PENERAPAN DATA *COUNT* DENGAN MENGGUNAKAN REGRESI *HURDLE POISSON* (Studi Kasus: Banyak Kematian Ibu di Provinsi Jawa Barat Tahun 2015)

Redicha Julianda H, Nar Herrhyanto dan Bambang Avip P. M.

Departemen Pendidikan Matematika, Fakultas FPMIPA, Universitas Pendidikan Indonesia
Email: redicha.juliandahrp@gmail.com

Abstrak : Regresi *Poisson* merupakan suatu model regresi yang dapat memodelkan bentuk hubungan antara variabel respon yang berbentuk data *count* dengan variabel prediktor. Pada regresi *Poisson* harus memenuhi asumsi bahwa nilai varians sama dengan nilai mean (*equidispersi*). Namun, dalam beberapa kasus, terdapat data yang memiliki banyak nilai nol (*excess zero*) pada variabel respon yang mengakibatkan nilai varians tidak sama dengan nilai mean. Nilai varians lebih besar dari nilai mean disebut *overdispersi*. Sehingga Model regresi *Poisson* menjadi tidak cocok untuk memodelkan data tersebut. Alternatif yang dapat digunakan apabila terjadi *overdispersi* akibat banyaknya data bernilai nol (*excess zero*) pada variabel respon adalah model regresi *Hurdle Poisson*. Model *Hurdle Poisson* terdapat dua bagian. Bagian pertama, model untuk data biner yang bernilai nol atau nilai positif yang ditaksir dengan menggunakan model Logit. Bagian kedua, model untuk data yang bernilai positif saja yang ditaksir dengan menggunakan model *Truncated Poisson*. Regrsi *Hurdle Poisson* kemudian diterapkan dalam memodelkan faktor-faktor yang mempengaruhi banyaknya kasus kematian ibu, dengan faktor-faktor tersebut ialah: persentase persalinan oleh tenaga medis, persentase K4, persentase wanita kawin dengan umur perkawinan pertama dibawah 18 tahun, persentase ibu hamil yang mendapatkan Fe1, persentase ibu hamil yang mendapatkan Fe3, persentase ibu nifas mendapatkan vitamin A, persentase komplikasi kebidanan yang ditangani. Dilihat dari hasil pengujian, faktor-faktor yang mempengaruhi banyaknya kasus kematian ibu secara signifikan pada model regresi Hurdle Poisson bagian model Logit adalah persentase ibu hamil mendapatkan Fe1. Sedangkan pada bagian model *Truncated Poisson* adalah persentase wanita kawin dengan umur perkawinan pertama dibawah 18 tahun dan variabel persentase ibu hamil mendapatkan Fe3.

Kata kunci : regresi *Poisson*, *excess zero*, overdispersi, model regresi *Hurdle Poisson,* banyak kasus kematian ibu.

Abstract : *Poisson* Regression is a model of regression that can represent relation form between response variable in a form of count data with predictor variable. *Poisson* regression must satisfy assumption the variance equal the means (*equidispersion*). However, in any cases, there are some data that have *excess zero* in response variable resulting in the variance not equal the means. the variance is bigger than the means called *overdispersion*. That is to say, *Poisson* regression model is not appropriate to represent those data. The alternative is *Hurdle Poisson* regression. *Hurdle Possion* model is divided into two parts. Firstly, it is binary data model means zero or positive score predicted using Logit model. Secondly, modelling only for positive score data predicted using *Truncated Poisson* model. *Hurdle Poisson* regression is applied to represent factors affecting to many death case, where the factors are: The percentage of birth by medic, K4, married woman with the first marriage phase under 18 years old, pregnant woman getting Fe1, pregnant woman getting Fe3, woman of childbirth-phase getting A vitamin, and obstretical complication. In this stance, the data shows that factors affect to the amount of woman death case significantly exist in Logit Model of *Hurdle Poisson* regression is the percentage of pregnant woman getting Fe1. Menawhile, for *Truncated Poisson* model is the percentage of married woman in the first marriage phase under 18 years old and variable of pregnant woman percentage getting Fe3.
Key Words : *Poisson* Regression, *excess zero*, overdispersion, *Hurdle Poisson* regression model, the amount of woman death case.

1. PENDAHULUAN

Angka kematian merupakan indikator *outcome* pembangunan kesehatan (Badan Pusat Statistik, 2015). Salah satu indikator yang digunakan untuk mengukur status kesehatan ibu pada suatu wilayah adalah angka kematian ibu (AKI). AKI merupakan salah satu indikator yang peka terhadap kualitas dan aksesibilitas fasilitas pelayanan kesehatan (Kementerian Kesehatan, 2015) Angka Kematian Ibu (AKI) menurut WHO (*World Health Organization*) merupakan kematian seorang wanita ketika hamil atau dalam waktu 42 hari setelah berakhirnya kehamilan, tanpa memperhatikan waktu dan lokasi kehamilan, dari segala penyebab yang terkait atau kehamilan itu sendiri tetapi tidak dari kecelakaan atau insidental.

