

PAPER • OPEN ACCESS

Modeling of Maternal Mortality and Infant Mortality Cases in East Kalimantan using Poisson Regression Approach Based on Local Linear Estimator

To cite this article: Darnah *et al* 2019 *IOP Conf. Ser.: Earth Environ. Sci.* **243** 012023

View the [article online](#) for updates and enhancements.



IOP | ebooks™

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

Start exploring the collection - download the first chapter of every title for free.

Modeling of Maternal Mortality and Infant Mortality Cases in East Kalimantan using Poisson Regression Approach Based on Local Linear Estimator

Darnah¹, M I Utoyo², N Chamidah^{3*}

¹ Department of Mathematics, Faculty of Mathematics and Natural Science, Mulawarman University, Samarinda, Indonesia

^{2,3} Department of Mathematics, Faculty of Sciences and Technology, Airlangga University, Surabaya, Indonesia

Email : nur-c@fst.unair.ac.id

Abstract. The Poisson regression is the popular regression model on the discrete response variables. The regression function in the Poisson regression model can be estimate by using both the parametric and the nonparametric approaches. In this research, we study the Poisson regression by using the nonparametric regression approach based on local linear estimator for modeling the maternal mortality and the infant mortality cases in East Kalimantan Indonesia. According to the model, we get that the model works nicely to explain the relationship between the maternal mortality case and the shaman number, and between the infant mortality case and the shaman number in East Kalimantan. Also, we obtain their determination coefficient (R²) values of 0.929094 and 0.9625168, respectively.

1. Introduction

The quality of health service delivery of a region can be measure from the maternal mortality and the infant maternal cases. If the maternal mortality and the infant maternal cases are the high, then the quality of health services in the area is bad. Conversely, if the maternal mortality and the infant mortality cases are the low, then the quality of health services in the area is good. The maternal mortality and the infant mortality cases can also be an indicator of the welfare level of a region. The maternal mortality is the death of woman while pregnant or within 42 days of termination of pregnancy, irrespective of the duration and site of the pregnancy, from any cause related to or aggravated by the pregnancy or its management but not from accidental or incidental causes. The infant mortality is the death of an infant before his or her first birthday or infant under one year of age [1]. According to the result of the Survey of Population Survey (SUPAS) in 2015, the maternal mortality in Indonesia is the higher than neighboring countries in the ASEAN region, which is 305 per 100.000 live births and the infant mortality is 22 per 1.000 live births [2]. In East Kalimantan, the maternal mortality in 2013 is 113 cases and the infant mortality is 889 cases [1]. Generally, the maternal mortality and the infant mortality cases in Indonesia and East Kalimantan is the high. One of the efforts to lower the maternal mortality and the infant mortality is to analyze the cause.

Regression model is a statistical model tool used to estimate the relationship among the predictor and response variables. Predictive power is one of the most important characteristics of regression models [3]. The Poisson regression model has been widely used for public health in recent years to analyze the cause of a disease. For instance, [4] proposed model to know the factors influence of



filariasis in East Java, Indonesia and [5] used Poisson regression for modeling of AIDS disease in Kelantan Malaysia and then comparing it with the Binomial Negative regression. The regression function in the Poisson regression model can be estimate using the parametric approach and the nonparametric approach. In the parametric approach, the relationship between the response variable and the predictor variable is assume to follow a certain curve so that the modeling based on the curve form of the regression function. Research [6] develop a weather-based transmission line failure rate model by using Ordinary Least Square (OLS) based on polynomial regression technique. In fact, often the relationship between predictor variables and response variables does not follow a particular pattern but tends to be irregular, so the parametric approach is less appropriate to use.

The nonparametric regression approach becomes an alternative to this condition because nonparametric regression has high flexibility by not assuming the form of regression function but estimating regression function based on data pattern. [7]-[11] have studied nonparametric regression model by using Local Polynomial estimator. Nonparametric regression model based on Spline estimator have been studied by [12]-[15], and based on Kernel estimator by [16][17].

This research uses the nonparametric regression approach based on local linear estimator. The local linear estimator is an extension of the scatterplot smoothing technique into a linear model based on the likelihood function. The Poisson regression approach based on local linear estimator is used for modeling of maternal mortality and infant mortality in East Kalimantan. Based on the model, we determine the relationship between the maternal mortality cases and the shaman number, and also the relationship between the infant mortality cases and the shaman number in East Kalimantan.

