

# Aplikasi *Generalized Poisson Regression* dalam Mengatasi Overdispersi pada Data Jumlah Penderita Demam Berdarah Dengue

ARWINI ARISANDI, ERNA TRI HERDIANI, SITTI SAHRIMAN

Program Studi Statistika, Departemen Matematika FMIPA, Universitas Hasanuddin

## ABSTRAK

Asumsi dasar dalam regresi Poisson yaitu nilai variansi data sama dengan nilai mean data. Namun, asumsi tersebut umumnya tidak terpenuhi, misalnya terdapat kasus overdispersi. Overdispersi dalam regresi Poisson terjadi apabila nilai variansinya lebih besar daripada nilai meannya. Jika terjadi overdispersi pada data, maka model regresi Poisson kurang akurat digunakan karena berdampak pada nilai standard error dari taksiran parameter yang dihasilkan cenderung menjadi underestimate sehingga kesimpulan yang diperoleh menjadi kurang valid. Dalam penelitian ini, kasus overdispersi dapat diatasi dengan model *generalized Poisson regression*. Hasil penelitian menunjukkan bahwa nilai AIC minimum diberikan oleh model *generalized Poisson regression*. Sehingga dalam penelitian ini disimpulkan bahwa pada penelitian terhadap data yang mengalami overdispersi pada Jumlah Penderita DBD di Kota Makassar tahun 2016, pemodelan regresi *generalized Poisson* mampu mengatasi terjadinya overdispersi yang terjadi pada pemodelan regresi Poisson. Nilai  $R^2$  yang dimiliki sebesar 67% yang artinya jumlah penderita DBD ditentukan oleh persentase tempat-tempat umum memenuhi syarat kesehatan, persentase penduduk yang memiliki akses air minum layak, persentase rumah tangga berperilaku hidup bersih dan sehat dan persentase rumah yang memenuhi syarat kesehatan. Selebihnya 33% ditentukan oleh faktor lain.

Kata Kunci: Overdispersi, *generalized Poisson regression*, AIC.

## 1. PENDAHULUAN

Analisis regresi linier merupakan metode analisis yang digunakan untuk mengukur pengaruh variabel prediktor terhadap variabel respon. Variabel respon dalam regresi linier merupakan data kontinu, seperti jarak rumah ke kampus, berat badan seorang anak, dan lain-lain. Namun, variabel respon umumnya dibatasi hanya pada data hitung, seperti data jumlah anak dalam suatu keluarga, jumlah korban kecelakaan, dan lain-lain. Karena keterbatasan nilai dari variabel respon, maka analisis regresi linier kurang tepat lagi digunakan untuk menaksir parameter untuk kasus data hitung. Sehingga untuk mendapatkan suatu model dari data hitung digunakan regresi Poisson.

Model regresi Poisson merupakan salah satu dari model *Generalized Linear Model* (GLM). Terdapat asumsi equidispersi dalam regresi Poisson yaitu nilai variansi data sama dengan nilai mean data. Namun, asumsi tersebut umumnya tidak terpenuhi, misalnya terdapat kasus overdispersi. Overdispersi dalam regresi Poisson terjadi apabila nilai variansinya lebih besar daripada nilai meannya. Jika terjadi overdispersi pada data, maka model regresi Poisson kurang akurat dalam pemodelan data karena berdampak pada nilai *standard error* dari taksiran parameter yang dihasilkan cenderung menjadi *underestimate* sehingga kesimpulan yang diperoleh menjadi kurang valid (McCullagh dan Nelder, 1989). Dalam perkembangannya, kasus overdispersi dapat diatasi dengan beberapa model regresi seperti model binomial negatif (Wedderburn, 1974) dan model *generalized Poisson* (Consul, 1989).

Beberapa peneliti telah menerapkan model regresi untuk data hitung yang mengalami overdispersi, diantaranya Gardner *et al.* (1995) dalam penelitiannya menjelaskan permasalahan yang terjadi ketika variabel respon merupakan data hitung dengan menggunakan regresi Poisson pada data jumlah kekerasan oleh orang yang berpenyakit mental. Hasilnya menunjukkan bahwa overdispersi dapat diatasi dengan model regresi binomial negatif (Gardner

dkk., 1995). Selain itu, Famoye *et al.* (2004) dalam penelitiannya membandingkan model regresi *generalized* Poisson dan binomial negatif pada data yang mengalami overdispersi yaitu data jumlah kasus kecelakaan. Hasilnya menunjukkan bahwa model regresi *generalized* Poisson lebih baik digunakan dibandingkan dengan model regresi binomial negatif (Famoye dkk., 2004). Selanjutnya Grover *et al.* (2015) membandingkan model regresi *generalized* Poisson dan regresi binomial negatif pada data yang mengalami overdispersi yaitu jumlah peningkatan sel CD4 pada penderita AIDS. Hasilnya menunjukkan bahwa model regresi *generalized* Poisson lebih baik digunakan dibandingkan dengan model regresi binomial negatif (Grover dkk., 2015). Pada penelitian ini, model regresi untuk mengatasi overdispersi adalah model regresi *generalized* Poisson.

