SAINTIFIK:
Jurnal Matematika, Sains, dan Pembelajarannya **Vol.9, No.2, Juli 2023, pp. 291~298**
ISSN 2407-4098 (print) DOI:10.31605/saintifik.v9i2.409
ISSN 2622-8904 (online)

Analysis Of Overdispersion Data Using Poisson Invers Gaussian Regression Model

Marsono
Badan Pusat Statistik Provinsi Sulawesi Barat
e-mail: marsono@bps.go.id

Abstrak

Salah satu pelanggaran asumsi dalam model regresi Poisson pada kebanyakan data count (cacahan) adalah ditemukan kasus overdispersi, dimana varians variabel respon lebih besar dari rata-ratanya. Untuk mengatasi kasus overdispersi, dibentuk model regresi yang merupakan perpaduan antara distribusi Poisson dengan beberapa distribusi lain. Poisson Invers Gaussian (PIG) merupakan salah satu model regresi dari model campuran untuk mengatasi overdispersi. Penelitian ini membandingkan model regresi Poisson Invers Gaussian dengan model regresi Poisson bertujuan mengetahui variabel-variabel yang mempengaruhi jumlah kasus bayi mati di Provinsi Jawa Timur pada tahun 2019. Penaksiran parameter regresi PIG dilakukan dengan metode Maximum Likelihood Estimator (MLE). Berdasarkan nilai AIC paling kecil diketahui model regresi Poisson Invers Gaussian lebih baik dari model regresi Poisson.

Kata Kunci— *Overdispersi, Poisson Invers Gaussian, Model Poisson, Kematian Bayi*

1. PENDAHULUAN

Untuk memodelkan hubungan antara suatu (satu atau lebih) variabel dependen atau respon dengan satu atau lebih variabel independen atau prediktor, umumnya digunakan model regresi linear klasik atau regresi *Ordinary Least Square* (OLS). Dalam regresi OLS terdapat beberapa asumsi yang harus dipenuhi. Dalam kenyataan di lapangan, tidak semua jenis data bisa memenuhi asumsi tersebut, misalnya pada kasus data cacahan (*count*). Pemodelan data cacahan menggunakan regresi OLS tidak dapat dilakukan karena melanggar dua asumsi yang disyaratkan dalam regresi OLS, yaitu error mengikuti distribusi normal (normalitas) dan varians harus konstan (homokedastisitas). Perkembangan dalam pemodelan data memunculkan pemodelan untuk data cacahan dengan *Generalized Linear Models* (GLMs). GMLs merupakan generalisasi dari model regresi normal klasik atau regresi OLS dari berbagai asumsi yang ketat dan menyediakan metode analisis bagi data tidak normal (De Jong dan Heller, 2008). Regresi Poisson adalah salah satu anggota keluarga dari GLMs yang berasal dari distribusi poisson. Distribusi poisson merupakan distribusi diskrit dengan nilai variabel random berupa bilangan bulat positif sehingga menjadi pilihan yang baik untuk pemodelan data cacahan. Distribusi poisson hanya ditentukan oleh satu parameter yang mendefinisikan baik *mean* maupun varians dari distribusi tersebut, sehingga dalam regresi Poisson terdapat asumsi yang harus terpenuhi yaitu *mean* dan varians variabel respon harus sama (*equidispersion*).

