

PEMODELAN REGRESI *HURDLE* POISSON DALAM MENGATASI *EXCESS ZEROS* UNTUK KASUS PENYAKIT TETANUS NEONATORUM PADA NEONATAL DI JAWA TIMUR

Cylvia Evasari Margaretha¹, Dwi Ispriyanti², Tatik Widiharini³

^{1,2,3}Departemen Statistika FSM Universitas Diponegoro

dwiispriyanti@yahoo.com

ABSTRACT

Tetanus Neonatorum is one of the infectious diseases that occur in newborns caused by *Clostridium Tetani* bacteria through cuts or scratches. The number of Tetanus Neonatorum cases in East Java Province in 2017 is discrete data Poisson distribution with a proportion of zero value of 73,7 percent. The amount of zero value data can result in overdispersion where the variance is greater than the mean. To overcome this problem, Hurdle Poisson regression model is a solution. To estimation of regression parameters for Hurdle Poisson regression is using the Maximum Likelihood Estimation (MLE) method and Broyden Fletcher Goldfarb Shanno (BFGS) iteration. Hurdle Poisson regression produces predictor variables that affect the number of Tetanus Neonatorum cases in East Java Province in the logit model are the percentage of pregnant women administered the K4 program, population density per km² and in the truncated Poisson model are the percentage of labor assisted by health workers the percentage of pregnant women administered the K4 program, population density per km² with the Akaike Information Criterion (AIC) value of 78,422.

Keywords: Tetanus Neonatorum, Excess Zeros, Overdispersion, Hurdle Poisson Regression

1. PENDAHULUAN

Tetanus Neonatorum merupakan salah satu penyakit yang mengakibatkan terjadinya kematian neonatal yang disebabkan oleh bakteri *Clostridium Tetani* melalui luka atau goresan (Dinkes Jatim, 2018). Provinsi Jawa Timur merupakan salah satu provinsi dengan kasus Tetanus Neonatorum tertinggi di Indonesia yaitu sebanyak 24 kasus yang dilaporkan dan 4 kasus diantaranya meninggal (Dinkes Jatim, 2018). Terjadinya kasus Tetanus Neonatorum dapat dipengaruhi oleh faktor medis maupun faktor nonmedis antara lain kurangnya edukasi ibu hamil tentang pentingnya pelayanan kesehatan ibu hamil dan vaksinasi tetanus toksoid, kurang tersedianya fasilitas persalinan dan tenaga medis sehingga banyak persalinan dilakukan di rumah serta penggunaan alat-alat yang tidak steril dalam penanganan tali pusat, adat istiadat setempat serta kondisi lingkungan yang buruk.

Data dari kasus penyakit Tetanus Neonatorum merupakan data diskrit berdistribusi Poisson. Faktor-faktor yang diduga mempengaruhi banyaknya neonatal yang menderita penyakit Tetanus Neonatorum dapat dianalisis menggunakan metode regresi Poisson. Model regresi Poisson memiliki keterbatasan pada asumsi *equidispersion* yaitu rata-rata sama dengan varian. Banyak kasus yang melanggar asumsi *equidispersion* yaitu disaat varian lebih besar dari rata-rata (*overdispersion*) atau varian lebih kecil dari rata-rata (*underdispersion*). Salah satu penyebab terjadinya overdispersi adalah banyaknya nilai nol yang berlebih pada variabel respon (*excess zeros*). Untuk dapat mengatasi masalah overdispersi akibat *excess zeros* maka digunakan Regresi *Hurdle* Poisson (Winkelmann, 2008). Model regresi *Hurdle* Poisson adalah model gabungan yang terdiri dari model logit dan model *truncated Poisson* yang berdasar pada asumsi nilai nol (*zero counts*) dan nilai

positif (*positive counts*) pada variabel respon. Dalam penelitian ini, model regresi *Hurdle Poisson* diterapkan dalam kasus penyakit Tetanus Neonatorum pada neonatal di Jawa Timur tahun 2017.

