

PENERAPAN REGRESI *ZERO INFLATED GENERALIZED POISSON* (ZIGP) PADA DATA OVERDISPERSION

Ni Wayan Amanda Dewi Sulistyaningsih^{1§}, I Komang Gde Sukarsa², Ni Luh Putu Suciptawati³

¹Program Studi Matematika, Fakultas MIPA – Universitas Udayana [Email: dewiamanda32@gmail.com]

²Program Studi Matematika, Fakultas MIPA – Universitas Udayana [Email: gedesukarsa@unud.ac.id]

³Program Studi Matematika, Fakultas MIPA – Universitas Udayana [Email: suciptawati@unud.ac.id]

[§]Corresponding Author

ABSTRACT

Zero Inflated Generalized Poisson (ZIGP) is a regression model used to analyze Poisson distributed discrete data which contains mostly zero and tends to experience overdispersion (varians value greater than the mean value). The purpose of this research is to find out the best model and the factors which influence the maternal mortality in Bali Province in year 2016 by using ZIGP regression model. The data used in this research was data from health profile Bali Province with the object totally 57 district rate data has proportion of zeros value more than 50% on the response variable. The analysis result of ZIGP data on maternal mortality cannot modeled using the ZIGP so ZIGP regression model became ZIP model. The best model which resulted from ZIP regression got one free variable which have significant impact towards the total number of maternal mortality. This significant variabel is the percentage of mother did visiting to KI(X_1).

Keywords: *Maternal mortality, Overdispersion, Zero Inflated Generalized Poisson (ZIGP), Zero Inflated Poisson (ZIP).*

1. PENDAHULUAN

Analisis regresi adalah salah satu model yang digunakan untuk mengetahui hubungan sebab akibat antar variabel, yaitu variabel bebas (X) dengan variabel respons (Y). Variabel respons dan variabel bebas memiliki hubungan yang dapat dinyatakan dalam persamaan matematis yaitu persamaan regresi. Persamaan regresi dapat berbentuk persamaan regresi linear maupun persamaan regresi non-linear. Persamaan regresi linear digunakan untuk menganalisis variabel respons yang berupa peubah acak kontinu dan mengikuti distribusi normal. Namun banyak ditemukan variabel respons yang tidak berdistribusi normal dan tidak linear dalam parameter. Untuk mengatasi hal tersebut dikembangkan *generalized linear model* (GLM). GLM digunakan sebagai perluasan model regresi umum dengan variabel responsnya berdistribusi keluarga eksponensial, meliputi distribusi normal, binomial, Poisson, binomial negatif, eksponensial, gamma, dan

invers normal (Myers *et al.*, 2010).

Regresi *Poisson* digunakan apabila variabel respons adalah peubah acak diskret. Analisis data menggunakan model regresi *Poisson* harus memenuhi asumsi seperti nilai varians sama dengan rata-rata dari variabel respons tersebut atau *equidispersi* (Myers *et al.*, 2010). Namun, dalam kenyataannya sering terjadi pelanggaran asumsi tersebut, yaitu nilai variansnya lebih besar dari nilai rata-rata yang dinamakan *overdispersi*. *Overdispersi* pada regresi *Poisson* dapat dilihat dari nilai taksiran dispersi yaitu nilai devians dan *pearson chi square* dibagi dengan derajat bebasnya bernilai lebih besar dari satu. Jika terjadi *overdispersi* pada data, maka regresi *Poisson* kurang akurat digunakan untuk analisis karena akan berdampak pada nilai *standard error* menjadi lebih kecil dari nilai sesungguhnya (*underestimate*), sehingga kesimpulan yang diperoleh menjadi tidak valid (McCullagh & Nelder, 1989).

Salah satu penyebab terjadinya overdispersi yaitu adanya nilai nol yang berlebih pada variabel respon (*excess zeros*). Untuk mengatasi masalah overdispersi tersebut dapat digunakan beberapa model regresi antara lain model regresi *zero inflated poisson* (ZIP) dengan proporsi nol sekitar 63,7%, dan *zero inflated generalized poisson* (ZIGP) dengan proporsi nol 65,7%. Model regresi *zero inflated generalized poisson* (ZIGP) merupakan perluasan dari model regresi *Poisson* dan merupakan gabungan dari model regresi *zero inflated poisson* (ZIP) dan model regresi *generalized poisson* (GP) (Famoye dan Singh, 2006).