Namun dalam kenyataannya sering terjadi pelanggaran asumsi tersebut dimana varians lebih kecil dari *mean* (*underdispersion*) atau varians lebih besar dari *mean* (*overdispersion*). Pada kebanyakan data *count* terkadang ditemukan kasus overdispersi (Consul dan Famoye, 1992). Dalam praktiknya, data cacahan sering menunjukkan varians yang cukup besar karena banyak mengandung nilai nol (*extra zeros*) atau sebaran yang lebih besar dari nilai-nilai pada data atau keduanya (Hu, Pavlicova dan Nunes, 2011). Data dengan jumlah nol yang banyak bisa disebabkan karena nol struktural (*structural zeros*) atau karena nol sampel (*sampling zeros*). Lebih lanjut, Hu, Pavlicova dan Nunes memberi contoh jumlah perilaku seksual beresiko tinggi dalam membandingkan kedua hal tersebut. Nilai nol pada jumlah perilaku seksual beresiko tinggi bisa disebabkan karena seseorang yang tidak memiliki pasangan sehingga memang belum pernah melakukan hubungan seksual yang disebut nol struktural (*structural zeros*) atau bisa juga karena seseorang yang sudah memiliki pasangan dan belum pernah melakukan seksual beresiko tinggi yang disebut sebagai nol sampel (*sampling zeros*).

Kasus overdispersi bila diabaikan bisa mengakibatkan terjadinya *underestimate* pada estimasi *standard error*, sehingga dapat mengakibatkan kesalahan pada pengambilan keputusan beberapa uji hipotesis, misalnya suatu variabel prediktor berpengaruh signifikan ketika pada kenyataan tidak berpengaruh signifikan (Hilbe, 2007). Dalam mengatasi kasus overdispersi, dibentuk beberapa pemodelan yang merupakan perpaduan antara distribusi Poisson dengan beberapa distribusi baik diskrit maupun kontinu (*mixed poisson distribution*). *Mixed poisson distribution* tersebut merupakan solusi alternatif untuk kasus overdispersi, tetapi hanya beberapa distribusi yang sering digunakan dalam penelitian dikarenakan penghitungannya yang rumit. Salah satunya adalah distribusi *Poisson Invers Gaussian* (PIG) yang merupakan *mixed poisson distribution* dengan random efek yang memiliki distribusi Invers Gaussian. Distribusi ini pertama kali diperkenalkan oleh Holla pada tahun 1966 (Karlis dan Nikoloulopoulos, 2005). Distribusi PIG sendiri merupakan bentuk dari distribusi *Siche (SI)* dengan dua parameter. SI disebut sebut sebagai model yang lebih baik dari model binomial negatif, terutama untuk data yang overdispersi yang tinggi dan cenderung menceng kanan (*highly skewed to the right*). Akan tetapi, penghitungannya lebih rumit karena memiliki tiga parameter pada fungsi kepadatan peluangnya. Sebagai bentuk dari distribusi SI adalah distribusi PIG yang digunakan dalam memodelkan data cacahan yang menceng kanan serta memiliki ekor yang sedikit lebih panjang. Distribusi PIG memiliki bentuk fungsi likelihood yang *closed form* dan penghitungannya lebih mudah sehingga banyak penelitian yang melibatkan data cacahan banyak yang menggunakan model ini (Stasinopoulus & Rigby, 2007).

Willmot (1987) menunjukkan potensi dari pemodelan dengan regresi poisson inverse gaussian sebagai alternatif dari regresi binomial negatif pada data klaim asuransi mobil. Enam set data klaim asuransi mobil disajikan dengan karakteristik hampir 80 persen data yang mengandung nol, dan menghasilkan kesimpulan bahwa pemodelan dengan regresi PIG merupakan model yang lebih baik dibandingkan model regresi binomial negatif. Penelitian lainnya yang menggunakan model regresi PIG adalah penelitian dari Shoukri, Asyali dan Vandorp (2004) yang menggunakan model tersebut pada data jumlah kasus penyakit mastitis pada sapi perah di Ontario Canada dan menghasilkan kesimpulan bahwa regresi PIG merupakan model yang lebih baik dari model regresi binomial negatif. Selanjutnya dalam beberapa penelitian keselamatan di jalan raya seperti pemodelan data kecelakaan dan penelitian asuransi motor, model regresi PIG sering digunakan sebagai alternatif dari model regresi binomial negatif (Deniz, Ghitany dan Gupta, 2014). Pemodelan dengan regresi PIG juga digunakan pada penelitian Zha, Lord dan Zou (2014) pada kasus jumlah kecelakaan motor yang terjadi di dua tempat berbeda yaitu di Texas dan Washington. Dalam penelitian tersebut data jumlah kecelakaan motor menceng kanan dengan ekor yang sedikit panjang serta 37 persen data mengandung nilai nol. Dengan nilai *Akaike Information Criterion* (AIC) dan *Bayesian Information Criterion* (BIC) yang lebih kecil, menunjukkan bahwa pemodelan dengan regresi PIG lebih baik untuk memodelkan kasus jumlah kecelakaan motor di dua tempat tersebut.

