CAUCHY –Jurnal Matematika Murni dan Aplikasi
Volume 7(3) (2022), Pages 384-393
p-ISSN: 2086-0382; e-ISSN: 2477-3344

Bayesian Hurdle Poisson Regression for Assumption Violation

Nur Kamilah Sa'diyah*, Ani Budi Astuti, Maria Bernadetha T. Mitakda

Department of Statistics, Faculty of Mathematics and Natural Science, Brawijaya University, Indonesia

Email: nurkamilahs@student.ub.ac.id

ABSTRACT

Violation of the Poisson regression assumption can cause the model formed will produce an unbiased estimator. There is a good method for estimating parameters on small sample sizes and on all distributions, namely the Bayesian method. The number of death due to chronic Filariasis data violates the Poisson regression assumption (overdispersion and response variable did not follow Poisson distribution), so it is modeled with the Bayesian Hurdle Poisson Regression. With the Bayesian method, convergence is fullfilled when 300000 iterations and 7 thin are performed. In addition to presenting an alternative method for estimating the Hurdle Poisson Regression parameter, the model obtained can be used by the Government in efforts to mitigate disease disasters through efforts to prevent, control, and handle cases of Filariasis. The results showed that in the logit model only the percentage of households that have access to proper sanitation in 34 Provinces in Indonesia had a significant effect on the number of death due to chronic Filariasis cases in 34 Provinces in Indonesia (Y). The Truncated Poisson model resulted in all predictor variables having a significant effect on the number of death due to chronic Filariasis cases.

Keywords: bayesian; filariasis; hurdle; overdispersion; poisson

INTRODUCTION

An important assumption in Poisson regression analysis is that the response variable in the form of count distribute Poisson, does not occur multicollinearity in the predictor variable, and occurs equidispersion (the mean of the data is equal to its variance). However, in certain cases, the assumption of conformity of Poisson's distribution and equidispersion is not fullfilled. This can cause the model formed will produce an unbiased estimator [1].

Equidispersion violations or often known as overdispersion (variance greater than the mean) can be overcome with Zero Inflated model and Hurdle model. The handling of overdispersion in this study uses the Hurdle Poisson model because Hurdle model better than the Zero Inflated Model [2]. The parameter estimation method often used in the Poisson Hurdle model is Maximum Likelihood Estimation (MLE). However, MLE cannot estimate parameters on small sample sizes and on certain distributions. There is a good method for estimating parameters on small sample sizes and on all distributions, namely the Bayesian method. The advantage of the Bayesian method is that it can estimate parameters for extremely small observations and can be used for all distributions [3].

The application of the Bayesian method to overdispersion data has been carried out to analyze the number of Filariasis sufferers in Papua Province, using the Bayesian Zero

Submitted: March 05, 2022 Reviewed: April 23, 2022 Accepted: April 26, 2022
DOI: http://dx.doi.org/10.18860/ca.v7i3.15549

Inflated Poisson model [4]. In this study will model data on the number of death from chronic Filariasis cases in Indonesia that violate the assumption of equidispersion and suitability of Poisson distribution with Bayesian Hurdle Poisson regression.

Filariasis or also known as elephant foot disease is believed to have existed since B.C. because in 1501-1480 BC found an ancient relief in a cemetery temple. Queen Hatshepsut in Thebet, Egypt who depicts the princess Punt suffering from Filariasis on her legs [5]. Filariasis in Indonesia is one of the endemic diseases (a disease that continues to infect certain regions) and was first reported by Haga and Van Eecke in 1889 in Jakarta caused by Brugaria Malayi [6]. Acute clinical symptoms of Filariasis disease include inflammation and swelling of the lymph canal accompanied by fever, headache, weak feeling and the onset of abscesses/ulcers while symptoms Chronic clinical is the occurrence of enlargement that persists in the legs, arms, breasts and genitals of women and men [7]. One of the efforts to inhibit the transmission of Filariasis disease is to Mass Preventive Drug Delivery (MPDD) Filariasis implemented by endemic Districts/Cities of Filariasis [5]. The success of the Filariasis control program can be known by looking at the number of districts/cities that managed to reduce the number of microphilia to <1% [8].

This study discusses the influence of the number of chronic cases of Filariasis in 34 Provinces in Indonesia (X_1), the number of Districts/Cities succeeded in reducing mikrophilia <1% in 34 Provinces in Indonesia (X_2), The number of Districts/Cities still carry out Mass Preventive Drug Delivery (MPDD) Filariasis in 34 Provinces in Indonesia (X_3), population density in 34 Provinces in Indonesia (X_4), and the percentage of households that have access to proper sanitation in 34 Provinces in Indonesia (X_5) against the number of deaths from chronic Filariasis in 34 Provinces in Indonesia (Y).

