© 2018 IJSRSET | Volume 4 | Issue 9 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
Themed Section : Engineering and Technology

Comparison Poisson Regression And Spatial Autoregressive (SAR) Poisson (Case Study : In Preventing The Malnutrition Factors In Java Island)

Ayu Sofia*, Budi Susetyo, Muhammad Nur Aidi
Department of Statistics, Bogor Agricultural University, Bogor, West Java, Indonesia

ABSTRACT

Indonesia is one of the developing countries whose population is always increasing every year as well as causing the population density in Indonesia to increase also. Based on the results of the population census in 2000 and 2010, Java island is the island with the highest population density. It can be caused by the number of high malnutrition events as well. Malnutrition is a major factor in infant under-five mortality. Based on world health statistics, malnutrition is the cause of one-third of child mortality under 5 years. The percentage of malnutrition sufferer from Java Island is 4,34%. Therefore it can be said that the incidence suffering from malnutrition in children under five in Java is a rare event, so it follows the distribution of Poisson. Where the Poisson distribution can be said as a very rare occurrence occurs. The result of the Moran's I tests states that there are spatial dependencies on severe malnutrition's rate of children under five years in Java Island. Therefore, Spatial Autoregressive (SAR) Poisson method is used in this experiment. The aims of this study is to compare factors that influence the malnutrition of children under five in Java Island using Poisson Regression and Spatial Autoregressive (SAR) Poisson. The result Variables which significantly affect severe malnutrition on Java Island through Poisson Regression and SAR Poisson method are the number of household with clean healthy living behavior (X_2), the number of houses with good health status (X_3), the number of household with access to source of clean water (X_4), maternal and child health center with active status (X_5), Vitamin A supplementation coverage(X_7) and the number of babies given breastfeeding exclusively [X]_8)The differences are variable the number of babies with low birth weight (X 1) is not significantly affecting severe malnutrition through SAR Poisson method.

Keywords : Malnutrition, Poisson Regression, Spatial Autoregressive (SAR) Poisson.

I. INTRODUCTION

Malnutrition is one of the serious problems, especially in developing countries such as Indonesia. As one of the countries with a very diverse population complexity, Indonesia is faced with the dynamics of malnutrition. Malnutrition is a major factor in infant and under-five mortality. Based on world health statistics, malnutrition is the cause of one-third of child mortality under 5 years. The symptoms of malnutrition usually occur gradually. Children under five who suffer from malnutrition in the early stages will lose weight until at a point of weight loss that is drastically far from the weight of a normal child in general. Nutritionists say if within 6 months of weight the child does not increase, it is an early sign of malnutrition. Furthermore, after the occurrence of weight loss, it will be followed by a decline in the functions of vital organs in children. In this phase, the child usually will be very susceptible to disease. This is called the critical phase where many children are not able to survive and will eventually die.

Indonesia is one of the developing countries whose population is always increasing every year as well as causing the population density in Indonesia to increase also. The density of the population itself is the number of population per unit area. The increasing of the population density, the more population problems that can arise. Java Island is the island with the highest population density in Indonesia. Based on the population census in 2010 (SP2010), five provinces with the highest population density in Indonesia are Jakarta,West Java, DI Yogyakarta, Banten and Central Java with a population density of 14,469 people/km2,1.217 people/ km2, 1104 people/ km2, 1100 people/ km2, 987 people/ km2, respectively (BPS 2010). Because of population density in each province of Java is high enough, it will allow the number of high malnutrition events as well.

In Indonesia, despite the national decline, but regions, there are some provinces that have recorded high rates of malnutrition in children under five. Central Java and East Java provinces for 6 consecutive years (2005-2010) fall into the category of 10 provinces with the highest malnutrition cases. Even in 2006, Central Java Province contributed the highest malnutrition rate of children under five in the national scale, which was 10376 cases, although in 2011 the number could be reduced to 3187 cases. When grouped by island region, it appears that the Java-Bali region is the largest contributor to the national figure (64.6%) (Gizinet 2011).

