Censored Hurdle Negative Binomial Regression (Case Study: Neonatorum Tetanus Case in Indonesia)

1Riza Yuli Rusdiana, 2Ismaini Zain, 3Santi Wulan Purnami

1Master’s Program Student of Statistics Department, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia. 2, 3Lecturer of Statistics Department, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia.

1rizayuli.rusdiana@yahoo.co.id, 2ismaini_z@statistika.its.ac.id, 3santi_wp@statistika.its.ac.id

Abstract. Hurdle negative binomial model regression is a method that can be used for discrete dependent variable, excess zero and under- and overdispersion. It uses two parts approach. The first part estimates zero elements from dependent variable is zero hurdle model and the second part estimates not zero elements (non-negative integer) from dependent variable is called truncated negative binomial models. The discrete dependent variable in such cases is censored for some values. The type of censor that will be studied in this research is right censored. This study aims to obtain the parameter estimator hurdle negative binomial regression for right censored dependent variable. In the assessment of parameter estimation methods used Maximum Likelihood Estimator (MLE). Hurdle negative binomial model regression for right censored dependent variable is applied on the number of neonatorum tetanus cases in Indonesia. The type data is count data which contains zero values in some observations and other variety value. This study also aims to obtain the parameter estimator and test statistic censored hurdle negative binomial model. Based on the regression results, the factors that influence neonatorum tetanus case in Indonesia is the percentage of baby health care coverage and neonatal visits.

Keywords: Hurdle negative binomial, neonatorum tetanus, right censored

1. Introduction
In the neonatal period, neonatorum tetanus is a major cause in the case of infant mortality. It is caused by Clostridium tetani which can lead to death. Stiff muscle pain is tetanus characteristic which is caused by neurotoxin and produced by Clostridium tetani in wound anaerobic. World Health Organization (WHO) in 1988 and UNICEF through the World Summit for Children in 1990 invited the world to eliminate neonatorum tetanus in 2000. However, this target was not achieved because it has not found an effective operational strategy. In 1999, developing countries were invited UNICEF, WHO and UNFPA to achieve the target of maternal and neonatal tetanus elimination in 2005, which later shifted to 2015. Elimination is achieved if the number of neonatorum tetanus case less than 1 case per 1000 live births. In 2008, WHO estimated 59000 death of newborns caused by neonatorum...
tetanus. There are 46 countries that have not been eliminated neonatorum tetanus, one of which is Indonesia[1].

The occurrence of neonatorum tetanus case is caused by non-medical factors and medical factors. Medical factors include standard prenatal care (lack of antenatal care to pregnant women, the lack of knowledge of pregnant women about the importance of immunization against tetanus toxoid), perinatal care (lack of availability of maternity facilities and medical personnel so many births do at home and use tools that are not sterile, including the handling of the umbilical cord) and neonatal care (neonatal born in a state of non-sterile, high prematurity, etc.), while for non-medical factors related to local customs [2].

Hurdle negative binomial model is two-component models with a truncated negative binomial component for positive counts and a hurdle component that models the zero counts. Among generalized liner models for handling underdispersion or overdispersion, the hurdle negative binomial regression model is one of the few that can accommodate both underdispersion and overdispersion [3]. In many applications, count data are often censored a specific point from above (right) or below (left) or a combination of them (interval). Right censored is those where observations are clustered at an upper threshold. The previous research [4] suggested the use of censored hurdle negative binomial model. They applied the hurdle negative binomial model to a data set on fish from UCLA Academic Technology Services website and compared it to the negative binomial regression models.

The number of neonatorum tetanus case, measured by the number of neonatorum tetanus case per province is an example of an event count. The basic model to establish is the Poisson regression model. However, this model rarely fits data because of overdispersion. Count data such as neonatorum tetanus case often show high incidence of zero counts. In such cases, hurdle negative binomial model may be more appropriate. Neonatorum tetanus case data contains zero value on most observations and some other value has non-negative integer value. This kind of data is called censored data. Type of censor that will be examined in this study is right censored. We argue that a two-part model is suited to describe the processes of (i) whether neonatorum tetanus case occurred; and (ii) the number of cases if neonatorum tetanus case is occurred. This paper, therefore used hurdle negative binomial model with right censored to establish the determinants of neonatorum tetanus case among province in Indonesia.

Our analysis also explored parameter estimation on hurdle negative binomial with dependent variable right censored using maximum likelihood estimation.