The results of this study can be utilized for many things, namely (1) Through the Bayesian Hurdle Poisson regression model that is built can be identified factors that affect the number of cases of chronic Filariasis death in Indonesia, so that this information can be utilized for appropriate policy making for the central and local governments and related agencies in order to mitigate the disaster of chronic Filariasis disease in Indonesia through prevention efforts, control, and handling of the case. (2) By using Bayesian parameter estimation approach, it is very useful and superior in various data challenge cases, namely for various sample sizes (any sample) small or large and various distributions (any distribution) with a data driven concept.

METHODS

This study uses secondary data from the Indonesian Health Profile in 2020, namely the number of cases of chronic Filariasis in 2020 with five predictor variables and one response variable [9]. The first step that must be done is testing the Poisson regression assumption (Poisson distribution suitability, non-multicollinearity, and overdispersion testing). The variables used in this study are the number of chronic cases of Filariasis in 34 Provinces in Indonesia (X_1), the number of Districts/Cities succeeded in reducing mikrophilia <1% in 34 Provinces in Indonesia (X_2), The number of Districts/Cities still carry out Mass Preventive Drug Delivery (MPDD) Filariasis in 34 Provinces in Indonesia (X_3), population density in 34 Provinces in Indonesia (X_4), and the percentage of households that have access to proper sanitation in 34 Provinces in Indonesia (X_5) against the number of deaths from chronic Filariasis in 34 Provinces in Indonesia (Y).

Poisson Regression Assumption

Poisson distribution suitability was tested with the Kolmogorov-Smirnov. Kolmogorov-Smirnov test statistics for testing the suitability of the Poisson distribution are presented in equation (1)[10].

$$D = maximum\left|F_N\left(y_{(i)}\right) - P\left(y_{(i)}, \lambda\right)\right| \tag{1}$$

If $D > D_{(n,\alpha)}$ or $p_{value} < 0.05$, so we can conclude that response variable does not follow a Poisson distribution. Assumption of non-multicollinierity was tested with the VIF criteria. If the VIF_j exceeds 10, non-multicollinierity assumption is not fulfilled [11]. The third assumption test that must be done is the overdispersion test. The overdispersion test is carried out by calculating Pearson Chi Square divided by the degrees of freedom of residual based on the formula (2).

$$\chi^2_{Pearson} = \sum_{i=1}^{n} \frac{(y_i - \hat{\mu}_i)^2}{\hat{\mu}_i} \tag{2}$$

where:
$\hat{\mu}_i = \hat{\lambda}_i = exp\left(\hat{\beta}_0 + \hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2} + \cdots + \hat{\beta}_k X_{ik}\right)$
$df = n - p$
n : number of observations
p : number of parameters $(k + 1)$
If $(\chi^2_{Pearson}/df) > 1$ then it can be said that observations contain overdispersion [12].

Bayesian Method

Suppose there are parameters θ to be estimated. In Bayesian method, parameters θ treated as variable will have value in the domain $f(\theta)$. The prior distribution is the initial information to form the posterior. With prior information combined with data, calculating the posterior will be easier. Based on the Bayesian method, the posterior distribution is proportional (comparable) to the combination of the prior distribution and the likelihood function based on equation (3) [13].

$$f(\theta|y) \propto f(y|\theta)f(\theta) \tag{3}$$

where:
$f(y|\theta)$: likelihood function
$f(\theta)$: prior distribution function
$f(\theta|y)$: posterior distribution function

Bayesian Hurdle Poisson Regression

There are three important components in Bayesian method, namely (1) the likelihood function of the HPR model, (2) the prior distribution and (3) the posterior distribution. The likelihood function of the HPR model is as presented in equation (4).

$$f(Y|\beta, \delta) = \prod_{\substack{i=1 \\ y_i=0}}^{n} \frac{1}{1+exp(\boldsymbol{X}^T\boldsymbol{\delta})} \times \prod_{\substack{i=1 \\ y_i>0}}^{n} \frac{\left[exp\left(-exp(\boldsymbol{X}^T\boldsymbol{\beta})\right)\right]\left[exp(\boldsymbol{X}^T\boldsymbol{\beta})\right]^{y_i}}{\left(1-\left[exp\left(-exp(\boldsymbol{X}^T\boldsymbol{\beta})\right)\right]\right)y_i!} \tag{4}$$

The prior distribution for β and δ is assumed to be normally distributed with the mean and variance σ^2 with the form as shown in equation (5).

