Eksakta Berkala Ilmiah Bidang
VOLUME 21 NO 01 2020, pp 54-58
ISSN : Print 1411-3724 — Online 2549-7464
DOI : https://doi.org/10.24036//eksakta/vol21-iss1/222

Eksakta
Berkala Ilmiah Bidang MIPA
http://www.eksakta.ppj.unp.ac.id/index.php/eksakta

Article

Generalized Poisson Regression Type-II at Jambi City Health Office

Corry Sormin*, Gusmanely Z, Nurhidayah

Department of Mathematics, Faculty of Science and Technology, Universitas Jambi, Jambi, Indonesia

Article Info

Article history :

Received April 13, 2020
Revised April 20, 2020
Accepted April 22, 2020
Published April 30, 2020

Keywords :

generalized poisson regression, distribution, analysis

Abstract. One statistical analysis is regression analysis. One regression that has the assumption of poisson distribution is poisson regression which has the assumption of poisson distribution. Neonatal deaths are still very rare, so the proper analysis is used, namely Generalized Poisson Regression. This regression method is specifically used for Poissson distributed data. The stages that will be carried out in this research are Poisson distribution test and equidispersion assumption, parameter estimation, model feasibility test and best model selection. Data from the Jambi City Health Office in 2018 showed that the Generalized Poisson Regression regression alleged had a variable number of first trimester visits. The number of pregnant women getting Tetanus Diptheria immunization ; the estimated number of neonatal infants with complications ; the number of infants receiving Hepatitis B immunization was less than twenty-four hours; and the number of infants receiving BCG immunizations.

Corresponding Author :
Corry Sormin
Department of Mathematics, Faculty of Science and Technology
Universitas Jambi, Jambi, Indonesia
Email : corry.sormin@unja.ac.id

1. Introduction

Statistics is one of the disciplines of science whose application has become widespread among the people. Starting from conducting research and proposing hypotheses, statistics have a role in it. One analysis of statistics is regression analysis, which is data analysis that utilizes the relationship between two or more variables. The purpose of regression analysis is to investigate the relationship between the response variable and the predictor variable and predict one of the other variables. In general, regression analysis has a response variable which is a continuous variable that follows the normal distribution. But there is also a regression analysis that has a response variable not a continuous variable but a discrete variable one of which is Poisson regression analysis[1-4].

Poisson regression analysis has several assumptions including the response variable following the Poisson distribution. This means that the data used in Poisson regression analysis is data with rare events. The next assumption is equidispersion, a condition in which the mean (mean) and variance of the response variable are the same. However, not all data meet the assumption of equidispersion so that if the data have conditions where the mean and variance are not the same then Poisson regression analysis should not be used[1-6].

The analysis used if the second assumption is not met, that is, if the mean is greater or smaller than the value of variance, it is often referred to as underdispersion or overdispersion, so analyzes such as Negative Binomial Regression and Generalized Poisson Regression can be used. Both of these regression analyzes have response variables in the form of discrete variables. Generalized Poisson regression analysis is analogous to Poisson regression analysis only that Generalized Poisson regression assumes the random component is generalized Poisson distribution[1, 7-12].

Often the data obtained has a variance of the response variable greater than the average value (overdispersion), so the assumption of equi-dispersion (the mean and variance of the response variable is equal) in the Poisson regression is not met. This shows that to overcome the problem of Underdispersion the Generalized Poisson Regression method is used.

