left(\sum_{j=1}^m y_{ij}\right)-1\right]+m\right\}!}{\left\{\left[\left(\sum_{j=1}^m y_{ij}\right)-1\right]-m\right\}!m!} (2z_i)^{-m} \tag{16}$$

Equations (9)–(11) equate to zero and produce non-explicit form. Hence, an iterative method needs to be applied for estimating parameters. The iterative method used is the Newton–Raphson iteration method while using the following algorithm:

- *Step 1.* Determine the initial value for parameter $\hat{\theta}^{(0)} = \left[\hat{\beta}_1^{T(0)} \ \hat{\beta}_2^{T(0)} \ \dots \ \hat{\beta}_j^{T(0)} \ \dots \ \hat{\beta}_m^{T(0)} \ \hat{\tau}^{(0)} \right]^T$. The initial value of parameter $\hat{\theta}^{(0)}$ is obtained while using the separate univariate Poisson regression. The initial value for overdispersion parameter τ used the average of the observed overdispersion based on the variance of PIGD [9].
- *Step 2.* Determine the gradient vector $\mathbf{g}(\hat{\theta}^{(r)})$, which is the elements consist of the first derivative of the log-likelihood function, $\mathbf{g}^T(\hat{\theta}^{(r)}) = \left[\frac{\partial l(\hat{\theta})}{\partial \hat{\beta}_1} \ \frac{\partial l(\hat{\theta})}{\partial \hat{\beta}_2} \ \dots \ \frac{\partial l(\hat{\theta})}{\partial \hat{\beta}_j} \ \dots \ \frac{\partial l(\hat{\theta})}{\partial \hat{\beta}_m} \ \frac{\partial l(\hat{\theta})}{\partial \hat{\tau}} \right]_{\hat{\theta}=\hat{\theta}^{(r)}}$.
- *Step 3.* Determine the Hessian matrix $\mathbf{H}(\hat{\theta}^{(r)})$ where the elements consist of the second derivative of the log-likelihood function, as follows

$$\mathbf{H}(\hat{\theta}^{(r)}) = \begin{bmatrix} \frac{\partial^2 l(\hat{\theta})}{\partial \hat{\beta}_1 \partial \hat{\beta}_1} & \dots & \frac{\partial^2 l(\hat{\theta})}{\partial \hat{\beta}_1 \partial \hat{\beta}_j} & \dots & \frac{\partial^2 l(\hat{\theta})}{\partial \hat{\beta}_1 \partial \hat{\beta}_m} & \frac{\partial^2 l(\hat{\theta})}{\partial \hat{\beta}_1 \partial \hat{\tau}} \\ & \ddots & \vdots & \vdots & \vdots & \vdots \\ & & \frac{\partial^2 l(\hat{\theta})}{\partial \hat{\beta}_j \partial \hat{\beta}_j} & \dots & \frac{\partial^2 l(\hat{\theta})}{\partial \hat{\beta}_j \partial \hat{\beta}_m} & \frac{\partial^2 l(\hat{\theta})}{\partial \hat{\beta}_j \partial \hat{\tau}} \\ & & & \ddots & & \\ & & & & \frac{\partial^2 l(\hat{\theta})}{\partial \hat{\beta}_m \partial \hat{\beta}_m} & \frac{\partial^2 l(\hat{\theta})}{\partial \hat{\beta}_m \partial \hat{\tau}} \\ & & & & & \frac{\partial^2 l(\hat{\theta})}{\partial \hat{\tau}^2} \end{bmatrix}$$

symmetric

- *Step 4.* Start the Newton–Raphson iteration using the following formula,

$$\hat{\theta}^{(r+1)} = \hat{\theta}^{(r)} - \mathbf{H}^{-1}(\hat{\theta}^{(r)}) \mathbf{g}(\hat{\theta}^{(r)}) \tag{17}$$

with $\hat{\theta}^{(r)} = \left[\hat{\beta}_1^T \ \hat{\beta}_2^T \ \dots \ \hat{\beta}_j^T \ \dots \ \hat{\beta}_m^T \ \hat{\tau} \right]^T$ and $r = 0, 1, 2, \dots, r^*$.

- *Step 5.* The iteration will stop if $\|\hat{\theta}^{(r+1)} - \hat{\theta}^{(r)}\| \leq \varepsilon$, with ε is a very small value and it will produce the estimator value for each parameter.

3.2. Factorial Simplification in the Third Modification of BESSEL Function

The factorial function in the third modification of Bessel function, as performed in Equations (15) and (16) results in the large value by the calculation in programming. These results cause the program unable to complete, so that parameter estimates cannot be obtained. Therefore, the factorial simplification in Bessel function is needed. Based on Equations (15) and (16), let $\sum_{j=1}^m y_j = N$. Because

the factorial function only works for these conditions, $N > 0$, $m > 0$, and $n > m$, accordingly $(N + m)!$ contains the whole of $(N - m)!$ and the factorial function can be simplified, as below:

$$\frac{(N+m)!}{(N-m)!} = \frac{1 \times 2 \times 3 \times \dots \times (N-m-1) \times (N-m) \times (N-m+1) \times \dots \times (N+m)}{1 \times 2 \times 3 \times \dots \times (N-m-1) \times (N-m)} = \prod_{a=1}^{2m} (N-m+a) \tag{18}$$