Salah satu metode statistika yang banyak digunakan untuk membuat model dan mengetahui bentuk hubungan antara satu atau beberapa variabel prediktor terhadap variabel respon adalah analisis regresi. Pada model regresi linear tidak bisa mengatasi permasalahan ketika data yang variabel responnya berbentuk diskrit. *Generalized Linear Model* (GLM) dapat digunakan sebagai solusi untuk mengatasi masalah tersebut.

Salah satu model regresi yang termasuk kedalam penerapan GLM adalah regresi *Poisson*. Karena variabel respon berupa kasus kematian ibu merupakan kejadian yang sangat langka tetapi pasti terjadi dalam selang waktu tertentu, maka kejadian tersebut berkaitan dengan kejadian distribusi *Poisson*. Variabel respon pada model regresi *Poisson* diasumsikan berdistribusi *Poisson*. Regresi *Poisson* digunakan untuk menganalisis hubungan antara satu atau lebih variabel prediktor dengan variabel respon dimana variabel responnya berupa data diskrit (data cacah). Di dalam prakteknya ada 3 kemungkinan yang terjadi dalam data cacah yaitu nilai varians sama dengan nilai mean (*equidispersi*), nilai varians lebih besar dari nilai mean (*overdispersi*) dan nilai varians lebih kecil dari nilai mean (*underdispersi*). Karena pada regresi *Poisson* diasumsikan bahwa nilai varians sama dengan mean nya, maka dalam regresi *Poisson* tidak cocok digunakan untuk mengatasi data yang mengalami *overdispersi* maupun *underdispersi*. *Overdispersi* yang terjadi dalam permasalahan tersebut dikarenakan banyaknya data yang bernilai nol (*excess zeros*), maka regresi *Poisson* menjadi tidak sesuai jika terdapat banyak data bernilai nol yang mengakibatkan terjadinya *overdispersi* (*excess zeros*) (Zia, 2016). Beberapa model yang mampu dapat mengatasi permasalahan *overdispersi* akibat banyak nya data yang bernilai nol (*excess zeros*) adalah *Zero Inflated* model dan *Hurdle* model. Model *Hurdle* bersifat fleksibel (Jinnan, 2016).

2. METODOLOGI

Metode penelitian yang digunakan pada skripsi ini merupakan metode penelitian yang melakukan kajian teoritis tentang regresi *Hurdle Poisson* dan menerapkan data banyak kasus kematian ibu di Provinsi Jawa Barat Tahun 2015 menggunakan regrsi *Hurdle Poisson*.

2.1 Regresi *Poisson*

Model regresi *Poisson* merupakan bagian dari *Generalized Linear Model* (GLM). Di dalam komponen *Generalized Linear Model* (GLM) terdapat sebuah fungsi penghubung (*link function*) yang menghubungkan komponen acak dengan komponen sistematik. Pada regresi *Poisson*, *link function* yang digunakan adalah *link function* log natural. *Link function* yang digunakan adalah *ln link g(μ)*, karena rata-rata dari variabel responnya akan berbentuk fungsi eksponensial dan akan menjamin bahwa nilai variabel yang ditaksir dari variabel responnya akan bernilai non negatif, sehingga dapat ditulis sebagai berikut :

$$g(\mu_i) = \ln \mu_i = \beta_0 + x_{i1}\beta_1 + x_{i2}\beta_2 + \cdots + x_{ip}\beta_p = \beta_0 + \sum_{j=1}^{p} x_{ij}\beta_j \qquad (2.1)$$

Sehingga biasa disebut model Loglinear dengan bentuk persamaan sebagai berikut:

$$\mu_i = \exp\left(\beta_0 + x_{i1}\beta_1 + x_{i2}\beta_2 + \cdots + x_{ip}\beta_p\right) = \exp\left(\beta_0 + \sum_{j=1}^{p} x_{ij}\beta_j\right) \qquad (2.2)$$

Dalam model regresi *poisson* terdapat beberapa asumsi-asumsi, dimana Y diasumsikan berdistribusi *Poisson*. Distribusi *Poisson* memiliki nilai varians dan nilai mean sama, yaitu:

$var(Y_i) = E(Y_i) = \mu_i$

2.2 Pendeteksian *Overdispersi*

Sebelum membuat model regresi *Hurdle Poisson*, akan dilihat terlebih dahulu apakah data tersebut mengalami overdispersi atau tidak. Cara untuk mendeteksi overdispersi pada model regresi *Poisson* adalah dengan menggunakan statistik *deviance*, dimana nilai statistik *deviance* yang dibagi dengan derajat bebas nya. Nilai hasil bagi tersebut harus mendekati 1, jika nilai hasil bagitersebut lebih dari 1 maka terdapat *overdispersi*. Langkah-langkah yang akan dilakukan untuk melakukan pengujian tersebut adalah sebagai berikut:

1. Perumusan Hipotesis

   $H_0 : \hat{\phi} = 0$ (Tidak terdapat *overdispersi* pada model regresi *Poisson*)

   $H_1 : \hat{\phi} > 0$ (Terdapat *overdispersi* pada model regresi *Poisson*)

2. Statistik Uji

   Statistik uji yang digunakan adalah:

   $$\hat{\phi} = \frac{\text{nilai deviance}}{\text{n} - \text{p}}$$

   $$D = 2\sum_{i=1}^{n}\left\{y_i log \frac{y_i}{\mu_i} - (y_i - \mu_i)\right\} \quad (2.3)$$

   Keterangan:

   D : nilai deviance

   $\hat{\phi}$: parameter dispersi

   y_i : nilai variabel respon dari pengamatan ke-i

   μ_i : taksiran rata-rata banyak kasus ke-i pada model regresi *Poisson*

   n : banyaknya pengamatan

   p : banyaknya parameter

3. Kriteria Pengujian

   Dengan mengambil taraf nyata sebesar α, maka H_0ditolak jika $\hat{\phi} > 1$

4. Kesimpulan

   Interpretasi ditolak atau diterimanya H_0

2.3 Model *Hurdle*

Penggunanan model *Hurdle* yang paling penting adalah ketika data nya bersifat diskrit (*count* data). Salah satu pendekatan untuk mengatasi masalah *overdispersi* adalah model *Hurdle. Overdispers* terjadi ketika nilai varians lebih besar dari nilai means. Salah satu penyebab yang terjadi *overdispersi* yaitu terdapat banyak nya nilai nol (*excess zero*) pada variabel respon (Zorn, 1996). Model *Hurdle* mampu mengatasi kasus *excess zeros* dengan cara membagi dua model ke dalam dua bagian yaitu:

1. Model untuk data biner yang bernilai nol (*zero counts*) atau nilai positif (*positive counts*), dimana model tersebut ditaksir dengan model logit.
2. Model untuk data yang bernilai positif (*positive counts*) saja, dimana model tersebut ditaksir dengan model *Truncated Poisson*.

Misalkan $k_1(0)$ adalah nilai probabilitas (peluang) ketika variabel respon (Y) bernilai sama dengan 0 dan $k_2(y), y = 1,2,3, \ldots$ adalah sebuah fungsi probabilitas (peluang) ketika variabel respon (Y) bernilai positif. Sehingga, fungsi probabilistiknya (peluang) dari model *Hurdle* adalah sebagai berikut (Saffari. dkk, 2012):

$$P(Y_i = y) = \begin{cases} k_1(0), \ y = 0 \\ \big(1 - k_1(0)\big)k_2(y), y = 1,2, \ldots \end{cases} \tag{2.4}$$

Pada tahun 1986 Mullahy mendiskusikan tentang kedua model *Hurdle* yang terbentuk pada fungsi probabilitas (peluang) untuk nilai non-negatif (Zorn, 1996), misalkan f_1 dan $f_2$dengan melihat bentuk umum model *Hurdle* di atas, maka $k_1(0) = f_1(0)$ dan
$k_2(y) = \frac{f_2(y)}{(1-f_2(0))}$. Pada kasus k_2, akan dilakukan normalisasi, karena f_2 berlaku untuk nilai non-negatif ($y = 0,1,2, \ldots$) sedangkan k_2 berlaku untuk nilai positif ($y = 1,2, \ldots$). Ini berarti perlu dilakukan untuk memotong fungsi peluang f_2.

$$f_2(0) + f_2(1) + f_2(2) + \cdots = 1, \ \ y = 0,1,2, \ldots$$

$$f_2(1) + f_2(2) + \cdots = 1 - f_2(0), \ \ y = 1,2,3, \ldots$$

$$\frac{f_2(1)}{1 - f_2(0)} + \frac{f_2(2)}{1 - f_2(0)} + \cdots = 1$$

$$\frac{f_2(y)}{1 - f_2(0)}, \ \ y = 1,2, \ldots$$

Sehingga menurut asumsi Mullahy, distribusi peluang dari model *Hurdle* adalah sebagai berikut:

$$P(y = 0) = f(y = 0) = f_1(0) \tag{2.5}$$

$$P(y = y) = f(y = y) = \frac{1-f_1(0)}{1-f_2(0)} f_2(y) = \theta f_2(y) \ , y = 1,2,3, \ldots \tag{2.6}$$

