APPLICATION OF NEGATIVE BINOMIAL REGRESSION ANALYSIS TO OVERCOME THE OVERDISPERSION OF POISSON REGRESSION MODEL FOR MALNUTRITION CASES IN INDONESIA

Yudi Setyawan[1], Kris Suryowati[2*], Dita Octaviana[3]
[1,2,3]Department of Statistics IST AKPRIND
Jl. Bimsakti No. 3, Pengok. Yogyakarta 55222, Indonesia

***e-mail**: suryowati@akprind.ac.id*

ABSTRACT

Indonesia is one of the developing countries that is struggling to eradicate the malnutrition problem. Malnutrition that occurs over a long period of time can have an impact on the deaths of sufferers and decrease human quality of life. This study aims to model the case of malnutrition that occurred in Indonesia Provinces during 2015 and get the main factors that cause the malnutrition problem. Variables studied consist of Malnutrition (Y), Vitamin A consumption (X_1), Exclusive breastfeeding (X_2), Immunization (X_3), Water quality (X_4), Healthcare center (X_5), and Poverty level (X_6). Based on the Kolmogorov-Smirnov test, the results of malnutrition data in Indonesia Province in 2015 do not follow Poisson distribution because of overdispersion. The presence of overdispersion cases in the Poisson regression model will have an impact on the inappropriateness of inferences. An alternative model that accommodates this case is the negative binomial regression model. By using this model, factors that are considered influencing malnutrition cases in Indonesia provinces in 2015 are Immunization (X_3), Water quality (X_4), and Poverty level (X_6). The best model obtained from the negative binomial regression analysis is $\hat{\mu}_i = exp(2.5111 - 0.0338X_3 + 0.0295X_4 + 0.0576X_6)$.

Keywords: *Malnutrition, Poisson Regression, Overdispersion, Negative Binomial Regression.*

INTRODUCTION

Nowadays, malnutrition is one of the problems that is still a major concern for developing countries. A chronic malnutrition can affect the decreased quality of life such as stunting that can reduce a person's productivity and increase risks of non-communicable diseases such as diabetes and heart disease. This is called the double burden of malnutrition (World Bank, 2013). Besides above effects, malnutrition is a major cause of child morbidity and mortality (Tette, 2015).

The number of malnutrition cases for under five-year children found in Indonesia in 2015 amounted to 26,518 people or 0.11% of the total number of under five-year children in Indonesia (Ministry of Health the Republic of Indonesia, 2016). Whereas in 2014 cases of malnutrition found in Indonesia was recorded at 32,521 people or 0.14% of the total number of under five-year children in Indonesia. This figure shows that the case of malnutrition in Indonesia Provinces in 2015 has slightly decreased. A small decline indicates that the government's coping strategies to address malnutrition in Indonesia was not optimal. Therefore, a deep analysis to determine the factors that cause malnutrition problem in Indonesia is still needed. The factors consist of socio-economic and health-related related factors. Unemployment and a lack of knowledge about recommended infant and child feeding practices are some examples of socio-economic factors, while low birth weight, inadequate vitamin A supplementation are health-related factors affecting the malnutrition problem for under age children (Kadima, 2012).

Poisson regression analysis is an analysis used to describe the relationship between response variable (Y) in the form of discrete data and one or more predictor variables (X). In the Poisson regression analysis, the response variable (Y) is assumed to be Poisson distributed, so that the variance value for the response variable (Y) must equal to its mean or commonly called equidispersion (Olsson, 2002). However, it is often found that the variance of the response variable (Y) is much greater than its mean namely overdispersion. The existence of overdispersion cases may result in incorrect standard errors of parameter estimators that lead to incorrect assessment of the significance of individual regression parameters. Consequently, our interpretation of the regression model will also be incorrect. Inappropriate imposition of the Poisson may underestimate the standard errors and overstate the significance of the regression parameters. As a result, it can give misleading inference about the regression parameters (Ismail, 2007). Therefore, alternative models should be proposed when overdispersion problem occurs. The negative binomial regression model offers a good solution to this problem.

MATERIAL AND METHOD

This study used the method of negative binomial regression analysis to overcome the problem of overdispersion in the Poisson regression model for estimating malnutrition level (Y) in Indonesia provinces. The predictors used to estimate are Vitamin A consumption (X_1), Exclusive breastfeeding (X_2), Immunization (X_3), Water quality (X_4), Healthcare center (X_5), and Poverty level (X_6).

