Eksakta : Berkala Ilmiah Bidang MIPA
VOLUME 23 NO 01 2022, pp 12-17
ISSN : Print 1411-3724 — Online 2549-7464
DOI : https://doi.org/10.24036/eksakta/vol23-iss01/288

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

Article

Negative Binomial Regression in Underdispersion (Case Study: Neonatal Mortality in Jambi City)

Article Info

Article history :

Received August 26, 2021
Revised 10 September, 2021
Accepted 14 September, 2021
Published 30 March, 2022

Keywords :

Neonatal Mortality,
Negative Binomial Regression

Corry Sormin[1*], Gusmanely Z.[1]

[1]Department of Mathematics, Faculty of Science and Technology, Universitas Jambi

Abstract. Neonatalis is birth before 28 days of a baby. Factors that are considered to affect neonatal mortality include the number of visits in the 1st and 3 rd trimester, the number of pregnant women receiving Tetanus Diptheria Immunization, the estimated number of neonatal infants with complications, the number of infants receiving Hepatitis B Immunization for less than 24 hours, the number of infants receiving BCG Immunization and number of 1 and 3 neonatal visits. Neonatal mortality is still very rare so that the right analysis is used, namely Negative Binomial Regression. This research aim to investigate negative binomial regression in underdipersion on neonatal mortality at Jambi City. These two regression methods are specifically used for Poisson distributed data because they are rare. The stages of the research that will be carried out are the Poisson distribution test and the equidispersion assumption, parameter estimation, model feasibility test, and selection of the best model. The results obtained that the best model without the variable number of 3rd-trimester visits or without the variable number of infants who received BCG immunization with AIC was 36.3.

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

1. Introduction

The National Population and Family Planning Agency (BKKBN) noted that the neonatal mortality rate based on the results of the Indonesian Demographic and Health Survey (SKDI) decreased, from 32 per 1,000 live births in 2012 to 15 per 1,000 live births in 2017[1,2,3,4,5]. The endogenous infant mortality rate or Neonatal mortality is the number of deaths occurring in the first month expressed in per thousand live births after the baby is born [6,7,8,9,10,11,12]. One of the causes of neonatal death is the factor brought by the child from birth obtained from the parents when fertilization occurs or during pregnancy[13,14].The existence of endogenous factors associated with pregnancy, causing neonatal death [15,16,17,18,19,20].

To assess the extent of the factors that cause neonatal death, an analysis is needed using negative binomial regression [21,21,23,24,25,26,27]. Often the data obtained has a variance from the response variable that is greater than mean (overdispersion) so that the assumption of equidispersion (mean and variance of the response variable is the same) in Poisson regression is not fulfilled. This shows that to overcome the problem of Overdispersion, the Negative Binomial Regression method is used.

2. Review Literature

2.1 Negative Binomial Regression

The Negative Binomial has the following probability density form [28,29,30,31]

$$f(y)=\frac{\Gamma\left(y+\frac{1}{\alpha}\right)}{\Gamma\left(\frac{1}{\alpha}\right)\Gamma(y+1)}\left(\frac{\frac{1}{\alpha}}{\mu+\frac{1}{\alpha}}\right)^{\frac{1}{\alpha}}\left(\frac{\mu}{\mu+\frac{1}{\alpha}}\right)^{y}$$

with $\mu>0$, $\frac{1}{\alpha}>0$, and $y=0,1,2,\ldots$. In general, the model of the Binomial Regression model is defined as follows:

$$Y_i=\exp\left(\beta_0+\beta_1X_{1i}+\cdots+\beta_pX_{pi}\right)+\varepsilon_i$$

Estimation in this regression uses *Maximum Likelihood Estimation* with likelihood function:

$$L\left(\frac{1}{\alpha},\mu\right)=\prod_{i=1}^{n}\frac{\Gamma\left(y+\frac{1}{\alpha}\right)}{\Gamma\left(\frac{1}{\alpha}\right)\Gamma(y+1)}\left(\frac{\frac{1}{\alpha}}{\mu+\frac{1}{\alpha}}\right)^{\frac{1}{\alpha}}\left(\frac{\mu}{\mu+\frac{1}{\alpha}}\right)$$