2. Methods

In this research, we use the secondary data, i.e., the maternal mortality and infant mortality cases in East Kalimantan 2015 as a response variable recorded by Health Department of East Kalimantan and the predictor variable is the shaman number recorded by Statistical Bureau (BPS) of East Kalimantan. East Kalimantan province is divided into ten regencies. The data of shaman number, maternal and infant mortalities are given in Table 1.

Table 1. Shaman number, maternal and infant mortalities in ten regencies of East Kalimantan

Number	Regencies	Maternal Mortality (y_1)	Infant Mortality (y_2)	Shaman number (x)
1	Mahakam Ulu	1	17	69
2	North Penajam Paser	3	50	118
3	Berau	6	86	229
4	Paser	8	80	292
5	Bontang	8	36	20
6	Balikpapan	9	78	60
7	West Kutai	10	64	249
8	East Kutai	12	87	166
9	Samarinda	14	53	79
10	Kutai Kartanegara	29	211	519

For analyzing the data, we create R-code through the following steps. The first step is estimate parameter of local linear model by using weighted kernel locally maximum likelihood method. It can be done by iteratively reweighted least squares (IRLS). The second step is determine optimum

bandwidth by using cross validation (CV) criterion. The CV criterion that has minimum value is the best model. The last step is interpretation of result based on the best model.

Poisson Regression

Let $Y_i (i=1,2, \dots, n)$ is the random variable for count data that follow Poisson distribution with probability density function:

$$P(Y_i = y_i) = f(y_i) = \frac{\exp(-\mu_i) \mu_i^{y_i}}{y_i!}, y_i = 0, 1, 2, \dots \quad (1)$$

where $\mu_i = \exp(\mathbf{x}_i^T \boldsymbol{\beta})$; $i = 1, 2, \dots, n$, $\mathbf{x}_i$ is a vector of covariates and $\boldsymbol{\beta}$ is a vector of regression parameters.

The conditional mean and variance of the distribution as given by [18]:

$$E(Y_i) = \text{Var}(Y_i) = \mu_i. \quad (2)$$

Based on (1), we have the likelihood and log likelihood functions:

$$\ell(\boldsymbol{\beta}) = \prod_{i=1}^n \frac{\exp(-\mu_i) \mu_i^{y_i}}{y_i!}; \quad (3)$$

and

$$\ln \ell(\boldsymbol{\beta}) = \sum_{i=1}^n [y_i \ln \mu_i - \mu_i - \ln(y_i!)] \quad (4)$$

respectively.

The Poisson regression estimator is derived by solving the first order condition for the log likelihood function:

$$\sum_{i=1}^n [y_i - \mu_i] \mathbf{x}_i = \mathbf{0} \quad (5)$$

The estimates of $\boldsymbol{\beta}$ can be solved by using the Iteratively Weighted Least Square (ILWS) procedure.

Local Linear Estimator for Poisson Regression

If Y_i is the poisson random variable and X is a covariate, then the conditional mean of Y_i is

$$E(Y | X = x_i) = \mu_i = \exp(m(x_i)), \quad (6)$$

where $m(x_i)$ is an unknown function and will be estimated by using local maximum likelihood.

In equation (6), $m(x_i)$ function is approximated by Taylor expansion of degree one called as local linear estimator. Then, for x_i in a neighborhood of x_0 , we have:

$$m(x_i) \approx \exp \left[\mu_0(x_0) + \mu_1(x_0)(x_i - x_0) \right] \quad (7)$$

Therefore, we obtain the local log likelihood function:

$$l(x_0, \mu_1 | \mathbf{x}, y, h) = \sum_{i=1}^n \left\{ (m(x_i) y_i) - \exp(m(x_i)) - \ln(y_i!) \right\} \times K \frac{(x_i - x_0)}{h} \quad (8)$$

where K is kernel function and h is bandwidth. The bandwidth h controls smoothness of the fit. If h is too small, the fit becomes too noisy and the variance increases. Conversely, if h is too large, the fit

becomes over-smoothed and important feature may be distorted or lost completely. That is, the fit will have large bias. The bandwidth must be chosen to compromise this bias-variance trade-off [19].

