PEMODELAN ANGKA KEMATIAN BAYI DI PROVINSI JAWA BARAT MENGGUNAKAN METODE REGRESI POISSON *INVERSE GAUSSIAN (PIG)*

Wisnu Sofyan[1], Indah Manfaati Nur[2], Tiani Wahyu Utami[3]
[123]Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Muhammadiyah Semarang
Email: wisnusofyan23@gmail.com

Abstrak

Angka Kematian Bayi (AKB) adalah salah satu contoh data *count* (data cacahan). Pemodelan data *count* dapat menggunakan regresi poisson. Terdapat asumsi yang harus dipenuhi jika menggunakan regresi poisson yaitu mean dan varians harus sama, sedangkan pada kasus data cacahan asumsi ini sering tidak terpenuhi. Regresi poisson *inverse gaussian* (PIG) merupakan salah satu bentuk regresi dari mixed poisson yang dirancang untuk data cacahan dengan kasus overdispersi dan telah digunakan pada beberapa penelitian yang menggunakan data cacahan. Oleh karena itu pemodelan angka kematian bayi di Provisni Jawa Barat pada tahun 2017 dapat digunakan pemodelan dengan regresi PIG. Adapun variabel yang di gunakan dalam penelitian ini adalah Angka Kematian Bayi (AKB) di Provinsi Jawa Barat, sedangkan untuk variabel penyertaanya yaitu presentase berat badan bayi lahir rendah (PBBLR), presentase persalinan dalam kurun waktu tertentu yang ditolong oleh tenaga kesehatan (PPSLN), persentase penduduk miskin menurut Kabupaten/Kota (PPM), persentase rumah tangga yang berprilaku bersih dan sehat menurut Kabupaten/Kota (PHBS), presentase kehamilan resiko tinggi (PKRT), persentase pelayanan kesehatan bayi yang ada di Kabupaten/Kota (PPKB) dari website BPS dan Profil Kesehatan Provinsi Jawa Barat dengan hasil analisis dan pembahasan yang dilakukan maka diperoleh model regresi Poisson Inverse Gaussian dengan variabel yang signifikan terhadap model yaitu presentase berat badan lahir rendah (PBBLR) dan peresentasi ibu hamil resiko tinggi (PKRT). Penambahan 1 rasio pada peresentase berat badan bayi lahir rendah (PBBLR) akan sebbanding dengan kenaikan rata-rata jumlah angka kematian bayi di Provinsi Jawa Barat sebesar 1.14140. Daan dengan penambahan 1 rasio pada peresentase ibu hamil resiko tinggi (PKRT) maka akan sebanding dengan kenaikan rata-rata jumlah angka kematian bayi di Provinsi Jawa Barat sebesar 1.24485.

Kata Kunci: *Angka Kemattian Bayi, Ditribusi Poisson, Overdispersi, Regresi PIG*

Abstract

Infant Mortality Rate (IMR) is one example of count data (data count). Count data modeling can use Poisson regression. There are assumptions that must be met when using Poisson regression, namely that the mean and variance must be the same, whereas in the case of enumerated data this assumption is often not fulfilled. Gaussian inverse poisson regression (PIG) is a form of mixed poisson regression designed for count data with overdispersion cases and has been used in several studies using counted data. Therefore, modeling the infant mortality rate in West Java Province in 2017 can be used modeling with PIG regression. The variables used in this study are the Infant Mortality Rate (IMR) in West Java Province, while the accompanying variables are the percentage of low birth weight babies (PBBLR), the percentage of deliveries in a certain period of time assisted by health workers (PPSLN), percentage of poor population by Regency / City (PPM), percentage of households with clean and healthy behavior according to Regency / City (PHBS), percentage of high risk pregnancies (PKRT), percentage of infant health services in the Regency / City (PPKB) from the website BPS and the Health Profile of West Java Province with the results of the analysis and discussion carried out, the Poisson Inverse Gaussian regression model was obtained with significant variables to the model, namely the percentage of low birth weight (PBBLR) and the percentage of high-risk pregnant women (PKRT). The addition of 1 ratio to the percentage of low birth weight (PBBLR) will be proportional to the increase in the average number of infant mortality rates in West Java Province of 1.14140. And with the addition of 1 ratio to the percentage of high risk pregnant women (PKRT), it will be *proportional to the increase in the average number of infant mortality rates in West Java Province of 1.24485.*

Keywords: *Infant Mortality Rate, Poisson Distribution, Overdispersion, PIG Regression*

PENDAHULUAN

Data cacahan *(count data)* adalah data yang menggambarkan sejumlah kejadian yang terjadi pada suatu kurun waktu tertentu. Pemodelan data cacahan tidak dapat dilakukan dengan menggunakan regresi OLS (*Ordinary Least Square*), karena pemodelan data cacahan akan melanggar dua asumsi yang disyaratkan dalam regresi OLS yaitu, error mengikuti distribusi normal (normalitas) dan memiliki sifat homokedastisitas (varians konstan). Pelanggaran asumsi varians konstan terjadi karena dalam data cacahan sering terjadi peningkatan varians bersyarat sebagai akibat peningkatan nilai prediktor. Inilah yang disebut sebagai heterokedastisitas yang dapat mengakibatkan standar error dan uji signifikansi yang bias bila menerapkan regresi OLS (Coxe, West dan Aiken, 2009).

