Article

A Comparison of Bivariate Zero-Inflated Poisson Inverse Gaussian Regression Models with and without Exposure Variables

Ermawati Ermawati 1,2, Purhadi Purhadi 1,* and Santi Puteri Rahayu 1

1 Faculty of Science and Data Analytics, Department of Statistics, Institut Teknologi Sepuluh Nopember, Surabaya 60111, Indonesia; ermawati@uin-alauddin.ac.id (E.E.); santi_pr@statistika.its.ac.id (S.P.R.)
2 Faculty of Science and Technology, Department of Mathematics, Universitas Islam Negeri Alauddin Makassar, Makassar 92118, Indonesia
*
Correspondence: purhadi@statistika.its.ac.id

Abstract: In this paper, we focus on the comparison of the bivariate zero-inflated Poisson inverse Gaussian regression (BZIPIGR) type II model in two cases: with and without exposure variables. The BZIPIGR type II model is applied to analyze the occurrence of maternal and early neonatal mortality in South Sulawesi Province, Indonesia using 2019 data, which contain many zero values and have the issue of overdispersion in the response variable. Furthermore, to analyze the number of deaths in various areas, the exposure variable is considered. The maximum likelihood estimation (MLE) is used in parameter estimation, which involves numerical iteration and application of the Berndt–Hall–Hall–Hausman (BHHH) algorithm. Sum square error (SSE) serves as the criterion of model selection when exposure variables are included. The existence of exposure variables strongly affects the model’s accuracy, especially using the BZIPIGR type II model. According to the SSE and RMSE values, the BZIPIGR type II model with exposure variables performs better than the model without exposure variables in estimating parameter values. All predictors with exposure variables in this study had a significant influence on the number of maternal and early neonatal mortalities.

Keywords: BZIPIGR type II; overdispersion; excess zero; exposure variable; MLE; BHHH; number of mortalities

1. Introduction

Count data that follow a Poisson distribution can be modeled by Poisson regression. Poisson regression requires the assumption that the mean and variance of the response variables must be equal (equidispersion), i.e., one parameter gives both the mean and variance of the distribution. In reality, this assumption may not hold, either with a variance larger than the mean (overdispersion) or otherwise (underdispersion) [1–5]. Such a violation can result in errors in decision making in hypothesis testing due to the occurrence of underestimates [6–8]. A model was developed based on mixed Poisson distribution to overcome the overdispersion problem, constituting a combination of Poisson distribution and other distributions, both discrete and continuous, including Poisson Inverse Gaussian (PIG) distribution [9].

Poisson regression can be used to model count data, but between events within a certain period, the data must be independent. For example, maternal mortality at a certain time does not affect maternal mortality at that time or another time. However, the number of mothers who die within a certain period in a given area may be influenced by other factors; for example, more births results in a higher occurrence of maternal mortality in an area. Thus, when comparing the number of mortalities between one area and another, it is necessary to consider the existence of other variables that can affect the number of occurrences, referred to as the exposure variable [10]. Frome and Morris (1989) [11] conducted research involving exposure variables with Poisson regression modeling. Violation of the equidispersion assumption, especially overdispersion within
the Poisson distribution, can be caused by the number of zero values (excess zero) in the response variable [11–13].

Phang and Loh recommend a zero-inflated Poisson inverse Gaussian regression (ZIP-IGR) model for data with excess zeros and overdispersion [14]. Zero-inflated data contains many zeros in the response variable. Hilbe conducted simulations with data containing excess zeros and overdispersion. Several models were included in this simulation, such as zero-inflated negative binomial regression (ZINBR), zero-inflated hurdle negative binomial regression (ZIHNBR) and zero-inflated generalized regression (ZIGPR). The simulation obtained the lowest AIC and BIC values with the ZIPIGR model, so the conclusion is that the ZIPIGR model is preferable to the other models [5]. Several other researchers have also used ZIPIGR to model with count data containing excess zeros and overdispersion [15,16].

The ZIPIGR model is limited to univariate cases; therefore, if the relevant application has more than one response variable, excess zeros or overdispersion, this model cannot be used. In this study, we used two correlated response variables: the number of maternal and early neonatal mortalities. Maternal mortality markedly affects the survival of the child, especially in cases of babies younger than one month old (early neonatal), where the basic needs of the baby are dependent on the mother, such as breast feeding. More than 50% of response variables in this study contained zero values and exhibited overdispersion. Therefore, a bivariate zero-inflated Poisson inverse Gaussian regression (BZIPIGR) model was developed.

The BZIPIGR type I model consists of two probability functions that accommodate the state of the response-variable pairs: all response variables are zero, and all response variables are not zero. However, not all cases consist of two response conditions. In the case of the number of maternal and early neonatal mortalities in South Sulawesi Province in 2018, there were four combinations of zero values in the response variable. Therefore, to identify the factors that significantly affected the number of maternal and early neonatal mortalities, a BZIPIGR type II model was developed to accommodate all possible combinations of zero values in the response variable. This is similar to the formation of the probability function in the BZIGPR model [17].

Several researchers have used Newton-Raphson iterations to maximize the likelihood function using the MLE method when the derivative of the likelihood function for the model parameters was in non-closed form. This algorithm requires a second derivative to form a Hessian matrix. Therefore, for a more complex probability distribution function, it is difficult to obtain a second derivative. According to Nugraha, the first-derivative approach can be used to inform the Hessian matrix using a Berndt–Hall–Hall–Hausman (BHHH) iteration [18]. The complexity of the probability function in the BZIPIGR type II model makes it difficult to obtain the second derivative. Therefore, we used the BHHH iteration method as an alternative to overcome this problem.

2. Bivariate Zero-Inflated Poisson Inverse Gaussian Regression Type II (BZIPIGR)

The BZIPIGR type II model is derived from the ZIPIGR model. The ZIPIGR model is a mixed of zero-inflated (ZI) and Poisson inverse Gaussian (PIG) distributions [16]. The zero value in the ZIPIG distribution can be divided into two sources, namely structural zero (zero state) with probability \( p_0 \) and sampling zero (mixed Poisson distribution state or PIG state) with probability \( 1 - p_0 \). If there are two random variables, \( Y_1 \) and \( Y_2 \), with BZIPIG distribution with parameters \( (\eta_1, \gamma_1, \tau) \) where \( h = 1, 2 \) or \( (Y_1, Y_2) \sim BZIPIG(\eta_1, \eta_2, \tau, \gamma_1, \gamma_2) \), and where \( \eta_1 \) and \( \eta_2 \) are the parameters of the Poisson distribution; \( \tau \) is the dispersion parameter, and \( \gamma_1, \gamma_2 \) is the parameter of the zero-inflated distribution, the probability function of \( Y_1 \) and \( Y_2 \) can be obtained with:

\[
(Y_1, Y_2) \sim (0, 0) \text{ with } p_0 \text{ probability} \\
\sim (\text{PIG}(\eta_1, \tau), 0) \text{ with } p_1 \text{ probability} \\
\sim (0, \text{PIG}(\eta_2, \tau)) \text{ with } p_2 \text{ probability} \\
\sim (\text{BPIG}(\eta_1, \eta_2, \tau)) \text{ with } p_3 \text{ probability}
\]
where \( p_3 = 1 - p_2 - p_1 - p_0 \). Hence, the joint probability function of \( Y_1 \) and \( Y_2 \) can be written as

\[
P(Y_1 = y_1, Y_2 = y_2) = \begin{cases} 
F_1 & \text{if } y_1 = 0, y_2 = 0 \\
F_2 & \text{if } y_1 = 0, y_2 > 0 \\
F_3 & \text{if } y_1 > 0, y_2 = 0 \\
F_4 & \text{if } y_1 > 0, y_2 > 0 
\end{cases}
\]

(1)

where

\[
F_1 = \gamma_1 \gamma_2 + \gamma_1 (1 - \gamma_2) \exp \left( \frac{1}{\tau} \left( 1 - \sqrt{1 + 2 \tau \eta_1} \right) \right) + \gamma_2 (1 - \gamma_1) \exp \left( \frac{1}{\tau} \left( 1 - \sqrt{1 + 2 \tau \eta_2} \right) \right) + (1 - \gamma_1)(1 - \gamma_2) \exp \left( \frac{1}{\tau} \left( 1 - \sqrt{1 + 2 \tau (\eta_1 + \eta_2)} \right) \right)
\]

(2)

\[
F_2 = \gamma_1 (1 - \gamma_2) \left( \frac{2}{\pi \tau} \sqrt{1 + \eta_2 \tau} \right) \frac{\sqrt{\eta_1} \exp \left( \frac{1}{2} \left( 1 - \sqrt{1 + 2 \tau \eta_1} \right) \right) K_{\frac{1}{2}} \left( \frac{1}{\tau} \sqrt{1 + 2 \tau \eta_1} \right)}{(\sqrt{2 \pi \eta_1 + 1})^\frac{1}{2} \eta_1!} + (1 - \gamma_1)(1 - \gamma_2) \left( \frac{2}{\pi \tau} \sqrt{1 + 2 \tau (\eta_1 + \eta_2)} \right) \frac{\sqrt{\eta_2} \exp \left( \frac{1}{2} \left( 1 - \sqrt{1 + 2 \tau (\eta_1 + \eta_2)} \right) \right) K_{\frac{1}{2}} \left( \frac{1}{\tau} \sqrt{1 + 2 \tau (\eta_1 + \eta_2)} \right)}{(\sqrt{1 + 2 \tau (\eta_1 + \eta_2)})^\frac{1}{2} \eta_2!}
\]