2. TINJAUAN PUSTAKA

2.1. Distribusi Poisson

Distribusi Poisson merupakan suatu distribusi yang menyatakan peluang banyaknya hasil percobaan yang terjadi dimana kejadian tergantung pada selang waktu tertentu atau di suatu daerah tertentu dengan hasil pengamatan berupa variabel diskrit. Misalkan $Y=0,1,2,\dots$ merupakan jumlah kejadian yang muncul dalam selang waktu dengan rata-rata μ dan Y adalah variabel acak Poisson, maka fungsi peluangnya sebagai berikut:

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

μ menyatakan rata-rata banyaknya sukses yang terjadi dalam selang waktu atau daerah tertentu dan $e = 2,71828$. Dalam distribusi Poisson, rata-rata dan varian mempunyai nilai sama dengan μ .

2.2. Uji Kecocokan Distribusi

Untuk mengetahui bahwa suatu data merupakan distribusi tertentu dapat digunakan uji Kolmogorov-Smirnov. Adapun hipotesisnya adalah sebagai berikut:

H_0 : data mengikuti distribusi Poisson

H_1 : data tidak mengikuti distribusi Poisson

Statistik uji: $T = \sup_x |S(x) - Fo(x)|$

dengan:

$S(x)$ = distribusi frekuensi kumulatif yang dihitung dari data sampel

$Fo(x)$ = distribusi frekuensi kumulatif yang dihipotesiskan

Kriteria ujinya adalah tolak H_0 pada taraf signifikansi jika $T > T^*_{(1-\alpha)}$ adalah nilai kritis yang diperoleh dari tabel *Kolmogorov-Smirnov* atau *nilai p-value* $< \alpha$ (Daniel,1989).

2.3. Regresi Poisson

Model Regresi Poisson adalah model regresi nonlinear yang digunakan untuk menganalisis data diskrit (count) yang menggambarkan hubungan antara variabel respon dan variabel prediktor dengan mengasumsikan variabel respon berdistribusi Poisson. Model Regresi Poisson merupakan penerapan dari *Generalized Linear Model* (GLM) sehingga diperoleh model regresi Poisson sebagai berikut:

$$\ln(\mu_i) = \eta$$

$$\ln(\mu_i) = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki}$$

$$\mu_i = e^{(x_i^T \beta)}$$

dengan:

μ_i = rata-rata jumlah peristiwa yang terjadi pada periode waktu tertentu

x_i = vektor dari variabel prediktor, $x_i = [1 \ x_{1i} \ x_{2i} \ \dots \ x_{ki}]^T$

β = vektor dari parameter regresi Poisson, $\beta = [\beta_0 \ \beta_1 \ \beta_2 \ \dots \ \beta_k]^T$

Penaksiran parameter dalam regresi Poisson menggunakan metode *Maximum Likelihood Estimation* (MLE) dengan fungsi log *likelihoodnya* adalah:

$$\ln L(\beta) = -\sum_{i=1}^n e^{x_i^T \beta} + \sum_{i=1}^n y_i x_i^T \beta - \sum_{i=1}^n \ln(y_i!)$$

2.4. Permasalahan pada Regresi Poisson

2.4.1 Overdispersi

Menurut Cameron & Trivedi (1998) suatu ciri dari distribusi Poisson adalah adanya keadaan equidispersi dimana nilai mean dan varian dari variabel respon bernilai sama. Namun terkadang ditemukan keadaan dimana nilai varian lebih besar dari nilai rata-rata yang disebut overdispersi. Asumsi equidispersi pada regresi Poisson dapat diukur menggunakan nilai *Deviance* dengan prosedur sebagai berikut:

Hipotesis

$H_0: \Phi \leq 1$ (tidak terjadi overdispersi)

$H_1: \Phi > 1$ (terjadi overdispersi)