Studi kasus yang digunakan pada penelitian ini adalah Angka Kematian Ibu Hamil di Provinsi Bali tahun 2016. Data menunjukkan terjadi peristiwa kematian ibu hamil tahun 2016 sebanyak 18 ibu di 57 Kecamatan (Dinas Kesehatan Provinsi Bali, 2016). Tujuan dari penelitian ini adalah untuk menentukan model terbaik ZIGP serta untuk mengetahui faktor-faktor yang mempengaruhi kematian ibu hamil.

Secara umum model regresi *Poisson* dapat ditulis sebagai berikut (Myers *et al.*, 2010)

$$Y_i = \mu_i + \varepsilon_i (i = 1, 2, 3, \dots, n) \quad (1.1)$$

dengan Y_i adalah jumlah kejadian dan μ_i adalah *mean* jumlah kejadian dalam periode t_i, μ_i .

Overdispersi pada data dapat dilihat dari taksiran dispersi yaitu nilai devians dan *pearson chi square* dibagi nilai derajat bebas bernilai lebih dari satu. Model yang dapat digunakan untuk mengatasi overdispersi di antaranya adalah model ZIGP dan ZIP.

Model ZIP adalah model campuran sederhana untuk data diskret dengan banyak peristiwa nol (Lambert, 1992). Fungsi peluang dari model ZIP adalah sebagai berikut:

$$P(Y_i = y_i) = \begin{cases} \omega_i + (1 - \omega_i)e^{-\mu_i}, & \text{untuk } y_i = 0 \\ (1 - \omega_i) \frac{e^{-\mu_i} \mu_i^{y_i}}{y_i!}, & \text{untuk } y_i = 1, 2, \dots; 0 < \omega_i < 1. \end{cases} \quad (1.2)$$

Dari fungsi peluang model regresi ZIP, diperoleh rata-rata dan varians sebagai berikut (Famoye dan Singh, 2006):

$$E(Y) = (1 - \omega_i)\mu, \quad (1.3)$$

$$Var(Y) = \mu(1 - \omega)(1 + \mu\omega). \quad (1.4)$$

model hubungan untuk μ dan ω adalah sebagai

berikut:

$$\ln(\mu) = X\beta \text{ dan } \text{logit}(\omega) = \log\left(\frac{\omega}{1-\omega}\right) = X\gamma. \quad (1.5)$$

Penaksiran parameter model regresi ZIP dilakukan dengan metode penaksir kemungkinan maksimum (Jansakul & Hinde, 2002). Penaksiran parameter dengan metode penaksir kemungkinan maksimum dihitung dengan memaksimalkan fungsi log *likelihood*-nya. Penjumlahan fungsi loglikelihood tidak dapat diselesaikan dengan metode numerik biasa, oleh karena itu dapat digunakan algoritma EM (*Expectation Maximization*).

Pengujian parameter regresi ZIP secara parsial dapat dilakukan dengan dua cara, yaitu pengujian parameter model log dan pengujian parameter model logit. Statistik yang digunakan adalah dengan menggunakan uji Wald (Myers *et al.*, 2010) sebagai berikut:

$$W = \left(\frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \right) \quad (1.6)$$

Untuk kriteria statistik ujinya adalah H_0 ditolak jika $W > X^2_{(db, \alpha)}$. Pemilihan model terbaik regresi ZIP dapat dilihat dari nilai AIC terkecil (Ismail & Jemain, 2005).