2. Review Literature

2.1. Generalized Poisson Regression

The functional form of the generalized Poisson distribution is expressed as Generalized Poisson P (GP-P), where P = 1 for the generalized Poisson distribution-1 and P = 2 for the generalized Poisson-2 distribution. The distribution of the generalized Poisson-1 opportunity is stated as: (Yang, 2009: 1515)

$$f(y) = \frac{\theta(\theta+\lambda y)^{y-1} \exp[-\theta-\lambda y]}{y!}, y = 0,1,2, \dots$$

where θ> 0 and max (-1, -θ / 4) <λ <1. The generalized Poisson distribution-1 will be the probability distribution of the generalized Poisson-2 if θ = μ / (1 + rμ) and λ = rμ / (1 + rμ) substituted in the equation which is stated as follows: (Almasi, 2016: 3)

$$f(y) = \left(\frac{\mu}{1+r\mu}\right)^{y} \frac{(1+ry)^{y-1}}{y!} \exp\left[\frac{-\mu(1+ry)}{1+r\mu}\right] \text{ for } \theta = \frac{\mu}{1+r\mu} > 0 \text{ and } \lambda = \frac{r\mu}{1+r\mu} > 0.$$

The generalized Poisson regression equation has the same regression equation as the Poisson regression which is defined as follows:

$$Y_i = \exp\left(\beta_0 + \beta_1 X_{1i} + \dots + \beta_p X_{pi}\right) + \varepsilon_i$$

The estimated parameter used is the Maximum Likelihood Estimation (MLE) with the estimation process with the likelihood function as follows: (Wang and Famoye, 1997: 277-278)

$$L(r,\mu) = \prod_{i=1}^{n} \left(\frac{\mu_i}{1+r\mu_i}\right)^{y_i} \frac{(1+ry_i)^{y_i-1}}{y_i!} \exp\left[\frac{-\mu_i(1+ry_i)}{1+r\mu_i}\right]$$

For the parameter r obtained:

$$\sum_{i=1}^{n}\left[y_i\left(\frac{-\mu_i}{1+r\mu_i}\right)+y_i\frac{(y_i-1)}{1+ry_i}-\left(\frac{y_i\mu_i-\mu_i^2}{(1+r\mu_i)^2}\right)\right]=0$$

For the β parameter obtained:

$$\sum_{i=1}^{n}\left[\frac{(y_i-\mu_i)\boldsymbol{X}_i^T}{(1+r\mu_i)^2}\right]=0$$

For parameter r as follows:

$$\frac{\partial \log L(r,\mu)}{\partial \beta}=\sum_{i=1}^{n}\left[\frac{(y_i-\mu_i)\boldsymbol{X}_i^T}{(1+r\mu_i)^2}\right]=0$$

The equation shows that the parameter r and the parameter β have a non-linear relationship so they have not provided a solution. As a result, another settlement is needed with an iterative numerical method.

3. Result and Discussion

3.1. Poisson Distribution Testing

The hypothesis used is the number of neonatal deaths not following the Poisson distribution. Based on SPSS output obtained α= 1,000 means that many neonatal deaths follow the Poisson distribution.

3.2. Assumption of Equidispersi

Generalized Poisson Regression analysis has different assumptions from Poisson Regression assumptions which must meet the equidispersion assumption. The analysis requires that the response variable does not have the same mean value as the variance value. Based on Descriptive Statistics obtained Mean = 0.25 and Variance = 0.197, meaning that the assumption of equidipersion is not fulfilled. As a result, the analysis of Generalized Poisson Regression can be continued.

3.3. Generalized Poisson Regression Analysis

Based on the estimation results, the estimated regression equation for the number of neonatal deaths per puskesmas is as follows:

Y ˆ = exp (-3.798370 + 0.005318X_1 + 0.001806X_2 + 0.002734X_3 — 1.635975X_4 + 0.007465X_5 — -0.000359X_6 + 0.229370X_7 + 0.004049X_8)

Before the next stage, the feasibility of the model will be tested first. The results of the likelihood ratio (G) of 18.3295 can be concluded that the Generalized Poisson Regression model can be used.

In determining the factors that affect neonatal mortality from the data obtained, the Remove method will be used in the regression analysis. First the variable will be issued which gives the greatest z value to the model.

