

Received: 6 Mei 2021

Revised: 16 Juni 2021

Accepted: 23 Juni 2021

Published: 30 Juni 2021

## Pemodelan Jumlah Kasus Malaria di Indonesia Menggunakan Generalized Linear Model

Vera Maya Santi <sup>1, a)</sup>, Abi Wiyono <sup>2, b)</sup>, Sudarwanto <sup>2, c)</sup><sup>1</sup> Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta<sup>2</sup> Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri JakartaE-mail: <sup>a)</sup> vmsanti@unj.ac.id, <sup>b)</sup> abiwiy09@gmail.com, <sup>c)</sup> sudarwanto@unj.ac.id

### Abstract

Generalized Linear Model (GLM) has been used for modelling various types of data where the distribution of response variables is an exponential family. Common examples include those for Binomial and Poisson response data. The GLM regression model determines the structure of the explanatory variable or covariate information, where the link function specifically determines the relationship between the regression model and the expected value of the observation. The purpose of this article is to obtain significant predictor variables for the model. Estimating the regression model parameters is using Maximum Likelihood Estimation. The number of malaria cases was identified as having a Poisson distribution. There are 3 predictor variables that have significant value on the regression model, that is proper sanitation access, healthy zoning and integrated vector control.

**Keywords:** Generalized Linear Model, Exponential family distribution, Link Function, Maximum Likelihood Estimation, GLM where the response variable is Poisson distribution

### Abstrak

*Generalized Linear Model (GLM)* telah banyak digunakan untuk memodelkan berbagai macam tipe data dimana distribusi dari variabel respon merupakan distribusi yang termasuk dalam distribusi keluarga eksponensial. Contoh umum dari distribusi keluarga eksponensial adalah distribusi Poisson dan Binomial. Model regresi GLM mendeskripsikan struktur dari variabel prediktor, sedangkan fungsi penghubung secara khusus mendeskripsikan hubungan antara model regresi dengan nilai ekspektasi dari variabel respon. Tujuan dari artikel ini adalah mendapatkan variabel-variabel prediktor yang berpengaruh signifikan terhadap model. Metode *Maximum Likelihood Estimation* digunakan untuk mencari estimasi dari nilai parameter regresi model. Jumlah kasus malaria di Indonesia diidentifikasi berdistribusi Poisson. Terdapat 3 variabel prediktor yang berpengaruh signifikan terhadap jumlah kasus malaria di Indonesia, yaitu persentase rumah tangga yang memiliki akses sanitasi layak, jumlah kabupaten/kota yang menyelenggarakan tatanan kawasan kesehatan dan jumlah kabupaten/kota yang melakukan pengendalian vektor terpadu.

**Kata-kata kunci:** *Generalized Linear Model*, Distribusi keluarga eksponensial, Fungsi penghubung, *Maximum Likelihood Estimation*, GLM untuk data variabel respon berdistribusi Poisson

## PENDAHULUAN

Model linier merupakan suatu model yang merepresentasikan hubungan antara variabel respon Y dan variabel-variabel prediktor X (Faraway dalam Jain et al., 2016). Terdapat 2 analisis dalam model linier yaitu analisis regresi linier sederhana dan analisis regresi linier berganda. Model regresi linier sederhana hanya memiliki satu variabel prediktor sedangkan pada model regresi linier berganda terdapat lebih dari satu variabel prediktor. Pada model regresi linier terdapat salah satu persyaratan yang harus terpenuhi yaitu uji normalitas, dimana asumsi sebaran data variabel respon berdistribusi Normal.