## 2. TINJAUAN PUSTAKA

### GENERALIZED LINEAR MODEL

*Generalized Linear Model* merupakan perluasan dari model regresi ketika variabel respon memiliki model distribusi eror selain distribusi normal. Seluruh model dalam GLM memiliki tiga komponen yaitu (Abdulkadir dkk., 2015):

1. Komponen acak, diidentifikasi oleh variabel respon ( $Y_i$ ) dengan  $Y_i$  ( $Y_1, Y_2, \dots, Y_n$ ) saling bebas dan diasumsikan memiliki distribusi peluang yang bergantung pada  $\theta$ . Fungsi kepadatan peluangnya merupakan keluarga eksponensial. Bentuk keluarga eksponensial adalah

$$f(y) = \exp[a(y)b(\theta) + c(\theta) + d(y)]$$

dengan  $a(\cdot)$ ,  $b(\cdot)$ ,  $s(\cdot)$ ,  $t(\cdot)$  adalah fungsi yang diketahui,  $s(y) = \exp d(y)$ ,  $t(\theta) = \exp c(\theta)$  dan  $\theta$  adalah parameter dispersi.

2. Komponen sistematis, meliputi variabel-variabel prediktor dari model. Dalam notasi matriks, ditulis  $\eta = X\beta$  dengan  $\eta$  merupakan vektor dari  $n$  prediktor linier  $\eta = (\eta_1, \eta_2, \dots, \eta_n)^t$  dan  $\beta$  adalah vektor dari  $p$  parameter  $\beta = (\beta_0, \beta_1, \beta_2, \dots, \beta_p)^t$  dan  $X$  adalah matriks berukuran  $(n \times p)$ .

3. Fungsi Penghubung (*link function*), yaitu fungsi yang menghubungkan ekspektasi dari variabel respon ( $Y$ ) dengan variabel-variabel prediktor melalui persamaan linier.

$$\mu_i = \exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_j x_{ij} + \dots + \beta_p x_{ip})$$

Menurut McCullagh dan Nelder (1989), distribusi sederhana yang digunakan untuk memodelkan data hitung adalah distribusi Poisson sehingga model regresi Poisson merupakan salah satu kasus khusus dari GLM. Nilai variansi dari model Poisson identik dengan nilai meannya sehingga nilai  $\theta = 1$  untuk parameter dispersinya dan fungsi variansinya adalah  $Var(Y) = \mu$ .

### REGRESI POISSON

Model regresi Poisson dapat digunakan untuk memodelkan hubungan antara variabel prediktor terhadap variabel respon yang diasumsikan berdistribusi Poisson (Myers, 1990). Pengujian kecocokan distribusi Poisson dilakukan sebagai berikut (Akinwande dkk., 2015).

$$H_0 : p_k = p_k^0 \text{ (data berdistribusi Poisson)}$$

$$H_1 : p_k \neq p_k^0 \text{ (data tidak berdistribusi Poisson)}$$

Statistik uji yang digunakan adalah:

$$X^2 = \sum_{k=0}^m \frac{(n_k - np_k)^2}{np_k}$$

dengan:

$p_k$  = peluang pengamatan untuk kategori  $k$

$p_k^0$  = distribusi peluang poisson

$n_k$  = banyaknya frekuensi data pada  $k$  kategori

$m$  = banyaknya  $k$  kategori

Kriteria pengujiannya yaitu tolak  $H_0$  apabila nilai  $\chi^2 > \chi^2_{(m-1, \alpha)}$ .

Fungsi peluang dari distribusi Poisson (Myers, 1990).

$$f(y) = \frac{e^{-\mu} \mu^y}{y!}; y = 0, 1, 2, \dots$$

dengan  $\mu > 0$  dan  $\mu$  merupakan mean dari variabel respon  $Y$  dan  $E(Y) = Var(Y) = \mu$ .