Regresi zero inflated generalized poisson (ZIGP) merupakan metode untuk menangani *overdispersion* dengan proporsi data bernilai nol adalah sekitar 65,7% (Famoye dan Singh, 2006). Model regresi ZIGP sebagai berikut.

$$P(Y_i = y_i | x_i, z_i) = \begin{cases} \pi_i + (1 - \pi_i)f(\mu_i, \omega, y_i), & y_i = 0 \\ (1 - \pi_i)f(\mu_i, \omega, y_i), & y_i > 0 \end{cases} \quad (1.7)$$

Dengan $f(\mu_i, \omega, y_i), y_i = 0, 1, 2, \dots$ adalah model regresi GP dengan $0 < \pi_i < 1$. Dalam fungsi $\mu_i = \mu_i(x_i)$ dan $\pi_i = \pi_i(z_i)$ memenuhi fungsi link sebagai berikut (Famoye dan Singh, 2006).

$$\log(\mu_i) = \sum_{j=1}^k x_{ij}\beta = \mathbf{x}^T\beta$$

dan

$$\text{logit}(\pi_i) = \log(\pi_i[1 - \pi_i])^{-1} = \sum_{j=1}^m z_{ij}\delta$$

Varians dan *mean* dari model ZIGP sebagai berikut:

$$E(y_i | x_i) = (1 - \pi_i)\mu_i(x_i) \quad (1.8)$$

dan

$$\begin{aligned} \text{Var}(y_i|x_i) &= (1 - \pi_i)[\mu_i^2 + \mu_i(1 + \alpha\mu_i)^2] \\ &\quad - (1 - \pi_i)^2\mu_i^2 \\ &= E(y_i|x_i)[(1 + \alpha\mu_i)^2 + \pi_i\mu_i] \end{aligned} \quad (1.9)$$

Model regresi ZIGP akan menjadi model regresi GP jika $\pi_i = 0$ dan akan menjadi model regresi ZIP jika $\omega = 0$. Jika matriks kovariat yang sama mempengaruhi π_i dan μ_i maka dapat ditulis bahwa π_i adalah fungsi dari μ_i dengan fungsi sebagai berikut.

$$\text{logit}(\pi_i) = \log\left(\frac{\pi_i}{1-\pi_i}\right) = -\tau \sum_{j=1}^k x_{ij}\beta_j = -\tau \mathbf{x}^T \boldsymbol{\beta} \quad (1.10)$$

Model regresi ZIGP dengan *log link* untuk μ_i dan untuk *logit link* untuk π_i akan dikembangkan dengan ZIGP(τ). Jika $\tau > 0$ maka kemungkinan *zero state* terjadi kecil dan jika $\tau < 0$ maka kemungkinan *zero state* terjadi lebih besar. Ketika $\tau = 0$ maka model ZIGP(τ) akan menjadi model ZIP(τ).

Penaksir parameter ZIGP digunakan *maximum likelihood estimator* (MLE) adalah salah satu metode penaksiran parameter yang dapat digunakan untuk menaksir parameter suatu model yang diketahui distribusinya. Metode MLE dilakukan dengan cara memaksimumkan fungsi *likelihood*. Fungsi *likelihood* regresi ZIGP sebagai berikut.

$$L(\boldsymbol{\beta}) = \begin{cases} \prod_{i=1}^n \frac{1}{1 + \mu_i^{-\tau}} + [\mu_i^{-\tau} + \exp\left(\frac{-\mu_i}{1 + \omega\mu_i}\right)] & , y_i = 0 \\ \prod_{i=1}^n \frac{1}{1 + \mu_i^{-\tau}} \left(\frac{-\mu_i}{1 + \omega\mu_i}\right)^{y_i} \frac{(1 + \omega y_i)^{y_i - 1}}{y_i} \exp\left[\frac{-\mu_i(1 + \omega y_i)}{1 + \omega\mu_i}\right] & , y_i > 0. \end{cases}$$

Persamaan tersebut dimaksimumkan dengan menggunakan maksimum *likelihood* untuk koefisien regresi dalam $\hat{\boldsymbol{\beta}}$. Pada kasus tertentu cara tersebut tidak menghasilkan suatu solusi yang eksplisit sehingga alternatif lain yang dapat digunakan untuk menaksir parameternya adalah dengan algoritma Newton Raphson. Menurut (Famoye dan Singh, 2006) metode Newton Raphson merupakan metode numerik untuk menyelesaikan persamaan non-linier secara iteratif. Dasar dari model ini adalah pendekatan deret Taylor linier.

Menurut Famoye dan Singh (2006), model regresi ZIGP akan menjadi model regresi ZIP ketika parameter *dispersion* $\omega = 0$. Oleh karena itu untuk melihat kesesuaian model

ZIGP, dilakukan pengujian seperti berikut.