$$f(\beta, \delta) = \prod_{j=0}^{k} \frac{1}{\sigma_\beta\sqrt{2\pi}} exp\left(-\frac{\left(\beta-\mu_\beta\right)^2}{2\sigma_\beta^2}\right) \times \prod_{j=0}^{k} \frac{1}{\sigma_\delta\sqrt{2\pi}} exp\left(-\frac{(\delta-\mu_\delta)^2}{2\sigma_\delta^2}\right) \tag{5}$$

The posterior distribution is obtained from the product of the likelihood function and the prior distribution in the form of an equation as presented in equation (6).

$$f(\beta,\delta|Y) \propto \prod_{\substack{i=1\\ y_i=0}}^{n} \frac{1}{1+exp(\boldsymbol{X}^T\boldsymbol{\delta})} \prod_{\substack{i=1\\ y_i>0}}^{n} \frac{\left[exp\left(-exp(\boldsymbol{X}^T\boldsymbol{\beta})\right)\right]\left[exp(\boldsymbol{X}^T\boldsymbol{\beta})\right]^{y_i}}{\left(1-\left[exp\left(-exp(\boldsymbol{X}^T\boldsymbol{\beta})\right)\right]\right)y_i!} \times$$
$$\prod_{j=0}^{k} \frac{1}{\sigma_\beta\sqrt{2\pi}} exp\left(-\frac{\left(\beta-\mu_\beta\right)^2}{2\sigma_\beta^2}\right) \prod_{j=0}^{k} \frac{1}{\sigma_\delta\sqrt{2\pi}} exp\left(-\frac{(\delta-\mu_\delta)^2}{2\sigma_\delta^2}\right) \quad (6)$$

The posterior distribution of the Bayesian Hurdle Poisson Regression model parameters has a complex function and requires a difficult integration process, so it is not easy to obtain analytically. Therefore, a numerical approach is needed using the Markov Chain Monte Carlo (MCMC) simulation method [14].

Bayesian Model Convergence Test

Convergence test method consists of trace plot, autocorrelation plot, ergodic mean plot, and Monte Carlo Error (MC Error) [15]. Convergence will be fullfilled if the trace plot does not form an ascending or descending pattern, the autocorrelation plot is close to one and the next lag is close to zero, after several iterations the ergodic mean plot is stable, or MC error is less than 5% of the standard deviation of each parameter.

RESULTS AND DISCUSSION

The results of the analysis begin with testing the Poisson regression assumption, then the parameter estimator of the Bayesian Hurdle Poisson regression.

Result of Poisson Regression Assumption Test

The first assumption in Poisson regression is the response variable in the form of count with Poisson distribution based on hypothesis.

H_0: The number of death due to chronic Filariasis cases follows a Poisson distribution *versus*

H_1: The number of death due to chronic Filariasis cases does not follows a Poisson distribution

The results of the Kolmogorov-Smirnov test with *Software* R showed that the p_{value} less than 2.2×10^{-16}. This suggests that the response variable did not follows a Poisson distribution. Then do fit distribution with EasyFit Software. Poisson distribution ranked third after uniform and geometric distribution. Since Poisson regression is the most common regression model for modeling response variable in the form of count, then no one has researched related *to uniform* regression and geometric regression, the study still uses Poisson's regression model, but uses the Bayesian method to estimate the parameters because they have advantages that can be applied to all distribution.

The next assumption is non-multocollinearity. The results of the multicollinearity test with the VIF_j are presented in Table 1.