Statistik uji

$$\Phi = \frac{D}{db}; \quad D = 2 \sum_{i=1}^n \{y_i \ln \frac{y_i}{\hat{\mu}_i}\}$$

dengan:

y_i = nilai peubah respon dari pengamatan ke- i

$\hat{\mu}_i$ = rata-rata penduga pada pengamatan ke- i

db = derajat bebas = $n-k-1$

k = banyaknya parameter termasuk konstanta

n = banyaknya pengamatan

Kriteria uji

H_0 ditolak jika nilai $\Phi > 1$

2.4.2 Excess Zeros

Salah satu permasalahan pada regresi Poisson adalah nilai nol yang berlebih (*excess zeros*). *Excess zeros* dapat dilihat pada proporsi variabel respon yang bernilai nol lebih besar dari data diskrit lainnya. Menurut Winkleman (2008) banyaknya proporsi data yang bernilai nol dari proporsi data lainnya dapat berakibat pada ketepatan (presisi) dalam pengambilan keputusan. Selain itu regresi Poisson juga menjadi tidak tepat digunakan karena *excess zeros* merupakan salah satu penyebab terjadinya overdispersi (Hinde dan Demetrio, 1998).

2.5. Regresi Hurdle Poisson

Menurut Mullahy (1986) kedua bagian model *Hurdle* terbentuk berdasarkan fungsi masa peluang pada bilangan bulat nonnegatif dengan dua bagian. Proses pertama pada model *Hurdle* adalah fungsi peluang dari proses biner menggunakan model logit. Kemudian fungsi peluang proses kedua model *Hurdle* bernilai positif sehingga diperlukan pemotongan nilai nol (*truncated*) pada proses kedua. Maka distribusi peluang dari model regresi *Hurdle* Poisson sebagai berikut:

$$P(Y = y_i) = \begin{cases} (1 - \pi_i) & , y_i = 0 \\ \pi_i \frac{e^{-\mu_i} \mu_i^{y_i}}{(1 - e^{-\mu_i})^{y_i} y_i!} & , y_i = 1, 2, \dots, n \end{cases}$$

dengan $\mu_i = \mu_i(x_i)$ dan $\pi_i = \pi_i(x_i)$, $0 < \pi_i < 1$

Model dari regresi *Hurdle* Poisson dapat dinyatakan sebagai berikut:

$$\ln(\mu_i) = \sum_{j=1}^k x_{ij} \beta_j = \mathbf{x}_i^T \boldsymbol{\beta} \quad \text{dan} \quad \text{logit } \pi_i = \ln\left(\frac{\pi_i}{1-\pi_i}\right) = \sum_{j=1}^k x_{ij} \delta_j = \mathbf{x}_i^T \boldsymbol{\delta}$$

dimana:

$\boldsymbol{\beta} = [\beta_0 \ \beta_1 \ \beta_2 \ \dots \ \beta_k]^T$ adalah vektor dari parameter yang akan ditaksir

$\boldsymbol{\delta} = [\delta_0 \ \delta_1 \ \delta_2 \ \dots \ \delta_k]^T$ adalah vektor dari parameter yang akan ditaksir

\mathbf{x}_i = vektor dari variabel prediktor, $\mathbf{x}_i = [1 \ x_{1i} \ x_{2i} \ \dots \ x_{ki}]^T$

Estimasi parameter regresi *Hurdle* Poisson dilakukan menggunakan metode *Maximum Likelihood Estimation* (MLE) dengan fungsi *likelihood* sebagai berikut:

$$\ln L(\boldsymbol{\delta}; \boldsymbol{\beta}) = \ln \left[\prod_{i=1}^n (1 - \pi_i) \prod_{i=1}^n (\pi_i) \frac{e^{-\mu_i} \mu_i^{y_i}}{(1 - e^{-\mu_i}) y_i!} \right]$$

dengan:

$$\mu_i = e^{\sum_{j=1}^k x_{ij} \beta_j} = e^{(\mathbf{x}_i^T \boldsymbol{\beta})} \text{ dan } \pi_i = \frac{e^{\sum_{j=1}^k x_{ij} \delta_j}}{1 + e^{\sum_{j=1}^k x_{ij} \delta_j}} = \frac{e^{(\mathbf{x}_i^T \boldsymbol{\delta})}}{1 + e^{(\mathbf{x}_i^T \boldsymbol{\delta})}}$$

Dari fungsi log *likelihood* akan dimaksimumkan secara terpisah menjadi fungsi log *likelihood* yang hanya memuat parameter model logit dan model *truncated* Poisson.