(3)

where \( K_{\frac{1}{2}} \) is the modified Bessel function of the third kind [19]. Therefore, Equation (1) becomes

\[
F_1 = \gamma_1 \gamma_2 + \gamma_1 (1 - \gamma_2) \exp \left( \frac{1}{\tau} \left( 1 - \sqrt{1 + 2 \tau \eta_2} \right) \right) + \gamma_2 (1 - \gamma_1) \exp \left( \frac{1}{\tau} \left( 1 - \sqrt{1 + 2 \tau \eta_1} \right) \right) + (1 - \gamma_1)(1 - \gamma_2) \exp \left( \frac{1}{\tau} \left( 1 - \sqrt{1 + 2 \tau (\eta_1 + \eta_2)} \right) \right)
\]

(4)

\[
F_3 = \gamma_2 (1 - \gamma_1) \left( \frac{2}{\pi \tau} \sqrt{1 + \alpha_1 \tau} \right) \frac{\eta_1 \exp \left( \frac{1}{2} \left( 1 - \sqrt{1 + 2 \tau \eta_1} \right) \right) K_{\frac{1}{2}} \left( \frac{1}{\tau} \sqrt{1 + 2 \tau \eta_1} \right)}{(\sqrt{2 \pi \eta_1 + 1})^\frac{1}{2} \eta_1!} + (1 - \gamma_1)(1 - \gamma_2) \left( \frac{2}{\pi \tau} \sqrt{1 + 2 \tau (\eta_1 + \eta_2)} \right) \frac{\eta_2 \exp \left( \frac{1}{2} \left( 1 - \sqrt{1 + 2 \tau (\eta_1 + \eta_2)} \right) \right) K_{\frac{1}{2}} \left( \frac{1}{\tau} \sqrt{1 + 2 \tau (\eta_1 + \eta_2)} \right)}{(\sqrt{1 + 2 \tau (\eta_1 + \eta_2)})^\frac{1}{2} \eta_2!}
\]

(5)

where \( K_{\frac{1}{2}} \) is the modified Bessel function of the third kind [20] and

\[
F_4 = \gamma_1 (1 - \gamma_2) \left( \frac{2}{\pi \tau} \sqrt{1 + 2 \tau \eta_1 \tau} \right) \frac{\eta_1 \eta_2 \exp \left( \frac{1}{2} \left( 1 - \sqrt{1 + 2 \tau (\eta_1 + \eta_2)} \right) \right) K_{\frac{1}{2}} \left( \frac{1}{\tau} \sqrt{1 + 2 \tau (\eta_1 + \eta_2)} \right)}{(\sqrt{1 + 2 \tau (\eta_1 + \eta_2)})^\frac{1}{2} \eta_1 \eta_2!}
\]

(6)

where \( K_{\frac{1}{2} + \frac{1}{2}} \) is the modified Bessel function of the third kind [20].

2.1. BZIPIGR Type II Model with Exposure Variables

The regression model for the BZIPIG distribution is divided into two equations: the regression model with exposure variable, as in Equation (7) [10]:

\[
\eta_{hi} = \eta_{hi} \exp \left( x_i^T \lambda_h \right); \ h = 1, 2; \ i = 1, 2, \ldots, n
\]

(7)
where $q_{hi}$ is the exposure variable for the $h$-th response variable and the $i$-th observation; $x$ is the vector of the predictor variable, and $\lambda$ is the model parameters. The model for the Zi distribution can be written as Equation (9):

$$\log \left( \frac{\gamma_h}{1-\gamma_h} \right) = \exp (x_i^T a_h)$$

(8)

$$\gamma_h = \frac{\exp (x_i^T a_h)}{1+\exp (x_i^T a_h)}$$

and

$$1 - \gamma_h = \frac{1}{1+\exp (x_i^T a_h)}$$

where $a$ is the vector of the zero-inflated parameters.

Based on the model in Equations (7) and (8), the probability function of the BZIPIGR type II model with exposure variable can be written as follows:

$$P(Y_{1i} = y_{1i}, Y_{2i} = y_{2i}) = \begin{cases} P_{1i} & ; \ y_{1i} = 0, \ y_{2i} = 0 \\ P_{2i} & ; \ y_{1i} = 0, \ y_{2i} > 0 \\ P_{3i} & ; \ y_{1i} > 0, \ y_{2i} = 0 \\ P_{4i} & ; \ y_{1i} > 0, \ y_{2i} > 0 \end{cases}$$

(9)

$$P_{1i} = \frac{\exp (x_1^T a_1) \left( \frac{2}{\sqrt{\pi}} \right) \left( \frac{1}{\sqrt{1+2Tq \exp (x_i^T a_1) + \exp (x_i^T a_2)}} \right)}{\left( 1 + \exp (x_1^T a_1) \right) \left( 1 + \exp (x_i^T a_2) \right) \left( \frac{2}{\sqrt{1+2Tq \exp (x_i^T a_1) + \exp (x_i^T a_2)}} \right)}$$

$$P_{2i} = \frac{\exp (x_1^T a_2) \left( \frac{2}{\sqrt{\pi}} \right) \left( \frac{1}{\sqrt{1+2Tq \exp (x_i^T a_1) + \exp (x_i^T a_2)}} \right)}{\left( 1 + \exp (x_1^T a_2) \right) \left( 1 + \exp (x_i^T a_2) \right) \left( \frac{2}{\sqrt{1+2Tq \exp (x_i^T a_1) + \exp (x_i^T a_2)}} \right)}$$

$$P_{3i} = \frac{\exp (x_1^T a_1) \left( \frac{2}{\sqrt{\pi}} \right) \left( \frac{1}{\sqrt{1+2Tq \exp (x_i^T a_1) + \exp (x_i^T a_2)}} \right)}{\left( 1 + \exp (x_1^T a_1) \right) \left( 1 + \exp (x_i^T a_2) \right) \left( \frac{2}{\sqrt{1+2Tq \exp (x_i^T a_1) + \exp (x_i^T a_2)}} \right)}$$

$$P_{4i} = \frac{\exp (x_1^T a_1) \left( \frac{2}{\sqrt{\pi}} \right) \left( \frac{1}{\sqrt{1+2Tq \exp (x_i^T a_1) + \exp (x_i^T a_2)}} \right)}{\left( 1 + \exp (x_1^T a_1) \right) \left( 1 + \exp (x_i^T a_2) \right) \left( \frac{2}{\sqrt{1+2Tq \exp (x_i^T a_1) + \exp (x_i^T a_2)}} \right)}$$

2.2. BZIPIGR Type II Model without Exposure Variable

The regression model for the BZIPIGR type II model without exposure variable, as in Equation (10) is [10]

$$\eta_{hi} = \exp (x_i^T a_h) ; \ h = 1, 2 ; i = 1, 2, \ldots, n$$

(10)

where $x$ is the vector of the predictor variable for the $i$-th observation, and $\lambda$ is the model parameters for the $h$-th response variable.

Based on the model in Equations (9) and (10), the probability function of the BZIPIGR type II model without exposure variables can be written as follows:

$$P(Y_{1i} = y_{1i}, Y_{2i} = y_{2i}) = \begin{cases} Q_{1i} & ; \ y_{1i} = 0, \ y_{2i} = 0 \\ Q_{2i} & ; \ y_{1i} = 0, \ y_{2i} > 0 \\ Q_{3i} & ; \ y_{1i} > 0, \ y_{2i} = 0 \\ Q_{4i} & ; \ y_{1i} > 0, \ y_{2i} > 0 \end{cases}$$

(11)
2.3. Parameter Estimation of BZIPGR Type II Model

Parameter estimation using the MLE method begins with taking $n$ random samples, namely $(Y_{1i}, Y_{2i}, X_{1i}, X_{2i}, \ldots, X_{ki})$, as there are four independent models in the BZIPGR type II model, namely $\{y_1 = 0, y_2 = 0\}$, $\{y_1 = 0, y_2 > 0\}$, $\{y_1 > 0, y_2 = 0\}$, and $\{y_1 > 0, y_2 > 0\}$, where the number of observations for each model is $n_1, n_2, n_3$ and $n_4$, respectively. The likelihood equation can be written as follows:

$$L(\lambda_1; \alpha_1; \tau) = \frac{1}{\prod_{i=1}^{n_1} P(Y_{2i} = y_{2i})} \frac{1}{\prod_{i=1}^{n_2} P(Y_{2i} = y_{2i})} \frac{1}{\prod_{i=1}^{n_3} P(Y_{2i} = y_{2i})} \frac{1}{\prod_{i=1}^{n_4} P(Y_{2i} = y_{2i})}$$