4. Conclusion

The estimated regression equation for the number of neonatal deaths per puskesmas is as follows: $\hat{Y} = \exp(-4.112862 + 0.000435X_1 - 0.00036X_3 + 0.002946X_4 + 0.002128X_5 + 0.001065X_6)$. The interpretation of the model, β_0 of -4.112862 means that the chance of neonatal death per puskesmas in Jambi City is 0.0164. The β_3 parameter of 0,00036 means that an increase in pregnant women receiving the Tetanus Diptheria immunization by 1 person will reduce the chance of 0.0164 neonatal deaths in Jambi City if the other variables are considered constant.

5. Acknowledgments

The researcher would like to thank for the financial assistance provided by Universitas Jambi so that the research went well.

References

[1] Hossain Z, Akter R, Sultana N, Kabir E.(2020). Modelling zero-truncated overdispersed antenatal health care count data of women in Bangladesh. PloS one;15:e0227824.
[2] Islam MA, Chowdhury RI.(2017). A generalized right truncated bivariate Poisson regression model with applications to health data. PloS one;12:e0178153.
[3] Islam MM, Alam M, Tariquzaman M, Kabir MA, Pervin R, Begum M, et al.(2013). Predictors of the number of under-five malnourished children in Bangladesh: application of the generalized poisson regression model. BMC public health;13:11.
[4] Okuno A, Shimodaira H.(2020). Hyperlink regression via Bregman divergence. Neural networks : the official journal of the International Neural Network Society;126:362-83.
[5] Wang W, Famoye F.(1997). Modeling household fertility decisions with generalized Poisson regression. Journal of population economics;10:273-83.
[6] Yang Z, Hardin JW, Addy CL, Vuong QH.(2007). Testing approaches for overdispersion in poisson regression versus the generalized poisson model. Biometrical journal Biometrische Zeitschrift;49:565-84.
[7] Almasi A, Rahimiforoushani A, Eshraghian MR, Mohammad K, Pasdar Y, Tarrahi MJ, et al.(2016). Effect of Nutritional Habits on Dental Caries in Permanent Dentition among Schoolchildren Aged 10-12 Years: A Zero-Inflated Generalized Poisson Regression Model Approach. Iranian journal of public health;45:353-61.
[8] Chen XD, Fu YZ, Wang XR.(2013). Local influence measure of zero-inflated generalized Poisson mixture regression models. Statistics in medicine;32:1294-312.
[9] Cui Y, Kim DY, Zhu J.(2006). On the generalized poisson regression mixture model for mapping quantitative trait loci with count data. Genetics;174:2159-72.
[10] Cui Y, Yang W.(2009). Zero-inflated generalized Poisson regression mixture model for mapping quantitative trait loci underlying count trait with many zeros. Journal of theoretical biology;256:276-85.
[11] Erhardt V, Bogdan M, Czado C.(2010). Locating multiple interacting quantitative trait Loci with the zero-inflated generalized poisson regression. Statistical applications in genetics and molecular biology;9:Article26.
[12] Farhadi Hassankiadeh R, Kazemnejad A, Gholami Fesharaki M, Kargar Jahromi S, Vahabi N.(2017). Assessment of length of stay in a general surgical unit using a zero-inflated generalized Poisson regression. Medical journal of the Islamic Republic of Iran;31:91.
[13] Sormin, C. 2013. *Aplikasi Regresi Poisson pada Faktor-Faktor yang Mempengaruhi Banyaknya Pasien Diabetes Mellitus*. Skripsi. Tidak diterbitkan. Fakultas Matematika dan Ilmu Pengetahuan Alam. Universitas Negeri Yogyakarta: Yogyakarta.

[14] Sormin, C. 2017. *Model Multilevel Regresi Poisson Tergeneralisasi Zero-Inflated. Tesis*. Tidak diterbitkan. Fakultas Matematika dan Ilmu Pengetahuan Alam. Universitas Gadjah Mada: Yogyakarta.

[15] Wang, W & Famoye, F. (1997). *Modeling household fertility decisions with generalized Poisson regression*, Journal of Population Economics, 10, 273-283.

[16] Yang, Z, dkk. (2009). *A score test for overdispersion in Poisson regression based on the generalized Poisson-2 model*, Journal of Statistical Planning and Inference 139, 1514-1521.