2.6. Metode Quasi Newton

Metode Quasi Newton merupakan modifikasi dari metode Newton Raphson dengan mengganti matriks Hessian pada metode Newton Raphson yang menggunakan nilai dari derivatif kedua fungsi log *likelihood* menjadi matriks Hessian yang dihitung menggunakan informasi yang didapatkan pada saat iterasi berlangsung. Algoritma Quasi Newton yang paling terkenal adalah metode Broyden Fletcher Goldfarb Shanno (BFGS) (Wei *et al*, 2006).

Untuk mengestimasi $\boldsymbol{\delta}$ dan $\boldsymbol{\beta}$ menggunakan metode BFGS didapatkan persamaan sebagai berikut:

$$\boldsymbol{\delta}^{(i+1)} = \boldsymbol{\delta}^{(i)} - \lambda^{(i)} (\mathbf{H}^{(i)})^{-1} \mathbf{g}^{(i)}$$

dan

$$\boldsymbol{\beta}^{(i+1)} = \boldsymbol{\beta}^{(i)} - \lambda^{(i)} (\mathbf{H}^{(i)})^{-1} \mathbf{g}^{(i)}$$

dengan:

$$i = 0, 1, 2, \dots, n$$

$\lambda^{(i)}$ merupakan fungsi yang dapat meminimumkan error yang akan terjadi dimana $\lambda^{(i)} = \min \ln L(\boldsymbol{\delta}^{(i)} - \lambda (\mathbf{H}^{(i)})^{-1} \mathbf{g}^{(i)})$ untuk parameter $\boldsymbol{\delta}$ dan $\lambda^{(i)} = \min \ln L(\boldsymbol{\beta}^{(i)} - \lambda (\mathbf{H}^{(i)})^{-1} \mathbf{g}^{(i)})$ untuk parameter $\boldsymbol{\beta}$.

\mathbf{g} merupakan derivatif pertama dari fungsi log *likelihood* yaitu:

$$\mathbf{g} = \begin{bmatrix} \frac{\partial \ln L(\boldsymbol{\delta})}{\partial \delta_0} \\ \frac{\partial \ln L(\boldsymbol{\delta})}{\partial \delta_1} \\ \vdots \\ \frac{\partial \ln L(\boldsymbol{\delta})}{\partial \delta_k} \end{bmatrix} \text{ untuk parameter } \boldsymbol{\delta} \text{ dan } \mathbf{g} = \begin{bmatrix} \frac{\partial \ln L(\boldsymbol{\beta})}{\partial \beta_0} \\ \frac{\partial \ln L(\boldsymbol{\beta})}{\partial \beta_1} \\ \vdots \\ \frac{\partial \ln L(\boldsymbol{\beta})}{\partial \beta_k} \end{bmatrix} \text{ untuk parameter } \boldsymbol{\beta}$$

\mathbf{H} merupakan matriks Hessian yang didapatkan dengan rumus:

$$\mathbf{H}^{(i+1)} = \mathbf{H}^{(i)} - \frac{\mathbf{H}^{(i)} \mathbf{s}^{(i)} (\mathbf{s}^{(i)})^T \mathbf{H}^{(i)}}{(\mathbf{s}^{(i)})^T \mathbf{H}^{(i)} \mathbf{s}^{(i)}} + \frac{\mathbf{y}^{(i)} (\mathbf{y}^{(i)})^T}{(\mathbf{s}^{(i)})^T \mathbf{y}^{(i)}}$$

dengan:

$$\mathbf{S}^{(i)} = \boldsymbol{\delta}^{(i+1)} - \boldsymbol{\delta}^{(i)} \text{ untuk parameter } \boldsymbol{\delta} \text{ dan } \mathbf{S}^{(i)} = \boldsymbol{\beta}^{(i+1)} - \boldsymbol{\beta}^{(i)} \text{ untuk parameter } \boldsymbol{\beta}$$