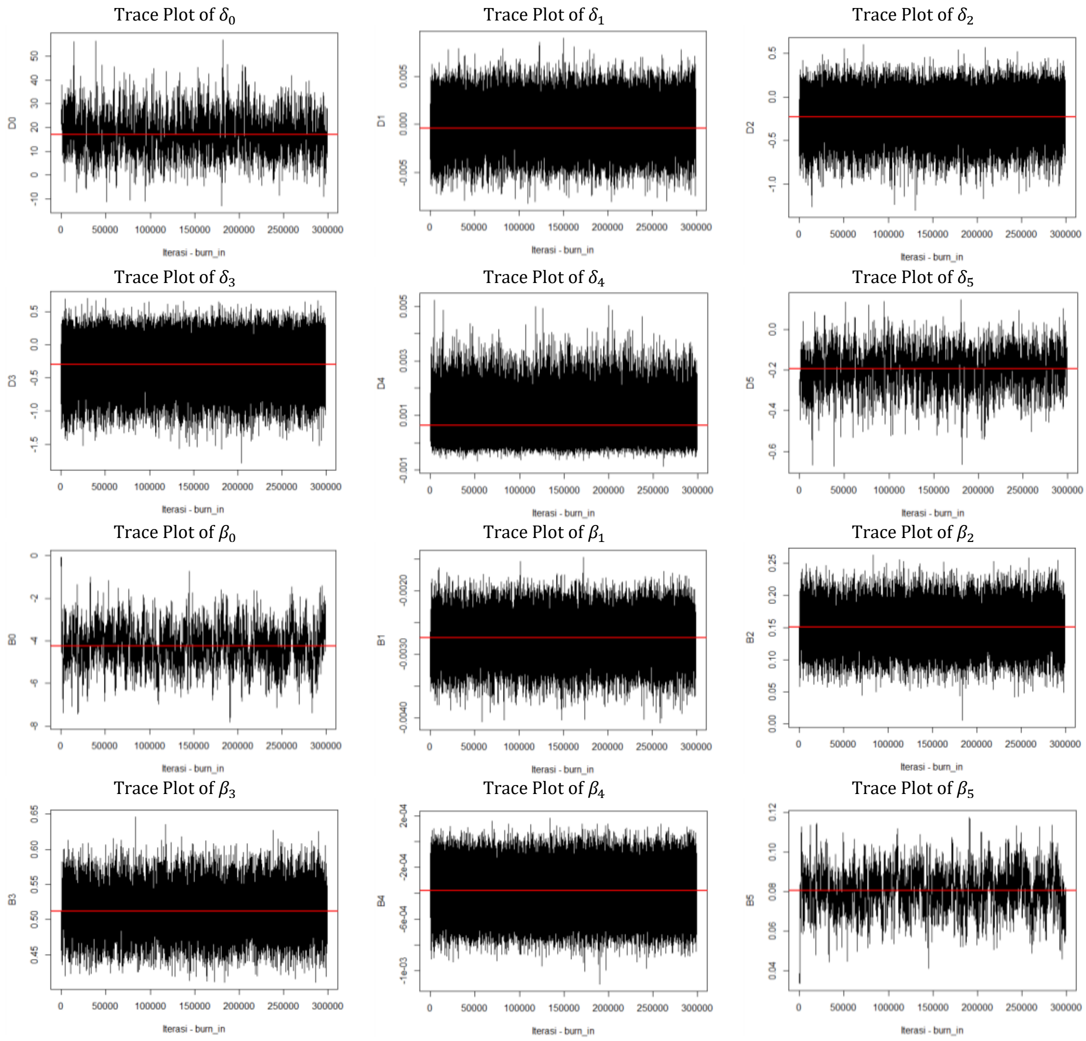
Table 1. VIF for All Predictors

Variable	VIF_j
X_1	4.782
X_2	1.530
X_3	3.872
X_4	1.173
X_5	3.162

Table 1 shows that the VIF of all predictor variables is less than 10, so it can be concluded that the non-multicollinearity assumption is fullfilled. The last assumption in Poisson regression is the occurrence of equidispersion. Overdispersion testing was carried out with $\chi^2_{Pearson}/df$. Data is said to contain overdispersion if $(\chi^2_{Pearson}/df) > 1$. The $\chi^2_{Pearson}/df = 212.549$, it can be concluded that the data contains overdispersion. Because the two Poisson regression assumptions are not fullfilled, then estimate the parameters with the Bayesian Hurdle Poisson regression model.

Result of Bayesian Model Convergence Test

In Bayesian method, parameters are generated using the Gibbs Sampling algorithm with 300000 iterations and 7 thin. It is important to check the convergence of the model parameters to check the accuracy of the parameter estimation using the Bayesian method. There are four methods for checking the convergence of parameters, namely (1) Trace Plot, (2) Autocorrelation Plot, and (3) Ergodic Mean Plot (4) MC error. Trace Plots for each parameter are presented in Figure 1.