According to Equation (18), the factorial function in Equation (15) can be written, as follows:

$$\sum_{m=0}^{\sum_{j=1}^m y_j} \frac{\left(\sum_{j=1}^m y_j + m\right)!}{\left(\sum_{j=1}^m y_j - m\right)! m!} (2z)^{-m} = \sum_{m=0}^{\sum_{j=1}^m y_j} \prod_{a=1}^{2m} \left(\sum_{j=1}^3 y_j - m + a\right) \frac{(2z)^{-m}}{m!} \tag{19}$$

with the same procedure, the factorial function in Equation (16) can be simplified, as below:

$$\sum_{m=0}^{\sum_{j=1}^3 y_j} \frac{\left(\sum_{j=1}^3 y_j - 1 + m\right)!}{\left(\sum_{j=1}^3 y_j - 1 - m\right)! m!} (2z)^{-m} = \sum_{m=0}^{\sum_{j=1}^3 y_j} \prod_{a=1}^{2m-1} \left(\sum_{j=1}^3 y_j - m + a\right) \frac{(2z)^{-m}}{m!} \tag{20}$$

Substitute Equations (19) and (20) to Equations (15) and (16) in order to calculate $M(y_i)$ in Equation (14). The result of $M(y_i)$ will be used in the first and second derivative of MPIGR log-likelihood function in order to obtain the parameter estimation.

3.3. Hypothesis Testing of MPIGR Model

The hypothesis testing of the MPIG regression model was undertaken by the Maximum Likelihood Ratio Test (MLRT) method both simultaneously and partially. Simultaneous hypothesis testing is performed in order to determine the significance of the regression parameters in the model simultaneously with the following hypothesis: the null hypothesis is $\beta_{j1} = \beta_{j2} = \dots = \beta_{jk} = \dots = \beta_{jp} = 0$ and $\tau = 0$ and the alternative hypothesis is at least one $\beta_{jk} \neq 0$ and $\tau \neq 0$, where $j = 1, 2, \dots, m$; and $k = 1, 2, \dots, p$.

Let Ω as a set of parameters under population with $\Omega = \{\beta_j, \tau; j = 1, 2, \dots, m\}$ and ω is a set of parameters under null hypothesis with $\omega = \{\beta_{0\omega j}, \tau_{\omega}; j = 1, 2, \dots, m\}$. The $L(\hat{\Omega})$ is the likelihood of the full model, which includes all of the predictor variables, and $L(\hat{\omega})$ is the likelihood of a saturated model without predictor variables. The likelihood function for each model, as follows:

$$L(\Omega) = \prod_{i=1}^n P(Y = y_i | x_i; \beta, \tau) \tag{21}$$

and

$$\begin{aligned} L(\omega) &= \prod_{i=1}^n P(Y = y_i | \beta_{0\omega}, \tau_{\omega}) \\ &= e^{\frac{n}{\hat{\tau}_{\omega}}} \left(\frac{2}{\pi}\right)^{\frac{n}{2}} \left(\frac{1}{\hat{\tau}_{\omega}}\right)^{\frac{n}{2}} \prod_{i=1}^n \left[K_s(z) \left(1 + 2\hat{\tau}_{\omega} \sum_{j=1}^3 q_j e^{\hat{\beta}_{0\omega j}} \right)^{-\frac{(2 \sum_{j=1}^3 y_{ij} - 1)}{4}} \frac{3}{\prod_{j=1}^3 \frac{(q_j e^{\hat{\beta}_{0\omega j}})^{y_{ij}}}{y_{ij}!}} \right] \end{aligned} \tag{22}$$

The test statistics for the hypothesis in the simultaneous test of MPIGR model is formulated, as below.

$$G^2 = -2 \log \left(\frac{L(\hat{\omega})}{L(\hat{\Omega})} \right) = 2(\log L(\hat{\Omega}) - \log L(\hat{\omega})) \quad (23)$$

The log-likelihood function in Equation (22) is maximized by determining the first-order partial derivative of the log-likelihood function with respect to the parameters of $\beta_{0\omega j}$ and τ_ω and the results are as follows.

$$\frac{\partial l(\omega)}{\partial \beta_{0\omega 1}} = \sum_{i=1}^n [y_{i1} - M(\mathbf{y}_i) e^{\beta_{0\omega 1}}] \quad (24)$$

Generally, we get:

$$\frac{\partial l(\omega)}{\partial \beta_{0\omega j}} = \sum_{i=1}^n [y_{ij} - M(\mathbf{y}_i) e^{\beta_{0\omega j}}] \quad (25)$$

$$\frac{\partial l(\omega)}{\partial \tau_\omega} = \sum_{i=1}^n \left\{ \frac{M(\mathbf{y}_i) \left(1 + \tau_\omega \sum_{j=1}^3 q_i e^{\beta_{0\omega j}} \right) - 1}{\tau_\omega^2} \right\} - \sum_{i=1}^n \sum_{j=1}^3 \frac{y_{ij}}{\tau_\omega} \quad (26)$$