There are several types of negative binomial model such as canonical negative binomial (NBC), linear negative binomial (NB1), and traditional negative binomial (NB2). The negative binomial regression model referring to NB2 can be formed from the Poisson-Gamma mixture distribution (Hilbe, 2011). The negative binomial regression model has the same usage as the Poisson regression model to analyze the relationship between a discrete response variable and one or more predictor variables. However, the negative binomial regression analysis has the dispersion parameter useful for describing the variation of the data. When the dispersion parameter approaches to zero, the data is said to be Poisson distributed.

1. **Poisson Regression**

Poisson regression is generally used to describe the relationship between response variable (Y) which is assumed Poisson distributed and predictor variables (X). If Y is a discrete data distributed Poisson with parameter μ>0, then the probability density function is

$$f(y_i) = \frac{e^{-\mu_i}\mu_i^{y_i}}{y_i!}, y_i = 0,1,2,\dots$$

It can be proved that $Var\ (y_i) = E(y_i) = \mu_i$. While the equation model on Poisson regression is as follows:

$$\mu_i = exp(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki}) + \varepsilon_i, \ \ i = 0,1,2,\ldots n$$

where,
μ_i is the expected value of the response variable (Y),
$\beta_0, \beta_1, \ldots, \beta_k$ are the coefficients of the regression parameters, and
ε_i is the error of the ith observation.

The estimation of the parameter coefficients $\beta_0, \beta_1, \ldots, \beta_k$ in the Poisson regression analysis is performed by the Maximum Likelihood Estimation (MLE) method, i.e. by maximizing its log-likelihood function with respect to the parameters to be estimated. This process is done iteratively using Newton-Raphson method. The likelihood function for Poisson regression is $L = \prod_{i=1}^{n} f(y_i; \mu_i)$, so that its log-likelihood function is:

$$l = \ln\ (L) = \ln\ (\prod_{i=1}^{n} \exp\ [y_i\ ln(\mu_i) - \mu_i - \ln(y_i!)]) = \sum_{i=1}^{n} [y_i\ ln(\mu_i) - \mu_i - \ln(y_i!)] \ .$$

By substituting $\mu_i = e^{\left(\beta_0 + \Sigma_{j=1}^{k} \beta_j X_{ji}\right)}$, then

$$l(y_i; \beta)\ = \sum_{i=1}^{n} \left\{ y_i \left(\beta_0 + \sum_{j=1}^{k} \beta_j X_{ji} \right) - e^{\left(\beta_0 + \Sigma_{j=1}^{k} \beta_j X_{ji}\right)} - ln(y_i!) \right\}.$$

The maximum likelihood estimator for $\beta_0, \beta_1, \ldots, \beta_k$ is obtained by solving equations:

$$\frac{\partial l(y_i; \beta)}{\partial \beta_0} = \sum_{i=1}^{n} \left\{ y_i - e^{\beta_0 + \Sigma_{j=1}^{k} \beta_j X_{ji}} \right\} = 0$$

$$\frac{\partial l(y_i; \beta)}{\partial \beta_1} = \sum_{i=1}^{n} \left\{ y_i x_{1i} - x_{1i} e^{\beta_0 + \Sigma_{j=1}^{k} \beta_j X_{ji}} \right\}\ = 0$$

...

$$\frac{\partial l(y_i; \beta)}{\partial \beta_k} = \sum_{i=1}^{n} \left\{ y_i x_{ki} - x_{ki} e^{\beta_0 + \Sigma_{j=1}^{k} \beta_j X_{ji}} \right\}\ = 0$$

The above nonlinear system of equations can be solved iteratively by using The Newton-Raphson method. This process involves complex numerical computation.

Overdispersion is a condition that occurs in Poisson regression analysis when the variance value of the response variable is greater than its mean. The existence of overdispersion cases may result in misleading conclusion, since the standard error from the estimated regression parameters generated becomes lower than it should be, the estimated value of the parameters that should not necessarily be significant will be considered significant. This will lead to incorrect prediction and interpretation of the model. Overdispersed conditions could be apparent or real (Hilbe, 2011). Apparent overdispersion can be caused by omitting important explanatory variables or by the existence of outliers. By including appropriate explanatory variables or adjusting outliers, apparent dispersion usually can be overcome. Sometimes adding interaction factors or transforming response or predictor variables is needed. When the real overdispersion persists, several methods can be used to resolve the problem such as the use of the generalized Poisson or negative binomial model.