$$\mathbf{y}^{(i)} = \mathbf{g}^{(i+1)} - \mathbf{g}^{(i)}$$

$\mathbf{H}^{(0)}$ = matriks identitas nxn

Persamaan akan terus berulang hingga didapatkan estimasi parameter yang konvergen yaitu pada saat $\|\boldsymbol{\delta}^{(i+1)} - \boldsymbol{\delta}^{(i)}\| \leq \varepsilon$ dan $\|\boldsymbol{\beta}^{(i+1)} - \boldsymbol{\beta}^{(i)}\| \leq \varepsilon$, dimana ε merupakan bilangan yang sangat kecil mendekati nol yaitu 10^{-5} .

2.7. Pengujian Parameter

2.7.1 Pengujian Serentak (Uji Rasio Likelihood)

Pengujian keseluruhan model regresi *Hurdle* Poisson dilakukan dengan menggunakan *Likelihood Ratio Test* (LRT). Langkah-langkah hipotesis parameter dapat dituliskan sebagai berikut:

Hipotesis

$$H_0: \delta_1 = \delta_2 = \dots = \delta_k = \beta_1 = \beta_2 = \dots = \beta_k = 0$$

H_1 : minimal ada satu $\delta_j \neq 0$ atau $\beta_j \neq 0$, dengan $j=1,2,\dots,k$

Statistik Uji

$$G^2 = -2(\ln L_0 - \ln L_1)$$

dengan:

L_0 = Nilai *likelihood* dari model tanpa variabel prediktor

L_1 = Nilai *likelihood* dari model dengan variabel prediktor

Kriteria Uji

H_0 ditolak jika $G^2 > \chi_{k;\alpha}^2$ atau $p\text{-value} < \alpha$ dengan k merupakan banyaknya parameter dibawah H_0 .

2.7.2 Pengujian Parsial (Uji Wald)

a. Pengujian signifikansi parameter model logit $\pi_i = \mathbf{x}_i^T \boldsymbol{\delta}$ dengan $i=1,\dots,n$

Hipotesis

$$H_0: \delta_j = 0$$

$H_1: \delta_j \neq 0$, dengan $j=1,2,\dots,k$

$$\text{Statistik Uji: } Z_{hitung} = \frac{\hat{\delta}_j}{SE(\hat{\delta}_j)}$$

Kriteria Uji

H_0 ditolak jika $|Z_{hitung}| > Z_{1-\alpha/2}$ atau $p\text{-value} < \alpha$.

b. Pengujian signifikansi parameter model $\ln(\mu_i) = \mathbf{x}_i^T \boldsymbol{\beta}$ dengan $i=1,\dots,n$

Hipotesis

$$H_0: \beta_j = 0$$

$H_1: \beta_j \neq 0$, dengan $j=1,2,\dots,k$

$$\text{Statistik Uji: } Z_{hitung} = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)}$$

Kriteria Uji

H_0 ditolak jika $|Z_{hitung}| > Z_{1-\alpha/2}$ atau $p\text{-value} < \alpha$.

2.8. Multikolinieritas

Uji multikolinieritas dilakukan menggunakan *Variance Inflation Factors* (VIF) sebagai berikut:

$$VIF = \frac{1}{1-R_i^2}$$

dengan R_i^2 adalah koefisien determinasi antar variabel prediktor. Nilai VIF lebih besar dari 10 menunjukkan adanya multikolinieritas.

2.9. Model Terbaik

Terdapat beberapa kriteria yang dapat digunakan untuk menentukan model terbaik, salah satunya adalah *Akaike Information Criteria* (AIC). Dengan statistik uji untuk AIC adalah:

$$AIC = -2 \ln \hat{L} + 2k$$

dengan $\ln \hat{L}$ adalah nilai likelihood dan k adalah banyaknya parameter. Menurut Widarjono (2017) model terbaik adalah model regresi yang mempunyai nilai AIC terkecil.