In the Negative Binomial Distribution, the connecting function commonly used is the *log link* of the form $g(\mu_i)=\log(\mu_i)=\boldsymbol{X}_i^T\boldsymbol{\beta}$ or in other words $\mu_i=\exp\left(\boldsymbol{X}_i^T\boldsymbol{\beta}\right)$. As a result, obtainde:

$$L(r,\beta)=\prod_{i=1}^{n}\frac{\Gamma\left(y_i+\frac{1}{\alpha}\right)}{\Gamma\left(\frac{1}{\alpha}\right)\Gamma(y_i+1)}\left(\frac{\frac{1}{\alpha}}{\mu_i+\frac{1}{\alpha}}\right)^{\frac{1}{\alpha}}\left(\frac{\mu_i}{\mu_i+\frac{1}{\alpha}}\right)^{y_i}$$

After obtaining the likelihood function, the the logarithm of the function will be searched as below:

$$\log L(r,\beta)=\sum_{i=1}^{n}\log\left(\frac{\Gamma\left(y_i+\frac{1}{\alpha}\right)}{\Gamma\left(\frac{1}{\alpha}\right)\Gamma(y_i+1)}\right)+\frac{1}{\alpha}\log\left[\frac{\frac{1}{\alpha}}{\exp\left(\boldsymbol{X}_i^T\boldsymbol{\beta}\right)+\frac{1}{\alpha}}\right]+y_i\log\left[\frac{\exp\left(\boldsymbol{X}_i^T\boldsymbol{\beta}\right)}{\exp\left(\boldsymbol{X}_i^T\boldsymbol{\beta}\right)+\frac{1}{\alpha}}\right]$$

The estimation of the parameters α dan β will be quite complicated so that other solutions withiterative numerical methods will be needed to solve these non linear equations.

3. Results and Discussion

3.1. Poisson Distribution Testing

Before starting the model formation, it will first be tested whether the response variable in this case is the number of neonatal mortality not following the Poisson distribution or not. The hypothesis used is the number of neonatal mortality not following the Poisson distribution. Obtained $\alpha = 1.000$ means data follows the Poisson distribution.

3.2 Assumption of Equidispersion

Binomial Negative Regression analysis has different assumptions from the *Poisson Regression* which must meet the equidispersion assumption that is the mean is equal to the variance. The analysis requires that the response variable does not have the same mean value as the variance value. Obtained $Mean = 0.25$ and $Variance = 0.197$, this means that the equidispersion assumption is not met. As a result, the *Binomial Negative Regression* analysis can be continued.

3.3 *Binomial Negative Regression* Analysis

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

$$\hat{Y} = \exp(-3.79851 + 0.005323X_1 + 0.001796X_2 + 0.002733X_3 - 1.635959X_4 + 0.007466X_5 - 0.000363X_6 + 0.229378X_7 + 0.004048X_8)$$

The result of the likelihood ratio (G) of 18,33262 can be concluded that the *Negative Binomial Regression* model can be used.

Next, the **Remove** method will be used in the regression analysis to find a better model in the *Binomial Negative Regression* analysis. First, the variable that gives the largest z value to the model will be removed. The selection of the best model will be seen from the smallest AIC value which is summarized in the Table 1.

Table 1. The selection of the best model will be seen from the smallest AIC value which is summarized

Model	AIC
All Variables	38,3
Without Variable X_2	36,3
Without Variable X_4	36,3

The final result of the predictive regression equation for the number of neonatal mortality:

$$\hat{Y} = \exp(-3.779736 + 0.005897X_1 + 0.002751X_3 - 1.581273X_4 + 0.007014X_5 - 0.000377X_6 + 0.223X_7 + 0.003954X_8) \quad \textbf{(1)}$$

or the following equation:

$$\hat{Y} = \exp(-3.81735 + 0.00509X_1 + 0.0018X_2 + 0.00275X_3 - 1.63411X_4 + 0.007555X_5 + 0.2289X_7 + 0.00403X_8) \quad \textbf{(2)}$$

The interpretation of the model in equation (**1**), β_0 is -3.779736 which means that the probability of neonatal mortality in jambi city is 0,0228. Parameter β_6 of 0,000377 means the number of infants who received BCG immunization as much as one person will reduce the chance of 0,9996 neonatal mortality in Jambi City if other variables are considered constant. The interpretation of the model in equation (**2**), β_0 of -3.81735 means that the probability of neonatal mortality per public health center in Jambi City of 0,022.