Jumlah kasus bayi mati di suatu wilayah merupakan salah satu bentuk data cacahan, sehingga dalam pemodelannya bisa menggunakan regresi Poisson. Akan tetapi, data kematian bayi juga berpotensi terjadi overdispersi, sehingga pemodelannya tidak cukup menggunakan regresi Poisson. Pada penelitian ini akan dibandingkan pemodelan dengan menggunakan regresi Poisson dan regresi Poisson Invers Gaussian untuk mendapatkan model terbaik dari variabel-variabel yang mempengaruhi jumlah kasus bayi mati di Jawa Timur pada tahun 2019. Data yang digunakan merupakan data sekunder yang berasal dari profil kesehatan provinsi Jawa Timur tahun 2019.

2. METODE PENELITIAN

Data yang digunakan dalam penelitian ini merupakan data sekunder yang berasal dari profil kesehatan provinsi Jawa Timur tahun 2019. Jumlah sampel yang digunakan, yaitu sebanyak 38 kabupaten/ kota yang ada di provinsi Jawa Timur. Dengan variabel prediktor yang digunakan sebanyak 5 variabel. Berikut diberikan definisi masing-masing variabel respon dan prediktor.

Tabel 1. Variabel Respon dan Prediktor

Variabel	**Keterangan**
y	Jumlah Kematian Bayi
x_1	Persentase Kunjungan Ibu Hamil dengan K1
x_2	Persentase Kunjungan Ibu Hamil dengan K4
x_3	Persentase Persalinan Oleh Tenaga Kesehatan
x_4	Persentase Ibu Hamil Mendapat Tablet Tambah Darah (TTD)
x_5	Persentase Komplikasi Kebidanan yang Ditangani

Langkah-langkah untuk menentukan faktor-faktor yang mempengaruhi jumlah kematian bayi di provinsi Jawa Timur 2019 dengan model regresi Poisson Inverse Gaussian adalah sebagai berikut:

1. Melakukan pemeriksaan kasus multikolinieritas dengan menggunakan kriteria VIF.
2. Melakukan Pemodelan yang meliputi estimasi parameter dan pengujian parameter yang signifikan untuk model regresi Poisson dan regresi Poisson Inverse Gaussian.
3. Melakukan uji overdispersi
4. Melakukan pemodelan dengan menggunakan regresi Poisson Inverse Gaussian yang meliputi estimasi parameter dan pengujian hipotesis secara simultan dan parsial.
5. Membandingkan nilai AIC untuk mencari model terbaik.
6. Melakukan interpretasi model regresi Poisson Inverse Gaussian yang didapatkan.
7. Membuat kesimpulan dari hasil analisis.

3. HASIL DAN PEMBAHASAN

Regresi Poisson Inverse Gaussian (PIG) merupakan regresi yang dapat diaplikasikan pada data *count* yang mengalami overdispersi. Pada penelitian ini, pemodelan regresi PIG diaplikasikan pada data jumlah kasus bayi mati di propinsi Jawa Timur pada tahun 2019. Sebelum melakukan pemodelan regresi PIG, perlu dilakukan uji multikolinieritas. Uji multikolinieritas bertujuan untuk mengetahui apakah variabel-variabel prediktor berhubungan secara linier satu dengan lainnya (kolinieritas). Kriteria yang digunakan untuk memeriksa kolinieritas antar variabel-variabel prediktor, yaitu dengan melihat nilai *Variance Inflation Factors* (VIF) pada variabel-variabel prediktor. Berikut ini adalah nilai VIF untuk masing-masing variabel prediktor.