The log likelihood local function in equation (8) can be written:

$$l(.) = \sum_{i=1}^n \left\{ \left[\binom{x_i}{0} + \binom{x_i}{1} (x_i - x_0) \right] y_i - \exp \left[\binom{x_i}{0} + \binom{x_i}{1} (x_i - x_0) \right] - \ln(y_i!) \right\} K \left(\frac{x_i - x_0}{h} \right) \quad (9)$$

The local linear Poisson estimator is obtained by solving the first order condition for the local log likelihood function. Derivative the first order $l(.)$ with respect to $\beta_0(x_0)$ is:

$$\frac{\partial l(.)}{\partial \beta_0(x_0)} = 0$$

$$\sum_{i=1}^n \left\{ y_i - \exp \left[\beta_0(x_0) + \beta_1(x_0)(x_i - x_0) \right] \right\} K \left(\frac{x_i - x_0}{h} \right) = 0$$

$$\beta_0(x_0) = \ln \left| \frac{\sum_{i=1}^n \left\{ y_i K \left(\frac{x_i - x_0}{h} \right) \right\}}{\sum_{i=1}^n \left\{ \exp \left[\beta_0(x_0) + \beta_1(x_0)(x_i - x_0) \right] K \left(\frac{x_i - x_0}{h} \right) \right\}} \right| \quad (10)$$

Derivative the first order $l(.)$ with respect to $\beta_1(x_0)$ is:

$$\frac{\partial l(.)}{\partial \beta_1(x_0)} = 0$$

$$\sum_{i=1}^n \left\{ \left(x_i - x_0 \right) y_i - \left(x_i - x_0 \right) \exp \left[\beta_0(x_0) + \beta_1(x_0)(x_i - x_0) \right] \right\} \times K \left(\frac{x_i - x_0}{h} \right) = 0$$

$$\sum_{i=1}^n y_i (x_i - x_0) K \left(\frac{x_i - x_0}{h} \right) - \exp \left[\beta_0(x_0) \right] \sum_{i=1}^n \exp \left[\beta_1(x_0)(x_i - x_0) \right] \times (x_i - x_0) K \left(\frac{x_i - x_0}{h} \right) = 0 \quad (11)$$

By substituting equation (10) into equation (11), we get

$$\sum_{i=1}^n y_i (x_i - x_0) K \left(\frac{x_i - x_0}{h} \right) - \exp \left[\beta_0(x_0) \right] \sum_{i=1}^n \exp \left[\beta_1(x_0)(x_i - x_0) \right] (x_i - x_0) K \left(\frac{x_i - x_0}{h} \right) = 0$$

$$\sum_{i=1}^n \exp \left[\beta_1(x_0)(x_i - x_0) \right] (x_i - x_0) K \left(\frac{x_i - x_0}{h} \right) = 0 \quad (12)$$

Likewise the system in equation (5), the system in equation (12) is nonlinear concerning the parameters, it is necessary to use iterative algorithm. The *Iteratively Reweighted Least Squares* (IRLS) is widely methods to solve it.

3. Result and Discussion

We begin this section with scatter plot of the maternal mortality case by shaman number and the infant maternal case by shaman number. They are given in Fig. 1 and Fig. 2.

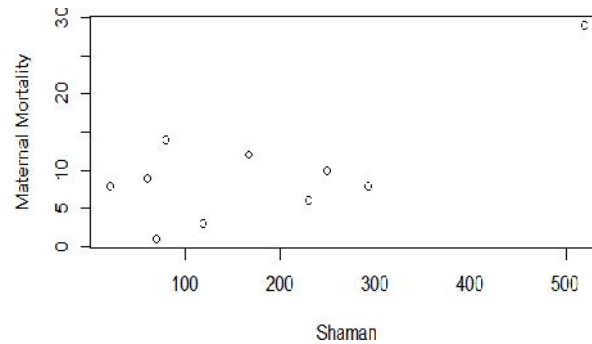


Figure 1. Plot of the Maternal Mortality Case by Shaman Number in East Kalimantan 2015