Trace Plot of δ_0

Trace Plot of δ_1

Trace Plot of δ_2

Trace Plot of δ_3

Trace Plot of δ_4

Trace Plot of δ_5

Trace Plot of β_0

Trace Plot of β_1

Trace Plot of β_2

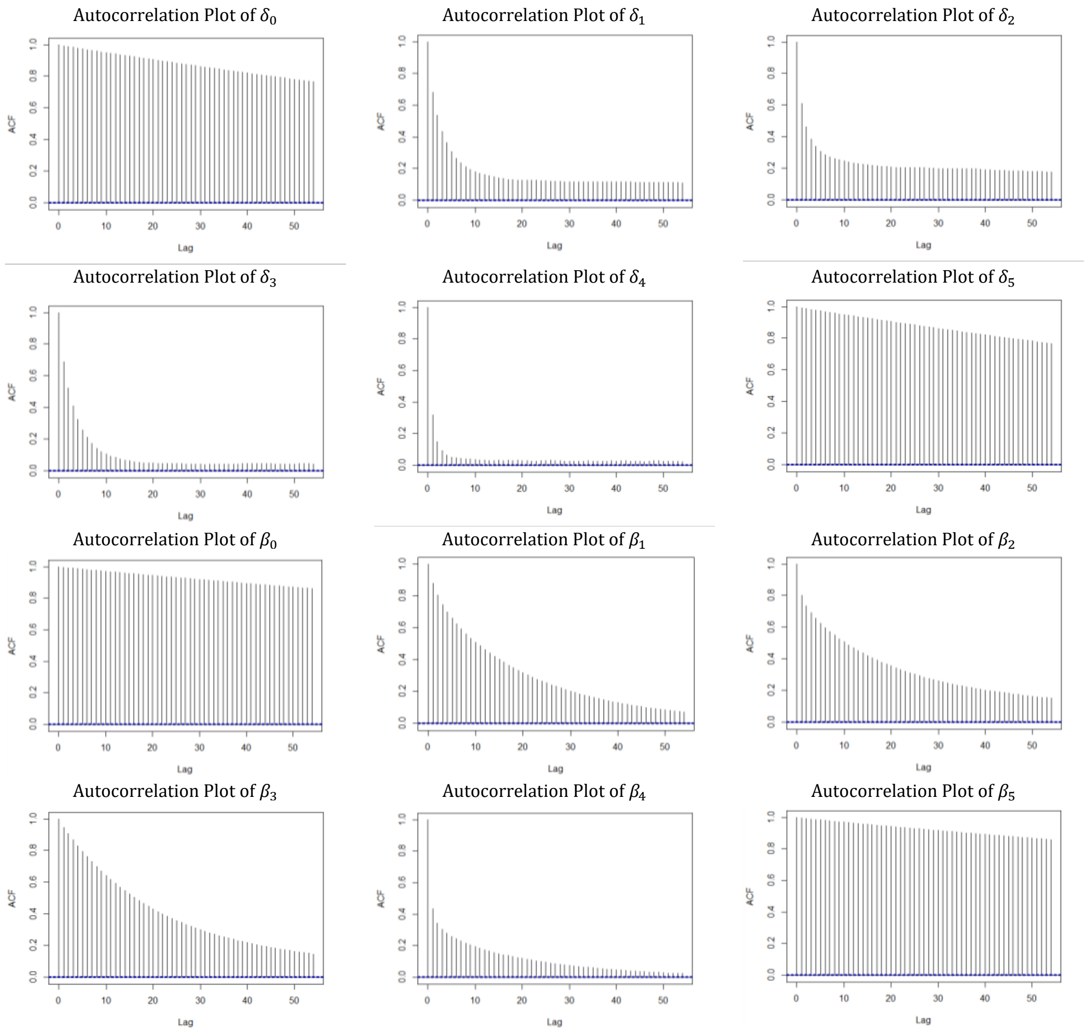
Trace Plot of β_3

Trace Plot of β_4

Trace Plot of β_5

Figure 1. Trace Plot for Bayesian Hurdle Poisson Regression Parameters

The Figure 1 shows that the trace plot is random when 300000 iterations are carried out and 7 thin. It can be concluded that the parameters are convergent, so the iteration is stopped. The second method used to check the convergence is the autocorrelation plot. The Figure 2 shows the autocorrelation plot for each parameter.



Autocorrelation Plot of δ_0 Autocorrelation Plot of δ_1 Autocorrelation Plot of δ_2

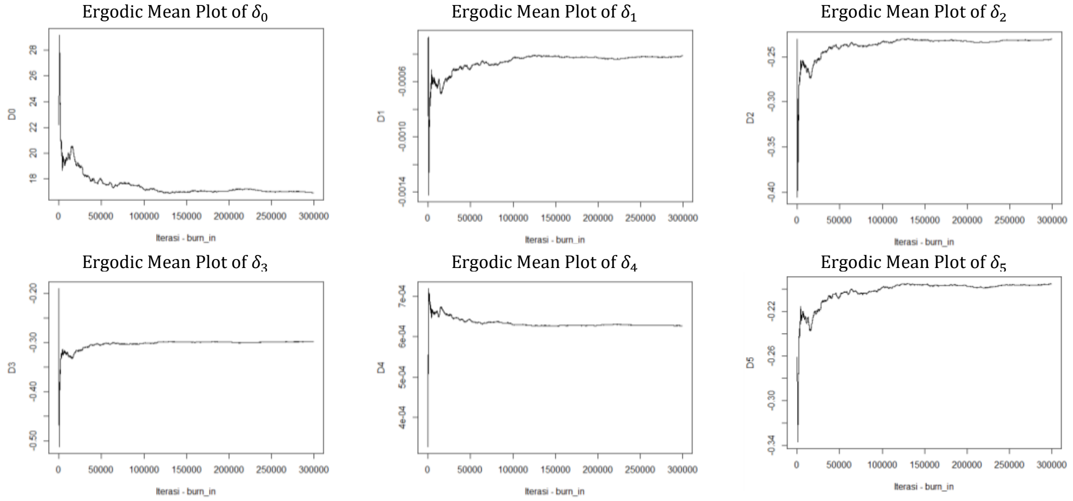
Autocorrelation Plot of δ_3 Autocorrelation Plot of δ_4 Autocorrelation Plot of δ_5