(12)

Then, the log likelihood function is determined from Equation (11), and the following equation is obtained:

$$\log L(\lambda_1; \alpha_1; \tau) = \log L_1(\lambda_1; \alpha_1; \tau) + \log L_2(\lambda_1; \alpha_1; \tau) + \log L_3(\lambda_1; \alpha_1; \tau) + \log L_4(\lambda_1; \alpha_1; \tau)$$

(13)

where

$$\log L_1(\lambda_1; \alpha_1; \tau) = \sum_{i=1}^{n_1} \left[ \log \left( \frac{1}{1 + \exp(x_i^T \alpha_1 + x_i^T \alpha_2)} \right) + \log \left( \frac{1}{1 + \sqrt{1 + 2 \tau \exp(x_i^T \lambda_1)}} \right) + \log \left( \frac{1}{1 - \sqrt{1 + 2 \tau \exp(x_i^T \lambda_1)}} \right) \right]$$

$$\log L_2(\lambda_1; \alpha_1; \tau) = \sum_{i=1}^{n_2} \left[ \log \left( \frac{1}{1 + \exp(x_i^T \alpha_1 + x_i^T \alpha_2)} \right) + \log \left( \frac{1}{1 + \sqrt{1 + 2 \tau \exp(x_i^T \lambda_1)}} \right) + \log \left( \frac{1}{1 - \sqrt{1 + 2 \tau \exp(x_i^T \lambda_1)}} \right) \right]$$

$$\log L_3(\lambda_1; \alpha_1; \tau) = \sum_{i=1}^{n_3} \left[ \log \left( \frac{1}{1 + \exp(x_i^T \alpha_1 + x_i^T \alpha_2)} \right) + \log \left( \frac{1}{1 + \sqrt{1 + 2 \tau \exp(x_i^T \lambda_1)}} \right) + \log \left( \frac{1}{1 - \sqrt{1 + 2 \tau \exp(x_i^T \lambda_1)}} \right) \right]$$

$$\log L_4(\lambda_1; \alpha_1; \tau) = \sum_{i=1}^{n_4} \left[ \log \left( \frac{1}{1 + \exp(x_i^T \alpha_1 + x_i^T \alpha_2)} \right) + \log \left( \frac{1}{1 + \sqrt{1 + 2 \tau \exp(x_i^T \lambda_1)}} \right) + \log \left( \frac{1}{1 - \sqrt{1 + 2 \tau \exp(x_i^T \lambda_1)}} \right) \right]$$
\[
\log(L_2(\lambda_2; \alpha_2; \tau)) = \sum_{i=1}^{n} \left( \log \left[ \frac{\exp(x_i^T \alpha_1) (\frac{3}{\sqrt{1 + 2\tau}} \exp(x_i^T \lambda_2))^{1/2} (\exp(x_i^T \lambda_2))^{1/2}}{(1 + \exp(x_i^T \lambda_2)) (1 + \exp(x_i^T \alpha_1))} \right] - 1 \right) y_{2i} > 0
\]

\[
y_{2i} = 0, y_{2i} > 0
\]

\[
\log(L_3(\lambda_3; \alpha_3; \tau)) = \sum_{i=1}^{n} \left( \log \left[ \frac{\exp(x_i^T \alpha_1) (\frac{3}{\sqrt{1 + 2\tau}} \exp(x_i^T \lambda_2))^{1/2} (\exp(x_i^T \lambda_2))^{1/2}}{(1 + \exp(x_i^T \lambda_2)) (1 + \exp(x_i^T \alpha_1))} \right] - \frac{1}{2} \log \left( \frac{\alpha_1^T}{\alpha_2^T} \right) \right)
\]

\[
y_{2i} = 0, y_{2i} > 0
\]

\[
\log(L_4(\lambda_4; \alpha_4; \tau)) = \sum_{i=1}^{n} \left[ - \log(1 + \exp(x_i^T \alpha_1)) + \log(1 + \exp(x_i^T \alpha_2)) + \frac{1}{2} \log \left( \frac{\alpha_1^T}{\alpha_2^T} \right) \right]
\]

\[
y_{2i} = 0, y_{2i} > 0
\]

If \( \theta = [\lambda_1^T \lambda_2^T \alpha_1^T \alpha_2^T \tau]^T \) is a parameter vector of the BZIPIGR type II model, then, to maximize the log likelihood function in Equation (12), these equations are derived for each parameter and equalized to zero. The first derivative obtained for each parameter is in non-closed form (Appendix A), so the maximization process continues with numerical iteration, which uses the Berndt–Hall–Hall–Hausman (BHHH) algorithm according to the following steps [21]:

Step 1. Determine the initial value of each parameter of the BZIPIGR type II model \( \theta = [\lambda_1^T \lambda_2^T \alpha_1^T \alpha_2^T \tau]^T \). The initial values of model parameters are obtained from the regression model parameter values using OLS, whereas the dispersion parameter is obtained from the variance of the PIG distribution.

Step 2. Define gradient vector

\[
\nabla(\theta^{(m)}) = \left[ \left( \frac{\partial L(\theta)}{\partial \lambda_1} \right)^T \left( \frac{\partial L(\theta)}{\partial \lambda_2} \right)^T \left( \frac{\partial L(\theta)}{\partial \alpha_1} \right)^T \left( \frac{\partial L(\theta)}{\partial \alpha_2} \right)^T \left( \frac{\partial L(\theta)}{\partial \tau} \right)^T \right]^T
\]

Step 3. Determine the Hessian matrix (\( H^* \)) for the BHHH algorithm

\[
H^*(\theta^{(m)}) = -\sum_{i=1}^{n} g_i(\theta^{(m)}) g_i(\theta^{(m)})^T
\]

where \( g_i(\theta^{(m)}) \) is the gradient vector at the i-th observation.

Step 4. Start the BHHH iteration using the following formula:

\[
\hat{\theta}^{(m+1)} = \hat{\theta}^{(m)} - H^{-1}(\hat{\theta}^{(m)}) \nabla(\theta^{(m)})
\]
The iteration of the BH algorithm starts at \( m = 0 \) and stops if \( \| \hat{\theta}^{(m+1)} - \hat{\theta}^{(m)} \| \leq \epsilon \) where \( \epsilon \) is a small positive close to 0.

### 2.4. Hypothesis Testing of the BZIPIGR Type II Model

Simultaneous hypothesis testing for the BZIPIGR model type II parameters can be performed using the maximum likelihood ratio test (MLRT) method. The hypothesis to be tested is as follows:

\[
H_0 : \lambda_{b0} = \lambda_{b1} = \ldots = \lambda_{bk} = a_{b0} = a_{b1} = \ldots = a_{bk} = \tau = 0; \ h = 1, 2; \ j = 0, 1, 2, \ldots, k
\]

\[
H_1 : \text{there is at least one of } \lambda_{bj} \neq 0, \ a_{bj} \neq 0, \text{or } \tau \neq 0; \ h = 1, 2; \ j = 1, 2, \ldots, k
\]

Before determining the test statistic to be used, the parameter set under the population \( (\Omega) \) and \( H_0(\omega) \) are defined as follows:

\[
\Omega = \{ \lambda_h, a_h, \tau | -\infty < \lambda_{bj} < \infty, -\infty < a_{bj} < \infty, \tau > 0; h = 1, 2; j = 0, 1, 2, \ldots, k \}
\]

\[
\omega = \{ \lambda_{b0}, a_{b0}, \tau | -\infty < \lambda_{b0} < \infty, -\infty < a_{b0} < \infty, \tau > 0 \}
\]

This simultaneous test uses test statistics

\[
G^2 = - \log \left( \Lambda^2 \right) = 2 \left[ \log L(\hat{\Omega}) - \log L(\hat{\omega}) \right].
\]

which are compared with \( \chi^2_{\delta(a-b)} \), where \( \delta \) is the level of significance; \( a \) is the number of parameters under the population, and \( b \) is the number of parameters under \( H_0 \). \( H_0 \) is rejected if \( G^2 > \chi^2_{\delta(a-b)} \).

Partial hypothesis testing to determine the significance of the effect of each predictor in the model is divided into two tests, namely a partial test of parameters \( \lambda_{bj} \) and \( a_{bj} \). The hypothesis statistical test for each parameter is as follows:

a. Partial test of parameter \( \lambda_{bj} \)

\[
H_0 : \lambda_{bj} = 0; \ h = 1, 2; \ j = 0, 1, 2, \ldots, k
\]

\[
H_1 : \lambda_{bj} \neq 0
\]

The test statistic is \( Z_{\lambda_{bj}} = \frac{\lambda_{bj}}{se(\hat{\lambda}_{bj})} \). Reject \( H_0 \) if \( |Z_{\lambda_{bj}}| > Z_{\delta/2} \), where \( \delta \) is the level of significance.

b. Partial test of parameter \( a_{bj} \)

\[
H_0 : a_{bj} = 0; \ h = 1, 2; \ j = 0, 1, 2, \ldots, k
\]

\[
H_1 : a_{bj} \neq 0
\]

The test statistic is \( Z_{a_{bj}} = \frac{a_{bj}}{se(\hat{a}_{bj})} \). Reject \( H_0 \) if \( |Z_{a_{bj}}| > Z_{\delta/2} \), where \( \delta \) is the level of significance.

c. Partial test of parameter \( \tau \)

\[
H_0 : \tau = 0;
\]

\[
H_1 : \tau \neq 0
\]

The test statistic is \( Z_{\tau} = \frac{\tau}{se(\hat{\tau})} \). Reject \( H_0 \) if \( |Z_{\tau}| > Z_{\delta/2} \), where \( \delta \) is the level of significance.