The percentage of malnourished children under five in Java is 4.34%. Therefore it can be said that the incidence of children under five suffering malnutrition in Java is a rare occurrence, so the incidence of children under five suffer from malnutrition following the Poisson distribution. In addition, the number of children with malnutrition in one region is thought to be influenced by the number of children with malnutrition around them. Therefore the aspect of location (spatial) also needs to be considered in determining factors that significantly affect the incidence of malnutrition in children in an area. A model that can explain the relationship between an area and its surrounding region is a spatial model. The purpose of this study was to compare factors that affect the malnutrition of children under five in Java using Poisson Regression and Spatial Autoregressive (SAR) Poisson.

II. METHODS AND MATERIAL

A. MATERIAL

The data that will be used in this study is secondary data which are published by Ministry of Health of the Republic of Indonesia with the tittle of Basic Health Reasearch (2012) from six provinces in Java Island that are DKI Jakarta, Banten, West Java, Central Java, East Java, and the Special Region of Yogyakarta. The analysis unit of this research is the regency/city in Java Island, which amounts is 117 districts / cities.

Tabel 1. Response Variable and Explanatory Variables

Types of variables	Variables	Names of Variables
Response Variable	Y	The number of malnutrition of children under five
Explanatory Variables	X_1	the number of babies with low birth weight
	X_2	the number of household with clean healthy living behavior
	X_3	the number of houses with good health status
	X_4	the number of household wih access to source of clean water
	X_5	maternal and child health center with active status
	X_6	children's health care coverage
	X_7	Vitamin A supplementation coverage
	X_8	the nurmber of babies given breastfeeding eclusively

B. METHODS

Stages of data analysis performed using software R.3.4.4. are as follows:

1. Description of data by using map on the response variable and explanatory variables. Response variable to see malnutrition of children under five in Java Island and explanatory variables to see the factors that cause malnutrition of children under five in Java Island.

2 Calculating the correlation between the explanatory variables and response variables and then checking the variables of the explanatory variables entered into the model.

3. Determining and predicting parameters for the Poisson Regression model.

4. Determining the spatial weighted matrix W * with queen contiguity method that the value is 0 or 1 that describes the neighboring structure for each unit.

5. Spatial spatial matrix is formed, then do the exploration of data to calculate spatial correlation between Regency/City in Java Island by using Moran Index.

The Moran Index with standardized spatial weighted matrix W :

$$I = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\,(x_j - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

which:

I : Moran's Indeks

n : Number of incident locations

x_i : value at location i

x_j : Value at location j

$\bar{x}$: Average of the number of variables

w_{ij} : Elements on standardized weights between region i and j

6. Predict SAR model parameters (spatial autoregressive model) with Newton-Raphson method.

-Determine $\hat{\beta}_{(0)}{}^{*}$ which $\beta_{(0)}{}^{*} = [\rho_0 \quad \beta_{00} \quad \cdots \quad \beta_{k0}]$, iterasi pada saat t=0 .

- Forming the gradient vector $g'_{t+1} = \left[\frac{\delta L(\beta^*)}{\delta\rho}, \frac{\delta L(\beta^*)}{\delta\beta}\right]$ and t representing the iteration number.

- Form a hessian matrix H:

$$H_{(k+1)x(k+1)} = \begin{bmatrix} \frac{\delta^2 \ln L(\beta^*)}{\delta\rho^2} & \frac{\delta^2 \ln L(\beta^*)}{\delta\beta_0\delta\rho} & \cdots & \frac{\delta^2 \ln L(\beta^*)}{\delta\beta_k\delta\rho} \\ & \frac{\delta^2 \ln L(\beta^*)}{\delta{\beta_0}^2} & \cdots & \frac{\delta^2 \ln L(\beta^*)}{\delta\beta_0\beta_k} \\ \vdots & \ddots & & \vdots \\ & \cdots & & \frac{\delta^2 \ln L(\beta^*)}{\delta{\beta_k}^2} \end{bmatrix}$$

- Input the value of $\hat{\beta}_{(0)}{}^{*}$ to the first elements vector g dan matrix H so it is obtained vector $g_{(0)}$ and $H_{(0)}$

- Performing iterations starting from t=0 in the equation $\hat{\beta}_{(t+1)}{}^{*} = \beta_t{}^{*} - H_t{}^{-1}g'_t$, the value $\beta_t{}^{*}$ represents a set of parameter estimators that converge on a t-iteration.

