Estimation Parameter of Generalized Poisson Regression Model Using Generalized Method of Moments and Its Application

Caecilia Bintang Girik Allo¹, Bambang Widjanarko Otok ¹*, Purhadi¹

¹Institut Teknologi Sepuluh Nopember (ITS), Surabaya 60111, Indonesia

*Corresponding author’s email: dr.otok.bw@gmail.com

Abstract. Poisson Regression is a standard model for data counts that can be used to determine the relationship between the response variable and predictors variables. Equidispersion is assumptions that must be met in Poisson Regression. Equidispersion is a condition that the variance of the response variable is equal with the average of the response variable. In real cases, these assumptions are often not met. In real cases, there are overdispersion and underdispersion cases. Generalized Poisson Regression (GPR) is one method that can handle cases of overdispersion and underdispersion. The GPR model is used to estimate regression parameters. Many articles proposed to use only Maximum Likelihood Estimation (MLE) to estimate the parameters of GPR. This article will develop the parameter estimation method of the GPR model, which is using the Generalized Method of Moments (GMM). GPR model is applied in the case of diarrhea in infants in Pasuruan Regency, East Java. The best model is chosen by the value of AICc. The smaller the value of AICc, the better the model. The best model is a model that includes exclusive breastfeeding, complete basic immunization, and healthy living behavior in the model.

Keywords: Overdispersion and Undersipersion, Generalized Poisson Regression (GPR), Generalized Method of Moments (GMM), Diarrhea

1. Introduction
Health problems in Indonesia are very complex. There are two kinds of diseases. They are infectious and non-infectious diseases. Acute Respiratory Tract Infection, diarrhea, malaria, and HIV/AIDS are examples of infectious diseases. Asthma, hypertension, coronary heart, and kidney are examples of non-infectious diseases. Diarrhea is one of infectious disease that can cause death. Period prevalence of diarrhea is 3.5% in 2013 [1]. East Java included in the top ten provinces the highest incidence of diarrhea was 3.8% [1]. Based on the number of patient visits to the Puskesmas in Pasuruan Regency, diarrhea is in the eight position of the top ten diseases in 2015 [2]. Infants are the age group that is very vulnerable to getting diarrhea. Therefore, it is necessary to know the factors that influence the incidence of diarrhea in infants.

Data in the form of the number of infants with diarrhea is count data. Poisson regression is a method that can be used to analyze count data. Poisson regression has an assumption that must be met, namely equidispersion. There is the development of PR model with regard to geographic aspects [3]. Equidispersion is a condition that the variance and the mean of the response variable are equal. In fact,
this assumption rarely met. Overdispersion is a condition that the variance is greater than the mean. Underdispersion is a condition that the variance is smaller than the mean. Generalized Poisson Regression (GPR) is one method that can handle the case with overdispersion and underdispersion.

Various developments and applications of Generalized Poisson Regression (GPR) have been carried out. Wang and Famoye (1997) applied the GPR model to analyze household fertility decisions [4]. They choose the best model based on AIC. Noriszura and Jemain (2005) applied Poisson Regression (PR) and GPR models on two types of motor insurance claim data in Malaysia [5]. The regression equation formed is used to make regression parameter estimates. One of parameter estimation methods is Generalized Method of Moments (GMM).

Mahpolah, et al (2018) used GMM to estimated parameters in the PR model and applied on the prevalence of Acute Respiratory Tract Infection (RTI) in South Kalimantan [6]. The result is estimate parameters using GMM better than MLE. Mahpolah, et al also applied GMM in logistic regression model and compare the result with MLE method [7]. This paper will use GMM to estimate the parameters of the GPR model. Furthermore, applying the GPR model in the case of infants suffering from diarrhea in Pasuruan Regency and choose the best model based on corrected AIC (AICc).

2. THEORETICAL REVIEW

In this section we will review some of the theories used.

2.1. Poisson Regression (PR)
A random variable \( Y \) has a Poisson Probability Distribution if only if

\[
 f(y) = \frac{e^{-\mu} \mu^y}{y!}, \quad y = 0, 1, 2, \ldots
\]

(1)

\( \mu \) is mean an event and \( \mu > 0 \) [8]. The mean and variance of Poisson Distribution is equal. This condition is called equidispersion.