Tabel 2. Nilai VIF

Nilai VIF				
x_1	x_2	x_3	x_4	x_5
2.619	2.526	2.910	1.601	1.925

Nilai VIF masing-masing variabel prediktor pada Tabel 2 menunjukkan bahwa nilai kurang dari 10, sehingga antar variabel prediktor dikatakan tidak terdapat kolinieritas. Karena tidak terdapat kolinieritas antar variabel prediktor, maka variabel-variabel prediktor dapat dimasukkan kedalam model.

Estimasi Parameter Model Regresi Poisson dan Regresi Poisson Invers Gaussian

Dalam penelitian ini, data jumlah kematian bayi di provinsi Jawa Timur tahun 2019 akan dimodelkan dengan menggunakan dua metode, yaitu regresi poisson sebagai pembanding ketika kasus overdispersi diabaikan dan regresi PIG untuk kasus overdispersi. Berdasarkan hasil analisis data yang telah dilakukan, berikut diperoleh hasil estimasi untuk masing-masing model yang ditunjukkan oleh Tabel 3 dan Tabel 4.

Tabel 3. Estimasi Parameter dalam Model Regresi Poisson

Parameter	**Nilai Parameter**	**Standar *Error***	**Nilai t**	***P-Value***	**Keputusan**
$\hat{\beta}_0$	2.267	0.456	4.575	**0.000**	Signifikan
$\hat{\beta}_1$	0.019	0.007	2.696	**0.007**	Signifikan
$\hat{\beta}_2$	-0.061	0.005	-12.541	**0.000**	Signifikan
$\hat{\beta}_3$	0.048	0.007	6.714	**0.000**	Signifikan
$\hat{\beta}_4$	0.007	0.003	1.995	**0.046**	Signifikan
$\hat{\beta}_5$	0.004	0.002	2.549	**0.011**	Signifikan

Berdasarkan hasil estimasi parameter yang diperoleh pada Tabel 3 dan Tabel 4, maka didapatkan hasil estimasi model untuk model regresi poisson sebagaimana persamaan (1) dan regresi PIG seperti pada persamaan (2).

$$\hat{y} = \exp\left(2.267 + 0.019x_1 - 0.061x_2 + 0.048x_3 + 0.007x_4 + 0.004x_5\right), \tag{1}$$

$$\hat{y} = \exp\left(2.451 + 0.049x_1 - 0.078x_2 + 0.045x_3 - 0.002x_4 - 0.001x_5\right). \tag{2}$$

Tabel 4. Estimasi Parameter dalam Model Regresi PIG

Parameter	**Nilai Parameter**	**Standar *Error***	**Nilai t**	***P-Value***	**Keputusan**
$\hat{\beta}_0$	2.451	3.014	0.813	0.422	Tidak Signifikan
$\hat{\beta}_1$	0.049	0.042	1.164	0.253	Tidak Signifikan
$\hat{\beta}_2$	-0.078	0.029	-2.704	**0.011**	Signifikan
$\hat{\beta}_3$	0.045	0.043	1.030	0.311	Tidak Signifikan
$\hat{\beta}_4$	-0.002	0.017	-0.103	0.919	Tidak Signifikan
$\hat{\beta}_5$	-0.001	0.009	-0.109	0.914	Tidak Signifikan
$\hat{\tau}$	-0.741	0.281	-2.636	**0.013**	Signifikan

Nilai AIC yang dihasilkan untuk model regresi poisson adalah sebesar 1050 sedangkan nilai AIC yang dihasilkan untuk model regresi PIG sebesar 401.68. Model terbaik yang diperoleh jika dipandang dari nilai AIC terkecil adalah model regresi PIG, akan tetapi pada model regresi PIG parameter yang signifikan hanya 2 parameter. Sedangkan, pada model regresi poisson semua parameter dalam model signifikan. Dalam penelitian ini, model yang akan dipilih adalah model dengan nilai AIC terkecil, yaitu model regresi PIG.