In this article, the main objective is to explain how we can use hurdle negative binomial model in right censored data. In section 2, the hurdle negative binomial regression model is defined and the likelihood function of hurdle negative binomial model in right censored data is formulated. In section 3, the parameter estimation is discussed using maximum likelihood method. In section 4, we applied hurdle negative binomial model in right censored data to an neonatorum tetanus data set.

2. Censored Hurdle Negative Binomial Model

In particular, a hurdle negative binomial model is mixed by a binary outcome of the count being below or above the hurdle, with a truncated negative binomial model for outcomes above the hurdle. That is why hurdle models sometimes are also called two-part models. The most important usage of a hurdle count data model is the hurdle at zero. The hurdle at zero formulation can account for excess zeros. It means that this model can be used in situations where there are many zeros at the dependent variable. In this case, the hurdle at zero defines a probability (P(Y = 0)) that is the first part of the two part models [4]. The truncated negative binomial model account for positive integer at the dependent variable. It is of the two part models defines a probability (P(Y = y_i), y_i = 1, 2, …)

Suppose a count dependent variable is a random variable and affected by p independent variables (x_1, x_2, ..., x_p). The hurdle negative binomial regression model derived by [3] is that the distribution of Y_1, conditional on independent variables (x_1, x_2, ..., x_p), and it is defined by
\begin{equation}
P(Y_i = y_i) = \begin{cases} \frac{e^{x_i^T \hat{\beta}}}{1 + e^{x_i^T \hat{\beta}}} & y_i = 0 \\ \frac{1}{1 + e^{x_i^T \delta}} g \left( \frac{y_i + 1}{\kappa} \right) \left( 1 + e^{x_i^T \kappa} \right)^{-\frac{1}{\kappa} - y_i} \left( e^{x_i^T \delta} \right)^{y_i} & y_i > 0 \end{cases}
\end{equation}

and

\begin{equation}
g = g(y_i; \kappa, \beta) = \frac{\Gamma \left( y_i + \frac{1}{\kappa} \right)}{\Gamma(y_i + 1) \Gamma \left( \frac{1}{\kappa} \right)} \left( 1 + e^{x_i^T \kappa} \right)^{-\frac{1}{\kappa} - y_i} \left( e^{x_i^T \delta} \right)^{y_i}
\end{equation}

where \( x_i = [1 \ x_{i1} \ x_{i2} \ldots x_{ip}]^T \) is \((p+1) \times 1\) dimensional vector, \( \beta = [\beta_0 \ \beta_1 \ \beta_2 \ldots \ \beta_p]^T \) and \( \delta = [\delta_0 \ \delta_1 \ \delta_2 \ldots \ \delta_p]^T \) is a \((p+1) \times 1\) dimensional vector of regression parameters. The parameter \( \kappa \) is a measure of dispersion. When \( \kappa = 0 \), the HNB model reduces to the Poisson regression model. For \( \kappa > 0 \), the HNB model can used to fit overdispersed count data. When \( \kappa < 0 \), the HNB model can used to fit underdispersed count data.

Zero hurdle model with logit link function:

\begin{equation}
\logit(y_i) = \hat{\delta}_0 + \sum_{j=1}^{p} x_{ij} \hat{\delta}_j \quad i = 1, 2, \ldots n \quad \text{and} \quad j = 1, 2, \ldots p
\end{equation}

Truncated negative binomial model with log link function:

\begin{equation}
\log(\mu_i) = \hat{\beta}_0 + \sum_{j=1}^{p} x_{ij} \hat{\beta}_j \quad i = 1, 2, \ldots n \quad \text{and} \quad j = 1, 2, \ldots p
\end{equation}

For some observations in data set, the value of \( Y_i \) may be censored [4]. Censored data is data that contains zero values in some observations whereas for most other observations has particular value that varies. No censoring occurs for the \( i^{th} \) observation, \( Y_i = y_i \) and censoring occurs for the \( i^{th} \) observation, \( Y_i \) is at least equal to \( y_i \) (\( Y_i \geq y_i \)). When the data are censored, the distribution that applies to the sample data is derived by using binary variable \( d_i \), this variable is defined as:

\begin{equation}
d_i = \begin{cases} 1 & \text{if } Y_i \geq y_i \text{ (censored)} \\ 0 & \text{otherwise (uncensored)} \end{cases}
\end{equation}