Poisson Regression (PR) model can be seen at (2).

\[
 \ln(\mu_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik}
\]

(2)

\( i = 1, 2, \ldots, n \)

\( x_i = [x_{i1}, x_{i2}, \ldots, x_{ik}]^T \)

\( \beta = [\beta_0, \beta_1, \beta_2, \ldots, \beta_k]^T \).

2.2. Generalized Poisson Regression (GPR)

Poisson regression cannot be used in overdispersion or underdispersion data. Let \( Y_i \sim GP(\mu, \phi) \), \( i = 1, 2, \ldots, n \) then \( Y_i \) has a probability function [5]:

\[
 f(y_i | \mu, \phi) = \left[ \frac{\mu_i}{1 + \phi \mu_i} \right]^y_i (1 + \phi y_i)^{-1} \frac{1}{y_i!} \exp \left( -\mu_i (1 + \phi y_i) \right)
\]

(3)

with \( \mu_i = e^{\beta x_i} \), \( x_i \) is vector from predictor variables, and \( \beta \) is vector from regression parameter.

The mean and variance of Generalized Poisson Distribution can be seen in (4) and (5).

\[
 E(Y_i) = \mu_i
\]

(4)

\[
 V(Y_i) = \mu_i (1 + \phi \mu_i)
\]

(5)

Generalized Poisson Distribution have two parameters, namely \( \mu \) and \( \phi \). \( \phi \) is dispersion parameter. If \( \phi > 0 \) then overdispersion implies that the variance is greater than the mean. If \( \phi < 0 \) then underdispersion which means that the variance is smaller than the mean [9]. GPR model can be seen in (6).
\[
\ln(\mu_i) = x_i^T \beta = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik} 
\]
(6)

\[
\mu_i = e^{x_i^T \beta} = \exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik}) 
\]
(7)

2.3. Generalized Method of Moments (GMM)

GMM is an extension of the estimation of the Moment Method when the number of moment conditions is greater than the number of parameters that are being estimated [10]. The moment conditions is a function of the model parameters and data that has been determined so the value of the expectation of the function is zero on the real value of parameter. It can be written mathematically as follows:

\[
E(h(y, \theta_0)) = 0 
\]
(8)

Parameters estimator using GMM are obtained by minimizing the weighted square of sample moment condition. It can be seen in (9). Sample moment condition can be seen in (10).

\[
\hat{\theta} = \arg \min_{\theta} h_n(\theta) = \frac{1}{n} \sum_{i=1}^{n} h(y_i, \theta) 
\]
(9)

\[
W(\theta) \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} 
\]
(10)

The original version of GMM proposed by Hansen (1982) is called two-step GMM (2SGM) [11]. Various developments of Generalized Method of Moments (GMM) have been carried out. Hasen et al (1996) developed two-stage GMM (2SGMM) become an iterative GMM (ITGMM) [12]. The algorithm is as follows [13]:

1. Compute \( \theta^{(0)} = \arg \min_{\theta} h_n(\theta) \)
2. Compute the weighted matrix, \( W(\theta^{(0)}) \)
3. Compute \( \theta^{(i)} = \arg \min_{\theta} h_n(\theta) \)
4. If \( \| \theta^{(i)} - \theta^{(0)} \| < tol \) stops where \( tol \) can be set as small as we want, else \( \theta^{(i)} = \theta^{(i-1)} \) and go to 2.
5. Define the ITGMM estimator \( \hat{\theta} \) as \( \theta^{(i)} \).