Pemeriksaan Kasus Overdispersi

Pemeriksaan kasus overdispersi dilakukan untuk melihat apakah data jumlah kasus bayi mati di propinsi Jawa Timur pada tahun 2019 terdapat overdispersi atau tidak. Berikut diberikan hipotesis serta statistik uji yang digunakan.

$$H_0 : \operatorname{var}(y) = \mu$$

$$H_1 : \operatorname{var}(y) = \mu + \alpha . g(.).$$

Tabel 5. Uji Overdispersi

Nilai α	**Nilai** z	*P-Value*	VT
35.73	2.53	0.006	1395.751

Dengan menggunakan *package* AER pada *software* R, diperoleh nilai $\alpha = 35.73$ dengan nilai *p-value* sebesar 0.006 kurang dari tingkat signifikansi 5% sehingga H_0 ditolak. Dapat disimpulkan bahwa terdapat overdispersi pada data jumlah kematian bayi di provinsi Jawa Timur tahun 2019. Dengan cara yang lain, yaitu melihat nilai VT yang diperoleh berdasarkan persamaan (8) yang lebih besar dari 1. Sehingga metode yang tepat untuk melakukan pemodelan jumlah kematian bayi di provinsi Jawa Timur tahun 2019, yaitu regresi Poisson Inverse Gaussian (PIG).

Estimasi Parameter Model Regresi PIG

Selanjutnya, akan dilakukan estimasi parameter dengan menggunakan metode *backward* untuk model regresi PIG guna mendapatkan hasil model terbaik. Berikut diperoleh beberapa hasil kemungkinan model dalam regresi PIG yang ditunjukkan oleh Tabel 6.

Tabel 6. Estimasi Parameter Kemungkinan dalam Model Regresi PIG

Model	**Variabel dalam Model**	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\tau}$
M_1	x_1, x_2, x_3, x_4, x_5	2.451	0.049	-0.078	0.045	-0.002	-0.001	-0.741
M_2	x_2, x_3, x_4, x_5	400.208	-	-0.083	0.076	0.000	0.003	-0.704
M_3	x_1, x_3, x_4, x_5	0.676	0.051	-	-0.008	-0.022	0.013	-0.514
M_4	x_1, x_2, x_4, x_5	2.903	0.073	-0.066	-	0.000	-0.000	-0.697
M_5	x_1, x_2, x_3, x_5	2.774	0.044	-0.084	0.050	-	-0.001	-0.733
M_6	x_1, x_2, x_3, x_4	2.475	0.047	-0.076	0.044	-0.002	-	-0.741
M_7	x_1, x_2, x_3	2.747	0.042	-0.081	0.049	-	-	-0.735
M_8	x_1, x_2	2.965	0.073	-0.065	-	-	-	-0.697

Berdarkan hasil estimasi parameter pada Tabel 6, diperoleh estimasi model regresi PIG untuk beberapa kengkinan diantaranya.

$$M_1 = \hat{y} = \exp\left(2.451 + 0.049x_1 - 0.078x_2 + 0.045x_3 - 0.002x_4 - 0.001x_5\right),$$

$$M_2 = \hat{y} = \exp\left(400.208 + -0.083x_2 + 0.076x_3 - 0.000x_4 - 0.003x_5\right),$$

$$M_3 = \hat{y} = \exp\left(0.676 + 0.051x_1 - 0.008x_3 - 0.022x_4 - 0.013x_5\right),$$