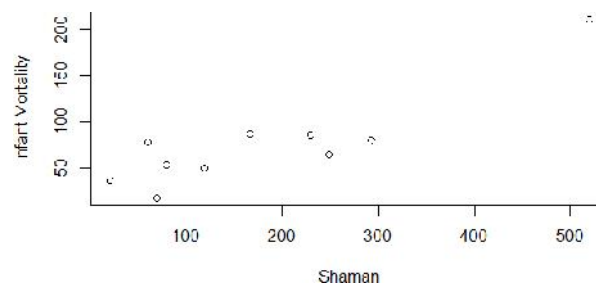


Figure 2. Plot of the Infant Mortality Case by Shaman Number in East Kalimantan 2015

The Fig. 1 and Fig.2 show that there are very high of both maternal and infant mortalities cases in East Kalimantan 2015 at the high of shaman number. The highest of both maternal and infant mortalities cases (i.e., 29 and 211 respectively) and the shaman number (i.e., 519) in the Kutai Kartanegara regency. Furthermore, we apply the local linear estimator for Poisson regression to the data. For choosing optimum bandwidth, we use cross validation (CV) criterion that has minimum value. Plot of CV versus bandwidth of both maternal and infant mortalities are showed in Figure 5 and Figure 6.

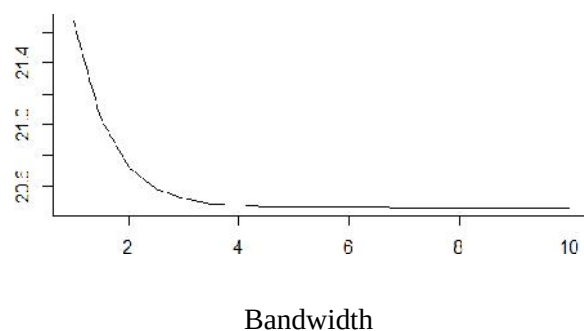


Figure 5. The Plot of CV Value Versus Bandwidth for the Maternal Mortality Case

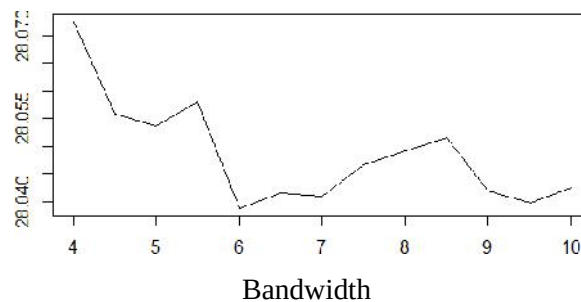


Figure 6. The Plot of CV Value Versus Bandwidth for the Infant Mortality Case

Based on **Figure 5**, we have the optimum bandwidth (h), minimum CV and R^2 value of the maternal mortality case, i.e., 10; 20.45757 and 0.929094, respectively. Also, based on **Fig. 6** we get the optimum bandwidth, minimum CV and R^2 value of the maternal mortality case, i.e., 6; 28.03881 and 0.9625168. By considering the R^2 value, it implies that the model works nicely for explaining the relationship between the maternal mortality and the infant mortality cases with the shaman number in East Kalimantan. The plot of observation (green) and estimation (red) to the maternal mortality and the infant mortality cases by shaman number is given in **Figure 7** and **Fig. 8**.

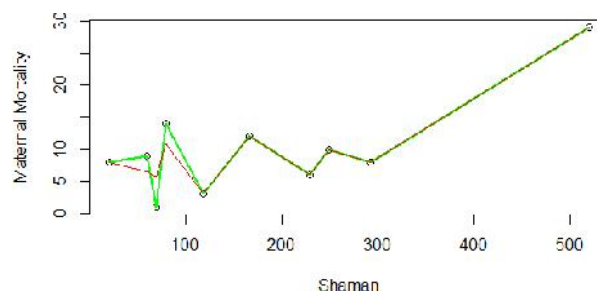


Figure 7. The Plot of Observations (Green) and Estimation (Red) for the Maternal Mortality Case