Penaksiran parameter model regresi Poisson dilakukan dengan menggunakan metode *Maximum Likelihood Estimation* (MLE) yaitu dengan memaksimalkan model fungsi *log likelihood*. Fungsi *log likelihood* dari regresi Poisson adalah

$$L(\beta) = \prod_{i=1}^n \left[ \frac{\exp(-\mu_i) \mu_i^{y_i}}{y_i!} \right]$$

$$\ln L(\beta) = \ln \left( \prod_{i=1}^n \left[ \frac{\exp(-\mu_i) \mu_i^{y_i}}{y_i!} \right] \right) = -\sum_{i=1}^n \mu_i + \sum_{i=1}^n y_i \ln \mu_i - \sum_{i=1}^n \ln(y_i!)$$

dengan  $\mu_i = \exp(x_i' \beta)$

Dengan memaksimalkan model fungsi *log likelihood*, akan diperoleh penaksir *maximum likelihood*, yaitu melalui turunan pertama terhadap  $\beta$  yang disamakan dengan nol dan diselesaikan dengan metode numerik yaitu menggunakan iterasi Newton-Raphson.

$$\frac{\partial \ln L(\beta)}{\partial \beta} = -\sum_{i=1}^n x_i' \cdot \exp(x_i' \beta) + \sum_{i=1}^n y_i x_i' = 0$$

## MULTIKOLINIERITAS

Pendeteksian adanya kasus multikolinieritas dapat diketahui melalui nilai *Variance Inflation Factor* (VIF) seperti berikut.

$$VIF_j = \frac{1}{1 - R_j^2}$$

dengan  $R_j^2$  merupakan koefisien determinasi. Jika nilai  $VIF \geq 10$  maka diasumsikan koefisien regresi yang dihasilkan memiliki error yang sangat besar sehingga terindikasi terjadi multikolinieritas yang signifikan antar variabel prediktor (Myers, 1990).

## OVERDISPERSI

Karakteristik distribusi Poisson adalah equidispersi yaitu nilai variansi sama dengan nilai mean. Namun, asumsi tersebut seringkali tidak dipenuhi karena nilai variansi lebih besar dari nilai mean atau disebut overdispersi. Pengujian hipotesis tentang kasus overdispersi dilakukan sebagai berikut.

$H_0 : \theta = 0$  (tidak terjadi overdispersi)

$H_1 : \theta \neq 0$  (terjadi overdispersi)

Statistik uji yang digunakan adalah (Bisri, 2015):

$$\hat{\sigma} = \frac{s^2}{df}; \quad X^2 = \sum_{i=1}^n \frac{(y_i - \hat{\mu}_i)^2}{\hat{\mu}_i} \sim \chi^2_{(n-p)}$$

dengan kriteria pengujianya yaitu tolak  $H_0$  apabila nilai  $\chi^2 > \chi^2_{(n-p)}$  atau  $p\text{-value} < \alpha$ . Taksiran dispersi diukur dengan nilai deviansi atau Pearson's Chi-Square yang dibagi derajat bebas, jika hasil pembagian menghasilkan nilai yang lebih besar 1 maka dapat disimpulkan data mengalami overdispersi.

## REGRESI GENERALIZED POISSON

Jika variabel acak  $Y$  berdistribusi *generalized* Poisson maka fungsi distribusi peluang dari  $Y$  adalah

$$f(y_i) = \left( \frac{\mu_i}{1+\theta\mu_i} \right)^{y_i} \frac{(1+\theta y_i)^{y_i-1}}{y_i!} \exp \left[ \frac{-\mu_i(1+\theta y_i)}{1+\theta\mu_i} \right], \quad y_i = 0, 1, 2, \dots$$

Penaksiran parameter model regresi *generalized* Poisson dilakukan dengan menggunakan metode MLE yaitu dengan memaksimumkan model fungsi *log likelihood*. Fungsi *log likelihood* dari regresi *generalized* Poisson adalah

$$L(\beta, \theta) = \prod_{i=1}^n \left[ \left( \frac{\mu_i}{1+\theta\mu_i} \right)^{y_i} \frac{(1+\theta y_i)^{y_i-1}}{y_i!} \exp \left[ \frac{-\mu_i(1+\theta y_i)}{1+\theta\mu_i} \right] \right]$$