$$M_4 = \hat{y} = \exp\left(2.903 + 0.073x_1 - 0.066x_2\right),$$

$$M_5 = \hat{y} = \exp\left(2.774 + 0.044x_1 - 0.084x_2 + 0.050x_3 - 0.001x_5\right),$$

$$M_6 = \hat{y} = \exp\left(2.475 + 0.047x_1 - 0.076x_2 + 0.044x_3 - 0.002x_4\right),$$

$$M_7 = \hat{y} = \exp\left(2.747 + 0.042x_1 - 0.081x_2 + 0.049x_3\right),$$

$$M_8 = \hat{y} = \exp\left(2.965 + 0.073x_1 - 0.065x_2\right).$$

Pengujian Parameter Secara Simultan

Pengujian parameter secara simultan dilakukan untuk masing-masing kemungkinan model dalam regresi PIG. Berikut diberikan nilai statsitik uji G sebagaimana persamaan

$$G = -2\ln\left(\frac{L(\hat{\omega})}{L(\hat{\Omega})}\right)$$

$$= 2\left(\ln L(\hat{\omega}) - \ln L(\hat{\Omega})\right)$$. yang ditunjukkan oleh Tabel 7.

Tabel 7. Pengujian Parameter Secara Simultan

Model	**Variabel dalam Model**	**Nilai Statistik** G	**Derajat Bebas**	**Nilai** χ^2_{Tabel}	**Keputusan**
M_1	x_1, x_2, x_3, x_4, x_5	389.67	32	46.194	Tolak H_0
M_2	x_2, x_3, x_4, x_5	390.85	33	47.400	Tolak H_0
M_3	x_1, x_3, x_4, x_5	396.87	33	47.400	Tolak H_0
M_4	x_1, x_2, x_4, x_5	390.86	33	47.400	Tolak H_0
M_5	x_1, x_2, x_3, x_5	389.68	33	47.400	Tolak H_0
M_6	x_1, x_2, x_3, x_4	389.68	33	47.400	Tolak H_0
M_7	x_1, x_2, x_3	389.70	34	48.602	Tolak H_0
M_8	x_1, x_2	390.86	35	49.802	Tolak H_0

Berdasarkan Tabel 7 di atas, dapat dilihat bahwa nilai statistik uji G untuk masing-masing model baik itu model M_1, M_2 sampai dengan model M_8 menghasilkan nilai statistik uji G yang lebih besar dari nilai χ^2_{Tabel}, sehingga dapat diambil putusan tolak H_0. Karena H_0, maka dapa diambil kesimpulan bahwa semua parameter signifikan secara simultan untuk semua model kemungkinan dalam regresi PIG.

Pengujian Parameter Secara Parsial dan Pemilihan Model Terbaik

Untuk melihat pengaruh masing-masing parameter terhadap model regresi PIG, maka dilakukan pengujian parameter secara parsial. Pengujian parameter secara parsial hanya diberikan untuk model regresi PIG yang memiliki nilai AIC terkecil. Nilai AIC terkecil untuk masing-masing kemungkinan model dalam regresi PIG diberikan oleh Tabel 8. Dalam penelitian ini, model terbaik yang yang dipilih adalah model dengan nilai AIC terkecil. Berdasarkan Tabel 8 di atas, model dengan nilai AIC terkecil adalah model M_8 dengan nilai AIC, yaitu sebesar 398.86, sehingga model terbaik yang dipilih adalah model M_8. Selanjutnya pengujian parameter secara parsial untuk M_8 diberikan oleh Tabel 9 dengan hipotesis pengujian.

Hipotesis untuk parameter model.

$H_0 : \beta_l = 0$

$H_1 : \beta_l \neq 0$ dengan $l = 1, 2, ..., k$.

Hipotesis untuk parameter dispersi.

$H_0 : \tau = 0$

$H_1 : \tau \neq 0$.