The negative binomial regression is an alternative solution used to overcome the problem when overdispersion occurs. In this regression analysis, the response variable is assumed to be negative binomial distributed, a distribution that can be approached by the Poisson-Gamma mixture distribution, that is a variable Y having probability density function as follows:

$$f(y_i; \mu_i, \alpha) = \frac{\Gamma\left(y_i + \frac{1}{a}\right)}{y_i! \, \Gamma\left(\frac{1}{a}\right)} \left(\frac{1}{a\mu + 1}\right)^{\frac{1}{a}} \left(\frac{a\mu_i}{a\mu_i + 1}\right)^{y_i},$$

$$= \begin{pmatrix} y_i + \frac{1}{\alpha} - 1 \\ \frac{1}{\alpha} - 1 \end{pmatrix} \left(\frac{1}{\alpha\mu_i + 1}\right)^{\frac{1}{\alpha}} \left(\frac{\alpha\mu_i}{\alpha\mu_i + 1}\right)^{y_i} \quad y_i = 0,1,2,\dots$$

where

$\mu_i = E(y_i)$,
$Var(y_i) = \mu + a\mu^2$,
a = dispersion parameter in the negative binomial regression model.

If a $\rightarrow$ 0 then Var (y_i) $\rightarrow$ μ, so that the negative binomial distribution converges to the Poisson distribution.
While the equation model in negative binomial regression is as follows:

$$\mu_i = exp(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki}) + \varepsilon_i, \quad i = 0,1,2,\dots n$$

where,

μ_i: the expected value of the response variable (Y),
$\beta_0, \beta_1, \dots, \beta_k$: the coefficients of the regression parameters, and
ε_i: the error of the ith observation.

The log-likelihood function for negative binomial regression is:

$$l(y_i; \mu_i, \alpha) = \sum_{i=1}^{n} y_i \, ln\left(\frac{\alpha\mu_i}{1 + \alpha\mu_i}\right) - \frac{1}{a} ln(a\mu_i + 1) + ln\Gamma(y_i + \frac{1}{\alpha}) - ln\Gamma(y_i + 1) - ln\Gamma\left(\frac{1}{\alpha}\right).$$

Let,

$$\beta = (\beta_0, \beta_1, \dots, \beta_k) \text{ and } \mu_i = exp(x_i'\beta),$$

The above log likelihood function become:

$$l(y_i; \beta, \alpha) = \sum_{i=1}^{n} y_i \, ln\left(\frac{\alpha exp(x_i'\beta)}{1 + \alpha exp(x_i'\beta)}\right) - \frac{1}{a} ln\left(\alpha exp(x_i'\beta) + 1\right) + ln\Gamma(y_i + \frac{1}{\alpha}) - ln\Gamma(y_i + 1)$$
$$- ln\Gamma\left(\frac{1}{\alpha}\right).$$

The Maximum Likelihood Estimation of the parameter coefficients $\beta_0, \beta_1, \dots, \beta_k$ in the negative binomial regression model is obtained by finding partial derivatives of the log-likelihood function with respect to each parameter to be estimated equated to zero [4], and then solved iteratively by using the Newton-Raphson method.

2. Methods

The data used in this study is secondary data obtained from the book of Indonesia Health Profile Year 2015 published by the Ministry of Health of the Republic of Indonesia. The variables used consist of Malnutrition (Y), Vitamin A consumption (X_1), Exclusive breastfeeding (X_2), Immunization (X_3), Water quality (X_4), Healthcare center (X_5), and Poverty level (X_6). While the steps of analysis in this study were as follows:

a. Descriptive analysis for response variables (Y) and predictor variables (X).
b. Assumption test for the distribution of the response variable (Y). The distribution of the response variable (Y) is assumed to be Poisson distributed and it will be tested by Kolmogorov-Smirnov test, i.e. using the statistic:

$$D = max \mid F_n(x_i) - F_0(x|i) \mid.$$

If the value of $D > D^*(\delta)$ (value obtained from the Kolmogorov-Smirnov table) or if the value of Asymp. Sig. (2-tailed) < α then the data is said not to be Poisson distributed.

c. Examination of mean and variance of the response variable (Y).

d. Overdispersion checking using quantity:

$$\phi = \frac{D}{df} = \frac{2\sum_{i=1}^{n}\left\{y_i \ln\left(\frac{y_i}{\mu_i}\right) - (y_i - \mu_i)\right\}}{n - k - 1}$$

if $\phi > 1$ then the Poisson regression model is overdispersed.

e. Establishment of negative binomial regression model.

f. Model Conformity Test

Test of conformity of the negative binomial regression model using the Deviance test:

$$D = 2\sum_{i=1}^{n}\left\{y_i \ln\left(\frac{y_i}{\mu_i}\right) - \left(y_i + \frac{1}{a}\right)\ln\left(\frac{ay_i + 1}{a\mu_i + 1}\right)\right\}$$

If the value of D < $\mathcal{X}^2{}_{(\alpha,\,df)}$, then the model is said to be feasible to use. Note:
α: significant level, *df: n-k*, *n:* the number of observations,
k: the number of parameters

g. Likelihood Ratio Test.