H_0 : Model ZIGP (τ) tidak sesuai

H_1 : Model ZIGP (τ) sesuai

Untuk mengetahui uji kesesuaian model menggunakan statistik uji sebagai berikut:

$$D = -2 \ln \left[\frac{L(\hat{\boldsymbol{\beta}}, \hat{\tau})}{L(\hat{\boldsymbol{\beta}}, \hat{\tau}, \hat{\omega})} \right] \quad (1.11)$$

Penolakan H_0 menunjukkan bahwa model regresi ZIP tidak tepat digunakan, sehingga dalam situasi seperti ini model regresi ZIGP lebih tepat digunakan dibandingkan model regresi ZIP.

Pengujian signifikan regresi ZIGP ada dua yaitu pengujian parameter secara serentak dan pengujian secara individu.

1. Pengujian parameter secara serentak

Pengujian parameter secara serentak dilakukan untuk mengetahui faktor yang berpengaruh signifikan terhadap model. Berikut ini merupakan hipotesis uji parameter secara serentak model regresi ZIGP.

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

$$H_1: \text{minimal ada satu parameter } \beta_k \neq 0, k = 1, 2, \dots, k$$

dengan statistik uji sebagai berikut.

$$G^2 = -2 \ln \left(\frac{l_0}{l_1} \right) \quad (1.12)$$

Dengan l_0 dan l_1 merupakan lambang *maximum log likelihood* di bawah H_0 dan secara keseluruhan ($H_0 \cup H_1$). Daerah penolakan pada pengujian parameter secara serentak adalah tolak H_0 jika $G^2 > X^2_{(p,\alpha)}$ atau jika $P - \text{value} < \alpha$.

2. Pengujian Parameter Secara Individu

Uji Wald merupakan salah satu pengujian yang dilakukan untuk menguji signifikansi parameter regresi secara individu. Berikut ini merupakan hipotesis uji parameter secara individu model regresi ZIGP.

$$H_0: \beta_k = 0$$

$$H_1: \beta_k \neq 0, k = 0, 1, 2, \dots, k$$

$$W = \frac{(\hat{\beta}_k)^2}{SE(\beta_k)^2} \quad (1.13)$$

Dengan $\hat{\beta}$ merupakan maksimum *likelihood estimator* parameter β . Daerah

penolakan pada pengujian ini adalah tolak H_0 jika $W_{hitung} > X^2_{(1,\alpha)}$ atau $Pvalue < \alpha$.

Pemilihan model terbaik (*Goodness of Fit*) Salah satu pemilihan model terbaik adalah dengan metode *Akaike's Information Criterion* (AIC). Nilai AIC dapat dihitung dengan menggunakan rumus berikut.

AIC (*Akaike's information Criterion*)

$$AIC = -2\log L(\hat{\beta}) + 2k \quad (1.14)$$

Dengan $L(\hat{\beta})$ merupakan nilai *likelihood* dan k merupakan jumlah parameter. Kriteria pemilihan model terbaik jika didapatkan nilai terkecil atau mendekati nol (Ismail & Jemain, 2005).

2. METODE PENELITIAN

Jenis dan sumber data Penelitian ini menggunakan data Angka Kematian Ibu Hamil pada tahun 2016. Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari Dinas Kesehatan Provinsi Bali dari seluruh Kabupaten/kota di Provinsi Bali yang telah terbagi menjadi 57 Kecamatan di Provinsi Bali.

Variabel yang digunakan pada penelitian ini adalah jumlah kematian ibu hamil (Y), Persentase ibu hamil yang melaksanakan kunjungan K1 (X_1), Persentase ibu hamil yang melaksanakan kunjungan K4 (X_2), Persentase ibu hamil yang mendapatkan tablet tambahan darah Fe1 (X_3), Persentase ibu hamil yang mendapatkan tablet tambahan darah Fe3 (X_4), Persentase ibu hamil mendapatkan imunisasi TT5 (X_5). Berikut adalah langkah-langkah yang dilakukan untuk mencapai tujuan dalam penelitian ini.