Berdasrkan Persamaan (2.5) dan (2.6) maka dapat ditulis distribusi peluang dari model *Hurdle* sebagai berikut:

$$P(Y_i = y) = \begin{cases} f_1(0), & y = 0 \\ \theta f_2(y), & y = 1,2,3, \ldots \end{cases}$$

2.4 Model *Hurdle Poisson*

Misalkan $Y_i, i = 1,2,3, \ldots, n$ adalah sebuah nilai non-negatif dari variabel acak, dan misalkan $Y_i = 0$ adalah observasi dengan frekuensi nilai 0 yang terlalu banyak, sehingga tidak bisa ditangani dengan menggunakan model regresi *Poisson* biasa. Pandang bahwa model regresi *Hurdle Poisson* dengan variabel respon

$Y_i, (\ i = 1,2,3, \ldots, n)$ memiliki distribusi sebagai berikut (Saffari, Adnan, & Greene, 2012):

$$P(Y_i = y_i) = \begin{cases} 1 - \pi_i & , y_i = 0 \\ (\pi_i) \frac{e^{-\mu_i} \mu_i^{y_i}}{(1-\pi_i) y_i!} & , y_i > 0 \end{cases} \quad (2.7)$$

Model Logit dapat ditulis sebagai berikut:

$$Logit(\pi_i) = log\left(\frac{\pi_i}{1-\pi_i}\right) = \sum_{j=1}^{p} z_{ij} \delta_j \quad (2.8)$$

dimana:

z : vektor kovariat pada variabel prediktor

$z_i = \left[z_{i1} = 1, z_{i2}, \ldots, z_{ip}\right]$

δ : vektor kolom parameter koefisien regresi untuk model logit

$\delta = [\delta_1 \quad \delta_2 \quad \cdots \quad \delta_p]^T$

Model *Truncated Poisson* dapat ditulis sebagai berikut:

$$\log(\mu_i) = \sum_{j=1}^{p} x_{ij} \beta_j$$

$$\mu_i = exp\left(\sum_{j=1}^{p} x_{ij} \beta_j\right) \quad (2.9)$$

dimana:

x : vektor kovariat pada variabel prediktor

$x_i = \left[x_{i1} = 1, x_{i2}, \ldots, x_{ip}\right]$

β : vektor kolom parameter koefisien regresi untuk model *Truncated Poisson*

$\beta = [\beta_1 \quad \beta_2 \quad \cdots \quad \beta_p]^T$

Model peluang *Hurdle Poisson* yang terbentuk dari kombinasi untuk data bernilai 0 dan model *Truncated Poisson* untuk data yang bernilai positif saja adalah:

$$P(Y_i = y_i) = \begin{cases} \dfrac{1}{1 + \exp\left(\sum_{j=1}^{p} z_{ij} \delta_j\right)} & , y_i = 0 \\ \left[\dfrac{\exp\left(\sum_{j=1}^{p} z_{ij} \delta_j\right)}{1 + \exp\left(\sum_{j=1}^{p} z_{ij} \delta_j\right)}\right] \left[\dfrac{\left[exp\left(\sum_{j=1}^{p} x_{ij} \beta_j\right)\right]^{y_i}}{\left[exp\left(exp\left(\sum_{j=1}^{p} x_{ij} \beta_j\right)\right) - 1\right] y_i!}\right] & , y_i > 0 \end{cases}$$

Metode penaksiran yang digunakan dalam metode *Hurdle Poisson* ini adalah metode kemungkinan maksimum. Fungsi kemungkinan dari model regresi *Hurdle Poisson* adalah sebagai berikut :

$$L(\delta, \beta) = \prod P(Y_i = y_i)$$

$$L(\delta, \beta) = I_{y_i=0} \frac{1}{1+\exp\left(\sum_{j=1}^{p} z_{ij} \delta_j\right)} I_{y_i>0} \left[\frac{\exp\left(\sum_{j=1}^{p} z_{ij} \delta_j\right)}{1+\exp\left(\sum_{j=1}^{p} z_{ij} \delta_j\right)}\right] \left[\frac{\left[exp\left(\sum_{j=1}^{p} x_{ij} \beta_j\right)\right]^{y_i}}{\left[exp\left(exp\left(\sum_{j=1}^{p} x_{ij} \beta_j\right)\right) - 1\right] y_i!}\right]$$

Metode maksimum kemungkinan, taksiran parameter tidak bisa menghasilkan taksiran secara eksplisit dan tidak menghasilkan taksiran parameter yang eksplisit, tetapi menghasilkan persamaan nonlinear yang cukup kompleks, sehingga perlu algoritma khusus untuk mendapatkan nilai taksiran parameter. Salah satu algoritma yang dapat digunakan adalah algoritma *Broyden-Fletcher-Goldfarb-Shanno*.