2.4. Nelder Mead Iteration

This study will use numerical method to get the ITGMM estimator. The selected numerical method is Nelder Mead. Nelder Mead method was first introduced by Nelder and Mead in 1965 [14]. Ten years later, Olsson and Nelson (1975) developed the method [15]. Lagarias et al (1998) discussed about the convergence properties of Nelder Mead [16]. Let there is a function \( g(\theta) \), \( \theta \in \mathbb{R}^m \). We want to minimize the function. We need \( m+1 \) points, suppose \( \theta_0, \theta_1, \ldots, \theta_{m+1} \). Nelder Mead algorithm as follows:

1. Order \( g(\theta_0) \leq g(\theta_1) \leq \ldots \leq g(\theta_{m+1}) \)
2. Calculate the mean \( \bar{\theta} = \frac{1}{m} \sum_{j=0}^{m} \theta_j \)
3. Reflection

\[
\theta_r = \bar{\theta} + \eta (\theta - \theta_{m+1}) \quad \eta \text{ is reflection coefficient and } \eta > 0. \quad \text{If } g(\theta_r) < g(\theta_m) \text{ then } \theta_{m+1} \text{ is replaced by } \theta_r \text{, else as follows:}
\]
a. If \( g(\theta_r) < g(\theta_l) \) then compute \( \theta_\gamma = \theta + \gamma(\theta_r - \theta) \), \( \gamma \) is expansion coefficient and \( \gamma > 1 \). Furthermore, if \( g(\theta_l) < g(\theta_r) \) then \( \theta_{m+1} \) is replaced by \( \theta_r \), else \( \theta_{m+1} \) is replaced by \( \theta_l \).

b. If \( g(\theta_m) \leq g(\theta_l) < g(\theta_{m+1}) \) then compute \( \theta_\zeta = \theta - \zeta(\theta_n - \theta) \), \( \zeta \) is contraction coefficient and \( 0 < \zeta < 1 \). Furthermore, if \( g(\theta_\zeta) \leq g(\theta_l) \) then \( \theta_{m+1} \) is replaced by \( \theta_\zeta \), else go to 4.

c. If \( g(\theta_l) \geq g(\theta_{m+1}) \) then compute \( \theta_\zeta = \theta \). Furthermore, if \( g(\theta_\zeta) < g(\theta_{m+1}) \) then \( \theta_{m+1} \) is replaced by \( \theta_l \), else go to 4.

4. Shrink
   Replace all value with \( \theta_i = \theta_i + \delta(\theta_l - \theta_i) \), \( 0 < \delta < 1 \) and \( 2 \leq i \leq m+1 \).

The iteration stops when \( \sqrt{\frac{1}{m} \sum_{i=1}^{m} (g(\theta_i) - g(\theta))} < \varepsilon \) where \( \varepsilon \) can be set as small as we want.

2.5. Corrected Akaike Information Criterion (AICc)

Akaike Information Criterion (AIC) is the model suitability criterion in estimating the model statistically. The calculation of AIC value as follows:

\[
\text{AIC} = -2\text{ln}(\text{maximum likelihood}) + 2p
\]

where \( p \) is number of parameters. When the sample is small, AICc can be used as the model suitability. The calculation of AICc value can be seen in (12) where \( n \) is number of samples.

\[
\text{AICc} = \text{AIC} + \frac{2p(p+1)}{n-p-1}
\]

3. Methodology

In this section we will show description of the data and variables used.

3.1. Data

The data that used in this study is secondary data obtained from twentyfour sub-districts in Pasuruan Regency, East Java. The data measurement in 2017. Each sub-district is recorded number of infants, number of infants affected by diarrhea, number of infants who get exclusive breastfeeding, number of infants who get complete basic immunization, and number of households who healthy living behavior.

3.2. Research Variables

There are three predictor variables, namely \( X_1, X_2, X_3 \). This study has one response variable, namely \( Y \). The research variables are as follows:

1. \( Y \) : number of infants affected by diarrhea
2. \( X_1 \) : percentage of infants who get exclusive breastfeeding
3. \( X_2 \) : percentage of infants who get complete basic immunization
4. \( X_3 \) : percentage of households who healthy living behavior.