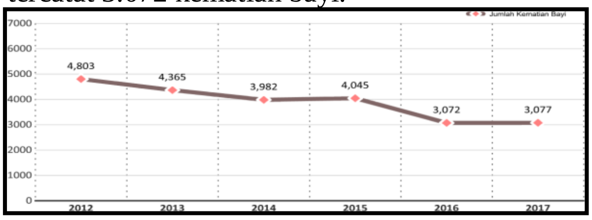
Pengembangan dalam pemodelan data cacahan memunculkan pemodelan data cacahan dengan *Generalized Linear Models* (GLMs). GLMs merupakan generalisasi dari model regresi normal klasik atau regresi OLS dari asumsi yang ketat dan menyediakan metode analisis bagi data tidak normal (DeJong dan Heller, 2008). 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 besardari *mean* (*overdispersion*). Pada kebanyakan data *count* terkadang ditemukan kasus overdispersi (Consul dan Famoye, 1992).

Berbagai eksperimen, seringkali data penelitian yang dimiliki berupa data cacahan. Model yang dapat digunakan untuk data cacahan diantaranya adalah model regresi Poisson. Regresi Poisson termasuk kedalam *Generalized Linear Model* (GLM) (Kusuma, Komalasari, Hadijati, 2013). Pada model regresi ini, peubah respon berupa data cacahan yang mengikuti distribusi Poisson. Distribusi Poisson sering digunakan untuk kejadian kejadian yang jarang terjadi dengan data berupa cacahan yang mempunyai nilai non negatif (Kismiantini, 2008).

Model regresi Poisson mensyaratkan equidispersi yaitu harus memenuhi asumsi bahwa nilai varinasi dari variabel respon harus sama dengan nilai rata-ratanya (Cahyandari, 2014). Sementara dalam penelitian regresi poisson biasanya terjadi pelanggaran terhadap asumsi tersebut. Dimana nilai variansi lebih kecil daripada rata-ratanya yang disebut dengan *underdispersi* dan ketika nilai variansinya lebih besar dari rata-ratanya yang disebut dengan *Overdispersi*. Apabila data dengan kondisi overdispersi tetap dianalisis menggunakan analisis regresi Poisson, maka akan ada informasi yang hilang akibat tidak termodelkannya parameter disperse dalam model regresi yang terbentuk. Parameter dispersi adalah parameter yang muncul akibat tidak terjadinya kondisi ekuidispersi (Keswari, Sumarjaya, Suciptawati, 2014).

Kasus *overdispersi* bila diabaikan bisa mengakibatkan terjadinya *underestimate* pada estimasi standar 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). 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*). Namun 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. Akan tetapi, distribusi PIG memiliki bentuk fungsi likelihood yang *close form* dan penghitungannya lebih mudah sehingga banyak penelitian yang melibatkan data cacahan banyak yang menggunakan model ini (Stasinopoulus dan Rigby, 2007).

Angka Kematian Bayi (AKB) atau *Infant Mortality Rate (IMR)* merupakan indikator yang sangat sensitif terhadap upaya pelayanan kesehatan terutama yang berhubungan dengan bayi baru lahir perinatal dan neonatal. AKB menggambarkan besarnya risiko kematian bayi (<1 tahun) dalam 1.000 kelahiran hidup. Berdasarkan kesepakatan international AKB merupakan indikator yang menggunakan konsep *rate*, meskipun dalam kenyataannya hanya *ratio*. Berdasarkan publikasi BPS, AKB Provinsi Jawa Barat sejak tahun 2007 sampai dengan 2012 cenderung mengalami penurunan. AKB berhasil diturunkan sebesar 9 poin (range 39 – 30/1.000 kelahiran hidup). Untuk AKB ditahun 2012, BPS melakukan publikasi berdasarkan SDKI 2012, di mana Provinsi Jawa Barat mempunyai AKB sebesar 30/1.000 kelahiran hidup. Berdasarkan pencatatan dan pelaporan, di Provinsi Jawa Barat tahun 2017 terdapat 3.077 bayi meninggal meningkat 5 orang dibanding tahun 2016 yang tercatat 3.072 kematian bayi.



Sumber : Tabel Profil Kesehatan Provinsi Jawa Barat Tahun 2017

Gambar 1.1 Angka Kematia Bayi di Provinsi Jawa Barat Tahun 2012-2017

Berdasarkan pada grafik Proporsi Kematian Bayi pada tahun 2017 sebesar 3,4/1000 kelahiran hidup, menurun 0,53 poin dibanding tahun 2016 sebesar 3,93/1000 kelahiran hidup. Dari kematian bayi sebesar 3,4/1.000 kelahiran hidup, terdapat angka kematian neonatal (bayi berumur 0-28 hari) sebesar 3,1/1.000 kelahiran hidup atau 84,63 % kematian bayi berasal dari bayi usia 0-28 hari, dengan demikian disarankan dalam penangan AKB lebih difokuskan pada Bayi Baru Lahir. Angka Kematian Bayi sebesar 3,4/1000 kelahiran hidup, sudah melampaui target MDGs yang pada tahun 2015 harus sudah mencapai 17/1.000 kelahiran hidup.