1. Melakukan pengumpulan dan penyajian data, serta mendeskripsikan data penelitian menggunakan statistik deskriptif.
2. Melakukan uji overdispersi. Uji overdispersi pada data dapat diukur dengan taksiran dispersi, apabila nilai devians dan pearson *chi-square* dibagi dengan derajat bebasnya bernilai lebih dari 1, maka data mengalami overdispersi. Jika data mengalami overdispersi maka dilanjutkan dengan langkah 3. Namun jika data tidak

mengalami overdispersi maka data dapat dianalisis menggunakan regresi Poisson.

3. Menaksir parameter model regresi ZIGP . Untuk menaksir parameter pada regresi ZIGP menggunakan algoritma Newton-Raphson.
4. Melakukan uji kesesuaian model regresi ZIGP. Menentukan apakah model yang diperoleh sesuai atau tidak dengan model regresi ZIGP. Dengan hipotesis-hipotesis sebagai berikut
 H_0 : Model ZIGP(τ) tidak sesuai
 H_1 : Model ZIGP(τ) sesuai
5. Melakukan uji signifikan pada model regresi ZIGP, dengan menggunakan uji parameter secara serentak dan pengujian parameter secara individu.
6. Menentukan model terbaik dari regresi ZIGP dengan menggunakan model *Akaike's Information Criterion* (AIC), $AIC = -2\log L(\hat{\beta}) + 2k$.

Langkah-langkah tersebut dapat dibantu dengan menggunakan *software* Minitab 17 dan *Software RStudio*.

3. HASIL DAN PEMBAHASAN

3.1 Deskripsi data

Deskripsi jumlah kematian ibu hamil (Y) memiliki nilai minimum 0, nilai maksimum 3, dan banyaknya pengamatan adalah 57. Jumlah total kematian ibu hamil pada tahun 2016 sebanyak 18 orang. Jumlah data yang bernilai nol (tidak ada ibu hamil yang meninggal) sebanyak 46 Kecamatan. Terdapat faktor-faktor yang diduga memengaruhi kasus Kematian Ibu Hamil di Provinsi Bali yaitu Deskripsi data pada variabel bebasnya dapat dilihat dari Tabel 1

Tabel 1. Deskripsi Data Variabel Bebas (X_1, X_2, X_3, X_4, X_5)

Variabel bebas	Min	Maks	Rata-rata	Varians
X1	70,0%	100%	91,54%	55,70%
X2	70,30%	100%	89,06%	93,37%
X3	64,50%	100%	93,21%	67,27%
X4	62,60%	100%	90,83%	71,07%
X5	25,20%	100%	70,58%	423,01%

Sumber: data diolah (2018)

Berdasarkan statistik deskriptif pada Tabel 1. Berturut-turut diperoleh nilai rata-rata ibu yang melakukan kunjungan K1, K4, mendapatkan tablet tambah darah Fe1, Fe3 dan ibu mendapatkan imunisasi TT5 yaitu sebesar 91,54%, 89,06%, 93,21%, 90,83%, dan 70,58% yang menunjukkan sebagian besar ibu hamil di provinsi Bali telah melaksanakan kunjungan K1, K4, mendapatkan tablet tambah darah Fe1, Fe3 dan ibu mendapatkan imunisasi TT5. Di sisi lain, nilai varians yang mendekati 100% menunjukkan bahwa terdapat perbedaan yang signifikan antara nilai rata-rata dan banyaknya ibu hamil yang melaksanakan kunjungan K1, K4, mendapatkan tablet tambah darah Fe1, Fe3 dan ibu mendapatkan imunisasi TT5. Hal ini menunjukkan ketimpangan yang cukup signifikan antar Kecamatan di provinsi Bali dalam hal ibu hamil yang melakukan kunjungan K1, K4, mendapatkan tablet tambah darah Fe1, Fe3 dan ibu mendapatkan imunisasi TT5.

Nilai varians tertinggi sebesar 93,21% yaitu ibu hamil melaksanakan kunjungan K4 (X_2) menunjukkan bahwa ketimpangan tertinggi dari nilai rata-rata terjadi pada ibu hamil yang melakukan kunjungan K4. Hal ini berarti sebagian besar Kecamatan di Provinsi Bali dengan ibu hamil yang melakukan kunjungan K4 jauh dari nilai rata-rata 89,06%.