4. Result and Discussion

Multicollinearity between predictor variables needs to be checked. From Table 1, there is no predictor variable that has VIF value more than 10 indicates that there is no multicollinearity problem. Furthermore, the value of deviance, \( D \), is used to check overdispersion on data. If the deviance value is divided by the degrees of freedom greater than one, then overdispersion occurs in the data. We get \( D = 3044.41 \) with \( df = 20 \), so the conclusion is overdispersion occurs in data.
Table 1. VIF of Predictor Variables

| Variable | VIF |
|----------|-----|
| $X_1$    | 1.092 |
| $X_2$    | 1.083 |
| $X_3$    | 1.013 |

4.1. Estimation Results of GPR Model

The moment condition for GPR model can be seen at (13).

$$
\mathbf{h}(y_i, \mathbf{x}_i, \theta) = \begin{bmatrix}
\mathbf{x}_i \left( y_i - e^{x_i \beta} \right) \\
\left( y_i - e^{x_i \beta} \right)^2 - \left( e^{x_i \beta} \left( 1 + \phi e^{x_i \beta} \right)^2 \right)
\end{bmatrix}
$$

(13)

where $\mathbf{x}_i = [1 \ x_{i1} \ x_{i2} \ x_{i3}]^T$ and $\beta = [\beta_0 \ \beta_1 \ \beta_2 \ \beta_3]^T$. From (13), we can get the sample moment condition that can be seen in (14).

$$
\mathbf{h}_n(y, \mathbf{x}, \theta) = \frac{1}{n} \begin{bmatrix}
\sum_{i=1}^{n} y_i - e^{x_i \beta} \\
\sum_{i=1}^{n} y_i - e^{x_i \beta} \\
\sum_{i=1}^{n} y_i - e^{x_i \beta} \\
\sum_{i=1}^{n} y_i - e^{x_i \beta}
\end{bmatrix} - \begin{bmatrix}
\sum_{i=1}^{n} \left( y_i - e^{x_i \beta} \right)^2 - \left( e^{x_i \beta} \left( 1 + \phi e^{x_i \beta} \right)^2 \right) \\
\sum_{i=1}^{n} \left( y_i - e^{x_i \beta} \right)^2 - \left( e^{x_i \beta} \left( 1 + \phi e^{x_i \beta} \right)^2 \right) \\
\sum_{i=1}^{n} \left( y_i - e^{x_i \beta} \right)^2 - \left( e^{x_i \beta} \left( 1 + \phi e^{x_i \beta} \right)^2 \right) \\
\sum_{i=1}^{n} \left( y_i - e^{x_i \beta} \right)^2 - \left( e^{x_i \beta} \left( 1 + \phi e^{x_i \beta} \right)^2 \right)
\end{bmatrix}
$$

(14)

The initial value used is the estimation coefficient from Poisson Regression with $\varepsilon = tol = 10^{-30}$. The result can be seen in Table 2. There are seven models that can be formed from existing predictor variables. Best model is chosen based on AICc. The result can be seen in Table 3. From Table 3 we get the best model is model that contain all parameter.

Table 2. Estimation Result of GPR

| Parameter | Estimator | SE  |
|-----------|-----------|-----|
| $\phi$    | 0.060657  | 0.011756 |
| $\beta_0$ | 7.6378    | 0.74642 |
| $\beta_1$ | -0.01049  | 0.0056114 |
| $\beta_2$ | -0.006475 | 0.0026211 |
| $\beta_3$ | -0.027393 | 0.0011426 |
| Table 3. AICc of GPR Model |
|---------------------------|
| Model                | AICc   |
| $X_1, X_2, X_3$       | 38302.72 |
| $X_1, X_2$            | 40282.14 |
| $X_1, X_3$            | 39356.31 |
| $X_2, X_3$            | 39135.62 |
| $X_1$                | 42277.21 |
| $X_2$                | 39898.25 |
| $X_3$                | 39528.64 |

So, we get GPR model can be seen in (15) or (16).

\[
\ln \hat{\mu} = 7.6378 - 0.0105X_1 - 0.0065X_2 - 0.0274X_3 \quad (15)
\]

\[
\hat{\mu} = e^{7.6378-0.0105X_1-0.0065X_2-0.0274X_3} \quad (16)
\]

The interpretation of each variables as follows:

1. If the all variables are constant and every one percent increase in the percentage of infants give exclusive breast feeding, then the average number of infants affected by diarrhea will decrease by 1.0105 times from the average number of infants affected by diarrhea originally.
2. If the all variables are constant and every one percent increase in the percentage of infants who get complete basic immunization, then the average number of infants affected by diarrhea will decrease by 1.0065 times from the average number of infants affected by diarrhea originally.
3. If the all variables are constant and every one percent increase in the percentage of households who healthy living behavior the average number of infants affected by diarrhea will decrease by 1.0277 times from the average number of infants affected by diarrhea originally.