Autocorrelation Plot of β_0 Autocorrelation Plot of β_1 Autocorrelation Plot of β_2

Autocorrelation Plot of β_3 Autocorrelation Plot of β_4 Autocorrelation Plot of β_5

Figure 2. Autocorrelation Plot for Bayesian Hurdle Poisson Regression Parameters

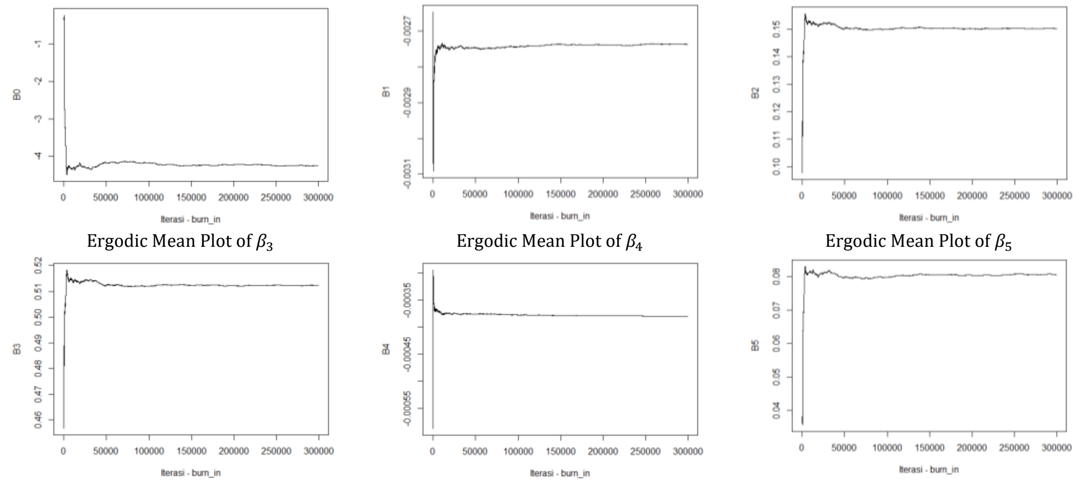
The Figure 2 shows that the first lag in the autocorrelation plot is close to one and the next lag is close to zero, so the convergence of parameters is fulfilled. The third method used to check convergence is the ergodic mean plot. Convergence will be fullfilled if after several iterations the ergodic mean plot is stable. The Figure 3 shows the ergodic mean plot for each parameter.



Ergodic Mean Plot of δ_0 Ergodic Mean Plot of δ_1 Ergodic Mean Plot of δ_2

Ergodic Mean Plot of δ_3 Ergodic Mean Plot of δ_4 Ergodic Mean Plot of δ_5

Ergodic Mean Plot of β_0 Ergodic Mean Plot of β_1 Ergodic Mean Plot of β_2



Ergodic Mean Plot of β_3 Ergodic Mean Plot of β_4 Ergodic Mean Plot of β_5

Figure 3. Ergodic Mean Plot for Bayesian Hurdle Poisson Regression Parameters

The Figure 3 shows that after 300000 iterations and 7 thin the ergodic mean plot is stable. It can be concluded that the parameters are convergent. In addition to using plots, convergence checks can also be done by comparing the MC error with 5% standard deviation for each parameter. The MC error for each parameter of the Bayesian Hurdle Poisson regression model are presented in the Table 2.