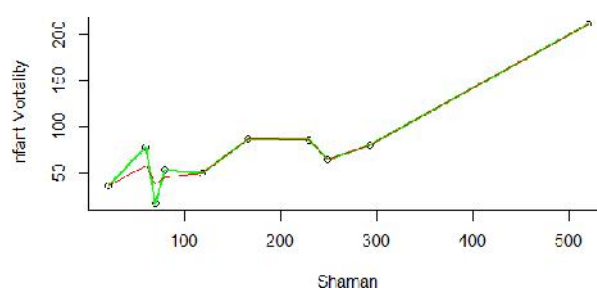


Figure 8. The Plot of Observations (Green) and Estimation (Red) for the Infant Mortality Case

The estimated model for explaining the relationship between the maternal mortality case and the shaman number in KutaiKartanegara regency East Kalimantan is

$$\hat{m}(x) = \exp \left[3.367296 - (3.658021e - 19)(x - 519) \right], 509 < x < 529$$

It means in every addition 10 shamans, will decrease 10 maternal mortality cases.

Meanwhile, the estimated model for explaining the relationship between the infant mortality case and the shaman number in KutaiKartanegara regency East Kalimantan is

$$\hat{m}(x) = \exp \left[5.351858 - (4.021859e - 19)(x - 519) \right], 513 < x < 525$$

It means in every addition 10 shaman, will increase 10 infant mortality cases.

4. Conclusion

The modeling of both the maternal mortality case and infant mortality case with the shaman number in East Kalimantan by using Poisson regression model approach based on local linear estimator give two estimated models of exponential functions which satisfy the goodness of fit criterion, i.e., R^2 values tend to one. In this paper, we discussed the maternal mortality case and infant mortality case partially. So, we use the single response model. Therefore, in the future, we will estimate the maternal mortality case and infant mortality case simultaneously by using the bi-response Poisson regression model approach based on local linear estimator.

Acknowledgement

The authors thank to the Ministry of Research, Technology and Higher Education of the Republic of Indonesia for the Doctoral Research Grant with contract number 91/UN17.41/KL/2018.

References

- [1] Dinkes Kaltim., 2013 Profil Kesehatan Kalimantan Timur, 16-20.
- [2] BPS, 2015 *Profil Penduduk Indonesia Hasil Supas*, BPS.
- [3] K. Varada Rajkumar, Yesubabu Adimulam and K. Subrahmanyam, 2018 *International Journal of Engineering & Technology*, **7**
- [4] Darnah, 2016. *AIP Conference Proceeding*, 1723:030006.
- [5] M. Asrul and N. N. Naing, 2012 *IOSR Journal of Mathematics*, **3**
- [6] M.K.N Arshad, N. Aminudin, M. Marsadek, S.Z.M Noor, R.H Salimin and D. Johar, 2018 *International Journal of Engineering & Technology*, **7**
- [7] N. Chamidah, I. N. Budiantara, S. Sunaryo, and I. Zain, 2012 *Journal of Mathematics and Statistics, Sciences Publication*, **8**
- [8] E. T. Astuti, I. N. Budiantara A Sunaryo and M. Dokhi, 2013 *International Journal of Applied Mathematics and Statistics*, **33**.
- [9] A. Baillo and A. Grane, 2009 *Journal of Multivariate Analysis*.
- [10] Z. Jin & W. He, 2016 *Journal of Multivariate Analysis*, **147**.
- [11] N. Chamidah and M. Rifada, 2016 *Far East Journal of Mathematical Sciences*, **99**.
- [12] B. Lestari, I. N. Budiantara, S. Sunaryo, and M. Mashuri, Oct. 12-13 *Proceeding of the Indo MS International Conference on Mathematics and Its Application*.
- [13] B. Lestari, I. N. Budiantara, S. Sunaryo, and M. Mashuri, 2010 *Journal of Basic Science*, **11**.
- [14] I. N. Budiantara, and D. T. J. Purnomo, 2011 *International Journal of Basic & Applied Sciences IJBS-IJENS*, **11**.
- [15] N. Chamidah and Eridani, 2015 *Int. J. Appl. Math*, **53**.
- [16] C. Tenreiro, 2013 *Statistical Journal*, **11**.
- [17] N. Chamidah and T. Saifuddin, 2013 *Applied Mathematical Science*, **7**.
- [18] J. A. Santos and M. M. Neves, 2008 *A Local Maximum Likelihood Estimator for Poisson Regression*, **68**.
- [19] H. Liang and D. Z. Cheng, 2005 *Annals of Biomedical Engineering*, **33** 1