- If it has not reached the convergent parameter estimator, then in step 2 do again until it reaches convergence. Convergent criteria are obtained when the root feature of the Fisher information matrix is positive.

7. Test the significance of parameters by using Wald test.

$$G_p = \left\{\frac{\hat{\rho}_0}{\widehat{se}(\widehat{\rho_0})}\right\}^2$$

Statistics G_p will follow the distribution of χ^2 with degrees of freedom equal to 1. The decision criterion is to reject H_0 , if $G_p > \chi^2{}_{(\alpha;1)}$. The hypothesis for the coefficient parameter β_k (fleiss et.al . 2003) adalah :

$H_0 : \beta_k = 0$

$H_1 : \beta_k \neq 0$

with Wald test statistic:

$$G_\beta = \left\{\frac{\hat{\beta}_0}{\widehat{se}(\widehat{\beta_0})}\right\}^2$$

Statistics G_β will follow the distribution of χ^2 with degrees of freedom equal to 1. The decision criterion is to reject H_0, if $G_\beta > \chi^2{}_{(\alpha;1)}$.

8. Test the model's goodness by calculating the determinant coefficient based on devians ($R^2{}_{DEV}$), corrected by degree of freedom ($R^2{}_{DEV,db}$), and based on sum of squares ($R^2{}_{jk}$).

The formula for $R^2{}_{DEV}$:

$$R^2{}_{DEV} = 1 - \frac{\ln L(y) - \ln L(\hat{\mu})}{\ln L\,(y) - \ln L(\bar{y})}$$

The formula for $R^2{}_{DEV,db}$:

$$R^2{}_{DEV,db} = 1 - \frac{(n-k-1)^{-1}[\ln L(y) - \ln L(\hat{\mu})]}{(n-1)^{-1}[\ln L\,(y) - \ln L(\bar{y})]}$$

The formula for $R^2{}_{jk}$:

$$R^2{}_{jk} = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{\mu}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

9. Comparing results from Poisson and SAR Regression (spatial autoregressive model)
10. Conclusion.

III. RESULTS AND DISCUSSION

A. Data Exploration

Based on the picture below, almost all districts in Java Island have malnutrition incidence rate under 150 from 10000 children under five. The lowest malnutrition rate of the whole province in Java is represented by Central Java Province. The districts in Java that have the highest incidence of malnutrition of children under five years old are shown on map with yellow color that is Situbondo Regency with malnutrition incidence of toddlers amounted to 293 children from 10000 children under five. The second highest district is the district with a pink color on the map of the Ngawi Regency with the number of malnutrition infants of 199 infants from 10,000 children under five.

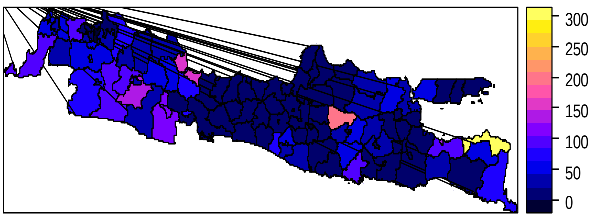

Figure 1. The Distribution of Underfive Malnutrition in Java

The multicollinearity test results using the VIF criterion.

Table 2. VIF values of explanatory variables

Peubah	Nilai VIF
X_1	2.658511
X_2	4.113674
X_3	2.862542
X_4	1.508496
X_5	1.339068
X_6	22.006635
X_7	23.018779
X_8	1.270805

Based on Table 2 , the VIF values that obtained is not all explanatory variables have values less than 10. It means that there is multicollinearity for this malnourished data. Variables contained multicollinearity are X6 and X7. So it is necessary to select the variables that will be used in the model. The results of the selection of variables showed that the number of the number of babies with low birth weight (X_1) , the number of household with clean healthy living behavior (X_2), the number of houses with good health status (X_3) , the number of household wih access to source of clean water (X_4) , maternal and child health center with active status (X_5) , Vitamin A supplementation coverage (X_7) and the nurmber of babies given breastfeeding eclusively (X_8) is the variables to be used in this study.

Comparison of Moran Index values, expected values of E (I) and Var (I) variance values are presented in Table 3.