Table 2. MC Error for Bayesian Hurdle Poisson Regression Parameters

Model	Parameter Estimator	Standard Deviation	5% Standard Deviation	MC *Error*	Decision
Logit	$\hat{\delta}_0$	7.278704	0.363935	0.255295	Convergence
	$\hat{\delta}_1$	0.001624	8.12×10^{-5}	2.17×10^{-5}	Convergence
	$\hat{\delta}_2$	0.175980	0.008799	0.003009	Convergence
	$\hat{\delta}_3$	0.242849	0.012142	0.002269	Convergence
	$\hat{\delta}_4$	0.000538	2.69×10^{-5}	3.59×10^{-6}	Convergence
	$\hat{\delta}_5$	0.084389	0.004219	0.002958	Convergence
Truncated Poisson	$\hat{\beta}_0$	0.918011	0.045901	0.040426	Convergence
	$\hat{\beta}_1$	0.000272	1.36×10^{-5}	3.33×10^{-6}	Convergence
	$\hat{\beta}_2$	0.025134	0.001257	0.000488	Convergence
	$\hat{\beta}_3$	0.026595	0.00133	0.000515	Convergence
	$\hat{\beta}_4$	0.000131	6.54×10^{-6}	9.12×10^{-7}	Convergence
	$\hat{\beta}_5$	0.010116	0.000506	0.000445	Convergence

Based on Table 2, MC error on all parameters is less than 5% standard deviation, then the convergence is met. Based on the four methods of checking the convergence, the results are the same, namely the convergence is fulfilled when 300000 and 7 thin amere performed.

Parameter Estimation Results of Bayesian Hurdle Poisson Regression Model

After the convergence is fullfilled, we can calculate the parameter estimator obtained from the sample generation using Gibbs Sampling. The parameter estimator is the average of the sample generation results for each parameter which is shown in Table 3. Testing the Bayesian model parameters using a confidence interval by looking at the lower limit of the 2.5% percentile and the upper limit of the 97.5% percentile. If it contains zero in that range, the decision to accept H_0 or the j^{th} predictor variable has no significant effect to the response variable.

Table 3. Parameter Estimator of Bayesian Hurdle Poisson Regression

Model	Parameter	Parameter Estimator	*Percentile* 2.5%	*Percentile* 97.5%	Decision
Logit	δ_0	16.6551	5.0846	28.4970	Reject H_0
	δ_1	0.0004	-0.0031	0.0022	Accept H_0
	δ_2	-0.2270	-0.5155	0.0572	Accept H_0
	δ_3	-0.2974	-0.6936	0.0984	Accept H_0
	δ_4	0.0006	-0.0001	0.0014	Accept H_0
	δ_5	-0.1922	-0.3257	-0.0549	Reject H_0
Truncated Poisson	β_0	-4.2404	-5.7179	-2.7044	Reject H_0
	β_1	-0.0027	-0.0032	-0.0023	Reject H_0
	β_2	0.1500	0.1086	0.1911	Reject H_0
	β_3	0.5121	0.4677	0.5551	Reject H_0
	β_4	-0.0004	-0.0006	-0.0002	Reject H_0
	β_5	0.0805	0.0636	0.0968	Reject H_0

Based on Table 3, the Bayesian Hurdle Poisson Regression model can be presented as follows

$$logit\ \hat{\pi}_i = 16{,}6551 - 0{,}1922X_{5i} \quad (6)$$

$$\ln \hat{\lambda}_i = -4{,}2404 - 0{,}0027X_{1i} + 0{,}1500X_{2i} + 0{,}5121X_{3i} - 0{,}0004X_{4i} + 0{,}0805X_{5i} \quad (7)$$

The interpretation of the logit model in equation (6), that is, every 1% increase in the percentage of households that have access to proper sanitation in 34 Provinces in Indonesia will increase the probability of the number of cases of death due to chronic Filariasis in 34 Provinces in Indonesia by exp(-0.1922) = 0.825 times of the original number of death from chronic Filariasis cases.

The interpretation of Poisson's truncated model in equation (7) is:

1. Every 1 person increase in the total number of chronic Filariasis cases in 34 Provinces in Indonesia will increase the average number of deaths due to chronic Filariasis in 34 Provinces in Indonesia by exp(-0.0027)=0.997≈1 person.
2. Every increase in 1 District/City that succeeds in reducing Microphilia <1% will increase the average number of cases of death due to chronic Filariasis in 34 Provinces in Indonesia by exp(0.1500)=1.16≈1 person.
3. Every increase in 1 District/City in Indonesia that is still implementing Mass Preventive Drug Delivery (MPDD) will increase the average number of cases of death due to chronic Filariasis in 34 Provinces in Indonesia by exp(0,5121)=1,669≈2 persons.
4. Every 1 person/km2 increase in population density in 34 Provinces in Indonesia will increase the average number of cases of death due to chronic filariasis in 34 Provinces in Indonesia by exp(-0.0004)=0.9996≈1 person.