Hence, the statistics G^2 for the MPIGR model determined by substituting Equations (9) and (22) to Equation (23) and the result is as follows:

$$G^2 = 2 \left[n \left(\frac{1}{\hat{\tau}} - \frac{1}{\hat{\tau}_\omega} + \frac{1}{2} \log \frac{\hat{\tau}_\omega}{\hat{\tau}} \right) + \sum_{i=1}^n \log \left(\frac{K_{S_i}(z_i)}{K_{S_i}(z_{\hat{\omega}i})} \right) + \sum_{i=1}^n \frac{\left(2 \sum_{j=1}^m y_{ij} - 1 \right)}{4} \log \left(\frac{1 + 2\hat{\tau}_\omega \sum_{j=1}^m q_i e^{\beta_{0\omega j}}}{1 + 2\hat{\tau} \sum_{j=1}^m q_i e^{\mathbf{x}_i^T \hat{\beta}_j}} \right) \right. \\ \left. + \sum_{i=1}^n \sum_{j=1}^m y_{ij} \mathbf{x}_i^T \hat{\beta}_j - \sum_{i=1}^n \sum_{j=1}^m y_{ij} \hat{\beta}_{0\omega j} \right] \quad (27)$$

The details of the statistics G^2 based on the likelihood ratio test method is described in Appendix A. The statistics G^2 follows the asymptotic of the Chi-square distribution, such that the significant level α reject the null hypothesis when G^2 value falls into the rejection region, i.e., when $G^2 > \chi_{\alpha, (p+1)k}^2$.

3.4. Application

The MPIGR model in this study is applied in order to model the number of infants, child, and maternal death. The data are collected from the Health Profile of Java, Indonesia, in 2017. There are six provinces with 119 cities or municipalities in Java Island. Because of the data limitation, Banten Province was not included in this study. Thus, this study only used 111 cities or municipalities.

The variables used in this study is consist of three correlated response variables, namely the number of infant mortality (Y_1), under-five children mortality (Y_2), and maternal mortality (Y_3). There are eight predictor variables, such as the percentage of antenatal care visit by pregnant women (X_1), the percentage of pregnant women who received Fe3 tablet (X_2), the percentage of complete neonatal visits (X_3), the percentage of Low Birth Weight (LBW) (X_4), the percentage of healthy house (X_5), the percentage of active integrated service post (X_6), the percentage of infants received vitamin A (X_7), and the percentage of births that are assisted by health workers (X_8). Banten Province does not have several predictor variables selected, which is quite important for modelling the response variable. Thus, Banten Province was excluded from the study on the consideration of the selected predictor variables.

Every region in Java Island has different characteristics. Therefore, the exposure variable is needed, because the city or municipality is worth comparing. The exposure variable used in this study is the number of live births of each city or municipality in Java.

Based on Table 1, the means of three response variables are 118.08, 20.41, and 16.4. Because the mean of Y_1 , Y_2 , and Y_3 differ greatly, we need to measure the spread of the data. The coefficient of variation (CoV) can be used to measure data distribution. The CoV of Y_1 , Y_2 , and Y_3 are 63.4, 425.6, and 89.1. The number of child mortality (Y_2) has the highest CoV, which means that variable Y_2 is more heterogenous than the other two variables. This evidence is also supported by histogram for Y_1 , Y_2 , and Y_3 in Figure 1. It shows that the Y_2 curve is quite skewed to the right than Y_1 and Y_3 .

Table 1. Description of the response variables.

Variable	Mean	SD	Coefficient of Variation	Min	Max
The number of infant mortality (Y_1)	118.08	72.885	63.4	7	403
The number of child mortality (Y_2)	20.41	36.554	425.6	0	278
The number of maternal mortality (Y_3)	16.4	12.233	89.1	0	59

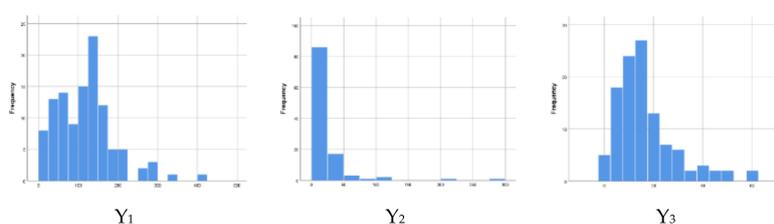


Figure 1. Histogram of response variables.

Table 2 displays the characteristics of each predictor that is presumed to influence the number of infants, child, and maternal deaths. The characteristics of each predictor are explained based on the mean and standard deviation for each province in Java Island. As a health indicator, these predictors are expected to meet the targets in order to improve the quality of health in Indonesia.

Table 2. Descriptive statistics of predictor variables that are based on city or municipality in Java.