$$\ln L(\beta, \theta) = \sum_{i=1}^n \left[ y_i \ln \left( \frac{\mu_i}{1+\theta\mu_i} \right) + (y_i - 1) \ln(1 + \theta y_i) - \ln(y_i!) + \left[ \frac{-\mu_i(1+\theta y_i)}{1+\theta\mu_i} \right] \right]$$

dengan  $\mu_i = \exp(x_i' \beta)$

Dengan memaksimumkan model *log likelihood* diperoleh penaksir *maximum likelihood* untuk  $\beta$  dan  $\theta$ , yaitu

$$\frac{\partial \ln L(\beta, \theta)}{\partial \beta} = \sum_{i=1}^n y_i x_i' - \sum_{i=1}^n \frac{y_i \theta \exp(x_i' \beta)}{1 + \theta \exp(x_i' \beta)}$$

$$+ \sum_{i=1}^n (1 + \theta y_i) \frac{-x_i' \exp(x_i' \beta) (1 + \theta \exp(x_i' \beta)) + (\exp(x_i' \beta))^2 \theta x_i'}{(1 + \theta \exp(x_i' \beta))^2} = 0$$

dan

$$\frac{\partial \ln L(\beta, \theta)}{\partial \theta} = - \sum_{i=1}^n \frac{y_i \mu_i}{1 + \theta \mu_i} + \sum_{i=1}^n \frac{y_i (y_i - 1)}{1 + \theta y_i} + \sum_{i=1}^n \frac{\mu_i (y_i - \mu_i)}{(1 + \theta \mu_i)^2} = 0$$

Persamaan tersebut memperlihatkan bahwa turunan pertama dari model fungsi *log likelihood* pada estimasi parameter menggunakan metode MLE tidak dalam bentuk linier. Oleh karena itu, prosedur estimasi yang dapat dilakukan dengan menggunakan iterasi numerik. Iterasi numerik yang digunakan dalam penelitian ini adalah *Iteratively Reweighted Least Squares* (IRLS).

## PENGUJIAN HIPOTESIS MODEL REGRESI

Pengujian signifikansi parameter model regresi dilakukan untuk mengetahui pengaruh variabel prediktor terhadap variabel respon. Dalam pengujian parameter model regresi Poisson dan *generalized* Poisson, metode yang digunakan adalah dengan metode *Maximum Likelihood Ratio Test* (MLRT) (Agresti, 2002). Pengujian signifikansi model regresi *generalized* Poisson dengan hipotesis sebagai berikut.

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$$

$$H_1 : \text{paling sedikit ada satu } \beta_j \neq 0, \quad j = 1, 2, \dots, p$$

Statistik uji:

$$D(\hat{\beta}) = -2 \ln \Lambda = -2 \ln \left( \frac{L(\hat{\omega})}{L(\hat{\eta})} \right)$$

dengan  $L(\hat{\omega})$  adalah nilai *likelihood* untuk model sederhana tanpa melibatkan variabel prediktor dan  $L(\hat{\Omega})$  adalah nilai *likelihood* untuk model lengkap dengan melibatkan variabel prediktor. Tolak  $H_0$  apabila  $D(\hat{\beta}) > \chi^2_{(\alpha;p)}$ . Jika  $H_0$  ditolak berarti paling tidak ada satu  $\hat{\beta}_j \neq 0$  yang menunjukkan bahwa  $x_j$  berpengaruh secara signifikan terhadap model. Pengujian dilanjutkan dengan uji secara parsial dengan hipotesis sebagai berikut.

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0 ; j = 1, 2, \dots, p$$

Statistik uji:

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

Kesimpulan: tolak  $H_0$  jika  $W_j \geq \chi^2_{(\alpha;1)}$  atau  $p - value < 0,05$ .

### PEMILIHAN MODEL TERBAIK

Pemilihan model regresi terbaik dapat dilihat dari nilai *Akaike Information Criterion* (AIC)

$$AIC = -2 L + k$$

dengan L adalah fungsi *likelihood* dari model hasil estimasi *maximum likelihood* dan k jumlah parameter dalam model. Jika model memiliki nilai *log-likelihood* tertinggi atau nilai AIC minimum maka model tersebut adalah model terbaik (Grover dkk., 2015).

### 3. METODE ANALISIS

Model *generalized Poisson regression* diterapkan pada data Jumlah Penderita DBD di Kota Makassar tahun 2016. Variabel prediktor yang diduga berpengaruh terhadap variabel respon adalah persentase tempat-tempat umum memenuhi syarat kesehatan ( $X_1$ ), persentase penduduk yang memiliki akses air minum layak ( $X_2$ ), persentase rumah tangga berperilaku hidup bersih dan sehat ( $X_3$ ) dan persentase rumah yang memenuhi syarat kesehatan ( $X_4$ ).