5. Every 1% increase in the percentage of households having access to proper sanitation in 34 Provinces in Indonesia will increase the average number of cases of death due to chronic Filariasis in 34 Provinces in Indonesia by exp(0.0805)=1.08≈ 1 person.

CONCLUSIONS

In the logit model, the percentage of households that have access to proper sanitation in 34 Provinces in Indonesia (X_5) has a significant effect on the number of cases of death due to chronic Filariasis in 34 Provinces in Indonesia (Y). Then in the Truncated Poisson model, all predictor variables, namely the number of all chronic cases of Filariasis in 34 Provinces in Indonesia (X_1), the number of district/cities managed to reduce microphilia <1% in 34 Provinces in Indonesia (X_2), the number of district/cities that are still implementing the Mass Preventive Drug Delivery (MPDD) for Filariasis in 34 Provinces in Indonesia (X_3), population density in 34 Provinces in Indonesia (X_4), as well as the percentage of households that have access to proper sanitation in 34 Provinces in Indonesia (X_5) have a significant effect on the number of deaths due to chronic Filariasis in 34 Provinces in Indonesia (Y).

REFERENCES

[1] D. W. Osgood, “Poisson Based Regression Analysis of Aggregate Crime Rates,” *Quant. Methods Criminol.*, vol. 16, no. 1, pp. 577–599, 2017, doi: 10.4324/9781315089256-23.

[2] W. T. Tedra, I. M. Rizki, and D. Prariesa, “Konsumsi Rokok Masyarakat Kota Bandung Tahun 2015 Dengan Model Hurdle Negatif Binomial (Hurdle-Nb),” *Forum Statistika dan Komputasi.,* vol. 15, no.1, pp. 18–27, 2015.

[3] A. Taufiq, A. B. Astuti, and A. A. Rinaldo Fernandes, “Geographically Weighted Regression in Cox Survival Analysis for Weibull Distributed Data with Bayesian Approach,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 546, no. 5, 2019, doi: 10.1088/1757-899X/546/5/052078.

[4] A. R. Maulana, S. Astutik, U. Brawijaya, and L. Belakang, “Penerapan Regresi Zero Inflated Poisson dengan Metode Bayesian,” *Prosiding Seminar Nasional Pendidikan Matematika,* vol. 1, pp. 226–233, 2016.

[5] G. Meliyanie and D. Andiarsa, “Program Eliminasi Lymphatic Filariasis di Indonesia,” *J. Heal. Epidemiol. Commun. Dis.*, vol. 3, no. 2, pp. 63–70, 2019, doi: 10.22435/jhecds.v3i2.1790.

[6] A. A. Arsin, Epidemiologi Filariasis di Indonesia, Makassar: Masagna Press, 2016.

[7] A. Ernawati, “Faktor Risiko Penyakit Filariasis (Kaki Gajah),” *J. Litbang Media Inf. Penelitian, Pengemb. dan IPTEK*, vol. 13, no. 2, pp. 105–114, 2017, doi: 10.33658/jl.v13i2.98.

[8] Kementerian Kesehatan RI, “Situasi Filariasis di Indonesia Tahun 2018,” *Infodatin Pusat Data dan Informasi Kementerian Kesehatan RI*. pp. 1&4, 2019, [Online]. Available: https://pusdatin.kemkes.go.id/download.php?file=download/pusdatin/infodatin/InfoDatin-Filariasis-2019.pdf.

[9] Kementerian Kesehatan RI, "Profil Kesehatan Indonesia 2020," 2021.

[10] F. Antoneli, F. M. Passos, L. R. Lopes, and M. R. S. Briones, “A Kolmogorov-Smirnov test for the molecular clock based on Bayesian ensembles of phylogenies,” *PLoS One*, vol. 13, no. 1, 2018, doi: 10.1371/journal.pone.0190826.

[11] D. N. Gujarati and D. C. Porter, Dasar-Dasar Ekonometrika, Edisi 5, Jakarta: Salemba Empat, 2012.

[12] A. Agresti, Categorical Data Analysis Second Edition, New York: John Wiley & Sons Inc, 2002.

[13] G. E. P. Box and G. C. Tiao, *Bayesian Inference in Statistical Analysis*. 1992.

[14] A. B. Astuti, N. Iriawan, Irhamah, H. Kuswanto, and L. Sasiarini, “Blood Sugar Levels of Diabetes Mellitus Patients Modeling with Bayesian Mixture Model Averaging,” *Glob. J. Pure Appl. Math.*, vol. 12, no. 4, pp. 3143–3158, 2016.

[15] I. Ntzoufras, "Bayesian modeling using WinBUGS", vol. 698. John Wiley & Sons, 2011.