Variable	Province				
	Jakarta	Yogyakarta	Central Java	West Java	East Java
The percentage of antenatal care visit by pregnant women	99.26 (6.10) ^a	90.92 (3.61)	92.86 (3.62)	97.44 (8.20)	89.34 (5.47)
The percentage of pregnant woman who received Fe3 tablet	95.14 (4.35)	88.05 (4.41)	92.85 (4.01)	95.88 (9.50)	88.37 (5.67)
The percentage of complete neonatal visits	95.44 (2.13)	77.32 (28.53)	92.97 (10.29)	94.30 (16.78)	96.34 (3.82)
The percentage of Low Birth Weight (LBW)	1.07 (1.45)	5.26 (1.14)	4.54 (0.93)	2.87 (1.66)	4.20 (1.41)
The percentage of healthy house	66.33 (18.85)	70.59 (17.97)	85.27 (14.17)	71.25 (15.84)	70.53 (16.40)
The percentage of active integrated service post	100 (0.00)	76.99 (9.36)	66.98 (18.88)	63.07 (20.58)	78.14 (14.53)
The percentage of infants received vitamin A	92.52 (8.13)	90.92 (16.40)	97.25 (8.43)	91.76 (16.01)	98.30 (7.97)
The percentage of births assisted by health workers	98.00 (5.56)	100.0 (0.00)	99.14 (1.56)	97.94 (8.39)	94.04 (4.16)
The number of live births ^b	34649 (23137)	8470 (4548)	15424 (7310)	33903 (25870)	15144 (10125)

a: Mean (Standard deviation); b: Exposure variable.

Predictors other than LBW are expected to have a high percentage, because these predictors are thought to reduce the number of infants, child, and maternal death. In comparison, LBW is expected to have a low percentage, because LBW is believed to be able to increase the number of infant deaths.

Based on Table 2, almost all of the provinces in Java reach an average of 80–90% for some predictors, except LBW. Still, the province with a low percentage, such as Yogyakarta, has the average of the ratio of complete neonatal visits at 77.32% that should be increased. Furthermore, for the percentage of active integrated service posts, other provinces except Jakarta have a percentage of 60–78%, which is a quite low average. The active integrated service post is a form of Community-Based Health Efforts to facilitate the public for infants, children, and maternal to get health services. Aside from active integrated service posts, the percentage of healthy homes in all provinces in Java has not reached 80% other than Central Java Province. Additional suggestions to the government, the role of the community

should be improved in order to encourage the people to get involved in the implementation of active integrated service posts, and the percentage of healthy homes.

The characteristics of each predictor provide a description and presumption about the predictors affect the number of infants, child, and maternal deaths in Java. Further analysis is required to obtain more accurate results. The MPIGR model is used to determine the predictors, since it significantly affects the number of infants, child, and maternal death in Java. Before applying the MPIGR, it is necessary to test the overdispersion assumption. Overdispersion occurs when the variance is higher than the mean. The overdispersion exists when the deviance value over the degree of freedom is higher than one, and the ratio of Pearson Chi-square value over the degree of freedom is higher than one.

Table 3 shows that all of the response variables suffer overdispersion, because the values of deviance/df are higher than one. Therefore, the MPIGR model should be used to model the data. The relationship between pair of response variables was measured by the Pearson's product-moment correlation. The coefficient of Pearson's correlation between variable Y_1 and Y_2 is 0.543 (p -value = 6.97×10^{-10}). The coefficient of Pearson's correlation between variable Y_1 and Y_3 is 0.587 (p -value = 1.29×10^{-11}). Otherwise, the coefficient of Pearson's correlation between variable Y_2 and Y_3 is 0.130 (p -value = 0.172). Even though there is one pair of the response variables that has significantly no correlation, we need to make sure whether there is dependency among the response variables in multivariate way. Therefore, we calculated the correlation using Bartlett's test. The result shows that $\chi^2(92.238) > \chi^2_{3,0.05}(7.815)$ and p -value (7.20×10^{-20}) $< \alpha$ (0.05). The decision is to reject the null hypothesis, stating that Pearson correlation matrix not equal to an identity matrix. Thus, the response variable can be used in multivariate analysis while using the MPIGR model.

Table 3. Overdispersion test.

Variable	Deviance	df	Deviance/df
Number of infant mortality (Y_1)	4462.60	102	43.75
Number of child mortality (Y_2)	2181.11	102	21.38
Number of maternal mortality (Y_3)	670.05	102	6.57

The significance of the simultaneous test shows that the statistics $G^2 = 39.86 \times 10^8$ is higher than $\chi^2_{0.05,25} = 37.653$; hence, the decision to reject the null hypothesis. It means that there is at least one predictor variable that significantly influences the number of infants, child, and maternal mortality.

The partial hypothesis testing is done in order to determine the significant predictor variables that are influencing the number of infants, child, and maternal mortality in Java. Table 4 shows the estimation results of MPIGR model parameters.

Table 4. Parameter estimation of each predictor at each response variables.