### 3. Application

The BZIPIGR Type II model was applied to analyze the number of maternal and early neonatal mortalities in South Sulawesi Province, Indonesia, in 2019. Modeling in this study was carried out for two cases: with and without exposure variables. The response variables used in this study were two correlated variables: the number of maternal
mortalities ($Y_1$) and the number of early neonatal mortalities ($Y_2$). There were six predictor variables: the percentage of visits by pregnant women ($X_1$), the percentage of pregnant women who received Fe3 tablets ($X_2$), percentage of births assisted by health workers ($X_3$), the percentage of mother who attended at least three postpartum maternal visits ($X_4$), the percentage of active integrated service posts ($X_5$) and the percentage of obstetric complications ($X_6$). The exposure variable used in this study was the number of live births in each sub-district ($q$).

Table 1 shows the descriptive statistics of the response variables and predictor variables. The average number of maternal and early neonatal mortalities was 0.587 and 0.533, respectively. This indicates that the maternal mortality rate was higher than the early neonatal mortality rate in six districts in South Sulawesi Province, Indonesia. Based on the coefficient of variance (CoV) value, maternal mortality was more homogeneous than early neonatal mortality because the CoV value of variable $Y_1$ was lower than that of variable $Y_2$.

Table 1. Descriptive statistics of variables.

| Variable ($n = 75$)                                      | Mean | SD   | Coefficient of Variance | Min | Max |
|--------------------------------------------------------|------|------|-------------------------|-----|-----|
| Number of maternal mortalities ($Y_1$)                 | 0.59 | 1.02 | 172.95                  | 0   | 6   |
| Number of early neonatal mortalities ($Y_2$)           | 0.53 | 1.14 | 214.3                   | 0   | 8   |
| Percentage of visits by pregnant women ($X_1$)         | 91.39| 8.14 | 8.91                    | 68.36| 100 |
| Percentage of pregnant women who received Fe3 tablets ($X_2$) | 82.31| 24.83| 30.16                   | 15.31| 127.76|
| Percentage of births assisted by health workers ($X_3$) | 91.1 | 9.15 | 10.05                   | 49.34| 100 |
| Percentage of mother who attended at least three postpartum maternal visits ($X_4$) | 76.54| 19.34| 25.26                   | 39.45| 100 |
| Percentage of active integrated service posts ($X_5$)  | 73.44| 35.71| 48.63                   | 0   | 100 |
| Percentage of obstetric complications ($X_6$)          | 72.03| 29.09| 40.39                   | 7.27 | 128.31|

Before applying the BZIPIGR type II model, it is necessary to test the overdispersion assumption and calculate the percentage of zero values in the response variable. Overdispersion occurs when the variance of the response variable is larger than the mean. This hypothesis can be tested with the following formula [22]:

$$D_h = 2 \sum_{i=1}^{n} Y_{hi} \log \left( \frac{Y_{hi}}{\hat{\eta}_{hi}} \right) - (Y_{hi} - \hat{\eta}_{hi})$$

where $\hat{\eta}_{hi} = \exp \left( \mathbf{x}_i^T \lambda_h \right); h = 1, 2; i = 1, 2, \ldots, n$. Overdispersion occurs if the deviance value divided by the degrees of freedom is larger than one. Table 2 shows that all response variables are overdispersed because the value of $D_h/df$ is greater than one, and there are excess zeros, with more than 61.33% of all observations on $Y_1$ and 68% on $Y_2$ equaling zero. Therefore, the BIZIPIGR type II model should be used to model these data.

Table 2. Overdispersion test and percentage of zero values.

| Variable                                   | % of Zero Value | $D_h$ | df | $D_h/df$ |
|--------------------------------------------|-----------------|-------|----|---------|
| Number of maternal mortalities ($Y_1$)     | 61.33           | 88.637| 68 | 1.303   |
| Number of early neonatal mortalities ($Y_2$)| 68              | 99.099| 68 | 1.457   |

3.1. Modeling the Number of Maternal and Early Neonatal Mortalities Using the BZIPIGR Type II Model Involving Exposure Variables

Based on the results of modeling using the BZIPIGR type II model, the statistical value, $G^2 = 1801.85$, was greater than $\chi^2_{0.95, 24} = 36.415$, so the conclusion was that at least one of the predictor variables influenced the number of maternal and early neonatal mortalities.
in South Sulawesi, Indonesia. The model and partial test statistics for each predictor are presented in Table 3.

### Table 3. Partial test statistics of models with exposure variables.

| Parameter | Estimate | Standard Error | Z     | p-Value |
|-----------|----------|----------------|-------|---------|
| \( \lambda_{01} \) | -7.308   | 0.002          | -3785.456 | \( p < 0.05 \) |
| \( \lambda_{11} \) | 0.005    | 0.002          | 2.860  | \( p < 0.05 \) |
| \( \lambda_{21} \) | -0.006   | 0.0003         | -20.465| \( p < 0.05 \) |
| \( \lambda_{31} \) | 0.043    | 0.002          | 27.838 | \( p < 0.05 \) |
| \( \lambda_{41} \) | -0.018   | 0.0003         | -56.475| \( p < 0.05 \) |
| \( \lambda_{51} \) | -0.002   | 0.0003         | -7.979 | \( p < 0.05 \) |
| \( \lambda_{61} \) | 0.005    | 0.0001         | 25.392 | \( p < 0.05 \) |
| \( \alpha_{01} \) | -5.031   | 0.0004         | -11,937.814| \( p < 0.05 \) |
| \( \alpha_{11} \) | 0.157    | 0.013          | 2.159  | \( p < 0.05 \) |
| \( \alpha_{21} \) | -0.004   | 0.003          | -1.539 | \( p > 0.05 \) |
| \( \alpha_{31} \) | -0.101   | 0.013          | -7.535 | \( p < 0.05 \) |
| \( \alpha_{41} \) | 0.013    | 0.003          | 4.272  | \( p < 0.05 \) |
| \( \alpha_{51} \) | 0.006    | 0.002          | 2.793  | \( p < 0.05 \) |
| \( \alpha_{61} \) | 0.009    | 0.002          | 4.985  | \( p < 0.05 \) |

Based on the results of the partial test in Table 3, all predictor variables were found to have a \( p \)-value < 0.05, so the conclusion was that all predictor variables had a significant effect, both for the variable \( Y_1 \) and \( Y_2 \), with a significance level of 0.05.

### 3.2. Modeling the Number of Maternal and Early Neonatal Mortalities Using the BZIPIGR Type II Model without Exposure Variables

Based on the results of modeling using the BZIPIGR type II model without exposure variables, the statistical value, \( \chi^2 = 776.754 \), was larger than \( \chi^2_{0.95;24} = 36.415 \), so the conclusion was that at least one of the predictor variables had an influence on the number of maternal and early neonatal mortalities in South Sulawesi, Indonesia. The model and partial test statistics for each predictor are presented seen in Table 4.

Based on the results of the partial test in Table 4, several predictor variables have a \( p \)-value > 0.05, namely the \( X_1 \), \( X_2 \), \( X_3 \) and \( X_4 \) variables for the variable \( Y_2 \), so the conclusion was that the effect of these four variables on the number of early neonatal mortalities (\( Y_2 \)) was not significant. The effect of all variables on the number of maternal mortalities (\( Y_1 \)) was significant.