3. ANALISIS DATA

3.1. Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari Profil Kesehatan Dinas Kesehatan Provinsi Jawa Timur tahun 2017 yang meliputi 38 kabupaten/kota dan dipublikasikan tahun 2018 dengan variabel yang digunakan:

Y : Banyaknya kasus penyakit Tetanus Neonatorum di Provinsi Jawa Timur tahun 2017

X_1 : Persentase persalinan ditolong oleh tenaga kesehatan.

X_2 : Persentase ibu hamil melaksanakan program K4.

X_3 : Kepadatan penduduk per km^2 .

3.2. Metode Analisis Data

Langkah dalam analisis data penelitian ini adalah:

1. Menguji asumsi apakah data berdistribusi Poisson pada variabel respon menggunakan uji Kolmogorov-Smirnov.
2. Melakukan uji Multikolinieritas pada variabel prediktor.
3. Mengestimasi parameter model regresi Poisson.
4. Melakukan uji overdispersi.
5. Melihat apakah data mengalami *excess zeros* atau tidak dengan melihat proporsi nilai nol lebih besar dari proporsi nilai lainnya pada variabel respon.
6. Mengestimasi parameter model regresi *Hurdle* Poisson.
7. Melakukan uji keseluruhan model regresi *Hurdle* Poisson.
8. Menguji signifikansi parameter secara individu model regresi *Hurdle* Poisson untuk parameter pada model logit dan model *truncated* Poisson.
9. Menentukan model akhir regresi *Hurdle* Poisson.

4. HASIL DAN PEMBAHASAN

4.1. Pengujian Distribusi Poisson pada Variabel Respon

Berdasarkan hasil uji Kolmogorov-Smirnov diperoleh nilai $T = 0,205$ atau $p\text{-value} = 0,082$. Nilai tersebut lebih kecil daripada nilai kritis pada tabel Kolmogorov-Smirnov dengan $\alpha = 5\%$ dan $n = 38$ yaitu $0,215$ sehingga dapat disimpulkan bahwa data variabel respon yaitu banyaknya kejadian Tetanus Neonatorum pada neonatal di Jawa Timur tahun 2017 mengikuti distribusi Poisson.

4.2. Uji Multikolinieritas

	X_1	X_2	X_3
VIF	1,3909	1,8138	1,3679

Nilai VIF pada semua variabel prediktor adalah kurang dari 10, sehingga dapat dikatakan bahwa antar variabel prediktor tidak saling berkorelasi atau tidak adanya multikolinieritas.

4.3. Pemodelan Regresi Poisson

Pemodelan jumlah kasus Tetanus Neonatorum pada neonatal di Jawa Timur tahun 2017 dengan menggunakan regresi Poisson didapatkan estimasi parameter sebagai berikut:

$$\ln(\mu_i) = 23,4309 - 0,0790X_{1i} - 0,2035X_{2i} + 0,0429X_{3i}$$

4.4. Pengujian Overdispersi

Berdasarkan hasil uji overdispersi diperoleh nilai $\Phi = 1,4627 > 1$ sehingga model regresi Poisson untuk kejadian Tetanus Neonatorum pada neonatal di Jawa Timur tahun 2017 mengalami overdispersi.

4.5. Excess Zeros pada Variabel Respon

Nilai	Jumlah	Persentase(%)
0	28	73,7
1	6	15,8
2	2	5,3
6	1	2,6
8	1	2,6
Total	38	100,0

Dapat terlihat bahwa nilai nol mempunyai proporsi sebesar 73,7% yang melebihi proporsi nilai pada variabel respon lainnya, sehingga variabel respon untuk kasus Tetanus Neonatorum pada neonatal di Jawa Timur tahun 2017 mengalami *excess zeros*.