Parameter	The Number of Infant Mortality				The Number of Under-Five Children Mortality				The Number of Maternal Mortality			
	Est	Se	Z	P	Est	Se	Z	P	Est	Se	Z	P
β_0	4.101	5.27×10^{-4}	-7.78×10^3	$p < 0.001$	-2.549	5.17×10^{-3}	4.92×10^2	$p < 0.001$	3.613	3.59×10^{-2}	-1.00×10^2	$p < 0.001$
β_1	-0.032	7.40×10^{-8}	4.31×10^5	$p < 0.001$	-0.072	7.24×10^{-7}	9.89×10^4	$p < 0.001$	0.014	1.06×10^{-5}	-1.34×10^3	$p < 0.001$
β_2	0.004	6.23×10^{-8}	6.16×10^4	$p < 0.001$	-0.018	6.15×10^{-7}	-2.87×10^4	$p < 0.001$	0.007	3.05×10^{-6}	2.31×10^3	$p < 0.001$
β_3	-0.003	5.16×10^{-9}	-5.33×10^5	$p < 0.001$	0.001	5.03×10^{-9}	-2.64×10^5	$p < 0.001$	-0.003	4.43×10^{-8}	-6.09×10^4	$p < 0.001$
β_4	-0.076	6.83×10^{-7}	1.11×10^5	$p < 0.001$	-0.449	1.41×10^{-5}	3.18×10^4	$p < 0.001$	-0.119	3.54×10^{-5}	3.36×10^3	$p < 0.001$
β_5	-0.002	3.88×10^{-9}	4.88×10^5	$p < 0.001$	0.005	2.04×10^{-8}	2.24×10^5	$p < 0.001$	-0.005	3.99×10^{-7}	1.28×10^4	$p < 0.001$
β_6	-0.005	3.67×10^{-9}	1.37×10^6	$p < 0.001$	0.019	2.53×10^{-8}	-7.59×10^5	$p < 0.001$	-0.013	2.60×10^{-7}	5.34×10^4	$p < 0.001$
β_7	0.004	6.67×10^{-9}	-6.03×10^5	$p < 0.001$	-0.005	2.02×10^{-8}	-2.57×10^5	$p < 0.001$	0.006	8.32×10^{-7}	-7.68×10^3	$p < 0.001$
β_8	0.041	3.26×10^{-8}	1.24×10^6	$p < 0.001$	0.144	1.94×10^{-6}	7.42×10^4	$p < 0.001$	3.613	3.59×10^{-2}	-1.00×10^2	$p < 0.001$

The estimate of the dispersion parameter (τ) is 0.493 with its $p < 0.001$. Based on the empirical results summarized in Table 3, all of the predictor variables have a significant effect on the three responses. The MPIGR model of these three responses and eight predictors can be written in the following equations:

$$\begin{aligned}\hat{\mu}_{1i} &= q_i \exp(4.101 - 0.032X_{1i} + 0.004X_{2i} - 0.003X_{3i} - 0.076X_{4i} - 0.002X_{5i} - 0.005X_{6i} + 0.004X_{7i} + 0.041X_{8i}) \\ \hat{\mu}_{2i} &= q_i \exp(-2.549 - 0.072X_{1i} - 0.018X_{2i} + 0.001X_{3i} - 0.449X_{4i} + 0.005X_{5i} + 0.019X_{6i} - 0.005X_{7i} + 0.144X_{8i}) \\ \hat{\mu}_{3i} &= q_i \exp(3.613 + 0.014X_{1i} + 0.007X_{2i} - 0.003X_{3i} - 0.019X_{4i} - 0.005X_{5i} - 0.013X_{6i} + 0.006X_{7i} - 0.013X_{8i})\end{aligned}$$

We use the Mean Squared Error (MSE) to measure the difference in the average squared between the estimated and the actual value in order to determine whether the model fits the data well. The Root Mean Squared Error (RMSE) reveals the estimates of standard deviation of each response, where their standard deviation of response observations are reported in Table 1. The MSE and RMSE for the response variables are tabulated in Table 5. It is shown that the RMSE values are close the standard deviation of each response. This empirical results also prove that the predicted responses are relatively very close to the observations values.

Table 5. The Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) for each response variable.

	Y_1	Y_2	Y_3
MSE	4821.05	707.31	104.17
RMSE	69.43	26.59	10.21

In addition to RMSE written in Table 5, the scatter plots of true and prediction values of each response are exhibited in order to show that the MPIGR model is good at predicting the observed data. To support this result, Figure 2 is displayed to see how spread out the residuals are. Based on Figure 2, the fitting values for Y_1 and Y_3 are better than those of Y_2 . This empirical results, of course, can be improved. These findings become the big concerns for the next research that are possibly related to the spatial dependencies among the responses that are discussed in the coming section.

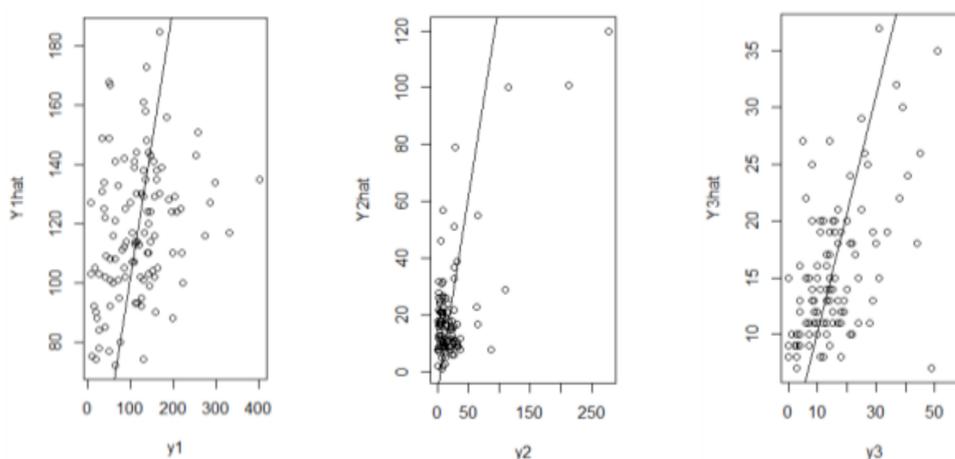


Figure 2. The plot of the actual and the estimate values.