### 3.3. Comparison of the BZIPIGR Type II Model with and without Exposure Variables

The comparison of the BZIPIGR type II model with and without the exposure variables was performed based on the sum square error (SSE) value between the two models. The SSE values of these two models are presented in Table 5.
Table 4. Partial test statistics without exposure variables.

| Parameter | Estimate | Standard Error | Z     | p-Value  |
|-----------|----------|----------------|-------|----------|
| \( \lambda_{01} \) | -11.274 | 0.119 | -94.127 | \( p < 0.05 \) |
| \( \lambda_{11} \) | -0.113 | 0.026 | -4.312 | \( p < 0.05 \) |
| \( \lambda_{21} \) | 0.047 | 0.008 | 5.948 | \( p < 0.05 \) |
| \( \lambda_{31} \) | 0.229 | 0.027 | 8.438 | \( p < 0.05 \) |
| \( \lambda_{41} \) | -0.037 | 0.005 | -7.323 | \( p < 0.05 \) |
| \( \lambda_{51} \) | -0.077 | 0.008 | -11.437 | \( p < 0.05 \) |
| \( \lambda_{61} \) | 0.055 | 0.005 | 12.159 | \( p < 0.05 \) |
| \( \alpha_{01} \) | -12.243 | 0.538 | -22.760 | \( p < 0.05 \) |
| \( \alpha_{11} \) | 0.109 | 0.009 | 11.254 | \( p < 0.05 \) |
| \( \alpha_{21} \) | 0.096 | 0.003 | 35.748 | \( p < 0.05 \) |
| \( \alpha_{31} \) | -0.138 | 0.009 | -14.912 | \( p < 0.05 \) |
| \( \alpha_{41} \) | 0.073 | 0.004 | 18.467 | \( p < 0.05 \) |
| \( \alpha_{51} \) | -0.084 | 0.003 | -30.094 | \( p < 0.05 \) |
| \( \alpha_{61} \) | 0.087 | 0.003 | 26.189 | \( p < 0.05 \) |

Number of early neonatal mortalities \( (Y_2) \)

| Parameter | Estimate | Standard Error | Z     | p-Value  |
|-----------|----------|----------------|-------|----------|
| \( \lambda_{02} \) | -4.917 | 0.296 | -16.627 | \( p < 0.05 \) |
| \( \lambda_{12} \) | 0.054 | 0.036 | 1.485 | \( p > 0.05 \) |
| \( \lambda_{22} \) | -0.009 | 0.023 | -0.421 | \( p > 0.05 \) |
| \( \lambda_{32} \) | -0.046 | 0.037 | -1.231 | \( p > 0.05 \) |
| \( \lambda_{42} \) | 0.020 | 0.031 | 0.655 | \( p > 0.05 \) |
| \( \lambda_{52} \) | -0.013 | 0.005 | -2.713 | \( p < 0.05 \) |
| \( \lambda_{62} \) | 0.051 | 0.005 | 10.056 | \( p < 0.05 \) |
| \( \alpha_{02} \) | 2.252 | 0.396 | 8.218 | \( p < 0.05 \) |
| \( \alpha_{12} \) | -0.159 | 0.023 | -6.904 | \( p < 0.05 \) |
| \( \alpha_{22} \) | -0.059 | 0.013 | -4.626 | \( p < 0.05 \) |
| \( \alpha_{32} \) | 0.383 | 0.035 | 10.827 | \( p < 0.05 \) |
| \( \alpha_{42} \) | -0.235 | 0.026 | -9.089 | \( p < 0.05 \) |
| \( \alpha_{52} \) | -0.021 | 0.003 | -6.562 | \( p < 0.05 \) |
| \( \alpha_{62} \) | 0.048 | 0.004 | 12.068 | \( p < 0.05 \) |

Table 5. SSE and RMSE values of models with and without exposure variables (real data).

| BZIPGR Type II Model | SSE | RMSE |
|----------------------|-----|------|
| Model with exposure variables | 218.8805 | 1.7811 |
| Model without exposure variables | 620.6658 | 2.9992 |

The above real data analysis obtained similar results to the simulation data analysis. Data were generated as much as \( n = 75 \) and \( n = 100 \), with 50% of response variables containing zero values. Data were analyzed using the BZIPGR type II model under two conditions: with and without exposure variables. The simulation SSE and MSE values are presented in Table 6.

Table 6. SSE and RMSE value of models with and without exposure variables (simulation data).

| \( N \) | BZIPGR Type II Model | SSE | RMSE |
|---------|----------------------|-----|------|
| \( N = 75 \) Model with exposure variables | 314.7814 | 2.1359 |
| Model without exposure variables | 369.6821 | 2.3147 |
| \( N = 100 \) Model with exposure variables | 584.6494 | 2.9109 |
| Model without exposure variables | 4241.603 | 7.8404 |

As shown in Tables 5 and 6, the SSE and RMSE values of the BZIPGR type II model with exposure variables are less than those of the model without exposure variables. Therefore, the model with exposure variables performed better than the model without exposure variables.
4. Discussion

Based on a comparison of the SSE and RMSE values in Tables 5 and 6, the model with exposure variables performed better than the model without exposure variables. This discussion focuses on the BZIPGR type II model with exposure variables. There are two BZIPGR model types II, the log model and the logit model. The log model states that the probability of $Y_i$ from the PIG state is significantly affected by a variable, whereas the logit model states that the probability of $Y_i$ from the zero state is significantly influenced by a variable. The two models for each response variable are as follows:

a. Regression model for the number of maternal mortalities:

$log(\gamma_{1i}) = -7.308 + 0.005X_1i - 0.006X_2i + 0.043X_3i - 0.018X_4i - 0.002X_5i + 0.005X_6i \quad (14)$

$logit(\gamma_{1i}) = -5.031 + 0.157X_1i - 0.004X_2i - 0.101X_3i + 0.013X_4i + 0.006X_5i + 0.009X_6i \quad (15)$

b. Regression model for the number of early neonatal mortalities:

$log(\gamma_{2i}) = -3.532 - 0.036X_1i + 0.066X_2i - 0.117X_3i + 0.007X_4i + 0.008X_5i \quad (16)$

$logit(\gamma_{2i}) = -10.340 - 0.143X_1i - 0.017X_2i + 0.144X_3i - 0.116X_4i - 0.009X_5i - 0.016X_6i \quad (17)$

Equations (13) and (14) show the model for the number of maternal mortalities. In this equation, if other variables are held constant, then every 1% increase in variable $X_1$ results in an increase in the number of maternal mortalities by 0.005 times, with the same explanation for variables $X_3$ and $X_6$. For variable $X_2$, if the other variables are held constant, then every 1% increase in variable $X_2$ results in a decrease in the number of maternal mortalities by 0.006 times, with the same explanation for variables $X_4$ and $X_5$. However, for the logit model in Equation (14), the coefficient values of $X_1$, $X_3$ and $X_6$ are positive. The means that, if other predictor variables are held constant, with every increase in the percentage of visits by pregnant women, the probability that maternal mortality will not occur increases, with the same explanation for $X_3$ (percentage of births assisted by health workers) and $X_6$ (percentage of obstetric complications).

Based on the log and logit models in Equations (15) and (16), the predictor variable that affects the PIG state also affects the zero state (ZI). Based on the coefficient values of the predictor variables, the influence of the predictor variable on the response variable can be interpreted. For example, $X_2$ (percentage of pregnant women who received Fe3 tablets) in Equation (15) has a positive value, which means that the number of early neonatal mortalities increases with the percentage of pregnant women who received Fe3 tablets. Variables $X_4$ (percentage of mother who attended at least three postpartum maternal visits), $X_5$ (percentage of active integrated service posts) and $X_6$ (percentage of obstetric complications) increase with the number of early neonatal mortalities. That means that if the other predictor variables are held constant, then every 1% increase in the $X_2$ variable (percentage of births assisted by health workers) increases the number of maternal mortalities by 0.066 times the number of early neonatal mortalities, with the same explanation for variables $X_4$, $X_5$ and $X_6$. For the logit model in Equation (16), where the coefficient value of the $X_1$ is negative, every increase in the percentage of visits by pregnant women reduces the likelihood of early neonatal mortalities not occurring if other predictor variables are held constant, with the same explanations for $X_2$, $X_4$ and $X_6$.

5. Conclusions

The BZIPGR model is derived from the Poisson regression model and can be used when the equidispersion assumption is violated and excess zeros are present. A violation of this assumption occurs when asymmetry exists between the mean and variance of the response variable, especially when the variance is larger than the mean (overdispersion). The BZIPGR type I model only accommodates the possibility that all response variables are zero or that all response variables are non-zero. Therefore, the BZIPGR type II model was developed to overcome other possibilities of the zero value of the response variable, namely that one of the response variables is zero. The BZIPGR type II model accommodates all possible combinations of zero values in the response variable. Thus, the BZIPGR type II
model accounts for bivariate cases that experience overdispersion and excess zeros, such as maternal and early neonatal mortalities in South Sulawesi Province, Indonesia, in 2019.

According to the SSE and RMSE values of real data in Table 5, the BZIPIGR type II model with exposure variables performed better than the model without exposure variables in estimating the parameter values. This conclusion can be generally accepted based on the results of the analysis using the simulation data in Table 6, where the same results are obtained, namely lower SSE and RMSE values in the BZIPIGR type II model with exposure variables. The existence of the exposure variables in the BZIPIGR type II model strongly affects the model accuracy. The empirical results of modeling the number of maternal and early neonatal mortalities at the sub-district level in the province of South Sulawesi, Indonesia with exposure variables using the BZIPIGR type II model showed that all predictor variables used in this study, i.e., percentage of visits by pregnant women, percentage of pregnant women who received Fe3 tablets, percentage of births assisted by health workers, percentage of mothers who attended at least three postpartum maternal visits, percentage of active integrated service posts and percentage of obstetric complications, had a significant effect on the response variable.

The limited availability of data at the sub-district level means it is possible that there are other predictor variables that were not included in the study. Therefore, for further research, the BZIPIGR type II model can be applied to the same case by replacing or adding other variables or other cases that meet the assumptions of this model. The existence of differences in the characteristics of each region can cause differences in the incidence in each region; therefore, further research using the BZIPIGR type II model can add spatial effects to the model.