4.6. Pemodelan Regresi Hurdle Poisson

Pemodelan kasus penyakit Tetanus Neonatorum pada neonatal di Jawa Timur tahun 2017 dengan menggunakan regresi Hurdle Poisson menghasilkan estimasi parameter sebagai berikut:

$$\begin{aligned} \text{logit}(\pi_i) &= 20,3871 - 0,0300X_{1i} - 0,2253X_{2i} + 0,0563X_{3i} \\ \ln(\mu_i) &= 44,5239 - 0,3111X_{1i} - 0,1983X_{2i} + 0,0383X_{3i} \end{aligned}$$

4.7. Pengujian Parameter Model Regresi Hurdle Poisson

a. Uji Keseluruhan Model Regresi Hurdle Poisson

Berdasarkan hasil uji keseluruhan model regresi Hurdle Poisson diperoleh nilai $G^2 = 22,294 > \chi^2_{6;0,05} = 12,592$ atau $p\text{-value} = 0,0011 < \alpha = 0,05$ sehingga dapat disimpulkan bahwa paling sedikit ada satu variabel prediktor yang berpengaruh terhadap model regresi Hurdle Poisson untuk kasus Tetanus Neonatorum pada neonatal di Jawa Timur tahun 2017.

b. Uji Parsial Model Regresi Hurdle Poisson

Pengujian signifikansi parameter model $\text{logit} \pi_i = \mathbf{x}_i^T \boldsymbol{\delta}$ dengan $i=1, \dots, 38$

Parameter	Estimasi	Standar Error	Z_{hitung}	$P\text{-value}$	Keputusan
δ_1	-0,0300	0,11321	-0,265	0,7908	H_0 diterima
δ_2	-0,2253	0,11463	-1,966	0,0493	H_0 ditolak
δ_3	0,0563	0,02333	2,413	0,0158	H_0 ditolak

Pada taraf signifikansi $\alpha = 5\%$ dapat disimpulkan bahwa variabel yang berpengaruh terhadap model logit yaitu variabel persentase ibu hamil yang melaksanakan program K4 dan kepadatan penduduk per km^2 sedangkan variabel persentase persalinan ditolong tenaga kesehatan tidak berpengaruh terhadap model logit.

Pengujian signifikansi parameter model $\ln(\mu_i) = \mathbf{x}_i^T \boldsymbol{\beta}$ dengan $i=1, \dots, 38$

Parameter	Estimasi	Standar Error	Z_{hitung}	$P\text{-value}$	Keputusan
β_1	-0,3111	0,15789	-1,970	0,0488	H_0 ditolak
β_2	-0,1983	0,09626	-2,060	0,0394	H_0 ditolak

β_3	0,0383	0,01854	2,067	0,0387	H_0 ditolak
-----------	--------	---------	-------	--------	---------------

Pada taraf signifikansi $\alpha = 5\%$ dapat disimpulkan bahwa variabel yang berpengaruh terhadap model *truncated* Poisson yaitu variabel persentase persalinan ditolong tenaga kesehatan, persentase ibu hamil yang melaksanakan program K4, dan kepadatan penduduk per km^2 .

Setelah dilakukan pengujian keseluruhan model dan pengujian parameter secara individu maka dihasilkan model akhir dari regresi *Hurdle* Poisson adalah:

$$\begin{aligned} \text{logit}(\pi_i) &= 20,3871 - 0,0300X_{1i} - 0,2253X_{2i} + 0,0563X_{3i} \\ \ln(\mu_i) &= 44,5239 - 0,3111X_{1i} - 0,1983X_{2i} + 0,0383X_{3i} \end{aligned}$$

Berdasarkan hasil estimasi parameter pada model logit dapat diketahui bahwa setiap perubahan banyaknya persalinan ditolong oleh tenaga kesehatan sebesar 1 persen maka akan menurunkan peluang terjadinya kasus Tetanus Neonatorum sebesar $e^{-0,0300} = 0,9704$ kali tanpa melibatkan variabel lainnya. Setiap perubahan ibu hamil melaksanakan program K4 sebesar 1 persen maka akan menurunkan peluang terjadinya kasus Tetanus Neonatorum sebesar $e^{-0,2253} = 0,7983$ kali tanpa melibatkan variabel lainnya. Setiap perubahan kepadatan penduduk per km^2 sebesar 1 satuan maka akan menaikkan peluang terjadinya kasus Tetanus Neonatorum sebesar $e^{0,0563} = 1,0056$ kali tanpa melibatkan variabel lainnya.