*Generalized Linear Model* (GLM) adalah perluasan dari model regresi linier dimana sebaran variabel respon merupakan anggota dari keluarga eksponensial (Nelder dan Wedderburn, 1972; McCullagh dan Nelder, 1989). Tujuan dari *Generalized Linear Model* (GLM) yaitu untuk mengetahui pengaruh dari variabel prediktor terhadap variabel respon (McCulloch, 2001). Keunggulan GLM dari model linier terletak pada bentuk kurva atau distribusi variabel respon. Variabel respon pada GLM tidak diisyaratkan berbentuk kurva lonceng simetris (berdistribusi Normal), akan tetapi distribusi-distribusi yang termasuk dalam distribusi keluarga eksponensial, di antaranya distribusi Poisson, Gamma, Binomial, Invers Gaussian, Normal, dan Binomial Negatif (Wedderburn, 1974). Beberapa algoritma dan metode penduga parameter model GLM juga sudah dikembangkan seperti yang dikerjakan oleh Bradley (1973), Baker dan Nelder (1978), Green (1984), dan Wedderburn (1976).

*Generalized Linear Model* untuk data variabel respon berdistribusi Poisson mengartikan hubungan antara variabel respon Y data sebaran Poisson dengan variabel prediktor X yang berupa data diskrit atau kontinu (Obubu et al., 2018). Contoh umum dari kasus GLM untuk data variabel respon berdistribusi Poisson adalah count data (data jumlah).

Malaria adalah penyakit menular akut yang disebabkan oleh gigitan nyamuk *Anopheles* betina yang terinfeksi parasit dari genus *Plasmodium*. Terdapat 4 spesies penyebab malaria di antaranya *P. Falciparum*, *P. Vivax*, *P. Ovale* dan *P. Malaria* (Puasa et al. dalam Santi et al., 2019). Terdapat 219 juta kasus malaria ditemukan di 97 negara selama tahun 2017. Kasus ini meningkat dari 216 juta kasus yang ditemukan pada tahun 2016 dengan angka kematian 435 ribu orang per tahunnya (WHO dalam Sorontou, 2020).

Jumlah kasus malaria yang tergolong tinggi di Indonesia berasal dari bagian timur daerah Indonesia antara lain; Papua, Nusa Tenggara Timur dan Maluku Utara. Provinsi lain juga dilaporkan memiliki jumlah kasus malaria cukup tinggi diantaranya provinsi Bengkulu, Bangka Belitung, Kalimantan Tengah, Lampung dan Sulawesi Tengah (Ditjen dan Kemenkes dalam Santi et al., 2019). Terdapat berbagai faktor yang memengaruhi jumlah kasus positif malaria di Indonesia, yaitu kepadatan penduduk, rasio GINI, rumah tangga yang mendapat akses sanitasi layak dan lain-lain. Oleh karena itu, untuk menyelesaikan masalah dalam menduga variabel-variabel prediktor yang berpengaruh signifikan pada masalah tersebut, diselesaikan dengan memodelkan jumlah kasus malaria di Indonesia menggunakan *Generalized Linear Model*.

Penelitian mengenai *Generalized Linear Model* untuk data variabel respon berdistribusi Poisson sebelumnya pernah dilakukan oleh Santi et al. pada tahun 2019 dengan judul *Pemodelan Jumlah Kasus Malaria di Indonesia menggunakan Bayesian Generalized Linear Model*. Penelitian lain juga dilakukan oleh Obubu et al. Pada tahun 2018 dengan judul *Modelling Count Data; A Generalized Linear Model Framework*. Pada kajian ini, akan menerapkan *Generalized Linear Model* dalam memodelkan jumlah kasus malaria di Indonesia beserta permasalahan pelanggaran asumsi *equidispersi*.

## METODOLOGI

### Bahan dan Data

Pada penelitian ini, data yang digunakan adalah data sekunder jumlah kasus malaria di Indonesia pada tahun 2017. Data diperoleh dari publikasi kementerian kesehatan Indonesia (Untung et al., 2018). Variabel-variabel prediktor yang digunakan adalah sebagai berikut:

- Y = Jumlah kasus positif Malaria
- X1 = Kepadatan penduduk
- X2 = Rasio GINI
- X3 = Puskesmas yang mengimplementasikan pengendalian terpadu penyakit tidak menular
- X4 = Rumah tangga yang memiliki akses sanitasi layak
- X5 = Kabupaten/kota yang menyelenggarakan tatanan kawasan kesehatan
- X6 = Desa/kelurahan yang melaksanakan sanitasi total berbasis masyarakat
- X7 = Kabupaten/kota yang memiliki kebijakan perilaku hidup bersih dan sehat
- X8 = Kabupaten/kota yang melakukan pengendalian vektor terpadu
- X9 = Desa/kelurahan yang melaksanakan pos pembinaan terpadu penyakit tidak menular
- X10 = Tempat-tempat umum yang memenuhi syarat kesehatan
- JP = Jumlah penduduk setiap Provinsi (*offset*)