Tabel 8. Nilai AIC untuk Masing-masing Model

Model	**Variabel dalam Model**	Nilai AIC
M_1	x_1, x_2, x_3, x_4, x_5	403.67
M_2	x_2, x_3, x_4, x_5	402.85
M_3	x_1, x_3, x_4, x_5	408.87
M_4	x_1, x_2, x_4, x_5	402.86
M_5	x_1, x_2, x_3, x_5	401.68
M_6	x_1, x_2, x_3, x_4	401.68
M_7	x_1, x_2, x_3	399.70
M_8	x_1, x_2	**398.86**

Tabel 9. Pengujian Parameter secara Parsial M_8

Parameter	**Nilai Parameter**	**Standar *Error***	**Nilai t**	***P-Value***	Keputusan
$\hat{\beta}_0$	2.965	3.065	0.967	0.340	Gagal Tolak H_0
$\hat{\beta}_1$	0.073	0.028	2.555	**0.015**	Tolak H_0
$\hat{\beta}_2$	-0.065	0.019	-3.517	**0.001**	Tolak H_0
$\hat{\tau}$	-0.697	0.282	-2.469	**0.019**	Tolak H_0

Dengan melihat nilai *p-value* pada Tabel 9 di atas, maka dapat dilihat bahwa nilai *p-value* untuk parameter $\hat{\beta}_1$ dan $\hat{\beta}_2$ kurang dari nilai signifikansi $5\% = 0.05$ yang artinya bahwa varibel x_1 dan x_2 signifikan berpengaruh terhadap jumlah kematian bayi di Provinsi Jawa Timur tahun 2019. Begitupula dengan parameter $\hat{\tau}$ yang signifikan dan berrati bahwa data jumlah kematian bayi di provinsi Jawa Timur tahun 2019 mengandung overdispersi, sehingga metode yang tepat untuk pemodelan kasus tersebut adalah regresi PIG.

Interpretasi Model

Interpretasi model regresi PIG dilakukan dengan melihat masing-masing koefien dalam model yang memiliki nilai AIC terkecil, yaitu model M_8 yang diberikan oleh persamaan (3) berikut.

$$\mathrm{M}_8 = \hat{y} = \exp\left(2.965 + 0.073x_1 - 0.065x_2\right). \tag{3}$$

Berdasarkan model pada persamaan (3) di atas, maka dapat dilakukan interpretasi untuk masing-masing variabel dengan menghiting nilai $\exp\left(\hat{\beta}_j\right)$, $j = 1, 2$. Untuk variabel Persentase Kunjungan Ibu Hamil dengan K1 $\left(x_1\right)$ meiliki nilai $\exp\left(\hat{\beta}_1\right) = \exp\left(0.073\right) = 1.076$ yang berati bahwa untuk kenaikan satu persen variabel Persentase Kunjungan Ibu Hamil dengan K1 $\left(x_1\right)$ akan menaikan kasus jumlah kematian bayi di provinsi Jawa Timur tahun 2019. Hal ini berbanding terbalik dengan teori yang sebenarnya bahwa resiko jumlah kematian bayi seharusnya menurun jika Persentase Kunjungan Ibu Hamil dengan K1 $\left(x_1\right)$ semakin tinggi. Akan tetapi, untuk variabel Persentase Kunjungan Ibu Hamil dengan K4 $\left(x_2\right)$ memiliki nilai $\exp\left(\hat{\beta}_1\right) = \exp\left(0.065\right) = 1.067$ dan koefisien bernilai negatif yang berarti bahwa setiap kenaikan satu persen Persentase Kunjungan Ibu Hamil dengan K4 $\left(x_2\right)$ akan menurunkan kasus jumlah kematian bayi di provinsi Jawa Timur tahun 2019.