Menurut penelitian Fitriyeni (2015) dalam skripsi Indrian (2012) yang menyatakan tinggi rendahnya ibu hamil melaksanakan kunjungan K1, K4, mendapatkan tablet tambah darah Fe1, Fe3 dan ibu mendapatkan imunisasi TT5 dipengaruhi oleh faktor status ekonomi, tingkat pendidikan, dukungan keluarga dan sulitnya akses terhadap pelayanan kesehatan.

3.2 Pemeriksaan Uji Overdispersi

Taksiran dispersi diukur dari nilai devians dan nilai pearson *chi-square* dibagi derajat bebas bernilai lebih dari satu. Data jumlah kematian ibu hamil di Provinsi Bali tahun 2016 memiliki nilai devians dibagi derajat bebasnya sebesar 1,106269 dan nilai pearson *chi square* dibagi derajat bebasnya sebesar 1,440812.

Hasil tersebut menunjukkan nilai devians dan pearson *chi square* yang dibagi derajat bebasnya bernilai lebih besar dari satu, sehingga disimpulkan data mengalami overdispersi.

3.3 Model Regresi Zero Inflated Generalized Poisson (ZIGP)

Pada kasus kematian ibu hamil variabel bebasnya yaitu persentase ibu hamil melaksanakan kunjungan K1 (X_1), persentase ibu hamil melaksanakan kunjungan K4 (X_2), Persentase ibu hamil yang mendapatkan tablet tambahan darah Fe1 (X_3), Persentase ibu hamil yang mendapatkan tablet tambahan darah Fe3 (X_4), dan Persentase ibu hamil mendapatkan imunisasi TT5 (X_5).

Sebelum membuat model, maka perlu dilakukan penaksiran parameter, pengujian kesesuaian model, dan pengujian parameter secara serentak dan secara individu. Estimasi parameter model regresi *zero inflated generalized Poisson* dengan bantuan *software RStudio* dapat dilihat pada Tabel 2.

Berdasarkan Tabel 2. dapat dilihat bahwa hasil analisis regresi *Zero Inflated Generalized Poisson (ZIGP)* menunjukkan bahwa tidak ada variabel yang signifikan terhadap tingkat kematian ibu hamil, sehingga ada indikasi model ZIGP tidak sesuai. Langkah selanjutnya dilakukan pengujian kesesuaian model ZIGP.

3.4 Pengujian Kesesuaian Model ZIGP

Untuk menguji kecocokan regresi *Zero Inflated Generalized Poisson (ZIGP)* dengan data, digunakan statistik uji Devians dan Pearson *chi-square* dengan hipotesis sebagai berikut:

H_0 : Model ZIGP (τ) tidak sesuai

H_1 : Model ZIGP (τ) sesuai.

Tingkat signifikan yang digunakan adalah 0,05, H_0 akan ditolak jika nilai devians lebih besar dari tabel *chi-square* ($D > \chi^2_{(0,05;51)} = 68,67$).

Table 2. Nilai Estimasi Parameter Model Regresi *Zero Inflated Generalized Poisson* dengan Seluruh Variabel Bebas

Koefisien	Variabel	Estimasi Parameter	z value	P-value
β_0	Intercept	8,16329	1,9984	0,0457
β_1	Kunjungan K1	-0,06456	- 1,7855	0,0742
β_2	Kunjungan K4	-0,02924	- 0,8637	0,3877
β_3	Tambah Darah Fe1	-0,02133	- 0,3481	0,7278
β_4	Tambah Darah Fe3	-0,00304	0,0523	0,9583
β_5	Imunisasi TT5	0,01573	0,9504	0,3419
Log-likelihood	Log-likelihood	-36,1		
AIC	AIC	86		
Omega	Omega	0,00000		

Sumber: data diolah (2018)

Berdasarkan luaran *software Rstudio* diperoleh devians nilai devians (D) = 61,95108 < 68,67, sehingga terima H_0 Oleh karena itu model ZIGP tidak sesuai. Dengan demikian model ZIGP akan menjadi model *Zero Inflated Poisson* (ZIP) (Famoye and Singh, 2006).