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

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Dataset not publicly available.

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

Appendix A

The joint probability function of $Y_1$ and $Y_2$ can be written as in the following equation:

$$P(Y_1 = y_1, Y_2 = y_2) = \begin{cases} F_1 & \text{if } y_1 = 0, y_2 = 0 \\ F_2 & \text{if } y_1 = 0, y_2 > 0 \\ F_3 & \text{if } y_1 > 0, y_2 = 0 \\ F_4 & \text{if } y_1 > 0, y_2 > 0 \end{cases}$$

where

$$F_1 = \gamma_1 \gamma_2 + \gamma_1(1 - \gamma_2) \exp \left( \gamma_2 \left( 1 - \sqrt{1 + 2 \tau \eta_2} \right) + \gamma_1(1 - \gamma_2) \exp \left( \gamma_1 \left( 1 - \sqrt{1 + 2 \tau \eta_1} \right) \right) \right)$$

$$F_2 = \gamma_2 \left( 1 - \gamma_2 \right) \left( \frac{2}{\pi \tau} \sqrt{1 + \eta_2} \right)^{\frac{1}{2}} \eta_2^{\frac{1}{2}} \exp \left( \frac{1}{2} \left( 1 + 2 \tau \eta_2 \right) \right) K_{y_2} \frac{\left( \frac{1}{2} \sqrt{1 + 2 \tau \eta_2} \right)}{\left( 1 - \gamma_1 \right) \left( 1 - \gamma_2 \right)}$$

$$+ \gamma_1 \left( 1 - \gamma_2 \right) \left( \frac{2}{\pi \tau} \sqrt{1 + 2 \tau \eta_1} \right)^{\frac{1}{2}} \eta_1^{\frac{1}{2}} \exp \left( \frac{1}{2} \left( 1 + 2 \tau \eta_1 \right) \right) K_{y_2} \frac{\left( \frac{1}{2} \sqrt{1 + 2 \tau \eta_1} \right)}{\left( 1 - \gamma_1 \right) \left( 1 - \gamma_2 \right)}$$

and

$$F_3 = \gamma_1 \gamma_2 + \gamma_1(1 - \gamma_2) \exp \left( \gamma_1 \left( 1 - \sqrt{1 + 2 \tau \eta_1} \right) \right)$$

$$+ \gamma_2 \left( 1 - \gamma_2 \right) \left( \frac{2}{\pi \tau} \sqrt{1 + \eta_2} \right)^{\frac{1}{2}} \eta_2^{\frac{1}{2}} \exp \left( \frac{1}{2} \left( 1 + 2 \tau \eta_2 \right) \right) K_{y_2} \frac{\left( \frac{1}{2} \sqrt{1 + 2 \tau \eta_2} \right)}{\left( 1 - \gamma_1 \right) \left( 1 - \gamma_2 \right)}$$

$$+ \gamma_1 \left( 1 - \gamma_2 \right) \left( \frac{2}{\pi \tau} \sqrt{1 + 2 \tau \eta_1} \right)^{\frac{1}{2}} \eta_1^{\frac{1}{2}} \exp \left( \frac{1}{2} \left( 1 + 2 \tau \eta_1 \right) \right) K_{y_2} \frac{\left( \frac{1}{2} \sqrt{1 + 2 \tau \eta_1} \right)}{\left( 1 - \gamma_1 \right) \left( 1 - \gamma_2 \right)}$$

$$+ \frac{1}{2} \left( 1 + 2 \tau \eta_1 \right) + \frac{1}{2} \left( 1 + 2 \tau \eta_2 \right)$$
\[ F_3 = \gamma_2(1 - \gamma_1) \left( \frac{2}{\pi^3} \sqrt{1 + \alpha_1^2} \right)^2 \sum_{i=1}^{n_1} \eta_i \exp \left( \frac{i}{2} \sqrt{1 + 2\tau \eta_i} \right) \left( \frac{1}{\sqrt{2\eta_i^2 + 1}} \right)^{y_1} + \right.
\]
\[ (1 - \gamma_1)(1 - \gamma_2) \left( \frac{2}{\pi^3} \sqrt{1 + 2\tau (\eta_1 + \eta_2)} \right)^{y_1} \cdot \frac{1}{\sqrt{2\eta_1^2 + 1}} \]
\[ F_4 = \left( 1 - \gamma_1 \right) (1 - \gamma_2) \left( \frac{2}{\pi^3} \sqrt{1 + 2\tau (\eta_1 + \eta_2)} \right)^{y_1} \cdot \frac{1}{\sqrt{2\eta_1^2 + 1}} \]

The First Derivative of Likelihood functions for parameters of the BZIPIGR type II model. The likelihood function of the BZIPIGR type II is

\[ L(\lambda, \alpha, \tau; h = 1, 2) = \prod_{i=1}^{n} P(y_{hi} | \lambda, \alpha, \tau; h = 1, 2) \]

The first derivative of the first log likelihood

\[ \log L(\lambda, \alpha, \tau; h = 1, 2) = \log L_1(\lambda, \alpha, \tau; h = 1, 2) + \log L_2(\lambda, \alpha, \tau; h = 1, 2) + \log L_3(\lambda, \alpha, \tau; h = 1, 2) + \log L_4(\lambda, \alpha, \tau; h = 1, 2) \]

if \( \theta = [ \lambda_1^T, \lambda_2^T, \alpha_1^T, \alpha_2^T, \tau_1^T ] \) is the vector parameters of the BZIPIGR model. Therefore, the log-likelihood function in Equation (10) is derived from each parameter and equated to zero, as follows:

\[ \frac{\partial \log L(\cdot)}{\partial \theta} = \frac{\partial \log L_1(\cdot)}{\partial \theta} + \frac{\partial \log L_2(\cdot)}{\partial \theta} + \frac{\partial \log L_3(\cdot)}{\partial \theta} + \frac{\partial \log L_4(\cdot)}{\partial \theta} \]

The first derivative of the first log likelihood (\( \log L(\cdot) \)) function to parameter \( \lambda_1 \) is

\[ \frac{\partial \log L_1(\cdot)}{\partial \lambda_1} = \sum_{i=1}^{n_1} \frac{-\eta_i \exp(x_i^T \lambda_1)x_i}{A_{11} + A_{12} + A_{13} + A_{14}} \left( \frac{A_{13}}{\sqrt{1 + 2\tau (\eta_i \exp(x_i^T \lambda_1) + q_{2i} \exp(x_i^T \lambda_2))}} \right) + \]

where

\[ A_{11} = \exp(x_i^T \alpha_1 + x_i^T \alpha_2), \quad A_{12} = \exp \left( x_i^T \alpha_1 + \frac{1}{\tau} \left( 1 - \sqrt{1 + 2\tau q_{2i} \exp(x_i^T \lambda_2)} \right) \right) \]

\[ A_{13} = \exp \left( \frac{1}{\tau} \left( 1 - \sqrt{1 + 2\tau (q_{1i} \exp(x_i^T \lambda_1) + q_{2i} \exp(x_i^T A_2)))} \right) \right) \]

\[ A_{14} = \exp \left( x_i^T \alpha_2 + \frac{1}{\tau} \left( 1 - \sqrt{1 + 2\tau q_{1i} \exp(x_i^T \lambda_1)} \right) \right) \]

\[ \frac{\partial \log L_2(\cdot)}{\partial \lambda_1} = \sum_{i=1}^{n_2} \frac{\beta_{i1} \beta_{i2} \beta_{i3} \beta_{i4}}{A_{21} A_{22} A_{23} A_{24} A_{25}} \left( 1 + 2\tau (q_{1i} \exp(x_i^T \lambda_1) + q_{2i} \exp(x_i^T \lambda_2)) \right)^{y_1} \cdot \frac{1}{\sqrt{A_{21} A_{22} A_{23} A_{24} A_{25}}} \]