### Metode Penelitian

Langkah-langkah yang dilakukan dalam pengolahan data sekunder untuk mendapatkan variabel-variabel prediktor yang signifikan terhadap variabel respon dalam kasus Jumlah kasus malaria di Indonesia menggunakan *Generalized Linear Model* adalah sebagai berikut :

- a. Identifikasi variabel respon berdistribusi Poisson
- b. Melakukan uji Multikolinieritas
- c. Menduga parameter regresi ( $\beta$ ) GLM untuk data variabel respon berdistribusi Poisson dengan menggunakan metode MLE
- d. Melakukan pengujian signifikansi parameter regresi GLM untuk data variabel respon berdistribusi Poisson. Pengujian parameter regresi pada GLM untuk data variabel respon berdistribusi Poisson dilakukan menggunakan uji serentak dan uji parsial
- e. Mendapatkan model GLM untuk data variabel respon berdistribusi Poisson
- f. Menguji syarat *equidispersi*  
Menguji syarat *equidispersi* dilakukan dengan menghitung hasil pembagian antara *Residual Deviance* dengan derajat bebasnya. Apabila hasil pembagian tersebut memiliki hasil 1 disebut *equidispersi*, maka model terhindar dari kasus *overdispersi*. Apabila hasil pembagian tersebut  $> 1$  maka terjadi kasus *overdispersi*, sehingga untuk penyelesaiannya menggunakan analisis alternatif, salah satunya menggunakan regresi Binomial Negatif (Melliana et al., 2013)
- g. Memilih model terbaik  
Memilih model yang tepat dengan melihat nilai *standard error* dari pendugaan parameter model dan kriteria AIC.

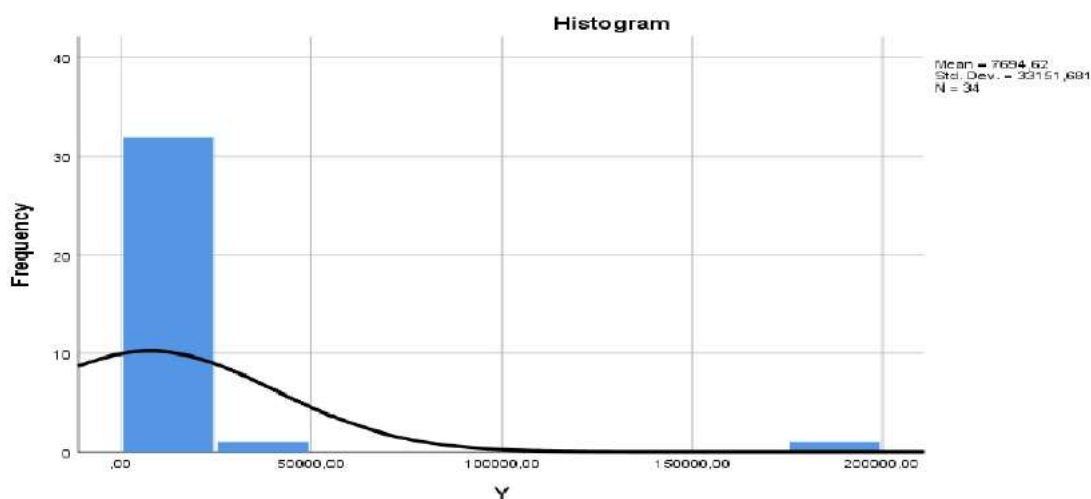
**HASIL DAN PEMBAHASAN**

Data variabel respon Y yaitu jumlah kasus positif Malaria di 34 provinsi di Indonesia tahun 2017 disajikan pada TABEL 1. Data respon diasumsikan memiliki distribusi Poisson karena merupakan data cacahan.