3.5 Analisis Regresi Zero Inflated Poisson (ZIP)

Berdasarkan analisis regresi *Zero Inflated Generalized Poisson* (ZIGP) disimpulkan bahwa data jumlah kematian ibu hamil (Y) tidak dapat dimodelkan dengan menggunakan model regresi *Zero Inflated Generalized Poisson* (ZIGP). Dengan demikian model ZIGP akan tereduksi menjadi model *Zero Inflated Poisson* (ZIP). Analisis regresi *Zero Inflated Poisson* (ZIP) dapat digunakan untuk menangani kasus *zero inflation* pada regresi Poisson jika asumsi overdispersi tidak terpenuhi. Langkah selanjutnya adalah menganalisis model regresi ZIP dengan bantuan *software Rstudio*.

Analisis regresi dengan ZIP akan menghasilkan dua model yakni *count model* untuk memprediksi data cacah dan *zero inflated model* untuk memprediksi probabilitas terjadinya *zero response* pada data. Pemodelan regresi ZIP pada kasus kematian ibu hamil di-

lakukan dengan memodelkan dari keseluruhan kombinasi variabel bebas dan dari model-model tersebut dipilih model yang signifikan dan memiliki nilai AIC terkecil.

Berdasarkan hasil seluruh kombinasi model dari variabel bebas diperoleh model yang memiliki AIC yang paling kecil. Model terbaik yang digunakan ditentukan dengan mempertimbangkan model yang signifikan dan memiliki AIC terkecil. berikut hasil estimasi parameter model regresi ZIP dengan bantuan *software Rstudio* pada Tabel 3.

Tabel 3. Estimasi Parameter Model Regresi ZIP

Variabel	Estimasi	Standar Error	z value	P-value
Count model				
Intercept	-3,61284	3,46884	-1,042	0,298
K1(X1)	-0,04217	0,04038	1,044	0,029
Zero inflated model				
Intercept	-26,5558	13,9174	-1,908	0,056
K1(X1)	-0,3021	0,1562	2,011	0,044
Log likelihood	-33,57			
AIC	75,13			

Sumber: data diolah (2018)

Berdasarkan Tabel 3. dapat disimpulkan bahwa kematian ibu hamil pada *count model* dan *zero inflated* dipengaruhi oleh variabel cakupan pemeriksaan kehamilan K1 X_1 yang berpengaruh signifikan. Model regresi ZIP pada kasus kematian ibu hamil menghasilkan nilai log-likelihood -33,57. model log dan logit regresi ZIP pada kasus kematian ibu hamil diatas dituliskan sebagai berikut.

$$\log(\mu_i) = -3,61284 - 0,04217X_1$$

$$\mu_i = \exp(-3,61284 - 0,04217K_1)$$

dan

$$\text{logit}(\omega) = \exp(-26,5558 - 0,3021X_1)$$

$$\omega = (-26,5558 - 0,3021K_1)$$

dengan X_1 adalah presentase ibu melaksanakan kunjungan K1. Interpretasi model yang terbentuk dari analisis regresi ZIP didasarkan pada nilai *ood rasio* dari masing-masing koefisien yang terlihat dari nilai $\exp(\hat{\beta})$. Nilai koefisien regresi ($\hat{\beta}$) beserta nilai $\exp(\hat{\beta})$ untuk model $\ln(\mu)$ dapat dilihat pada Tabel 4.

Tabel 4. Nilai Koefisien Regresi ($\hat{\beta}$) beserta nilai $\exp(\hat{\beta})$ model $\ln(\mu)$

Variabel bebas	Koefisien ($\hat{\beta}$)	$\exp(\hat{\beta})$
Intercept	-3,61284	0,0269
X_1	-0,04217	0,9587

Sumber: data diolah (2018)

Interpretasi dari nilai ood rasio variabel X_1 adalah hasil penelitian menunjukkan koefisien regresinya bernilai negatif, maka untuk setiap peningkatan 1% cakupan melaksanakan kunjungan K1 di satu Kecamatan di Provinsi Bali akan mengurangi rata-rata jumlah kematian ibu hamil sebesar 4,13%. Pengurangan sebesar 4,13% pada rata-rata jumlah kematian ibu hamil karena terjadinya peningkatan 1% ibu yang melaksanakan kunjungan didasarkan pada perhitungan berikut.

$$100(e^{0,9587(1)} - 1)\% = 100(0,9587 - 1)\% = -4,13\%$$

Nilai koefisien regresi ($\hat{\gamma}$) beserta nilai $\exp(\hat{\gamma})$ untuk model $\text{logit}(\omega)$ dapat dilihat pada Tabel 5.