where

\[ A_2 = A_{21} A_{22} A_{23} A_{24} A_{25} \]
\[
A_{21} = \frac{\exp(x_i^T \mathbf{a}_1)}{1 + \exp(x_i^T \mathbf{a}_1)}; A_{22} = (\frac{2}{\pi})^\frac{1}{2} y_{21}^{-1}; A_{23} = (1 + 2\pi q_{21} \exp(x_i^T \mathbf{a}_2))^{-\frac{1}{2}}; A_{24} = q_{21}^y \exp(x_i^T \mathbf{a}_2 y_{21} + \tau^{-1}); A_{25} = K_{32}(z_{21}); z_{21} = \frac{1}{2} \sqrt{1 + 2\pi q_{21} \exp(x_i^T \mathbf{a}_2)}
\]
\[
B_2 = B_{21} B_{22} B_{23} B_{24} B_{25}
\]
\[
A_{31} = \frac{\exp(x_i^T \mathbf{a}_3)}{1 + \exp(x_i^T \mathbf{a}_3)}; A_{32} = (\frac{2}{\pi})^\frac{1}{2} y_{31}^{-1}; A_{33} = (1 + 2\pi q_{31} \exp(x_i^T \mathbf{a}_3))^{-\frac{1}{2}}; A_{34} = q_{31}^y \exp(x_i^T \mathbf{a}_3 y_{31} + \tau^{-1}); A_{35} = K_{33}(z_{31})
\]
\[
B_3 = B_{31} B_{32} B_{33} B_{34} B_{35}
\]
\[
B_{31} = (\frac{2}{\pi})^\frac{1}{2} y_{11}^{-1}; B_{32} = (\frac{2}{\pi})^\frac{1}{2} y_{21}^{-1}; B_{33} = (1 + 2\pi q_{11} \exp(x_i^T \mathbf{a}_3) + q_{21} \exp(x_i^T \mathbf{a}_3))^{-\frac{1}{2}}; B_{34} = q_{11}^y \exp(x_i^T \mathbf{a}_3 y_{11} + \tau^{-1}); B_{35} = K_{33}(z_{32}); K_{y_{21}^{-1}}(\frac{1}{2} \sqrt{1 + 2\pi q_{11} \exp(x_i^T \mathbf{a}_3) + q_{21} \exp(x_i^T \mathbf{a}_3)}) = K_{33}(z_{32})
\]
\[
\frac{\partial \log L_2(\bullet)}{\partial \lambda_1^2} = \sum_{i=1}^{n_1} \left[ \frac{2q_{21} \exp(x_i^T \mathbf{a}_1) y_{11} x_i}{1 + 2\pi q_{21} \exp(x_i^T \mathbf{a}_2) + q_{21} \exp(x_i^T \mathbf{a}_2)} + y_{11} x_i + \frac{z_{21} K_{y_{21}^{-1} y_{21}}(1 + 2\pi q_{21} \exp(x_i^T \mathbf{a}_2) + q_{21} \exp(x_i^T \mathbf{a}_2))}{1 + 2\pi q_{21} \exp(x_i^T \mathbf{a}_2) + q_{21} \exp(x_i^T \mathbf{a}_2)} \right]
\]
\[
\frac{\partial \log L_3(\bullet)}{\partial \lambda_2^2} = \sum_{i=1}^{n_1} \left[ \frac{-q_{21} \exp(x_i^T \mathbf{a}_2) x_i}{\lambda_{11}^2 + \lambda_{12}^2 + \lambda_{13}^2 + \lambda_{14}^2} + \frac{A_{12}}{1 + 2\pi q_{21} \exp(x_i^T \mathbf{a}_2) + q_{21} \exp(x_i^T \mathbf{a}_2)} \right]
\]

The first derivative of the log \( L(\bullet) \) function to parameter \( \lambda_2 \) is obtained as
\[
\frac{\partial \log L_2(\bullet)}{\partial \gamma_i} = \sum_{i=1}^{n_2} (x_i - \frac{y_i}{2}) \left( A_{21} A_{22} \left[ A_{24} A_{25} \left( \frac{y_i}{2} - \frac{y_{2i}}{2} \right) (1 + 2\tau q_1 \exp(x_i^T A_2)) \right] - \frac{\tau_k}{2} \right) 2\tau q_i \exp(x_i^T A_2) + \\
y_{1i} = 0, y_{2i} > 0
\]

A_{23} A_{25} y_{2i} (q_i \exp(x_i^T A_2 y_{2i})) y_{2i} \exp(\frac{1}{\tau}) + \frac{A_{23} A_{24} q_{1i} \exp(x_i^T A_2) x_i}{1 + 2\tau q_i \exp(x_i^T A_2)}
\]

\[
\left(-z_{21} K_{y_{2i} - \frac{1}{2}}(z_{21}) + (y_{2i} - \frac{1}{2}) K_{y_{2i} - \frac{1}{2}}(z_{21}) \right) + B_{22} B_{22} \left[ B_{24} B_{23} \left( \frac{1}{2} - z_{2i} \right) y_{2i} \exp(x_i^T A_2) \right] + \\
B_{23} B_{25} y_{2i} (q_i \exp(x_i^T A_2 y_{2i})) y_{2i} \exp(\frac{1}{\tau}) + \left( 1 + 2\tau q_i (\exp(x_i^T A_1) + \exp(x_i^T A_2)) \right) 2\tau q_i \exp(x_i^T A_2)
\]

\[
\frac{\partial \log L_2(\bullet)}{\partial A_i} = \sum_{i=1}^{n_2} \left[ \left( \frac{1}{A_i + \alpha_i} \right) B_{21} B_{22} B_{24} B_{23} \left( \frac{1}{2} - z_{2i} \right) \left( y_{1i} + y_{2i} \right) \exp(x_i^T A_2) + \\
\frac{B_{33}}{1 + 2\tau q_i (\exp(x_i^T A_1) + \exp(x_i^T A_2))} \right]
\]

\[
\frac{\partial \log L_2(\bullet)}{\partial A_i} = \sum_{i=1}^{n_2} \left[ \left( \frac{1}{A_i + \alpha_i} \right) B_{21} B_{22} B_{24} B_{23} \left( \frac{1}{2} - z_{2i} \right) \left( y_{1i} + y_{2i} \right) \exp(x_i^T A_2) + \\
\frac{B_{33}}{1 + 2\tau q_i (\exp(x_i^T A_1) + \exp(x_i^T A_2))} \right]
\]

where

\[z_4 = \tau^{-1} \sqrt{1 + 2\tau q_i (\exp(x_i^T A_1) + \exp(x_i^T A_2))}\]

The first derivative of the log-likelihood function to parameter \( \alpha_1 \) is obtained as

\[
\frac{\partial \log L_1(\bullet)}{\partial \alpha_1} = \left[ \sum_{i=1}^{n_1} \left( \frac{A_{11} A_{12} - \exp(x_i^2 A_1)}{1 + \exp(x_i^2 A_1)} \right) \right] + \\
y_{1i} = 0, y_{2i} = 0
\]

\[
\sum_{i=1}^{n_2} \left( \frac{A_{12} - B_2 \exp(x_i^2 A_2)}{1 + \exp(x_i^2 A_2)} \right) - \sum_{i=1}^{n_3} \frac{\exp(x_i^2 A_2) (A_2 + B_2)}{(A_2 + B_2) (1 + \exp(x_i^2 A_2))} + \\
y_{1i} = 0, y_{2i} > 0
\]

The first derivative of the log-likelihood function to parameter \( \alpha_2 \) is
\[
\frac{\partial \log L_1(\bullet)}{\partial \alpha_i^2} = x_i \left[ \sum_{i=1}^{n_1} \left( \frac{\partial}{\partial \tau} \left( A_{11} + A_{14} \right) \right) \right] + \sum_{i=1}^{n_1} \left( \frac{\partial}{\partial \tau} \left( \frac{\partial}{\partial \tau} \left( A_{11} + A_{12} + A_{13} + A_{14} \right) \right) - \exp(x_i^T \alpha) \right) + \left( \frac{\partial}{\partial \tau} \left( \frac{\partial}{\partial \tau} \left( A_{11} + A_{12} + A_{13} + A_{14} \right) \right) \right)
\]

\[
\frac{\partial \log L_2(\bullet)}{\partial \alpha_i^3} = \sum_{i=1}^{n_1} \left( \frac{\partial}{\partial \tau} \left( \frac{\partial}{\partial \tau} \left( A_{21} \right) \right) \right) + \sum_{i=1}^{n_1} \left( \frac{\partial}{\partial \tau} \left( \frac{\partial}{\partial \tau} \left( A_{22} \right) \right) \right) + \sum_{i=1}^{n_1} \left( \frac{\partial}{\partial \tau} \left( \frac{\partial}{\partial \tau} \left( A_{24} \right) \right) \right) + \sum_{i=1}^{n_1} \left( \frac{\partial}{\partial \tau} \left( \frac{\partial}{\partial \tau} \left( A_{25} \right) \right) \right) + \sum_{i=1}^{n_1} \left( \frac{\partial}{\partial \tau} \left( \frac{\partial}{\partial \tau} \left( A_{26} \right) \right) \right)
\]

\[
\frac{\partial \log L_3(\bullet)}{\partial \alpha_i^3} = \sum_{i=1}^{n_1} \left( \frac{\partial}{\partial \tau} \left( \frac{\partial}{\partial \tau} \left( A_{31} \right) \right) \right) + \sum_{i=1}^{n_1} \left( \frac{\partial}{\partial \tau} \left( \frac{\partial}{\partial \tau} \left( A_{32} \right) \right) \right) + \sum_{i=1}^{n_1} \left( \frac{\partial}{\partial \tau} \left( \frac{\partial}{\partial \tau} \left( A_{33} \right) \right) \right) + \sum_{i=1}^{n_1} \left( \frac{\partial}{\partial \tau} \left( \frac{\partial}{\partial \tau} \left( A_{34} \right) \right) \right) + \sum_{i=1}^{n_1} \left( \frac{\partial}{\partial \tau} \left( \frac{\partial}{\partial \tau} \left( A_{35} \right) \right) \right)
\]