4. Discussion

The result of the significant partial test shows that at a significance level $\alpha = 0.05$, the predictor variables that significantly influence the number of infants, child, and maternal deaths are the percentage of antenatal care visit by pregnant women (X_1), the percentage of pregnant women who received Fe3 tablet (X_2), the percentage of complete neonatal visits (X_3), the percentage of Low Birth Weight

(LBW) (X_4), the percentage of healthy house (X_5), the percentage of active integrated service post (X_6), the percentage of infants received vitamin A (X_7), and the percentage of births that are assisted by health workers (X_8).

According to regression coefficient of the MPIGR model, the percentage of Low Birth Weight (LBW) (X_4) gave the greatest effect to the number of infant mortality (Y_1), the number of under-five children mortality (Y_2), and the number of maternal mortality (Y_3). However, it has inappropriate dependencies with Y_1 , Y_2 , and Y_3 . Based on these results, we need to look at the pattern of the relationship between the predictor and the response variables below.

The pattern of the relationship between the predictor and response variables leads to the conclusion that several predictors have an inappropriate relationship structure. The percentage of antenatal care visit by pregnant women (X_1) has a negative relationship with Y_1 and Y_2 , while it has positive or inappropriate relationship with Y_3 . The percentage of pregnant women who received Fe3 tablet (X_2) has negative dependencies with Y_2 , while it has a positive or inappropriate relationship with Y_1 and Y_3 . On the other side, the percentage of complete neonatal visits (X_3) has negative or appropriate dependencies with Y_1 and Y_3 . Meanwhile, the percentage of Low Birth Weight (LBW) (X_4) has positive dependencies with all of the response variables. However, conflicting results were obtained for all of the response variables.

The percentage of healthy house (X_5) has negative relationship with Y_1 and Y_3 , while it has positive or inappropriate sign with Y_2 . The percentage of active integrated service post (X_6) has negative dependencies with Y_1 and Y_3 , while it has a positive effect on Y_2 . The percentage of infants received vitamin A (X_7) has a positive relationship with Y_1 and Y_3 , while it has negative or appropriate dependencies with Y_2 . The latter one, the percentage of births assisted by health workers (X_8), has negative dependencies with Y_3 , while it has an inappropriate relationship with Y_1 and Y_3 . This finding means that, even though all predictor variables are statistically significant, not all of them have an appropriate relationship with all of the response variables.

The results of this study lead us to more deeply investigate the appropriate method for modeling the data. We assume that the spatial aspect needs to be added to the modeling. Thus, to support our assumptions, a spatial heterogeneity test was carried out in order to test the differences in the characteristics between one point of observation and another. The test statistics used the Glejser test with the null hypothesis (H_0) is no spatial heterogeneity and the alternative hypothesis (H_1) is that there is spatial heterogeneity. The results of the test is $G^2 = 1309.84$, which is higher than $\chi^2_{0.05,24} = 36.415$. Therefore, the decision is to reject H_0 , which means that the response variables have a spatial heterogeneity and can be modelled using a spatial model in future work. The local model, for example, geographically weighted regression, for MPIGR will be the big concern for our future research.

5. Conclusions

The Multivariate Poisson Inverse Gaussian Regression (MPIGR) is the development of the Poisson Inverse Gaussian univariate regression (PIGR) model. The MPIGR model that is proposed in this research accommodates the exposure variable within the model, as the quality of each observation unit level, to overcome the overdispersion problem. The parameter estimation is performed using the Maximum Likelihood Estimation (MLE), followed by the Newton–Raphson iteration. Meanwhile, to simultaneously and partially examine the MPIGR model, the Maximum Likelihood Ratio Test (MLRT) method is used.

The empirical results of the application, where the unit of observation is city or municipality, showed that all eight predictors affect the three response variables, i.e., number of infants, under-five children, and maternal mortality in Java, Indonesia. The predictors are the percentage of antenatal care visits of pregnant women, the percentage of pregnant women receiving Fe3 tablets, the percentage of neonatal visits complete, the percentage of low birth weight babies (LBW), the percentage of healthy

homes, the percentage of active integrated service post, the percentage of babies receiving vitamin A, and the percentage of deliveries assisted by health workers.

According to computational problems that were suffered by factorial calculation in Bessel function, the factorial simplification in the third modification of Bessel function was done to avoid the large numbers incomputable by a computer. The proposed simplification procedure is very crucial to make the fitting of the MPIGR models computable. The application of the MPIGR model to real data in this study is limited to three-count response variables. Based on the result of the spatial heterogeneity test, the spatial aspects are needed to model the data with the spatial model in future research.

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

Funding: This research and the APC were funded by the Ministry of Education and Culture (Kemendikbud) of the Republic of Indonesia with grant number 1279/PKS/ITS/2020.

Acknowledgments: All authors thank the editor and reviewers for the improvement of this paper through criticism and suggestions provided.