**TABEL 1.** Jumlah kasus positif Malaria di Indonesia tahun 2017

No.	Provinsi	Y	No.	Provinsi	Y
1	Aceh	299	18	Nusa Tenggara Barat	765
2	Sumatera Utara	2442	19	Nusa Tenggara Timur	30451
3	Sumatera Barat	512	20	Kalimantan Barat	152
4	Riau	211	21	Kalimantan Tengah	760
5	Jambi	187	22	Kalimantan Selatan	1135
6	Sumatera Selatan	911	23	Kalimantan Timur	1573
7	Bengkulu	1031	24	Kalimantan Utara	65
8	Lampung	4297	25	Sulawesi Utara	900
9	Kep. Bangka Belitung	95	26	Sulawesi Tengah	543
10	Kep. Riau	357	27	Sulawesi Selatan	1201
11	DKI Jakarta	111	28	Sulawesi Tenggara	596
12	Jawa Barat	328	29	Gorontalo	46
13	Jawa Tengah	904	30	Sulawesi Barat	147
14	Di Yogyakarta	86	31	Maluku	4019
15	Jawa Timur	107	32	Maluku Utara	957
16	Banten	42	33	Papua Barat	13706
17	Bali	33	34	Papua	192648

Pertama-tama identifikasi distribusi variabel respon Y dari data penelitian dengan melihat *histogram*. Berdasarkan GAMBAR 1, bentuk distribusi variabel respon Y tidak berbentuk kurva lonceng simetris (berdistribusi normal), melainkan kurva yang condong ke kanan. *Count data* (data jumlah) diidentifikasi memiliki sebaran Poisson (Obubu et al., 2018). Maka data variabel respon Y yaitu Jumlah Kasus Positif Malaria di Indonesia diidentifikasi berdistribusi Poisson.



GAMBAR 1. Histogram variabel respon Y

Langkah selanjutnya yaitu melakukan uji multikolinieritas. Pada model regresi disyaratkan tidak boleh terjadi korelasi antar variabel prediktor (*multikolinieritas*). Untuk mendeteksi ada tidaknya multikolinieritas mengacu pada kriteria nilai VIF. Jika nilai VIF (*Variance Inflation Factor*) < 10 maka tidak terjadi kasus multikolinieritas (Dobson dalam Obubu et al., 2018). Dengan bantuan *software R*, diperoleh nilai VIF sebagaimana diperlihatkan pada TABEL 2 di bawah ini.

TABEL 2. Uji Multikolinieritas

Model	VIF
Intercept	
X1	2.191
X2	1.424
X3	4.580
X4	1.407
X5	5.063
X6	2.462
X7	5.887
X8	1.721
X9	5.685
X10	1.771

Berdasarkan TABEL 2, semua variabel prediktor X mempunyai nilai VIF < 10, yang memiliki makna bahwa tidak terjadi kasus multikolinieritas antar variabel prediktor.

Langkah selanjutnya yaitu menduga parameter regresi ( $\beta$ ) dengan metode MLE, kemudian melakukan uji signifikansi parameter secara serentak dan parsial, diperoleh output seperti pada TABEL 3.

**TABEL 3.** Hasil uji signifikansi parameter regresi ( $\beta$ ) GLM data variabel respon berdistribusi Poisson

Parameter	Taksiran $\beta$	Std. Error	Z value	P value
(Intercept)	-8.770	0.012	-716.45	<2e-16 ***
X1	-2.352	0.060	-39.135	<2e-16 ***
X2	0.644	0.006	101.99	<2e-16 ***
X3	-0.587	0.013	-45.371	<2e-16 ***
X4	-0.580	0.005	-119.556	<2e-16 ***
X5	-0.996	0.012	-80.583	<2e-16 ***
X6	0.341	0.006	59.174	<2e-16 ***
X7	-0.120	0.013	-9.281	<2e-16 ***
X8	0.203	0.006	36.776	<2e-16 ***
X9	-0.632	0.016	-39.458	<2e-16 ***
X10	-0.290	0.008	-35.969	<2e-16 ***