The probability function of observed counts \( y \) for individual \( i \) is equal to

\begin{equation}
P(Y = y_i, d_i, x_i) = P(Y_i = y_i)^{1-d_i} \ P(Y_i \geq y_i)^{d_i}
\end{equation}

where

\begin{equation}
P(Y_i \geq y_i) = \sum_{\ell = y_i}^{\infty} P(Y_i = \ell) = 1 - \sum_{(\ell = 0)}^{y_i-1} P(Y_i = \ell)
\end{equation}

then HNB model in right censored data introduced by [4] is given by

\begin{align}
P(Y = y_i, d_i, x_i) &= P(Y_i = y_i)^{1-d_i} \ P(Y_i \geq y_i)^{d_i} \\
&= \left[ f_{y_i=0}(y_i; \kappa, \delta, \beta) I_{y_i>0} f(y_i; \kappa, \delta, \beta) \right]^{1-d_i} \sum_{\ell=y_i}^{\infty} f(\ell; \kappa, \delta, \beta)^{d_i}
\end{align}

The first term on the right gives the likelihood contribution of the noncensored observations, while the second term on the right gives the contribution of the censored observations.
3. Parameter Estimation

CHNB can be estimated by using maximum likelihood method. The maximum likelihood method is used to estimate the parameter $\kappa$, $\delta$ and $\beta$. The likelihood function of CHNB model is given by

$$L(\kappa, \delta, \beta | y_i) = \prod_{i=1}^{n} \left[ I_{y_i=0} f\left(0; \kappa, \delta, \beta\right) I_{y_i>0} f\left(y_i; \kappa, \delta, \beta\right) \right]^{1-d_i}$$

$$\sum_{l=y_i}^{n} f(l; \kappa, \delta, \beta)$$

The log-likelihood function of

$$l(\kappa, \delta, \beta | y_i) = \sum_{i=1}^{n} \left[ (1-d_i) \left[ I_{y_i=0} \ln \left( \pi_i \right) + I_{y_i>0} \left( \ln \left( 1-\pi_i \right) + \ln \left( g \right) \right) \right] - \ln \left( 1-\left( 1+e^{x_i^T \beta} \right)^{\pi_i} \right) \right] + \left( d_i \right) \ln \left( \sum_{l=y_i}^{n} f(l; \kappa, \delta, \beta) \right)$$

Estimation of $\kappa$, $\delta$ and $\beta$ are obtained by taking the partial derivations of the log-likelihood function and setting them equal to zero. Thus, we obtain

$$\frac{\partial l(\kappa, \delta, \beta | y_i)}{\partial \kappa} = \sum_{i=1}^{n} \left( 1-d_i \right) I_{y_i>0} \left[ g'_{\kappa} - \frac{\kappa^{-1} e^{x_i^T \beta} \left( 1+e^{x_i^T \beta} \kappa \right)^{-1} - \kappa^{-2} \ln \left( 1+e^{x_i^T \beta} \kappa \right)}{1 - \left( 1+e^{x_i^T \beta} \kappa \right)^{-1}} \right] = 0$$

$$\frac{\partial l(\kappa, \delta, \beta | y_i)}{\partial \delta} = \sum_{i=1}^{n} \left( 1-d_i \right) \left[ I_{y_i>0} \left( \frac{\pi_i}{\pi_i} - I_{y_i=0} \frac{\pi_i}{1-\pi_i} \right) + \frac{d_i}{F} F'_{\delta} \right] = 0$$

$$\frac{\partial l(\kappa, \delta, \beta | y_i)}{\partial \beta} = \sum_{i=1}^{n} \left( 1-d_i \right) \left[ I_{y_i>0} \left( \frac{g'_{\beta}}{g} - \frac{x_i e^{x_i^T \beta} \left( 1+e^{x_i^T \beta} \kappa \right)^{-1} \kappa^{-1} \ln \left( 1+e^{x_i^T \beta} \kappa \right)}{1 - \left( 1+e^{x_i^T \beta} \kappa \right)^{-1}} \right] + \frac{d_i}{F} F'_{\beta} \right] = 0$$

where

$$F = \sum_{i=1}^{n} f(l; \kappa, \delta, \beta) = \sum_{i=1}^{n} \frac{(1-\pi_i) g}{1 - \left( 1+e^{x_i^T \beta} \kappa \right)^{-1}}$$