2.5 Pengujian Parameter Model *Hurdle Poison*

a. Uji Serempak

Uji parameter secara serempak (*overall*) model Logit dengan uji *likelihood Ratio Test* sebagai berikut:

1. Perumusan hipotesis

   $H_0 := \beta_1 = \beta_2 = \beta_3 = \cdots = 0$ (semua variabel prediktor dalam model yang tidak berpengaruh)

   H_1 : Paling sedikit ada satu $\beta_j \neq 0$ dimana $j = 1,2,3,\ldots,p$ (paling sedikit ada satu variabel prediktor dalam model yang berpengaruh)

2. Statistik uji

   $$G^2 = -2\ln\left[\frac{L(\Omega_0)}{L(\Omega)}\right] \tag{2.10}$$

   Dengan

   $L(\Omega_0)$: fungsi likelihood untuk model yang tidak mengandung variabel prediktor

   $L(\Omega)$: fungsi likelihood untuk model yang mengandung semua variabel prediktor

3. Kriteria pengujian

   Dengan mengambil taraf nyata sebesar α, maka H_0 ditolak jika

   $G^2 > \chi^2_{\alpha,p}$

   atau

   p-value < α

4. Kesimpulan

   Interpretasi ditolak atau diterimanya H_0

b. Uji Parsial

Uji Parsial digunakan untuk menguji masing-masing parameter untuk mengetahui apakah suatu model mempunyai variabel signifikans untuk masuk ke model. Pengujian parameter parsial dilakukan pada dua model yaitu model logit dan model *Truncated Poisson*.

Uji parameter parsial model Logit dan model *Truncated Poisson* dengan uji *Wald Test* sebagai berikut:

1. Perumusan hipotesis

   $H_0 : \beta_j = 0$

   $H_1 : \beta_j \neq 0$

2. Statistik uji

   $$W_j = \left(\frac{\widehat{\beta_j}}{SE(\widehat{\beta_j})}\right)^2 \tag{2.11}$$

dengan:

$\widehat{\beta_J}$: penaksir dari parameter δ_j pada model Logit

$SE(\widehat{\beta_J})$: kekeliruan baku taksiran dari $\widehat{\delta_J}$ pada model Logit

3. Kriteria pengujian

   Dengan mengambil taraf nyata sebesar α, maka H_0ditolak jika

   $W_j > \chi^2_{\alpha,p}$

   atau

   p-value < α
4. Kesimpulan

   Interpretasi ditolak atau diterimanya H_0

3. HASIL DAN PEMBAHASAN

Berdasarkan tabel 3.1 yang menyajikan deskriptif statistik data banyak kematian ibu di Provinsi Jawa Barat tahun 2015 dan variabel-variabel yang digunakan dalam penelitian ini ditunjukkan bahwa variabel prediktor Persentase persalinan oleh tenaga medis (X_1), Persentase K4 (X_2), Persentase Wanita Kawin dengan Umur Perkawinan Pertama dibawah 18 Tahun (X_3), Persentase ibu hamil mendapatkan tablet Fe1 (X_4), Persentase ibu hamil mendapatkan tablet Fe3 (X_5), Persentase ibu nifas mendapatkan vitamin A (X_6), Persentase komplikasi kebidanan yang ditangani (X_7).

Tabel 3.1 Deskriptif Statistik Data Banyak Kematian Ibu di Provinsi Jawa Barat Tahun 2015 dan Variabel-Variabel dalam Penelitian

	N	Minimum	Maksimum	Mean	Variansi
Kematian Ibu	27	0	15	2,925926	11,91738
Persalinan Tenaga Medis	27	74,40	121,10	98,35556	139,1403
K4	27	79,00	121,00	98,51852	116,2477
Umur Perkawinan Pertama	27	15,10	55,15	38,84519	149,3329
Fe1	27	87,63	129,33	104,0807	135,3786
Fe3	27	79,74	121,19	96,45667	129,7695
Vitamin A	27	80,62	122,67	98,49593	127,1074
Komplikasi Kebidanan ditangani	27	32,60	150,70	100,4222	668,9203