4. KESIMPULAN

Berdasarkan analisis data yang telah dilakukan, maka dapat diambil kesimpulan bahwa model regresi PIG memberikan hasil model lebih baik dari model regresi Poisson dalam mengatasi data yang memiliki kasus overdispersi dengan melihat nilai AIC yang terkecil. Selanjutnya, dari beberapa kemungkinan model regresi PIG yang terbentuk, model terbaik yang diperoleh adalah model dengan nilai AIC terkecil, yaitu model M_8. Dengan variabel prediktor dalam model M_8 adalah Persentase Kunjungan Ibu Hamil dengan K1 $\left(x_1\right)$ dan Persentase Kunjungan Ibu Hamil dengan K4 $\left(x_2\right)$. Berdasarkan model M_8, maka dapat disimpulkan bahwa resiko jumlah kematian bayi di Provinsi Jawa Timur tahun 2019 akan mengalami kenaikan jika kunjungan ibu hamil hanya dengan K1 sedangkan jika kunjungan ibu hamil dilakukan juga untuk K4 maka akan menurunkan resiko jumlah kematian bayi di provinsi Jawa Timur tahun 2019. Ini berarti, resiko kematian bayi dipengaruhi kuantitas kunjungan ibu hamil ke posyandu. Semakin sering ibu hamil melakukan kunjungan ke posyandu untuk memeriksa kehamilannya, maka resiko kematian bayi di Provinsi Jawa Timur semakin rendah.

DAFTAR PUSTAKA

Cameron, A. C dan Trivedi, P. K. (1998), *Regression Analysis of Count Data*, Cambridge University Press, USA.

Consul, P. C. dan Famoye, F. (1992), "Generalized Poisson Regression Model", *Communications in Statistics - Theory and Methods*, Vol. 21, No. 1, hal. 89-109.

Déniz, E. G., Ghitany, M. E. dan Gupta, R. C. (2014), "Poisson-Mixed Inverse Gaussian Regression Model and Its Application", *Communications in Statistics - Simulation and Computation*, DOI: 10.1080/03610918.2014.925924.

Hilbe, J.M. (2007), *Negative Binomial Regression*, 1st edition, Cambridge University, Press., New York.

Hu, M. C., Pavlicova, M. dan Nunes, E. V. (2011), "Zero-Inflated and Hurdle Models of Count Data with Extra Zeros: Examples from an HIV-Risk Reduction Intervention Trial", *The American Journal of Drug and Alcohol Abuse*, Vol. 37, hal. 367-375.

Jong, P. dan Heller, G. Z. (2008), *Generalized Linear Models for Insurance Data*, 1st edition, Cambridge University, Press., New York.

Karlis, D. dan Nikoloulopoulos, E. (2005), "Mixed Poisson Distribution", *International Statistical Review*, Vol. 73, No. 1, hal. 35-58.

McCullagh, P. dan Nelder, J. A. (1989), *Generalized Linear Models 2nd Edition*. Chapman and Hall, London.

Ozdemir, T. dan Eyduran, E. (2005), "Comparison of Chi – Square and Likelihood Ratio Chi – Square Tests: Power of Test", *Journal of Applied Sciences Research, University of Yuzuncu*, Hal. 242-244.

Shoukri, M. M., Asyali, M. H., Vandorp, R. dan Kelton, R. (2004), "The Poisson Inverse Gaussian Regression Model in the Analysis of Clustered Counts Data", *Journal of Data Science*, Vol. 2, No. 2, Hal. 17-32.

Stasinopoulus, D. M. dan Rigby, R. A. (2007), "Generalized Additive Models for Location Scale and Shape", *Journal of Statistical Software*, Vol. 23, Hal.1-46.

Willmot, G. E. (1987), "The Poisson-Inverse Gaussian Distribution as An Alternative to the Negative Binomial", *Scandinavian Actuarial Journal*, Vol. 3, No. 4, hal. 113-127.

Zha, L., Lord, D. dan Zou, Y. (2014), "The Poisson Inverse Gaussian (PIG) Generalized Linear Regression Model for Analyzing Motor Vehicle Crash Data", *Journal of Transportation Safety and Security*, DOI:20.2080/19439962.2014.977502.