$g'_{\kappa}$ is first derivative $g$ against $\kappa$ and $g'_{\beta}$ is first derivative $g$ against $\beta$. $F'_{\kappa}$, $F'_{\delta}$ and $F'_{\beta}$ are first derivative $F$ against parameters $\kappa$, $\delta$ and $\beta$. The likelihood equations from (12) until (14) are non-linear in parameters $\kappa$, $\delta$ and $\beta$. These equations are solved simultaneously by using an iterative algorithm. It can be used to carry out Newton-Raphson method for solving these equations. On taking the second partial derivatives, we obtained Hessian matrix which is a square matrix and given by
The second partial derivative against parameter $\kappa$, $\delta$ and $\beta$ as follows:

$$
\frac{\partial^2 l(\kappa, \delta, \beta) | y_i)}{\partial \kappa^2} = \sum_{i=1}^{m} \left(1 - d_i \right) I_{y_i > 0} \left\{ \frac{g''_\kappa \left( g - g''_\kappa \right)}{g^2} - \frac{(A'_\kappa - B'_\kappa) C - C'_\kappa (A - B)}{C^2} D \right\} 
+ D'_\kappa \left( \frac{A - B}{C} \right) 
+ d_i \left\{ \frac{F'\beta F' \delta - F' \beta F' \delta}{F^2} \right\}
$$

(15)

$$
\frac{\partial^2 l(\kappa, \delta, \beta) | y_i)}{\partial \delta \partial \kappa} = \sum_{i=1}^{m} \left(1 - d_i \right) I_{y_i > 0} \left\{ \frac{g''_\kappa \left( g - g''_\kappa \right)}{g^2} - \frac{(A'_\kappa - B'_\kappa) C - C'_\kappa (A - B)}{C^2} D \right\} 
+ D'_\kappa \left( \frac{A - B}{C} \right) 
+ d_i \left\{ \frac{F'\beta F' \delta - F' \beta F' \delta}{F^2} \right\}
$$

(16)

$$
\frac{\partial^2 l(\kappa, \delta, \beta) | y_i)}{\partial \beta \partial \kappa} = \sum_{i=1}^{m} \left(1 - d_i \right) I_{y_i > 0} \left\{ \frac{g''_\beta \left( g - g''_\beta \right)}{g^2} - \frac{(A'_\beta - B'_\beta) C - C'_\beta (A - B)}{C^2} D \right\} 
+ D'_\kappa \left( \frac{A - B}{C} \right) 
+ d_i \left\{ \frac{F'\beta F' \delta - F' \beta F' \delta}{F^2} \right\}
$$

(17)

$$
\frac{\partial^2 l(\kappa, \delta, \beta) | y_i)}{\partial \beta \partial \delta} = \sum_{i=1}^{m} \left(1 - d_i \right) I_{y_i > 0} \left\{ \frac{g''_\delta \left( g - g''_\delta \right)}{g^2} - \frac{(A'_\delta - B'_\delta) C - C'_\delta (A - B)}{C^2} D \right\} 
+ D'_\kappa \left( \frac{A - B}{C} \right) 
+ d_i \left\{ \frac{F'\beta F' \delta - F' \beta F' \delta}{F^2} \right\}
$$

(18)

$$
\frac{\partial^2 l(\kappa, \delta, \beta) | y_i)}{\partial \beta \partial \delta} = \sum_{i=1}^{m} \left(1 - d_i \right) I_{y_i > 0} \left\{ \frac{g''_\delta \left( g - g''_\delta \right)}{g^2} - \frac{(A'_\delta - B'_\delta) C - C'_\delta (A - B)}{C^2} D \right\} 
+ D'_\kappa \left( \frac{A - B}{C} \right) 
+ d_i \left\{ \frac{F'\beta F' \delta - F' \beta F' \delta}{F^2} \right\}
$$

(19)

The expressions for other function provided in the appendix.

4. Application
This study was based on Indonesian health profile 2015 [5]. Data and information on Indonesian health profile comes from technical units in the Ministry of Health and other institutions which related with health data such as BPS and BKKBN. Data included in this analysis comprised of 33 province in
Indonesia. The independent variables are baby health care coverage, TT immunization of pregnant women, neonatal visits and TT immunization of women of fertile age. Baby health care coverage are handled by trained health care personnel throughout the health care facility. Newborn services are care counseling include ASI and umbilical cord care. TT immunization of pregnant women is given to pregnant women with at least two doses of tetanus toxoid called TT2+. In addition to the baby health care coverage, indicators that describe the health service for neonatal period is neonatal visit which requires that every newborn visit at least three times as standard in region. TT immunization of women of fertile age is given to women of fertile age with at least one doses of tetanus toxoid.