Tabel 4. Moran Index, expectation Moran Index and variance Moran Index value for malnutrition data of children under five in Java Island.

Nilai Indeks Moran	*E(I)*	*Var(I)*	Nilai-*p*
0.14276	-0.00862	0.00365	0.006163

B. Poisson Regression

Tabel 5. The expected value of the Poisson regression model parameters

Parameters		expected value	Pr(>|z|)
Intersep	$\boldsymbol{\beta_0}$	1.836x10^{-04}	<2e-16 *
X_1	$\boldsymbol{\beta_1}$	-7.137 x10^{-07}	<2e-16 *
X_2	$\boldsymbol{\beta_2}$	-1.415 x10^{-06}	<2e-16 *
X_3	$\boldsymbol{\beta_3}$	2.856 x10^{-06}	<2e-16 *
X_4	$\boldsymbol{\beta_4}$	-2.532 x10^{-04}	<2e-16 *
X_5	$\boldsymbol{\beta_5}$	9.839 x10^{-06}	<2e-16 *
X_7	$\boldsymbol{\beta_7}$	-2.171 x10^{-05}	<2e-16 *
X_8	$\boldsymbol{\beta_8}$	1.836x10^{-04}	<2e-16 *

note : * significant at 5%

Table 4 in the significance test showed that all the variables showed significant results. This model shows that the lower the number of household with clean healthy living behavior (X_2), the number of houses with good health status (X_3), maternal and child health center with active status (X_5), and and the nurmber of babies given breastfeeding eclusively (X_8) will increase the number of malnourished children under five in Java. In contrast, the increase in the number of babies with low birth weight (X_1) , the number of household with access to source of clean water (X_4), and Vitamin A supplementation coverage (X_7) will increase the number of malnutrition sufferers.

C. Spatial Autoregressive Poisson (SAR Poisson)

The parameter estimation on Spatial Autoregressive Poisson model coefficient is done by using maximum probability estimation method which the estimation process is done iteratively with newton-Raphson method. The iterative process will stop until the approximate difference of each convergent iteration.

Tabel 5. The expected value of the SAR Poisson model parameters

Parameter		Expected value	error	G Value	χ_1^2
Rho	$\boldsymbol{\rho}$	0.1	4.116 x10^{-08}	5.90e+12*	
Intersep	$\boldsymbol{\beta_0}$	5.06	6.263 x10^{-09}	6.52e+17*	
X_1	$\boldsymbol{\beta_1}$	-1.270 x10^{-05}	1.424 x10^{-05}	7.94e-01	
X_2	$\boldsymbol{\beta_2}$	-5.280 x10^{-07}	7.774 x10^{-08}	4.61e+01*	
X_3	$\boldsymbol{\beta_3}$	-1.327 x10^{-06}	7.500 x10^{-08}	3.13e+02*	3.841
X_4	$\boldsymbol{\beta_4}$	2.028 x10^{-06}	3.386 x10^{-08}	3.58e+03*	
X_5	$\boldsymbol{\beta_5}$	-5.555 x10^{-04}	1.102 x10^{-05}	2.53e+03*	
X_7	$\boldsymbol{\beta_7}$	1.049 x10^{-05}	9.313 x10^{-08}	1.27e+04*	
X_8	$\boldsymbol{\beta_8}$	-3.366 x10^{-05}	7.495 x10^{-07}	2.01e+03*	

note : * significant at 5%

Table 5 in the significance test showed that the variable the number of babies with low birth weight (X_1) was not significant, while for other variables showed significant results. This model shows that the number of household with clean healthy living behavior (X_2), the number of houses with good health status (X_3), maternal and child health center with active status (X_5), and and the nurmber of babies given breastfeeding eclusively (X_8) will increase the number of malnourished children under five in Java. In contrast, the increase the number of household with access to source of clean water (X_4),and Vitamin A supplementation coverage (X_7) will increase the number of malnutrition sufferers.

D. The comparison between Poisson Regression and Spatial Autoregressive (SAR) Poisson

Table 6. The R2 comparison between Poisson Regression and Spatial Autoregressive (SAR) Poisson

	R^2_{DEV}	R^2_{adj}	R^2_{jk}
Poisson Regresson	43%	42%	54%
SAR Poisson	41%	41%	55%

Based on table 6, it can be seen that the comparative analysis of Poisson Regression and Poisson SAR through the value of R^2 does not produce significant difference so that it can be concluded that Poisson Regression and Poisson SAR as well as in predicting factors that affect the malnutrition of children under five in Java.