Berdasarkan TABEL 3, dari sepuluh variabel prediktor seluruhnya berpengaruh signifikan terhadap variabel respon Y karena memiliki  $p\text{-value} < \alpha = 0.05$ . Maka berdasarkan output pada TABEL 3 diperoleh model GLM untuk data variabel respon berdistribusi Poisson yaitu

$$\mu = \exp(-0.8770 - 2.352X_1 + 0.644X_2 - 0.587X_3 - 0.580X_4 - 0.996X_5 + 0.341X_6 - 0.120X_7 + 0.203X_8 - 0.632X_9 - 0.290X_{10}).$$

Kemudian melakukan uji kesesuaian model (*Godness of fit*) dan lakukan uji *equidispersi* pada model GLM data variabel respon berdistribusi Poisson. *Overdispersi* dapat dilihat dari pembagian antara nilai *Residual Deviance* dengan derajat bebasnya, apabila hasil pembagian tersebut  $> 1$ , maka model dikatakan mengalami *overdispersi*. sehingga harus dilakukan analisis alternatif untuk mengatasi *overdispersi*, salah satunya melalui regresi Binomial Negatif (Melliana et al., 2013). *Overdispersi* menyebabkan nilai ragam menjadi *underestimate* yang mengakibatkan kecenderungan untuk menolak  $H_0$ , yang artinya parameter regresi ( $\beta$ ) yang seharusnya tidak signifikan menjadi signifikan, sehingga mengakibatkan pengambilan kesimpulan yang tidak valid (Melliana et al., 2013).

Uji kesesuaian model (*Godness of fit*) bertujuan untuk mengetahui model yang dihasilkan sesuai atau tidak dengan data yang diamati (Lindsey dalam Zahro et al., 2018). Kriteria pengujiannya yaitu tolak  $H_0$  jika  $G > \chi^2_{\{n-(p+1), \alpha\}}$  dimana  $n$  adalah banyaknya sampel,  $p$  adalah jumlah variabel prediktor dan  $\alpha$  adalah taraf signifikansi. Berdasarkan *output R*, nilai *Residual Deviance* pada model GLM untuk data variabel respon berdistribusi Poisson yaitu 59101, sedangkan nilai  $\chi^2_{(23, 0.05)}$  adalah 35.172, sehingga nilai  $G > \chi^2_{\{n-(p+1), \alpha\}}$ . Maka, Tolak  $H_0$  yang artinya model yang dihasilkan sesuai dengan data yang diamati.

Uji *equidispersi* pada model GLM data variabel respon berdistribusi Poisson. Berdasarkan *ouput R*, nilai *Residual deviance* pada model GLM untuk data variabel respon berdistribusi Poisson yaitu 59101, sedangkan derajat bebasnya adalah 23. Hasil bagi antara *Residual deviance* dengan derajat bebasnya yaitu  $59101 : 23 = 2569.609 > 1$ , sehingga model GLM data variabel respon berdistribusi Poisson mengalami kasus *overdispersi*.

Regresi Binomial Negatif adalah model campuran Poisson-Gamma yang digunakan untuk menganalisis data cacah apabila model GLM untuk data variabel respon berdistribusi Poisson mengalami kasus overdispersi (Hilbe dalam Melliana et al., 2013). Lakukan pendugaan parameter regresi ( $\beta$ ) menggunakan metode MLE, kemudian lakukan uji signifikansi parameter secara serentak dan parsial untuk model Regresi Binomial Negatif, diperoleh output seperti pada TABEL 4.