The simultaneous significance test for negative binomial regression parameters is:

$$LR = -2(\hat{l}_0 - \hat{l}_1)$$

If the value of LR> $\mathcal{X}^2{}_{(\alpha,\,df)}$ then it can be said that there are significant parameters to the model.

h. Wald Test

If the likelihood ratio test is rejected, then partial test for each regression parameters is done by Wald test. The Wald statistics is:

$$W_j = \left[\frac{\hat{\beta}_j}{se\left(\hat{\beta}_j\right)}\right]^2$$

If $W_j > \mathcal{X}^2_{\alpha,1}$ or if p-value < α then it can be said that the jth predictor variable is significant.

i. The predictor variables removal process

Predictor variables (X) that are considered insignificant to the model are removed one by one starting from the least significant variable to the model.

j. Selection of the Best Model

The best model selection is done by comparing the smallest AIC values of some models formed from the combination of predictor variables that are considered significant against the negative binomial regression model obtained from process No.9. (Zwilling, 2013).

RESULTS AND DISCUSSION

The average number of under five children affected by malnutrition cases (Y) for every 10,000 under five-year children in Indonesia Provinces in 2015 are 14.323≈14 children. Average percentage of Vitamin A coverage for under five year children (X_1) equals to 76.74%, average percentage of under five year children getting exclusive breastfeeding (X_2) is 57.89%, average percentage of under five year having complete immunization (X_3) is 82.41%, average of households having access to good water quality (X_4) is 68.88%, average number of health centers for every 30.000 population (X_5) is 1.863 ≈ 2 health centers, and average percentage of poverty level (X_6) is 11.70%. The following table presents the descriptive analysis used to see the characteristics of each variable used in this study.

Table 1. Descriptive of research data

Variable	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Malnutrition (Y)	34	1.00	84.00	14.3235	18.02308	324.832
Vitamin A consumption (X_1)	34	0.00	98.80	76.7441	26.46795	700.552
Exclusive breastfeeding (X_2)	34	0.00	86.90	57.9824	17.16696	294.705
Immunization (X_3)	34	47.30	99.80	82.4147	13.36917	178.735
Water quality (X_4)	34	41.08	93.40	68.8832	11.10864	123.402
Healthcare center (X_5)	34	0.58	5.20	1.8635	1.07289	1.151
Poverty level (X_6)	34	3.61	28.40	11.7015	6.18546	38.260

The first step taken in the Poisson regression analysis is Kolmogorov-Smirnov test used to see the distribution of malnutrition data in the Provinces of Indonesia. The result is as follows:

Table 2. Kolmogorov-Smirnov test to evaluate data distribution

		Y
N		34
Poisson Parameter[a,b]	Mean	14.3235
	Absolute	0.582
	Positive	0.582
	Negative	-0.194
Kolmogorov-Smirnov Z		3.391
Asymp. Sig. (2-tailed)		0.000

Based on the results of output in Table 2, the value of Asymp. Sig. (2-tailed) < 0.05. It can be concluded that with a confidence level of 95%, the malnutrition data is not Poisson distributed. The average value and variance of malnutrition data are shown in Table 3.

Table 3. Mean and variance of the response variable

Variable	N	mean	Variance
Malnutrition *(Y)*	34	14.3235	324.832

Based on Table 3 the variance of malnutrition data that occurred in the province of Indonesia in 2015 is much greater than its mean value, so it is indicated that data from Poisson regression analysis has overdispersion problem. Examination of overdispersion cases in Poisson regression can be done by looking at deviance value divided by its degree of freedom. If the value is greater than one, then the data is said to be overdispersion.

The results of output in Table 4 shows that $\phi = 8.98 > 1$, so it can be concluded that overdispersion occurs in the Poisson regression model. The existence of overdispersion cases may result in misleading conclusion due to underestimate standard errors and overstate the significance of the regression parameters. Therefore, the Poisson model is not applicable.