5. Summary

Generalized Poisson Regression (GPR) can be used when Poisson Regression assumption are not meet. There are many methods to estimate the parameter of GPR including Generalized Method of Moments (GMM). GMM can be used when the number of moment condition is greater than the number of parameters that are being estimated. The result of this study is all predictor variables can reduce the average of infants affected by diarrhea in Pasuruan Regency. This shows that exclusive breastfeeding, complete basic immunization, and healthy living behavior are very important.

References

[1] Kemenkes RI. (2013). *Riset Kesehatan Dasar 2013*. Jakarta: Kementerian Kesehatan.
[2] Dinas Kesehatan Kabupaten Pasuruan. (2016). *Profil Kesehatan Kabupaten Pasuruan Tahun 2015*. Pasuruan: Dinas Kesehatan Kabupaten Pasuruan.
[3] Triyanto, Purhadi., Otok, B. W., & Purnami, S. P. (2015). Parameter Estimation of Geographically Weighted Multivariate Poisson Regression. *Applied Mathematical Sciences, 9*, 4081-4093.
[4] Wang, W., & Famoye, F. (1997). Modeling Household Fertility Decisions with Generalized Poisson Regression. *Journal of Population Economics, 10*, 273-283.
[5] Noriszura, I., & Jemain, A. A. (2005). Generalized Poisson Regression: An Alternative for Risk Classification. *Jurnal Teknologi, 43*, 39-54.
[6] Mahpolah, Suharto, Wibowo, A., & Otok, B. W. (2018). The Estimation of Generalized Method Moment Poisson Regression Model on The Prevalence of Acute Respiratory Tract Infection (RTI) in South Kalimantan. *CAUCHY-Jurnal Matematika Murni dan Aplikasi, 5(3)*, 161-168.
[7] Mahpolah, Suharto, Wibowo, A., & Otok, B. W. (2017). Generalized Method Moment Logistic Regression Model on the Prevalence of Acute Respiratory Tract Infections. *Journal of Applied Environmental and Biological Sciences*, 7, 92-96.

[8] Cameron, A. C., & Trivedi, P. K. (1998). *Regression Analysis of Count Data*. New York: Cambridge University Press.

[9] Famoye, F., Wulu, J. T., & Singh, K. P. (2004). On The Generalized Poisson Regression Model with An Application to Accident Data. *Journal of Data Science*, 2, 287-295.

[10] Matyas, L. (1999). *Generalized Method of Moments Estimation*. New York: Cambridge University Press.

[11] Hansen, L. P. (1982). Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica*, 50, 1029-1054.

[12] Hansen, L. P., Heaton, J., & Yaron, A. (1996). Finite-Sample Properties of Some Alternative GMM Estimators. *Journal of Business and Economic Statistics*, 14, 262-280.

[13] Chausse, P. (2010). Computing Generalized Method of Moments and Generalized Empirical Likelihood with R. *Journal of Statistical Software*, 34(11), 1-35.

[14] Nelder, J., & Mead, R. (1965). A Simplex Method for Function Minimization. *The Computer Journal*, 7(4), 308-313.

[15] Olsson, D., & Nelson, L. (1975). The Nelder Mead Simplex Procedure for Function Minimization. *Technometrics*, 17, 45-51.

[16] Lagarias, J., Reeds, J., Wright, M., & Wright, P. (1998). Convergence Properties of Nelder Mead Simplex Algorithm in Low Dimensions. *SIAM J. Optim.*, 9, 112-147.