**TABEL 4.** Hasil uji signifikansi parameter regresi ( $\beta$ ) model regresi Binomial Negatif

Parameter	Taksiran $\beta$	Std. Error	Z value	P value
(Intercept)	6.968	0.177	39.43	<2e-16 ***
X1	-0.105	0.266	-0/396	0.692
X2	0.383	0.214	1.788	0.074
X3	-0.198	0.384	-0.516	0.606
X4	-0.912	0.213	-4.288	1.80E-05 ***
X5	-0.851	0.404	-2.106	0.035 ***
X6	-0.116	0.281	-0.414	0.679
X7	0.259	0.435	0.596	0.551
X8	0.831	0.235	3.532	0.000412 ***
X9	-0.102	0.428	-0.238	0.812
X10	-0.246	0.239	-1.033	0.302

Berdasarkan TABEL 4, parameter regresi ( $\beta$ ) yang signifikan yaitu  $\beta_0$ ,  $\beta_4$ ,  $\beta_5$  dan  $\beta_8$  karena memiliki  $p\text{-value} < \alpha = 0.05$ . Maka diperoleh model regresi Binomial Negatif yaitu

$$\mu = \exp ( 6.968 - 0.912 X4 - 0.851 X5 + 0.831 X8)$$

Langkah terakhir yaitu memilih model terbaik dengan melihat nilai AIC dan *standard error* terkecil. Dalam kasus GLM data variabel respon berdistribusi Poisson yang mengalami kasus *overdispersi*, dalam memilih model terbaik tidak dapat mengacu pada *standard error* terkecil karena *overdispersi* menyebabkan nilai ragam menjadi *underestimate* yang mengakibatkan kecenderungan untuk menolak  $H_0$ , yang artinya parameter regresi ( $\beta$ ) yang seharusnya tidak signifikan menjadi signifikan, sehingga mengakibatkan pengambilan kesimpulan yang tidak valid (Melliana et al., 2013).

**TABEL 5.** Pemilihan Model Terbaik

Model	Nilai AIC
GLM variabel respon berdistribusi Poisson	59401
Regresi Binomial Negatif	565.87

Berdasarkan TABEL 5, model yang memiliki nilai AIC terkecil yaitu model regresi Binomial Negatif, sehingga model yang dipilih untuk memodelkan Jumlah kasus Malaria di Indonesia yaitu

$$\mu = \exp ( 6.968 - 0.912 X4 - 0.851 X5 + 0.831 X8).$$

Berdasarkan model yang terbentuk dapat dijelaskan interpretasi dari model yaitu sebagai berikut:



1. Setiap penambahan 1 persen variabel X4 akan meningkatkan rata-rata variabel respon sebesar  $\exp(-0.912) = 0.402$  kali dari rata-rata variabel respon semula bila variabel lain konstan. Hal ini berarti setiap penambahan 1 persen dari persentase rumah tangga yang memiliki akses sanitasi layak akan setara dengan penurunan rata-rata jumlah kasus positif malaria di Indonesia sebesar 0:402 kali dari rata-rata semula bila variabel lain konstan.
2. Setiap penambahan 1 persen variabel X5 akan meningkatkan rata-rata variabel respon sebesar  $\exp(-0.851) = 0.427$  kali dari rata-rata variabel respon semula bila variabel lain konstan. Hal ini berarti setiap penambahan 1 persen dari jumlah kabupaten/kota yang menyelenggarakan tatanan kawasan kesehatan akan setara dengan penurunan rata-rata jumlah kasus positif malaria di Indonesia sebesar 0.427 kali dari rata-rata semula bila variabel lain konstan.
3. Setiap penambahan 1 persen variabel X8 akan meningkatkan rata-rata variabel respon sebesar  $\exp(0.831) = 2.290$  kali dari rata-rata variabel respon semula bila variabel lain konstan. Hal ini berarti setiap penambahan 1 persen dari jumlah kabupaten/kota yang melaksanakan pengendalian vektor terpadu akan setara dengan kenaikan rata-rata jumlah kasus positif malaria di Indonesia sebesar 2.290 kali dari rata-rata semula bila variabel lain konstan.
4. Estimasi rata-rata variabel respon adalah sebesar  $\exp(6.968) = 1062.309$  jika nilai X4, X5 dan X8 ketiganya bernilai 0. Hal ini berarti, estimasi rata-rata jumlah kasus positif malaria di Indonesia adalah sebesar  $1062.309 \approx 1063$  bila nilai persentase rumah tangga yang mempunyai akses sanitasi layak, jumlah kabupaten/kota yang menyelenggarakan tatanan kawasan kesehatan dan jumlah kabupaten/kota yang melaksanakan pengendalian vektor terpadu ketiganya bernilai 0.