Tahapan yang dilakukan dalam penelitian ini adalah sebagai berikut:

- 1) Melakukan pengujian distribusi Poisson.
- 2) Mengidentifikasi dan menyelesaikan adanya kasus multikolinieritas dengan melihat nilai VIF. Multikolinieritas yang signifikan terjadi jika nilai VIF lebih dari 10.
- 3) Melakukan pengujian overdispersi pada regresi Poisson.
- 4) Membentuk model regresi *generalized Poisson* jika terjadi overdispersi pada data dengan menggunakan MLE sebagai metode penduga parameternya dan melakukan pengujian signifikansi model dan signifikansi parameter.
- 5) Memilih model terbaik dengan melihat nilai AIC minimum.

### 4. HASIL DAN PEMBAHASAN

#### UJI DISTRIBUSI POISSON

Pengujian distribusi Poisson dilakukan untuk menguji suatu data berdistribusi Poisson atau tidak dengan melakukan uji kecocokan *Chi-Square* untuk distribusi Poisson. Uji kecocokan distribusi Poisson dapat dilihat pada Tabel 1.

**Tabel 1.** Uji Distribusi Poisson

$X^2$	$\chi^2_{(14;0,05)}$
7488,764	23,685

Hasil uji distribusi Poisson menunjukkan bahwa nilai statistik uji  $\chi^2 = 7488,764 > \chi^2_{(14;0,05)} = 23,685$  sehingga  $H_0$  ditolak. Artinya, data jumlah penderita DBD tidak berdistribusi Poisson. Hal ini menunjukkan bahwa terjadi pelanggaran asumsi equidispersi dalam regresi Poisson pada data jumlah penderita DBD.

## UJI MULTIKOLINIERITAS

Kasus multikolinieritas dapat diketahui dengan menggunakan nilai VIF.

**Tabel 2.** Nilai VIF Antar Variabel Prediktor

Variabel	$X_1$	$X_2$	$X_3$	$X_4$
VIF	1,089	1,042	1,046	1,001

Hasil dari uji VIF pada Tabel 2 menunjukkan bahwa semua variabel prediktor memiliki nilai kurang dari 10 sehingga dapat disimpulkan tidak terjadi multikolinieritas yang signifikan antar variabel prediktor.

## UJI OVERDISPERSI

Overdispersi dideteksi dengan nilai deviansi yang dibagi dengan derajat bebasnya.

**Tabel 3.** Nilai Deviansi Model Poisson

Deviansi	Df	$\theta$
146,450	41	3,572

Tabel 3 menjelaskan bahwa  $\theta = 3,572$  dari estimasi parameter model regresi poisson lebih dari 1 sehingga dapat disimpulkan bahwa model regresi Poisson pada data jumlah penderita DBD mengalami overdispersi. Oleh karena itu, kasus overdispersi diatasi dengan model *generalized Poisson regression*.

## MODEL GENERALIZED POISSON REGRESSION

Estimasi parameter model regresi *generalized Poisson* dilakukan dengan menggunakan metode MLE. Estimasi parameter yang diperoleh dilakukan pengujian signifikansi model dan parameter dari model regresi.

**Tabel 4.** Estimasi Parameter Model *Generalized Poisson Regression*

	Estimasi	Standard Error	$W_j$
$\hat{\beta}_0$	1,032	0,183	31,726
$\hat{\beta}_1$	0,216	1,753	0,015
$\hat{\beta}_2$	-0,012	0,331	0,001
$\hat{\beta}_3$	2,062	1,040	3,929
$\hat{\beta}_4$	-0,506	0,556	0,827
$D(\hat{\beta})$	= 242,976		

Tabel 4 menunjukkan bahwa nilai deviansi sebesar 242,976 lebih besar dari nilai *Chi-Square* yaitu 9,488 yang menyatakan terdapat paling sedikit satu variabel prediktor yang berpengaruh terhadap variabel respon. Variabel prediktor yang berpengaruh secara signifikan terhadap variabel respon yaitu  $X_3$ , karena nilai  $W_3 = 3,929 \geq 3,841$ . Model *generalized Poisson regression* yang terbentuk adalah sebagai berikut.

$$\hat{\mu}_i = \exp(1,032 + 0,216x_1 - 0,012x_2 + 2,062x_3 - 0,506x_4)$$

## PEMILIHAN MODEL TERBAIK

Model regresi terbaik adalah model yang memiliki nilai AIC minimum dan nilai  $R^2$  lebih besar antara model regresi Poisson dan *generalized* Poisson yang diperoleh.