Kasus kematian ibu yang paling banyak yaitu 15 kasus yang ditemukan di Kabupaten Cianjur dan paling sedikit adalah tidak ditemukannya kasus kematian ibu yang ditemukan di 5 Kabupaten/Kota yaitu Kabupaten Sumedang, Kabupaten Pangandaran, Kota Sukabumi, Kota Bekasi dan Kota Banjar. Untuk nilai kematian ibu memiliki rata-rata sebesar 2,925926 dan variansi 11,91738. Rata-Rata persentase persalinan oleh tenaga medis adalah 98,35556%. Rata-rata persentase K4 adalah 98,51852%. Rata-rata persentase wanita kawin dengan umur perkawinan pertama dibawah 18 tahun adalah 38,84519%. Rata-rata persentase ibu hamil yang mendapatkan cakupan Fe1 adalah 104,0807%. Rata-rata persentase ibu hamil yang mendapatkan cakupan Fe3 adalah 96,45667. Sedangkan rata-rata persentase ibu nifas mendapatkan vitamin A sebesar 98,49593%. Dan rata-rata persentase komplikasi kebidanan ditangani adalah 100,4222%.

Hasil pemodelan dengan Regresi *Poisson* menunjukkan bahwa seluruh variabel prediktor secara signifikan berpengaruh terhadap kematian ibu dengan tingkat signifikansi 5% yaitu Persentase K4 (X_2), Persentase ibu hamil mendapatkan tablet Fe1 (X_4), Persentase ibu hamil mendapatkan tablet Fe3 (X_5). Bentuk persamaan Regresi *Poisson* sebagai berikut :

Tabel 3.2 Taksiran Parameter Model Regresi *Poisson*

Parameter	Taksiran	p-value	$\alpha = 0,05$	Keterangan
β_0	-0,761673			
β_1	0,120601	0,00125	0,05	Signifikan
β_2	-0,044244	0,08946	0,05	Tidak signifikan
β_3	0.076561	0,0000156	0,05	Signifikan
β_4	0,001503	0,97092	0,05	Tidak signifikan
β_5	-0,135764	0,00125	0,05	Signifikan
β_6	0,057087	0,05287	0,05	Tidak signifikan
β_7	-0,016801	0,05790	0,05	Tidak signifikan

$$\mu = \exp(-0,761673 + 0,120601X_1 + 0.076561X_3 - 0,135764X_5)$$

Pada regresi *Poisson*, terdapat asumsi yang harus dipenuhi yaitu asumsi *equidispersi* atau nilai varians respon harus sama dengan nilai meannya. Akan tetapi, pada prakteknya asumsi tersebut sulit terpenuhi ketika berlebihnya variabel respon yang bernilai 0 dan dapat mengakibatkan *overdispersi* atau nilai varians respon lebih besar dari nilai mean. Setelah didapatkan model regresi *Poisson*, selanjutnya akan dilakukan pengujian *overdispersi*.

Tabel 3.3 Pengujian *Overdispersi*

	Nilai	Derajat bebas(db)	Statistik Uji
Deviance Residual	36,71	19	1,932105

Berdasarkan hasil dari tabel 3.3 bahwa nilai *Deviance Residual* sebesar 36,71 dan derajat bebasnya 19 , sehinga hasil bagi antara nilai *Deviance Residual* dengan derajat bebas sebesar 1.932105. Nilai tersebut lebih besar dari 1 (satu), maka dapat disimpulkan bahwa terjadi *overdispersi* pada data. karena terjadi masalah overdispersi maka dapat disimpulkan bahwa model regresi *Poisson* tidak cocok digunakan pada data kasus kematian ibu di Provinsi Jawa Barat tahun 2015. Model alternatif yang digunakan untuk memodelkan data tersebut, yaitu model regresi *Hurdle Poisson*.

Hasil Penaksiran parameter model logit pada regresi *Hurdle Poisson* disajikan pada tabel 3.4

Tabel 3.4 Taksiran Parameter Model Logit pada Regresi *Hurdle Poisson*

Parameter	Taksiran	p-value	$\alpha = 0{,}05$	Keterangan
δ_0	10,23365			
δ_1	0,67484	0,1664	0,05	Tidak signifikan
δ_2	-0,46603	0,0607	0,05	Tidak signifikan
δ_3	0,15182	0,1374	0,05	Tidak signifikan
δ_4	-0,80205	0,0363	0,05	Signifikan
δ_5	0,27458	0,4825	0,05	Tidak signifikan
δ_6	0,22069	0,6309	0,05	Tidak signifikan
δ_7	0,02445	0,6466	0,05	Tidak signifikan