#### KESIMPULAN DAN SARAN

Berdasarkan hasil analisis dan pembahasan yang telah dilakukan, pendugaan parameter regresi pada GLM data variabel respon berdistribusi Poisson dilakukan dengan metode MLE. Kemudian menggunakan regresi Binomial Negatif untuk mengatasi kasus *overdispersi*, dapat diketahui bahwa terdapat 3 variabel prediktor yang berpengaruh secara signifikan terhadap jumlah kasus positif Malaria di Indonesia yaitu persentase rumah tangga yang memiliki akses sanitasi layak (X4), jumlah kabupaten/kota yang menyelenggarakan tatanan kawasan kesehatan (X5) dan jumlah kabupaten/kota yang melakukan pengendalian vektor terpadu (X8).

Saran yang dapat diberikan yaitu pada penelitian ini dalam menduga parameter regresi ( $\beta$ ) penulis menggunakan metode MLE. Pada penelitian selanjutnya perlu dikaji lebih dalam mengenai penggunaan metode Bayesian. Saran kedua yaitu pada penelitian ini penulis menggunakan regresi Binomial Negatif dalam mengatasi kasus *overdispersi* pada GLM data variabel respon berdistribusi Poisson. Pada penelitian selanjutnya perlu dikaji lebih dalam mengenai regresi Quasi-Likelihood.

#### REFERENSI

- Baker. R.J.. and Nelder, J.A. (1978). The GLIM System, Release 3: Generalized linear Interactive Modelling. Numerical Algorithms Group, Oxford.
- Bradley. E.L. (1973). The equivalence of maximum likelihood and weighted least squares in the exponential family, J. Statist. Assoc., 68, 199–200.
- Green, P.J. (1984). Iteratively reweighted least squares for maximum likelihood estimation. and some robust and resistant alternatives, J. Roy. Statist. Soc., Ser. B, 46. 149–192 (with discussion).
- Hilbe, J.M 2011. “*Negative Binomial Regression Second Edition*. New York: Cambridge University”.
- Jain, S et al. 2016. “Regression Analysis - Its Formulation and Execution In Dentistry”, Journal of Applied Dental and Medical Sciences, 2
- McCullagh, P.. and Nelder. J.A. (1989). “Generalized linear Models”. Chapman and Hall, London.



- McCulloch CE, Searle SR. 2001. “*Generalized, Linear and, Mixed Models*”. New York: John Wiley and Sons, Inc
- Melliana, A et al. 2013. “The Comparison Of Generalized Poisson Regression And Negative Binomial Reression Methods in Overcoming Overdispersion.”, *International Journal of Scientific and Technology Research*, 2.
- Nelder, J.A, Wedderburn, R.W. M. 1972. “*Generalized Linear Models*”. New York: John Wiley and Sons, Inc
- Obubu, M et al. 2018. “Modelling Count Data; A Generalized Linear Model Framework”, *American Journal of Mathematics and Statistics*, 8(6).
- Santi, V.M et al. 2019. “Modelling of the Number of Malarias Suffers in Indonesia using Bayesian Generalized Linear Models”, *Journal of Physics: Conference Series*, 1402(7).
- Sorontou, Y 2020. “Uji Efektivitas Abate dan Kapur Kerang Anadonta Anatinae terhadap Larva Nyamuk Anopheles sp sebagai Vektor Malaria”.
- Untung, S et al. 2018. “Profil Kesehatan Indonesia Tahun 2017, Jakarta: Kementerian Kesehatan Republik Indonesia”.
- Wedderburn, R.W.M. (1974). “Quasilikelihood functions, generalized linear models and the Gauss-Newton method”. *Biometrika*, 61, 439–447.
- Wedderburn, R.W.M. (1976). On the existence and uniqueness of the maximum likelihood estimates for certain generalized linear models, *Biometrika*, 63, 27–32.