**Tabel 5.** Model Regresi Terbaik

Model	AIC	$R^2$
Regresi Poisson $\hat{\mu}_i = \exp(2,055 - 1,353x_1 - 0,226x_2 + 1,651x_3 - 0,454x_4)$	296,860	12%
Regresi <i>Generalized</i> Poisson $\hat{\mu}_i = \exp(1,032 + 0,216x_1 - 0,012x_2 + 2,062x_3 - 0,506x_4)$	254,976	67%

Tabel 5 menjelaskan bahwa nilai AIC minimum antara model regresi Poisson dan *generalized* Poisson adalah 254,976 ditunjukkan oleh regresi *generalized* Poisson:  $\hat{\mu}_i = \exp(1,032 + 0,216x_1 - 0,012x_2 + 2,062x_3 - 0,506x_4)$  sehingga hal ini menunjukkan bahwa overdispersi dapat teratasi dengan menggunakan model regresi *generalized* Poisson.

## 5. KESIMPULAN

Berdasarkan hasil yang diperoleh dan diuraikan pada pembahasan sebelumnya, dapat ditarik kesimpulan bahwa pada penelitian terhadap data yang mengalami overdispersi pada Jumlah Penderita DBD di Kota Makassar tahun 2016, pemodelan regresi *generalized* Poisson mampu mengatasi terjadinya overdispersi yang terjadi pada pemodelan regresi Poisson. Hal ini dibuktikan dengan nilai AIC minimum yaitu 254,976 sehingga diperoleh model regresi *generalized* Poisson sebagai berikut:

$$\hat{\mu}_i = \exp(1,032 + 0,216x_1 - 0,012x_2 + 2,062x_3 - 0,506x_4)$$

Model tersebut memperlihatkan nilai  $R^2$  sebesar 67% yang artinya jumlah penderita DBD ditentukan oleh persentase tempat-tempat umum memenuhi syarat kesehatan, persentase penduduk yang memiliki akses air minum layak, persentase rumah tangga berperilaku hidup bersih dan sehat dan persentase rumah yang memenuhi syarat kesehatan. Selebihnya 33% ditentukan oleh faktor lain.

## DAFTAR PUSTAKA

- McCullagh P, Nelder JA. 1989. *Generalized Linier Models Second Edition*. London: Chapman and Hall.
- Wedderburn RWM. 1974. Quasi-likelihood functions, Generalized linear models, and the Gauss-Newton method. *Biometrika*. 61:439-447.
- Consul PC. 1989. *Generalized Poisson Distribution: Properties and Applications*. New York: Marcel Dekker.
- Gardner W, Mulvey EP, dan Shaw EC. 1995. Regression Analyses of Counts and Rates: Poisson, Overdispersed Poisson, and Negative Binomial Models. *Psychological Bulletin*. 118(3):392-404.
- Famoye F, Wulu JT, Singh KP. 2004. On the Generalized Poisson Regression Model with an Application to Accident Data. *Journal of Data Science*. 2:287-295.
- Grover G, Vajala R, dan Swai PK. 2015. On the assessment of various factors effecting the improvement in CD4 count of aids patients undergoing antiretroviral therapy using Generalized Poisson regression. *Journal of Applied Statistics*. 42:1291-1305.
- Abdulkabir M, Edem UA, Tunde RS, Kemi BL. 2015. An Empirical Study of Generalized Linear Model for Count Data. *Applied & Computational Mathematics Journal*. 4:1-3.
- Myers R. 1990. *Classical and Modern Regression with Applications, second edition*. Boston: PWS-KENT Publishing Company.

- Akinwande MO, Dikko HG, Samson A. 2015. Variance Inflation Factor: As a Condition for Inclusion of Suppressor Variable(s) in Regression Analysis. *Scientific Research Publishing*. 5:754-767.
- Bisri H. 2015. Analisis Regresi Poisson Lagrange dan Regresi Binom Negatif pada Data Overdispersi (Studi Kasus Penderita Penyakit Demam Berdarah *Dengue* di Kabupaten Mojokerto) [skripsi]. Malang: Universitas Brawijaya.
- Agresti A. 2002. *Categorical Data Analysis Second Edition*. Canada: John Wiley & Sons, Inc., Hoboken, New Jersey.