Table 4. Overdispersion analysis

Residual Deviance (D)	df	ϕ
242.58	27	8.98

The next step is constructing negative binomial regression as an alternative model used when the data in Poisson regression analysis experienced overdispersion. The result is as follows:

Table 5. Parameter estimation to the negative binomial regression

Variable	Estimate	Std.Error	*P-value*
Malnutrition (Y)	1.0381	1.9969	0.6031
Vitamin A consumption (X_1)	-0.0015	0.0076	0.8432
Exclusive breastfeeding (X_2)	0.0079	0.0096	0.4086
Immunization (X_3)	-0.0273	0.0161	0.0897
Water quality (X_4)	0.0325	0.0152	0.0325
Healthcare center(X_5)	0.2053	0.2162	0.3423
Poverty level (X_6)	0.0571	0.0280	0.0418

Based on the output in Table 5. the general model for the negative binomial regression is:

$$\hat{\mu}_i = exp(1.0381 - 0.0015X_1 + 0.0079X_2 - 0.0273X_3 + 0.0325X_4 + 0.2053X_5 + 0.0571X_6)$$

To see the feasibility of the above negative binomial regression model, it can be checked from output result in Table 6. Based on the result of output in Table 6 obtained value D = 34.669 < χ^2 (0.05; 27) = 40.113. It can be concluded that with a confidence level of 95% negative binomial regression model obtained is feasible to use.

Table 6. Goodness of fit test for the model

Residual Deviance (D)	df
34.699	27

The likelihood ratio test is used to test the significance of the negative binomial regression parameters simultaneously. The result is as follows:

Table 7. Likelihood Ratio

$\hat{l}_0$	$\hat{l}_1$
-251.240	-229.575

Based on the result of output in Table 7 we get the value of LR is:

$$LR = -2(-251.240 - (-229.575)) = 43.33 > \chi^2{}_{0.05;27} = 40.113,$$

therefore with 95% level of confidence, there are parameters significantly affecting the variability of the response variable. The next step to carry out is Wald test to see the parameters considered significant to the model. The result is shown in Table 8. Based on the results of the output in Table 8 obtained value of P-value for Water Quality variable (X_4) of 0.0325 and for the variable Poverty ($X6$) of 0.0418. Both values of these variables show the value of P-value <0.05. Therefore, at the level of confidence of 95% it can be said that there is a significant influence between the variable Water Quality ($X4$) and Poverty ($X6$) on cases of malnutrition that occurred in the Province of Indonesia in the year 2015.

Table 8. Wald test

Variable	Estimate	Std. Error	P-value
Malnutrition (Y)	1.0381	1.9969	0.6031
Vitamin A consumption (X_1)	-0.0015	0.0076	0.8432
Exclusive breastfeeding (X_2)	0.0079	0.0096	0.4086
Immunization (X_3)	-0.0273	0.0161	0.0897
Water quality (X_4)	0.0325	0.0152	0.0325 *
Healthcare center (X_5)	0.2053	0.2162	0.3423
Poverty level (X_6)	0.0571	0.0280	0.0418 *

The next steps are the process of removing predictor variables that are considered insignificant to the model to get the new model. These steps are executed one by one, starting from the predictor variable having the highest p-value. The result of these steps is removing predictor variables (X_1), (X_2) and (X_5), and the following final results are obtained:

Table 9. Removal process of predictor variables

Variable	Estimate	P-value
Malnutrition (Y)	2.5111	0.0720
Immunization (X_3)	-0.0338	0.0040 *
Water quality (X_4)	0.0295	0.0399 *
Poverty level (X_6)	0.0576	0.0246 *

Based on the results of output in Table 9 it can be concluded that there are three significant variables to the model, i.e. (X_3), (X_4) and (X_6) variables. Negative binomial regression model using these three predictor variables is:

$$\hat{\mu}_i = exp(2.5111 - 0.0338X_3 + 0.0295X_4 + 0.0576X_6).$$

From these three variables, it can be formed several new models used to find the best model. Akaike Information Criterion (AIC) is used to get the best fitted model. There are 7 negative binomial models that can be formed. The variables included in each model and their AIC values are shown in Table 10 below.