4. Conclusion

The regression model for the number of neonatal mortality per public health center is as follows

$$\hat{Y} = \exp(-3.779736 + 0.005897X_1 + 0.002751X_3 - 1.581273X_4 + 0.007014X_5 - 0.000377X_6 + 0.223X_7 + 0.003954X_8)$$

5. Acknowledgement

The researcher would like to thank for the finansial assistance provided by Universitas Jambi so that the research will.

References

[1] Mahmud, A., Ekoriano, M., Titisari, A. S., Wijayanti, U. T., Sitorus, M. A., & Rahmadhony, A. (2021). Determinants of modern contraceptives use in Indonesia: a spatial analysis. *Systematic Reviews in Pharmacy*, *12*(3), 769-777.

[2] Utari, D., & Eryando, T. (2018). Meeting the unmet need with a fit model for contraception mix. *Indian Journal of Public Health Research and Development*, *9*(11), 538-543.

[3] Damayati, N., Heldayani, E., & Anggraini, W. (2020). Trends and fertility control in South Sumatera Province (Further Analysis of IDHS 2017). *Sumatra Journal of Disaster, Geography and Geography Education*, *4*(1), 85-90.

[4] Utami, F. P. (2020). The availability of family planning information enable the used of traditional contraceptive in Yogyakarta, *International Journal of Biomedicine and Public Health*, *3*(3).

[5] Ekawati, R., Rahayuwati, L., & Praptiwi, A. (2020). The potential variables of first child's environmental quality: a retrospective analysis from 1994 to 2012. *Indonesian Journal of Urban And Environmental Technology*, *3*(2), 209-219.

[6] Spencer, A., Pitt, M., Allen, M., & Mújica-Mota, R. E. (2019). The heterogeneous effects of neonatal care: a model of endogenous demand for multiple treatment options based on geographical access to care.

[7] Dhrifi, A. (2019). Health-care expenditures, economic growth and infant mortality: Evidence from developed and developing countries. *CEPAL Review* No. 125, August 2018, 69.

[8] Patel, K. K., & Gouda, J. (2018). Infant mortality in northern and southern regions of India: differentials and determinants. *Social Science Spectrum*, *3*(2), 81-92.

[9] Benshaul-Tolonen, A. (2019). Local industrial shocks and infant mortality. *The Economic Journal*, *129*(620), 1561-1592.

[10] Karlsson, L. (2018). Indigenous infant mortality by age and season of birth, 1800–1899: did season of birth affect children's chances for survival? *International journal of environmental research and public health*, *15*(1), 18.

[11] Blankenship, S. A., Brown, K. E., Simon, L. E., Stout, M. J., & Tuuli, M. G. (2020). Antenatal corticosteroids in preterm small-for-gestational age infants: A systematic review and meta-analysis. *American Journal of Obstetrics & Gynecology MFM*, 100215.

[12] Bois, A., Garcia-Roger, E. M., Hong, E., Hutzler, S., Irannezhad, A., Mannioui, A., & Tronche, S. (2019). Infant mortality across species. A global probe of congenital abnormalities. *Physica A: Statistical Mechanics and its Applications*, *535*, 122308.

[13] Lehtonen, L., Gimeno, A., Parra-Llorca, A., & Vento, M. (2017). Early neonatal death: a challenge worldwide. In *Seminars in Fetal and Neonatal Medicine* Vol. 22, No. 3, pp. 153-160.

[14] Zhou, M., Wang, H., Zeng, X., Yin, P., Zhu, J., Chen, W., & Liang, X. (2019). Mortality, morbidity, and risk factors in China and its provinces, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*, *394*(10204), 1145-1158.

[15] Jones, K., Robb, M., Murphy, S., & Davies, A. (2019). New understandings of fathers' experiences of grief and loss following stillbirth and neonatal death: a scoping review. *Midwifery*, *79*, 102531.

[16] Kokou-Kpolou, K., Megalakaki, O., & Nieuviarts, N. (2018). Persistent depressive and grief symptoms for up to 10 years following perinatal loss: Involvement of negative cognitions. *Journal of affective disorders*, *241*, 360-366.