Berdasarkan tabel 3.4 dapat diketahui bahwa variabel prediktor yang signifikan dengan tingkat signifikansi 5% pada model logit adalah Persentase ibu hamil mendapatkan tablet Fe1 (X_4). Model logit dapat dikatakan sebagai apakah suatu Kabupaten/Kota di Provinsi Jawa Barat memiliki peluang terjadinya kematian ibu ketika melahirkan dengan diketahui nilai-nilai variabel yang mempengaruhi secara signifikan. Persamaan model logit berdasarkan tabel 3.4 adalah:

$$\log\left(\frac{\pi}{1-\pi}\right) = 10{,}23365 - 0{,}80205 Z_4 \qquad (3.1)$$

Berdasarkan persamaan (3.1), hal ini berarti bahwa setiap penambahan satu persen ibu hamil mendapatkan Fe1, maka akan menurunkan peluang banyak kematian ibu sebanyak exp (0,80205) = 2,23 kali, dengan menganggap variabel lain konstan.

Hasil Penaksiran parameter model logit pada regresi *Hurdle Poisson* disajikan pada tabel 3.5.

Tabel 3.5 Taksiran Parameter Model *Truncated Poisson* pada Regresi *Hurdle Poisson*

Parameter	Taksiran	p-value	$\alpha = 0{,}05$	Keterangan
β_0	-1,341740			
β_1	0,098701	0,05194	0,05	Tidak signifikan
β_2	-0,009181	0,82312	0,05	Tidak signifikan
β_3	0,071861	0,00926	0,05	Signifikan
β_4	0,057737	0,38468	0,05	Tidak signifikan
β_5	-0,176918	0,00205	0,05	Signifikan
β_6	0,032389	0,29229	0,05	Tidak signifikan
β_7	-0,016413	0,15400	0,05	Tidak signifikan

Berdasarkan tabel 3.5 dapat diketahui bahwa variabel prediktor yang signifikan dengan tingkat signifikansi 5% pada model logit adalah Persentase ibu hamil mendapatkan tablet Fe1 (X_4). Model logit dapat dikatakan sebagai apakah suatu Kabupaten/Kota di Provinsi Jawa Barat memiliki peluang terjadinya kematian ibu dengan diketahui nilai-nilai variabel yang mempengaruhi secara signifikan. Persamaan model *Truncated Poisson* berdasarkan tabel 354 adalah:

$$\mu = \exp(-1{,}341740 + 0{,}071861X_3 - 0{,}176918X_5) \qquad (3.2)$$

Berdasarkan persamaan (3.2), hal ini berarti bahwa dapat diinterpretasikan untuk model *Truncated Poisson* sebagai berikut:

1. Variabel persentase wanita kawin dengan umur perkawinan pertama dibawah 18 tahun (X_3)
   Setiap penambahan satu persen wanita kawin dengan umur perkawinan pertama di bawah 18 tahun, maka akan meningkatkan rata-rata banyak kematian ibu sebanyak exp (0.071861) = 1,07 kali, dengan menganggap variabel lain konstan.
2. Variabel persentase ibu hamil mendapatkan Fe3 (X_5)
   Setiap penambahan satu persen ibu hamil mendapatkan Fe3, maka akan menurunkan rata-rata banyak kematian ibu sebanyak exp (0,176918) = 1,19 kali, dengan menganggap variabel lain konstan.

4. KESIMPULAN

Berdasarkan bahasan pada bab sebelumnya, diperoleh kesimpulan sebagai berikut :

1. Kajian teoritis model regresi *Hurdle Poisson,* Model regresi *Hurdle Poisson* merupakan model regresi yang digunakan ketika variabel respon merupakan data cacahan (*count*) dan dapat digunakan dalam keadaan overdispersi.
   Distribusi peluang model *Hurdle* sebagai berikut:
   $$P(Y_i = y) = \begin{cases} f_1(0), & y = 0 \\ \theta f_2(y), & y = 1,2,3,\ldots \end{cases}$$
   Distribusi peluang *Hurdle Poisson* dapat ditulis sebagai berikut:
   $$P(Y_i = y_i) = \begin{cases} \dfrac{1}{1+\exp\left(\sum_{j=1}^{p} z_{ij}\delta_j\right)} & , y_i = 0 \\ \left[\dfrac{\exp\left(\sum_{j=1}^{p} z_{ij}\delta_j\right)}{1+\exp\left(\sum_{j=1}^{p} z_{ij}\delta_j\right)}\right]\left[\dfrac{\left[exp\left(\sum_{j=1}^{p} x_{ij}\beta_j\right)\right]^{y_i}}{\left[exp\left(exp\left(\sum_{j=1}^{p} x_{ij}\beta_j\right)\right)-1\right] y_i!}\right] & , y_i > 0 \end{cases}$$
   Model regresi *Hurdle Poisson* terbagi menjadi dua bagian yaitu:
   a. Model untuk data biner yang bernilai nol (*zero counts*) atau nilai positif (*positive counts*), dimana model tersebut ditaksir dengan model logit.
      Model Logit dapat ditulis sebagai berikut:
      $$logit(\pi_i) = log\left(\frac{\pi_i}{1-\pi_i}\right) = \sum_{j=1}^{p} z_{ij}\delta_j$$
   b. Model untuk data yang bernilai positif (*positive counts*) saja, dimana model tersebut ditaksir dengan model *Truncated Poisson*.
      Model *Truncated Poisson* dapat ditulis sebagai berikut:

$$\log(\mu_i) = \sum_{j=1}^{p} x_{ij}\beta_j$$

$$\mu_i = exp\left(\sum_{j=1}^{p} x_{ij}\beta_j\right)$$

Penaksir parameter untuk model regresi *Hurdle Poisson* dapat diperoleh denggan menggunakan algoritma *Broyden-Fletcher-Goldfarb-Shanno* (BFGS). Pengujian signifikansi parameter secara serempak (*Overall*) dengan menggunakan uji rasio kemungkinan (*Likelihood Ratio Test*) dan pengujian signifikansi masing-masing parameter secara parsial dengan uji Wald.

2. Penerapan model regresi *Hurdle Poisson,* Pada data banyaknya kasus kematian ibu di Jawa Barat tahun 2015 diperoleh informasi bahwa terjadi *overdispersi*, karena banyaknya observasi pada variabel respon yang bernilai nol (*excess zeros*) maka dapat disimpulkan model yang cocok adalah model regresi *hurdle poisson* pada kasus tersebut.

   Model *Hurdle Poisson* terbagi menjadi dua bagian yaitu model Logit dan model *Truncated Poisson,* sehingga model tersebut dapat ditulis sebagai berikut:

   a. Model Logit:
   $\log\left(\frac{\pi}{1-\pi}\right) = 10{,}23365 + 0{,}67484Z_1 - 0{,}46603Z_2 + 0{,}15182Z_3 - 0{,}80205Z_4 + 0{,}27458Z_5 + 0{,}22069Z_6 + 0{,}02445Z_7$, tetapi setelah dilakukan serangkaian uji, hanya X_4 yang merupakan variabel yang signifikan yaitu variabel persentase ibu hamil mendapatkan Fe1. Sehingga model Logit yang diperoleh secara signifikan adalah sebagai berikut:

   $$\log\left(\frac{\pi}{1-\pi}\right) = 10{,}23365 - 0{,}80205X_4$$

   b. Model *Truncated Poisson*:
   $\log(\mu) = -1{,}341740 + 0{,}098701X_1 - 0{,}009181X_2 + 0{,}071861X_3 + 0{,}057737X_4 - 0{,}176918X_5 + 0{,}032389X_6 - 0{,}016413X_7$, tetapi setelah dilakukan serangkaian uji, ada dua variabel yaitu X_3 dan X_5 yang merupakan variabel yang signifikan yaitu variabel persentase wanita kawin dengan umur perkawinan pertama dibawah 18 tahun dan variabel persentase ibu hamil mendapatkan Fe3. Sehingga model *Truncated Poisson* yang diperoleh secara signifikan adalah sebagai berikut:

   $$\log(\mu) = -\,1{,}341740 + 0{,}071861X_3 - 0{,}176918X_5$$

5. DAFTAR PUSTAKA

Badan Pusat Statistik, P. (2015). *Indikator Kesejahteraan Rakyat Provinsi Jawa Barat.* Bandung: BPS Provinsi Jawa Barat.

Jinnan, I. (2016). *Model Regresi Poisson Hurdle untuk Mengatasi Overdispersi akibat Excess Zeros.* Skripsi Program Studi Statistika. Yogyakarta: Universitas Gajah Mada.

Kementerian Kesehatan, R. I. (2015). *Profil Kesehatan Indonesia 2014.* Jakarta: Kementerian Kesehatan.

Saffari, S., Adnan, R., & Greene, W. (2012). Parameter Estimation on Hurdle Poisson Regression Model with Censored Data. *Jurnal Teknologi* , 189-198.

Zia, R. (2016). *Pemodelan Regresi Hurdle Poisson untuk Menentukan Faktor yang Mempengaruhi Banyaknya Kasus Difteri di Jawa Barat.* Skripsi Program Studi Statistika. Bandung: Universitas Padjadjaran.

Zorn, C. (1996). *Evaluating Zero Inflated and Hurdle Poisson Specifications.* Ohio State University: Midwest Political Science Assosiation.