Tabel 5. Nilai Koefisien Regresi ($\hat{\gamma}$) beserta Nilai $\exp(\hat{\gamma})$ Model $\text{logit}(\omega)$

Variabel Bebas	Koefisien ($\hat{\gamma}$)	$\exp(\hat{\gamma})$
Intercept	-26,5558	$2,93 \times 10^{-12}$
X_1	-0,3021	0,7392

Sumber: data diolah (2018)

Interpretasi dari nilai ood rasio adalah hasil penelitian menunjukkan bahwa koefisien regresinya bernilai negatif, maka untuk setiap peningkatan 1% ibu yang melaksanakan kunjungan akan mengurangi probabilitas jumlah kematian ibu sebesar 26,08%. Pengurangan sebesar 26,08% pada probabilitas jumlah kematian ibu hamil karena terjadinya peningkatan 1% ibu yang melaksanakan kunjungan didasarkan pada perhitungan berikut:

$$100(e^{0,7392(1)} - 1)\% = 100(0,7392 - 1)\% = -26,08\%$$

4. KESIMPULAN DAN SARAN

Berdasarkan hasil dan pembahasan maka dapat diambil kesimpulan sebagai berikut:

1. Dari hasil analisis model regresi ZIGP diperoleh bahwa hasil analisis model regresi ZIGP tidak sesuai untuk memodelkan data angka kematian ibu hamil di Provinsi Bali. Dengan demikian angka kematian ibu hamil di Provinsi Bali dianalisis menggunakan metode regresi ZIP.

Dari hasil analisis ZIP diperoleh model terbaik sebagai berikut:

$$\log(\mu_i) = -3,61284 - 0,04217X_1$$

$$\text{logit}(\omega) = \exp(-26,5558 - 0,3021X_1)$$

2. Berdasarkan kriteria nilai AIC terkecil pada model ZIP maka faktor yang mempengaruhi kematian ibu hamil yaitu ibu melaksanakan kunjungan (K1).

Hal-hal yang disarankan peneliti berdasarkan penelitian jumlah kematian ibu hamil di Provinsi Bali adalah penelitian selanjutnya dapat membandingkan model regresi *Zero Inflated Generalized Poisson* (ZIGP) dengan model regresi *Zero Inflated Negative Binomial Poisson* (ZINB).

DAFTAR PUSTAKA

- Dinas Kesehatan Provinsi Bali. 2016. *Profil Kesehatan Provinsi Bali Tahun 2016*. Bali: Dinkes Provinsi Bali.
- Famoye, F. dan Singh, K. P., (2006), Zero Inflated Generalized Poisson Regression Model with an Application to Domestic Violence Data, *Journal of Data Science*, 4, hal.117-130.
- Ismail, N. & A.A. Jemain. 2005. Generalized Poisson Regression: An Alternative For Risk Classification. *Jurnal Teknologi*. Vol. 43, No.1, 39-50
- Jansakul, N & J.P. Hinde. 2002. Score Tests for Zero-Inflated Poisson Models. *Computational Statistics & Data Analysis*. Vol.40, No.1,75-96.
- Lambert, D. 1992. Zero Inflated Poisson Regression, with an Application to Detect in Manufacturing. *Technometrics*. Vol.34, No.1, 1-14.

- McCullagh, p. & Nelder, J. A. 1989.
Generalized Linier Models. Second edition.
London: Chapman and Hall
- Myers, R.H., Douglas C. Montgomery, G.
Geoffrey Vinning, & Timothy J,
Robinson.2010. *Generalized Linier Models
with Application in Engineering and The
Sciences*. Second edition. New Jersey: John
Wiley and Sons.
- Indrian Dini, Lisa. 2012. Faktor-Faktor Yang
Berhubungan Dengan Pemanfaatan
Pelayanan Antenatal Care Oleh Ibu Hamil
Di Wilayah Kerja Puskesmas Baserah Kab.
Kuansing Propinsi Riau Tahun 2012
[Skripsi] Padang. FKM UNAND. 2012