The first derivative of the first log-likelihood function \((\log L(\bullet))\) to parameter \(\tau\) is obtained as

\[
\frac{\partial \log L_1(\bullet)}{\partial \tau} = \sum_{i=1}^{n_1} \left( \frac{1}{A_{11} + A_{12} + A_{13} + A_{14}} \right) + \sum_{i=1}^{n_1} \left( \frac{\partial}{\partial \tau} \left( \frac{\partial}{\partial \tau} \left( A_{11} + A_{12} + A_{13} + A_{14} \right) \right) \right)
\]

where

\[
\frac{\partial A_{33}}{\partial \tau} = \left( \frac{1}{y_i} - y_i \right), \quad \frac{\partial A_{33}}{\partial \tau} = -\left( \frac{\partial}{\partial \tau} \left( \frac{\partial}{\partial \tau} \left( A_{31} \right) \right) \right)
\]

\[
\frac{\partial A_{35}}{\partial \tau} = \left( \frac{1}{y_i} - y_i \right) \right)
\]

\[
\frac{\partial B_{33}}{\partial \tau} = \left( \frac{1}{y_i} - y_i \right) \right)
\]

\[
\frac{\partial B_{35}}{\partial \tau} = \left( \frac{1}{y_i} - y_i \right) \right)
\]
where 
\[
\frac{\partial \Lambda_{32}}{\partial \tau} = \left( \frac{2}{\pi} \right)^{\frac{1}{2}} y_{11}^{-1} \left( -\frac{1}{2} \right) (\tau)^{-\frac{1}{2}}; \quad \frac{\partial \Lambda_{33}}{\partial \tau} = \frac{(1-y_{11}) y_{12} \exp(x_i^T A_1)}{(1+2\tau_{11} \exp(x_i^T A_1))^{\frac{1}{2}}}; \\
\frac{\partial \Lambda_{32}}{\partial \tau} = -q_{11} y_{11} (\exp(x_i^T A_1)) y_{12} \exp\left( \frac{1}{2} \right) \tau^{-2} \\
\frac{\partial \Lambda_{33}}{\partial \tau} = \frac{(1-y_{11}) y_{12} \exp(x_i^T A_1)}{(1+2\tau_{11} \exp(x_i^T A_1))^{\frac{1}{2}}}; \\
\frac{\partial \Lambda_{34}}{\partial \tau} = -q_{11} y_{11} (\exp(x_i^T A_1)) y_{12} \exp\left( \frac{1}{2} \right) \tau^{-2} \\
\frac{\partial \Lambda_{44}}{\partial \tau} = \left( -z_{31} K_{y_{11}+\frac{1}{2}(z_{31})} + (y_{11} - \frac{1}{2}) K_{y_{11}+\frac{1}{2}(z_{31})} \right) \tau \left( 1 + \frac{1+2\tau_{11} \exp(x_i^T A_1)}{\tau} \right) \\
\frac{\partial \Lambda_{45}}{\partial \tau} = \frac{(2/\pi)^{1/2} y_{11}^{-1} \left( -\frac{1}{2} \right) (\tau)^{-\frac{1}{2}}}{\tau}; \\
\frac{\partial \Lambda_{46}}{\partial \tau} = \frac{(1-y_{11}) y_{12} \exp(x_i^T A_1)}{(1+2\tau_{11} \exp(x_i^T A_1))^{\frac{1}{2}}}; \\
\frac{\partial \Lambda_{47}}{\partial \tau} = \frac{(1-y_{11}) y_{12} \exp(x_i^T A_1)}{(1+2\tau_{11} \exp(x_i^T A_1))^{\frac{1}{2}}}; \\
\frac{\partial \Lambda_{48}}{\partial \tau} = \left( -z_{32} K_{y_{11}+\frac{1}{2}(z_{32})} + (y_{11} - \frac{1}{2}) \right) K_{y_{11}+\frac{1}{2}(z_{32})} \tau \left( 1 + \tau (q_{11} + y_{11} \exp(x_i^T A_1) + q_{21} \exp(x_i^T A_2)) \right) \\
\frac{\partial \log L_d(e)}{\partial \tau} = \sum_{i=1}^{n_4} \left( \left( -\frac{1}{\tau^2} + \left( 1 - \frac{y_{11} + y_{22}}{2} \right) \right) \frac{2(y_{11} \exp(x_i^T A_1) + q_{21} \exp(x_i^T A_2))}{1+2\tau (q_{11} \exp(x_i^T A_1) + q_{21} \exp(x_i^T A_2))} - \frac{1}{\tau^2} + \frac{y_{11} > 0, y_{22} > 0}{K_{y_{11}+\frac{1}{2}(z_{34})}} \left( -\frac{z_{34} K_{y_{11}+\frac{1}{2}(z_{34})} \left( 1 + \tau (q_{11} \exp(x_i^T A_1) + q_{21} \exp(x_i^T A_2)) \right) \right) \right)
\]

The first derivative of the likelihood for the model without exposure variables is like the BZIPIGR type II model with exposure variables, but $q_i$ is not involved in the equation above.

References

1. 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. [CrossRef]
2. 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]
3. Famoye, F.; Sighn, K.P. Zero-Inflated Generalized Poisson Regression Model with an Application to Domestic Violence Data. *J. Data Sci.* 2006, 4, 117–130. [CrossRef]
4. Purhadi; Sari, D.N.; Aini, Q.; Irhamah. Geographically weighted Bivariate Zero Inflated Generalized Poisson Regression. *Heliyon* 2021, 7, e07491. [CrossRef][PubMed]
5. Cameron, A.C.; Trivedi, P.K. *Regression Analysis of Count Data*; Cambridge University Press: New York, NY, USA, 1998.
6. Hilde, J.M. Poisson Inverse Gaussian Regression. In *Modeling Count Data*; Cambridge University Press: New York, NY, USA, 2014. [CrossRef]
7. Ozmen, I.; Famoye, F. Count Regression Models with an Application to Zoological Data Containing Structural Zeros. *Data Sci. J.* 2007, 5, 491–502. [CrossRef]
8. McCullagh, P.; Nelder, J.A. *Generalized Linear Models*, 2nd ed.; Chapman and Hall: London, UK, 1989.
9. Karlis, D.; Xekalaki, E. A Simulation Comparison of Several Procedures for Testing the Poisson Assumption. *Statistician* 2000, 49, 355–382. [CrossRef]
10. Mardalena, S.; Purhadi; Purnomo, D.T.; Prastojo, D.D. Parameter Estimation and Hypothesis Testing of Multivariate Poisson Inverse Gaussian Regression. *Symmetry* 2020, 12, 1738. [CrossRef]
11. Frome, E.L.; Morris, M.D. Evaluating Goodness of Fit of Poisson Regression Model in Cohort Studies. *Am. Stat.* 1989, 43, 144–147.
12. Leffondré, K.; Jager, K.J.; Boucquemont, J.; Stel, V.S.; Heinze, G. Representation of exposures in regression analysis and interpretation of regression coefficients: Basic concepts and pitfalls. *Nephrol. Dial. Transplant.* 2014, 29, 1806–1814. [CrossRef][PubMed]
13. Long, D.L.; Preisser, J.S.; Herring, A.H.; Golin, C.E. A Marginalized Zero-Inflated Poisson Regression Model with Overall Exposure Effects. *NIH Public Access* 2014, 33, 5151–5165. [CrossRef][PubMed]
14. Phang, Y.N.; Loh, E.F. Statistical Analysis for Overdispersed Medical Count Data. *Int. J. Math. Comput. Sci.* 2014, 8, 292–294.
15. Silva, J.G. Zero-Inflated Mixed Poisson Regression Models. Master’s Thesis, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, 17 February 2017.

16. Gonzalves, J.N.; Souza, W.B. Flexible regression models for counts with high-inflation of zeros. Metron 2020, 78, 71–95. [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. Nugraha, J.; Guritno, S.; Haryatmi, S. Probit Model on Multivariate Binary Response Using Simulated Maximum Likelihood Estimator. J. Ilmu Dasar 2010, 11, 70–75. (In Indonesia)

19. Shoukri, M.M.; Asyali, M.H.; Kelton, D. The Poisson Inverse Gaussian Regression Model of Clustered count Data. J. Data Sci. 2014, 2, 17–32. [CrossRef]

20. Willmot, G.E. The Poisson-Inverse Gaussian distribution as an alternative to the negative binomial. Scand. Actuar. J. 1987, 1987, 113–127. [CrossRef]

21. Beren, C.; Thomrin, S.A.; Jaya, A.K. Estimasi Parameter Distribusi Exponentiated Modified Weibull Extension. In Proceedings of the National Conference Variansi, Makassar, Indonesia, 10 October 2016; ISBN 978-602-99837-2-2.

22. Mardalena, S.; Purhadi; Purnomo, D.T.; Prastyo, D.D. Bivariate Poisson Inverse Gaussian Regression Model with Exposure Variable: Infant and Maternal Death Case Study. J. Phys. Conf. Ser. 2021, 1752, 012016. [CrossRef]