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

Nomenclature

PIG	Poisson Inverse Gaussian
PIGD	Poisson Inverse Gaussian Distribution
PIGR	Poisson Inverse Gaussian Regression
MPIGD	Multivariate Poisson Inverse Gaussian Distribution
MPIGR	Multivariate Poisson Inverse Gaussian Regression
MLE	Maximum Likelihood Function
MLRT	Maximum Likelihood Ratio Test
CoV	Coefficient of Variation
LBW	Low Birth Weight

Appendix A

The statistical test of the hypothesis for MPIGR model based on the likelihood ratio test method is formulated as follows:

$$\begin{aligned}
 G^2 &= 2(\log L(\hat{\Omega}) - \log L(\hat{\omega})) \\
 &= 2 \left[\frac{n}{\hat{\tau}} - \frac{n}{2} \log \hat{\tau} + \frac{n}{2} \log \left(\frac{2}{\pi} \right) + \sum_{i=1}^n \log(K_{s_i}(z_i)) - \sum_{i=1}^n \left(\frac{2 \sum_{j=1}^3 y_{ij} - 1}{4} \right) \log \left(1 + 2\tau \sum_{j=1}^3 q_i e^{\mathbf{x}_i^T \hat{\beta}_j} \right) \right. \\
 &\quad + \log \left(\sum_{i=1}^n \sum_{j=1}^3 q_i \log(y_{ij}) \right) + \sum_{i=1}^n \sum_{j=1}^3 y_{ij} \mathbf{x}_i^T \hat{\beta}_j - \log \left(\sum_{i=1}^n \sum_{j=1}^3 y_{ij}! \right) - \left\{ \frac{n}{\hat{\tau}_\omega} - \frac{n}{2} \log \hat{\tau}_\omega + \frac{n}{2} \log \left(\frac{2}{\pi} \right) \right. \\
 &\quad + \sum_{i=1}^n \log(K_{s_i}(z_{i\hat{\omega}})) - \sum_{i=1}^n \left(\frac{2 \sum_{j=1}^3 y_{ij} - 1}{4} \right) \log K_{s_i}(z_i) \left(1 + 2\hat{\tau}_\omega \sum_{j=1}^3 q_i e^{\hat{\beta}_{0j\hat{\omega}}} \right) + \log \left(\sum_{i=1}^n \sum_{j=1}^3 q_i \log(y_{ij}) \right) \\
 &\quad \left. \left. + \sum_{i=1}^n \sum_{j=1}^m y_{ij} \hat{\beta}_{0j\hat{\omega}} - \sum_{i=1}^n \sum_{j=1}^m \log y_{ij}! \right\} \right]
 \end{aligned}$$

$$\begin{aligned}
G^2 &= 2 \left[\frac{n}{\hat{\tau}} - \frac{n}{\hat{\tau}_{\hat{\omega}}} + \frac{n}{2} \log \frac{\hat{\tau}_{\hat{\omega}}}{\hat{\tau}} + \log \sum_{i=1}^n \left(\frac{K_{S_i}(z_i)}{K_{S_i}(z_{i\hat{\omega}})} \right) + \sum_{i=1}^n \frac{\left(2 \sum_{j=1}^3 y_{ij} - 1 \right)}{4} \log \left(\frac{1 + 2\tau \sum_{j=1}^3 q_i e^{\hat{\beta}_{0j\hat{\omega}}}}{1 + 2\tau \sum_{j=1}^3 q_i e^{x_i^T \hat{\beta}_j}} \right) \right. \\
&\quad \left. + \sum_{i=1}^n \sum_{j=1}^3 y_{ij} x_i^T \hat{\beta}_j - \sum_{i=1}^n \sum_{j=1}^m y_{ij} \hat{\beta}_{0j\hat{\omega}} \right] \\
&= 2 \left[\frac{n}{\hat{\tau}} - \frac{n}{\hat{\tau}_{\hat{\omega}}} - \frac{n}{2} \log \hat{\tau} + \frac{n}{2} \log \hat{\tau}_{\hat{\omega}} + \frac{n}{2} \log \left(\frac{\hat{\tau}}{\pi} \right) - \frac{n}{2} \log \left(\frac{\hat{\tau}_{\hat{\omega}}}{\pi} \right) + \sum_{i=1}^n \log(K_{S_i}(z_i)) - \sum_{i=1}^n \log(K_{S_i}(z_{i\hat{\omega}})) \right. \\
&\quad \left. - \sum_{i=1}^n \left(\frac{2 \sum_{j=1}^3 y_{ij} - 1}{4} \right) \log \left(1 + 2\tau \sum_{j=1}^3 q_i e^{x_i^T \hat{\beta}_j} \right) + \sum_{i=1}^n \frac{\left(2 \sum_{j=1}^3 y_{ij} - 1 \right)}{4} \log \left(1 + 2\tau \sum_{j=1}^3 q_i e^{\hat{\beta}_{0j\hat{\omega}}} \right) \right. \\
&\quad \left. + \log \left(\sum_{i=1}^n \sum_{j=1}^3 q_i \log(y_{ij}) \right) - \log \left(\sum_{i=1}^n \sum_{j=1}^3 q_i \log(y_{ij}) \right) + \sum_{i=1}^n \sum_{j=1}^3 y_{ij} x_i^T \hat{\beta}_j - \sum_{i=1}^n \sum_{j=1}^m y_{ij} \hat{\beta}_{0j\hat{\omega}} \right. \\
&\quad \left. - \log \left(\sum_{i=1}^n \sum_{j=1}^3 y_{ij}! \right) + \log \left(\sum_{i=1}^n \sum_{j=1}^m y_{ij}! \right) \right] \\
&= 2 \left[\frac{n}{\hat{\tau}} - \frac{n}{\hat{\tau}_{\hat{\omega}}} - \frac{n}{2} \log \hat{\tau} + \frac{n}{2} \log \hat{\tau}_{\hat{\omega}} + \sum_{i=1}^n \log(K_{S_i}(z_i)) - \sum_{i=1}^n \log(K_{S_i}(z_{i\hat{\omega}})) \right. \\
&\quad \left. - \sum_{i=1}^n \frac{\left(2 \sum_{j=1}^3 y_{ij} - 1 \right)}{4} \log \left(1 + 2\tau \sum_{j=1}^3 q_i e^{x_i^T \hat{\beta}_j} \right) + \sum_{i=1}^n \frac{\left(2 \sum_{j=1}^3 y_{ij} - 1 \right)}{4} \log \left(1 + 2\tau \sum_{j=1}^3 q_i e^{\hat{\beta}_{0j\hat{\omega}}} \right) \right. \\
&\quad \left. + \sum_{i=1}^n \sum_{j=1}^3 y_{ij} x_i^T \hat{\beta}_j - \sum_{i=1}^n \sum_{j=1}^m y_{ij} \hat{\beta}_{0j\hat{\omega}} \right]
\end{aligned}$$