Table 10. Negative Binomial Regression Model

Model	Predictor	Parameter	AIC	Significant Parameters
1	X_3	β_0, β_3	244,28	β_3
2	X_4	β_0, β_4	256,00	-
3	X_6	β_0, β_6	245.66	β_6
4	X_3 , X_4	$\beta_0, \beta_3, \beta_4$	244,69	β_3
5	X_4 , X_6	$\beta_0, \beta_4, \beta_6$	246,38	β_6
6	X_3 , X_6	$\beta_0, \beta_3, \beta_6$	246,38	β_6
7	X_4 , X_3, X_6	$\beta_0, \beta_3, \beta_4, \beta_6$	241,29*	$\beta_3, \beta_4, \beta_6$

*) least AIC

Based on Table 10, the least AIC (=241.29) is obtained from model-7. Therefore, the best negative binomial regression model is:

$$\hat{\mu}_i = exp(2.5111 - 0.0338X_3 + 0.0295X_4 + 0.0576X_6).$$

CONCLUSION

Model estimators obtained from the results of Negative Binomial regression analysis to overcome the Poisson regression overdispersion problem in malnutrition case in Indonesia Provinces 2015 are as follows:

$$\hat{\mu}_i = exp(2.5111 - 0.0338X_3 + 0.0295X_4 + 0.0576X_6).$$

While the factors that significantly affect to malnutrition cases in Indonesia provinces in 2015 are the coverage of immunization in under five-year children (X_3), healthy drinking water (X_4), and the level of poverty (X_6). It means that by assuming other variables being fixed, for each percent increase in the immunization coverage will decreases the malnutrition level by 0.0332, for each percent increase access in healthy drinking water level will increase the malnutrition level by 0.0299, for each percent increase in the poverty level will increase the malnutrition level by 0.0593.

The result that malnutrition related to the coverage of immunization in infants confirms Abedi's work (Abedi, 2012) that fully immunized children had better nutrition status. The fact that unhealthy drinking water can be responsible for diarrhoeas and other waterborne diseases especially for under five-year children that lead to malnutrition problem is aligned with van der Hoek (Hoek, 2002) result. Furthermore, the result that poverty affected children's malnutrition has confirmed the study held by (Arif, 2012), (Hilbe, 2011), and (Vorster, 2010).

REFERENCES

Abedi, A., & Srivastasa, J. P. (2012). The Effect of Vaccination on Nutrition Status of Pre-School Children in Rural and Urban Lucknow. J. Acad. Indus. Res., Vol. 1(4).

Arif, G., Farooq, S., Nazir, S., & Satti, M. N. (2012). Child Malnutrition and Poverty: The Case of Pakistan. Islamabad, Pakistan: Pakistan Institute of Development Economics.

Filteau, S. (2000). Role of Breast-Feeding in Managing Malnutrition and Infectious Disease. Proceeding of the Nutrition Society, (p. Vol. 59 Issue 4). Cambridge.

Hilbe, J. (2011). Negative Binomial Regression 2nd edition. Cambridge, UK: Cambridge University Press.

Inguita, R., Islam, K. M., & Hossain, M. M. (2009). Analysis of the Relation between Rural Poverty, Malnutrition and Hunger, and the Interlinks with Agricultural Production Growth and Productivity Growth with Special Reference to Mozamique and Bangladesh. Helsinki: Discussion Papers No. 32 Department of Economics and Management University of Hensinki.

Ismail, N., & Jemain, A. A. (2007). Handling Overdispersion with Negative Binomial and Generalized Poisson Regression Models. Casualty Actuarial Social Forum, Winter 2007.

Kadima, Y. (2012). Factors Influencing Malnutrition among Children under 5 Years of Age in Kweneng, West Distrik of Botswana. Pretoria: Dissertation, University of South Africa.

Ministry of Health the Republic of Indonesia. (2016). Health Profile of Indonesia 2015. Jakarta.

Muchina, E. (2007). Maternal Breastfeeding Practices in Relation to Growth and Nutrition Status of Children Aged 0-24 Months in Nairobi, Kenya. Nairobi: Dissertation, University of Nairobi.

Olsson, U. (2002). Generalized Linear Model - An Applied Approach, Studenlitteratur. Lund, Sweden.

Tette, E. e. (2015). Factors Affecting Malnutrition in Children and the Uptake of Interventions to Prevent the Condition. BMC Pediatric.

Van der Hoek, W., Feenstra, S. G., & Konradsen, F. (2002). Availability of Irrigation Water for Demestic Use in Pakistan: Its Impact on Prevalence of Diarrhoea and Nutrition Status of Children. J. Health Popul. Nutr. 20, 77-84.

Vorster, H. (2010). The Link between Poverty and Malnurition: A South African Perspective, Heathh SA Gesondheid. Journal of Interdisciplinary Health Sciences, Vol. 15 No. 1.

Zwilling, M. (2013). Negative Binomial Regression. The Mathematica Journal.