Factors influencing the number of malnutrition patients in Java Island based on the Poisson regression model and the Spatial Autoregressive Poisson model yielded almost the same result. In the Poisson Regression model, all variables used in the model that are, the number of babies with low birth weight (X_1), the number of household with clean healthy living behavior (X_2), the number of houses with good health status (X_3), the number of household wih access to source of clean water (X_4) , maternal and child health center with active status (X_5) , Vitamin A supplementation coverage (X_7) and the nurmber of babies given breastfeeding eclusively (X_8) affect malnutrition children under five in Java Island. While in Spatial Autoregressive model resulted factors that influence spatially to malnutrition of under five in Java Island that are the number of household with clean healthy living behavior (X_2), the number of houses with good health status (X_3), the number of household wih access to source of clean water (X_4) , maternal and child health center with active status (X_5), Vitamin A supplementation coverage (X_7) and the number of babies given breastfeeding eclusively (X_8).

Any reduction the number of household with clean healthy living behavior, maternal and child health center with active status, the number of houses with good health status, and the number of exclusive breastfeeding will increase the expectation of the number of malnourished children under five. Each increase in the number of household wih access to source of clean water and Vitamin A supplementation coverage will increase the expectation value of malnourished children under five. The difference between the two models above is that the BBLR (Low Birth Weight Infant) weights do not affect spatially if tested using the Spatial Autoregressive Poisson model.

IV. CONCLUSION

The conclusions that can be formulated in this research are Poisson Regression and Poisson (SAR) as well as predicting factors influencing malnutrition of children under five years old in Java Island and the number of household with clean healthy living behavior (X_2), the number of houses with good health status (X_3), the number of household wih access to source of clean water (X_4) , maternal and child health center with active status (X_5) , Vitamin A supplementation coverage (X_7) and the nurmber of babies given breastfeeding eclusively (X_8) similarly affect malnutrition of under five children based on Poisson Regression model and Spatial Autoregressive Poisson. However, the difference of the the number of babies with low birth weight did not have a spatial effect if tested using the Spatial Autoregressive Poisson model.

V. REFERENCES

[1]. Prof. Ministry of Health of Republic of Indonesia. 2012. Profil Kesehatan DKI Jakarta 2012. DKI Jakarta (ID): KEMENKES.

[2]. Ministry of Health of Republic of Indonesia. 2012. Profil Kesehatan Banten 2012. Banten (ID): KEMENKES.

[3]. Ministry of Health of Republic of Indonesia. 2012. Profil Kesehatan Jawa Barat 2012. Jawa Barat (ID): KEMENKES.

[4]. Ministry of Health of Republic of Indonesia. 2012. Profil Kesehatan Jawa Timur 2012. Jawa Timur (ID): KEMENKES.

[5]. Ministry of Health of Republic of Indonesia. 2012. Profil Kesehatan Yogyakarta 2012. Yogyakarta (ID): KEMENKES.

[6]. Ministry of Health of Republic of Indonesia. 2012. Profil Kesehatan Jawa Tengah 2012. Jawa Tengah (ID): KEMENKES.

[7]. Fleiss JL, Levin B, Paik MC. 2003. Statistical Methods for Rates and Proportions. Ed ke-3. USA: Columbia University.

[8]. Fotheringham AS, Rogerson PA. 2009. Handbook of Spatial analysis. London:Sage Publications Ltd.

[9]. Gizinet. 2011. Laporan Kasus Gizi buruk balita 2010 : Menurun. http://gizi.depkes.go.id/laporan-kasus-gizi-buruk-2010-menurun (10 April 2013).

[10]. Kosfeld R. Spatial Econometric. 2006. URL: http://www.scribd.com.

[11]. Lambert DM, Brown JP, Florax RJGM. 2010. A Two-Step Estimator for a Spatial Lag Model of Counts: Theory, Small Sample Performance and application. USA: Dept. of Agricultural Economics Purdue University.