Berdasarkan hasil estimasi parameter pada model *truncated* Poisson dapat diketahui bahwa setiap perubahan banyaknya persalinan ditolong oleh tenaga kesehatan sebesar 1 persen maka akan menurunkan rata-rata kasus Tetanus Neonatorum sebesar $e^{-0,3111} = 0,7326$ kali tanpa melibatkan variabel lainnya. Setiap perubahan ibu hamil melaksanakan program K4 sebesar 1 persen maka akan menurunkan rata-rata kasus Tetanus Neonatorum sebesar $e^{-0,1983} = 0,8201$ kali tanpa melibatkan variabel lainnya. Setiap perubahan kepadatan penduduk per km^2 sebesar 1 satuan maka akan menaikkan rata-rata kasus Tetanus Neonatorum sebesar $e^{0,0383} = 1,0390$ kali tanpa melibatkan variabel lainnya.

4.8. Perbandingan model Regresi Poisson dengan Regresi Hurdle Poisson

Model	AIC
Regresi Poisson	82,554
Regresi <i>Hurdle</i> Poisson	78,422

Dilihat dari nilai AIC, dapat disimpulkan bahwa nilai AIC dari model regresi *Hurdle* Poisson lebih kecil dibandingkan dengan model regresi Poisson. Hal tersebut menunjukkan bahwa model regresi *Hurdle* Poisson lebih baik digunakan pada kasus Tetanus Neonatorum pada neonatal di Jawa Timur tahun 2017.

5. KESIMPULAN

Dari penelitian yang sudah dilakukan, maka dapat diperoleh kesimpulan sebagai berikut:

1. Model regresi *Hurdle* Poisson untuk jumlah kasus Tetanus Neonatorum pada neonatal di Jawa Timur tahun 2017 adalah sebagai berikut:

$$\begin{aligned} \text{logit}(\pi_i) &= 20,3871 - 0,0300X_{1i} - 0,2253X_{2i} + 0,0563X_{3i} \\ \ln(\mu_i) &= 44,5239 - 0,3111X_{1i} - 0,1983X_{2i} + 0,0383X_{3i} \end{aligned}$$

2. Faktor-faktor yang dapat mempengaruhi kasus Tetanus Neonatorum pada neonatal di Jawa Timur tahun 2017 pada model logit yaitu variabel persentase ibu hamil yang melaksanakan program K4 dan kepadatan penduduk per km^2 sedangkan pada model *truncated* Poisson yaitu persentase persalinan ditolong tenaga kesehatan, persentase ibu hamil yang melaksanakan program K4 dan kepadatan penduduk per km^2 .

DAFTAR PUSTAKA

- Cameron, A. C dan Trivedi, P. K. 1998. *Regression Analysis of Count Data*. Cambridge: Cambridge University Press.
- Daniel, W. W. 1989. *Statistika Nonparametrik Terapan*. Diterjemahkan oleh: Alex Tri Jantjono W. Jakarta: Gramedia. Terjemahan dari: *Applied Nonparametric Statistics*.
- [Dinkes Jatim] Dinas Kesehatan Provinsi Jawa Timur. 2018. *Profil Kesehatan Provinsi Jawa Timur Tahun 2017*. Surabaya: Dinkes Jatim.
- Hinde, J. dan Dementrio, C. 1998. *Overdispersion: Models and Estimation*. *Computational Statistic and Data Analysis*. 27: 151-170.
- Mullahy, J. 1986. *Specification and Testing of Some Count Data Models*. *Journal of Econometrics*. 33: 341-365.
- Widarjono, A. 2007. *Ekonometrika Teori dan Aplikasi untuk Ekonomi dan Bisnis*, Edisi kedua. Yogyakarta: Ekonisia.
- Winkelmann, R. 2008. *Econometric Analysis of Count Data 5th edition*. Berlin: Springer.
- Wei, Z. Li, G. dan Qi, L. 2006. *New Quasi-Newton Methods for Unconstrained Optimization Problems*. *Applied Mathematical and Computation*. 175: 1156-1188.