The dependent variable, the number of neonatorum tetanus case, is non-negative integer ranging from zero to twenty one in the sample. The number of neonatorum tetanus case in Indonesia is 53 cases with case fatality as many as 27 people. The distribution of the number of neonatorum tetanus case is 0.4 percent or 13 province from 33 province in Indonesia which exposed neonatorum tetanus case.

![Figure 1. Histogram of the number of neonatorum tetanus case](image)

Figure 1. Histogram of the number of neonatorum tetanus case

Figure 1 shows frequency distribution of the number of neonatorum tetanus case. The graph indicates a high frequency of zeros and long right tail giving evidence of overdispersion in the data.

| Variable                              | Mean | Variance |
|---------------------------------------|------|----------|
| Number of neonatorum tetanus case     | 1.61 | 17.25    |
| Baby health care coverage             | 74.96| 614.32   |
| TT immunization of pregnant woman    | 57.57| 481.04   |
| Neonatal visits                       | 68.95| 612.69   |
| TT immunization of women of fertile age| 2.06 | 13.51    |

Table 1. Descriptives statistics of variables
The sample variance of dependent variable is greater than the sample mean, therefore we expected that overdispersion occurs and dispersion parameter \( \kappa \) may be positive. In this study used right censored with censored point of 1, because we expected less than one case in each province or neonatorum tetanus case was eliminated in Indonesia. About 39.4% of the sample has dependent variable \( y_i \geq 1 \), and then we consider as censored. The dispersion parameter \( \kappa \) is positive 3.145 and 0.188 which also indicate that modeling overdispersed data. The AIC values for hurdle negative binomial and censored hurdle negative binomial models are 109.5 and 50.7, respectively, which also indicate that using censored hurdle negative binomial model is more appropriate than hurdle negative binomial model.

Table 2. Parameter Estimation of Regression

| Independent Variables | Censored | Hurdle Negative Binomial | Not Censored |
|-----------------------|----------|--------------------------|--------------|
|                       | Estimate | OR | P-value | Estimate | OR | P-value |
| \( \delta_0 \)        | 0.108    | 1.114 | 0.941 | 0.009 | 1.009 | 0.994 |
| \( \delta_1 \)        | 0.021    | 1.021 | 0.793 | 0.033 | 1.034 | 0.588 |
| \( \delta_2 \)        | -0.015   | 0.985 | 0.564 | -0.013 | 0.987 | 0.517 |
| \( \delta_3 \)        | 0.006    | 1.006 | 0.939 | -0.013 | 0.987 | 0.827 |
| \( \delta_4 \)        | 0.266    | 1.305 | 0.624 | -0.176 | 0.839 | 0.374 |
| \( \beta_0 \)         | -0.278   | 0.757 | 0.953 | -3.647 | 0.026 | 0.396 |
| \( \beta_1 \)         | -0.569   | 0.566* | 0.044 | -0.033 | 0.968 | 0.844 |
| \( \beta_2 \)         | -0.079   | 0.924 | 0.549 | 0.009 | 1.009 | 0.829 |
| \( \beta_3 \)         | 0.432    | 1.540** | <0.0001 | 0.076 | 1.079 | 0.677 |
| \( \beta_4 \)         | -0.159   | 0.853 | 0.834 | 0.074 | 1.077 | 0.514 |
| \( \kappa \)          | 0.188    | 1.207 | 0.374 | 3.145 | 23.219 | 0.699 |
| -2LL                   | 28.7     | 87.5 |
| AIC                    | 50.7     | 109.5 |

*significant at 5% level, **significant at 1% level

Table 2 summarizes the results censored hurdle negative binomial regression. The number of neonatorum tetanus case was found to be associated with baby health care coverage (\( X_1 \)) and neonatal visits (\( X_3 \)). The number of neonatorum tetanus case decreased with increasing baby health care coverage (\( X_1 \)) and neonatal visits (\( X_3 \)). There was no significant difference in neonatorum tetanus case with TT immunization of pregnant woman (\( X_2 \)) and TT immunization of women of fertile age (\( X_4 \)).