[17] Zeinalzadeh, A. H., Khodaei, R., & Heidarzadeh, M. (2017). Causes of neonatal mortality in the neonatal intensive care unit of Taleghani Hospital. *Iranian Journal of Neonatology IJN*, *8*(3), 58-61.

[18] Karimi, P., Mahmudi, L., Azami, M., & Badfar, G. (2019). Mortality in neonatal intensive care units in Iran: a systematic review and meta-analysis. *Iranian Journal of Neonatology IJN*, *10*(3), 70-80.

[19] Mangu, C. D., Rumisha, S. F., Lyimo, E. P., Mremi, I. R., Massawe, I. S., Bwana, V. M., ... & Mboera, L. E. (2021). Trends, patterns and cause-specific neonatal mortality in Tanzania: a hospital-based retrospective survey. *International Health*, *13*(4), 334-343.

[20] Baldwin, H. J., Patterson, J. A., Nippita, T. A., Torvaldsen, S., Ibiebele, I., Simpson, J. M., & Ford, J. B. (2017). Maternal and neonatal outcomes following abnormally invasive placenta: a population-based record linkage study. *Acta obstetricia et gynecologica Scandinavica*, *96*(11), 1373-1381.

[21] Tang, Y. (2018). Sample size for comparing negative binomial rates in noninferiority and equivalence trials with unequal follow-up times. *Journal of biopharmaceutical statistics*, *28*(3), 475-491.

[22] Zou, Y., Ash, J. E., Park, B. J., Lord, D., & Wu, L. (2018). Empirical Bayes estimates of finite mixture of negative binomial regression models and its application to highway safety. *Journal of Applied Statistics*, *45*(9), 1652-1669.

[23] Ardiles, L. G., Tadano, Y. S., Costa, S., Urbina, V., Capucim, M. N., da Silva, I., & Martins, L. D. (2018). Negative Binomial regression model for analysis of the relationship between hospitalization and air pollution. *Atmospheric Pollution Research*, *9*(2), 333-341.

[24] Yu, R., Wang, Y., Quddus, M., & Li, J. (2019). A marginalized random effects hurdle negative binomial model for analyzing refined-scale crash frequency data. *Analytic methods in accident research*, *22*, 100092.

[25] Caraka, R. E., Shohaimi, S., Kurniawan, I. D., Herliansyah, R., Budiarto, A., Sari, S. P., & Pardamean, B. (2018). Ecological show cave and wild cave: negative binomial gllvm's arthropod community modelling. *Procedia Computer Science*, *135*, 377-384.

[26] Yu, R., Wang, Y., Quddus, M., & Li, J. (2019). A marginalized random effects hurdle negative binomial model for analyzing refined-scale crash frequency data. *Analytic methods in accident research*, *22*, 100092.

[27] Rusli, R., Haque, M. M., King, M., & Voon, W. S. (2017). Single-vehicle crashes along rural mountainous highways in Malaysia: an application of random parameters negative binomial model. *Accident Analysis & Prevention*, *102*, 153-164.

[28] Khodadadi, A., Tsapakis, I., Das, S., Lord, D., & Li, Y. (2021). Application of different negative binomial parameterizations to develop safety performance functions for non-federal aid system roads. *Accident Analysis & Prevention*, *156*, 106103.

[29] Iqbal, W., Tang, Y. M., Chau, K. Y., Irfan, M., & Mohsin, M. (2021). Nexus between air pollution and NCOV-2019 in China: application of negative binomial regression analysis. *Process Safety and Environmental Protection*, *150*, 557-565.

[30] Shaon, M. R. R., Qin, X., Shirazi, M., Lord, D., & Geedipally, S. R. (2018). Developing a Random Parameters Negative Binomial-Lindley Model to analyze highly over-dispersed crash count data. *Analytic methods in accident research*, *18*, 33-44.

[31] Khattak, M. W., Pirdavani, A., De Winne, P., Brijs, T., & De Backer, H. (2021). Estimation of safety performance functions for urban intersections using various functional forms of the negative binomial regression model and a generalized Poisson regression model. *Accident Analysis & Prevention*, *151*, 105964.