References

1. Consul, P.; Famoye, F. Generalized poisson regression model. *Commun. Stat. Theory Methods* **1992**, *21*, 89–109. [\[CrossRef\]](#)
2. Holla, M.S. On a poisson-inverse gaussian distribution. *Metr. Int. J. Theor. Appl. Stat.* **1967**, *11*, 115–121. [\[CrossRef\]](#)
3. Dean, C.; Lawless, J.F.; Willmot, G.E. A mixed poisson-inverse-gaussian regression model. *Can. J. Stat.* **1989**, *17*, 171–181. [\[CrossRef\]](#)
4. Hilbe, J.M. Poisson Inverse Gaussian Regression. In *Modeling Count Data*; Cambridge University Press (CUP): Cambridge, UK, 2014; pp. 162–171.
5. Karlis, D.; Xekalaki, E. A Simulation Comparison of Several Procedures for Testing the Poisson Assumption. *J. R. Stat. Soc. Ser. D Stat.* **2000**, *49*, 355–382. [\[CrossRef\]](#)
6. Ouma, V.M. Poisson Inverse Gaussian (PIG) Model for Infectious Disease Count Data. *Am. J. Theor. Appl. Stat.* **2016**, *5*, 326. [\[CrossRef\]](#)
7. Xie, F.-C.; Wei, B.-C. Influence analysis for Poisson inverse Gaussian regression models based on the EM algorithm. *Metrika* **2007**, *67*, 49–62. [\[CrossRef\]](#)
8. Zha, L.; Lord, D.; Zou, Y. The Poisson inverse Gaussian (PIG) generalized linear regression model for analyzing motor vehicle crash data. *J. Transp. Saf. Secur.* **2014**, *8*, 18–35. [\[CrossRef\]](#)
9. Ghitany, M.E.; Karlis, D. An EM Algorithm for Multivariate Mixed Poisson. *Appl. Math. Sci.* **2012**, *6*, 6843–6856.
10. United Nations Inter-agency Group for Child Mortality Estimation (UN IGME). *Levels & Trends in Child Mortality: Report 2018, Estimates Developed by the United Nations Inter-Agency Group for Child Mortality Estimation*; United Nations Children's Fund: New York, NY, USA, 2018.
11. WHO; UNICEF; UNFPA; The World Bank. *Maternal Mortality: Level and Trends 2000 to 2017*; World Health Organization: Geneva, Switzerland, 2019; ISBN 978-92-4-151648-8.
12. Moucheraud, C.; Worku, A.; Molla, M.; Finlay, J.E.; Leaning, J.; Yamin, A.E. Consequences of maternal mortality on infant and child survival: A 25-year longitudinal analysis in Butajira Ethiopia (1987–2011). *Reprod. Health* **2015**, *12*, S4. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Miaou, S.P.; Lord, D. Modeling Traffic Crash–Flow Relationships for Intersections Dispersion Parameter, Functional Form, and Bayes Versus Empirical Bayes Methods. *Transp. Res. Rec. J. Transp. Res. Board* **2003**, *1840*, 31–40. [\[CrossRef\]](#)

14. Willmot, G.E. The Poisson-Inverse Gaussian distribution as an alternative to the negative binomial. *Scand. Actuar. J.* **1987**, *1987*, 113–127. [[CrossRef](#)]
15. Stein, G.Z.; Juritz, J.M. Bivariate compound poisson distributions. *Commun. Stat. Theory Methods* **1987**, *16*, 3591–3607. [[CrossRef](#)]
16. Stein, G.Z.; Zucchini, W.; Juritz, J.M. Parameter estimation for the Sichel distribution and its multivariate extension. *J. Am. Stat. Assoc.* **1987**, *82*, 938–944. [[CrossRef](#)]

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



© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).