Zero hurdle model:

\[
\pi_i = \frac{\exp(0.108 + 0.021X_1 - 0.015X_2 + 0.006X_3 + 0.266X_4)}{1 + \exp(0.108 + 0.021X_1 - 0.015X_2 + 0.006X_3 + 0.266X_4)}
\]

Truncated negative binomial model:

\[
\mu_i = \exp(-0.278 - 0.5691X_1 - 0.079X_2 + 0.432X_3 + 0.159X_4)
\]

The average of number of neonatorum tetanus cases expected will be decreased by 0.566 for a percent increase baby health care coverage and will be increased by 1.54 for a percent increase neonatal visit when other variables constant.
5. Conclusion
The result of parameter estimation by the Maximum Likelihood Estimation method is shaped so that the implicit Newton Raphson iteration process is required. Based on applied study CHNB regression is better than HNB regression. Results indicate that neonatorum tetanus cases in Indonesia is associated with baby health care coverage and neonatal visit.

Appendix
From (11) until (20), the expression for other function given by

\[ g'_{\pi} = g \left[ \Gamma'(y_i + \kappa^{-1}) - \Gamma'(\kappa^{-1}) + \kappa^{-2} \ln \left( 1 + e^{x_i^T \beta} \kappa \right) - \frac{y_i + \kappa^{-1} e^{x_i^T \beta}}{1 + e^{x_i^T \beta} \kappa} \right] \]

\[ g^*_{\pi} = g \left[ x_i \left( y_i - e^{x_i^T \beta} \right) \right] \]

\[ \pi_i = \frac{x_i e^{x_i^T \beta}}{(1 + e^{x_i^T \beta})^2} \]

\[ F_{\pi} = \sum_{i=1}^{\infty} (1 - \pi_i) \left[ \frac{g'_{\pi}}{\left( 1 + e^{x_i^T \beta} \kappa \right)^{-1}} + g \left( 1 + e^{x_i^T \beta} \kappa \right)^{-1} \ln \left( 1 + e^{x_i^T \beta} \kappa \right) - \kappa^{-1} e^{x_i^T \beta} \left( 1 + e^{x_i^T \beta} \kappa \right)^{-1} \right] \]

\[ F_{\beta} = \sum_{i=1}^{\infty} (1 - \pi_i) \left[ \frac{g'_{\beta}}{\left( 1 + e^{x_i^T \beta} \kappa \right)^{-1}} - \left( x_i e^{x_i^T \beta} \right) \left( 1 + e^{x_i^T \beta} \kappa \right)^{-1} \right] \]

\[ A = \kappa^{-1} e^{x_i^T \beta} \left( 1 + e^{x_i^T \beta} \kappa \right)^{-1} \]

\[ B = \kappa^{-2} \ln \left( 1 + e^{x_i^T \beta} \kappa \right) \]

\[ C = 1 - \left( 1 + e^{x_i^T \beta} \kappa \right)^{-1} \]

\[ D = \left( 1 + e^{x_i^T \beta} \kappa \right)^{-1} \]

\[ E = \left( 1 + e^{x_i^T \beta} \kappa \right)^{-x_i^T \beta} \]

\[ H = x_i e^{x_i^T \beta} \]

\[ P = \left( y_i + \kappa^{-1} \right) e^{x_i^T \beta} \]

\[ Q = 1 + e^{x_i^T \beta} \kappa^{-1} \]

References
[1]. Kemenkes. (2012). Buletin: Eleminasi Tetanus Maternal dan Neonatal. [http://www.depkes.go.id/resources/download/pusdatin/buletin/buletin-mnnte.pdf] diakses tanggal 11 Agustus 2016.

[2]. Handoko, M. (2011). Tetanus Neonatorum. Kepaniteraan Klinik Ilmu Penyakit Anak. [https://www.scribd.com/doc/51651896/Tetanus-neonatorum] diakses tanggal 10 Oktober 2016.

[3]. Pohlmeier, W. and Ulrich, V. (1995). An Econometric Model of The Two-Part Decision-Making Process In The Demand For Health Care. The Journal of Human Resources, 30, 339–361.
[4]. Saffari, S. E., Robiah, A. and Greene, W. (2012). *Hurdle Negative Binomial Regression Model with Right Censored Count Data*. Malaysia: Journal of Statistics and Operations Research Transactions, Vol.36(2):181-194.

[5]. Kemenkes. (2016). *Profil Kesehatan Indonesia 2015*. Jakarta: Kemenkes RI.