Pada penelitian ini yang akan diteliti yaitu jumlah kematian bayi di Provinsi Jawa Barat. Dimana jumlah kematian bayi merupakan salah satu indikator penting bagi pemerintah dalam mengevaluasi dibidang kesehatan. Mengetahui faktor-faktor yang mempengaruhi, tentu saja pemerintah memiliki gambaran langkah apa saja yang harus dilakukan dalam menekan angka kematian bayi.

Data kematian bayi merupakan data cacahan sehingga dalam pemodelannya bisa menggunakan regresi Poisson. Data kematian bayi juga berpotensi mengalami *overdispersi* sehingga dalam penanganannya diperlukan sebuah model regresi yang dapat dilakukan untuk data yang mengalami *overdispersi*. Berdasarkan penelitian yang telah dilakukan oleh beberapa peneliti, maka model regresi yang akan diterapkan pada penelitian ini yaitu model regresi Poisson Inverse Gaussian.

TINJAUAN PUSTAKA

1. Distribusi Poisson

Distribusi Poisson merupakan suatu distribusi untuk peristiwa yang probabilitas kejadiannya kecil, dimana kejadian tergantung pada selang waktu tertentu atau di suatu daerah tertentu dengan hasil pengamatan berupa variabel diskrit dan antar variabel prediktor saling independen. Selang waktu tersebut dapat berupa berapa saja panjangnya, misalnya semenit, sehari, seminggu, sebulan, bahkan setahun. Daerah tertentu yang dimaksudkan dapat berupa suatu garis, suatu luasan, suatu volume, atau mungkin sepotong bahan (Walpole, 1995).

2. Distribusi *Inverse* Gaussian

Distribusi inverse gaussian merupakan distribusi kontinu dengan fungsi kepadatan mirip dengan distribusi gamma tapi dengan kemencengan lebih besar dan keruncingan tajam. Inverse gaussian digunakan pada keadaan dengan kemencengan yang ekstrem. Nama inverse gaussian sendiri berasal dari fungsi kumulant yang memiliki hubungan invers dengan fungsi kumulant (logaritma natural dari fungsi MGF) distribusi normal/distribusi Gaussian (De Jong dan Heller, 2008).

3. Overdispersi

Model regresi poisson terdapat beberapa asumsi yang harus dipenuhi. Salah satunya adalah asumsi kesamaan antara rata-rata (mean) dan variansinya (variance) yang disebut dengan ekuidispersi (Darnah, 2011). Namun dalam analisis data statistika sering dijumpai data yang variansinya lebih kecil atau lebih besar dari rataannya. Keadaan ini disebut dengan underdispersi (underdispersion) atau overdispersi (*overdispersion*). Salah satu penyebab terjadinya *overdispersion* adalah terlalu banyak nilai nol (excess zero) pada variabel respon (Kusuma, Komalasari, Hadijati, 2013). *Overdispersi atau underdispersi* dapat menyebabkan taksiran parameter yang diperoleh tidak efisien. Penggunaan yang tidak tepat pada model regresi poisson (yang mengalami *overdispersi* atau *underdispersi*) dapat berakibat fatal dalam interpretasi model, terutama pada estimasi parameter model karena dapat menaksir *standard error* yang terlalu rendah dan dapat memberikan kesimpulan yang keliru tentang signifikan atau tidaknya parameter regresi yang terlibat (Darnah, 2011).

4. Distribusi Poisson Inverse Gaussian

Distribusi Poisson Inverse Gaussian merupakan salah satu distribusi mixed poisson yang ditentukan oleh dua parameter yaitu rata-rata (μ) sebagai parameter lokasi dan parameter disperse (τ) sebagai parameter bentuk (Herindrawati, Latra, Purhadi, 2017).

5. Regresi Poisson Inverse Gaussian

Model regresi Poisson Inverse Gaussian dengan fungsi kepadatan peluang sebagai berikut:

$$\mathrm{P}(\mathrm{Y}=\mathrm{y}|\mathrm{x}_{\mathrm{i}};\beta;\tau)=\left\{\frac{e^{x_i^T\beta y_i}e^{\frac{1}{\tau}}}{y_i!}\left(\frac{2}{\pi\tau}\right)^{\frac{1}{2}}(2e^{x_i^T\beta y_i}+1)^{-\frac{\left(y_i-\frac{1}{2}\right)}{2}}K_{si}(Z_i)\right\}$$

6. Estimasi Parameter Regresi Poisson Inverse Gaussian

Persamaan 1 merupakan model regresi PIG dan parameter β pada regresi PIG ditaksir dengan metode *maximum likelihood*. Langkahnya adalah dengan menentukan fungsi likelihood dari distribusi PIG sebagai berikut:

$$L(\beta;\tau)=\prod_{i=1}^{n}P(Y=y_i|x_i;\beta;\tau)$$

$$L(\beta;\tau)=\prod_{i=1}^{n}\left\{\frac{\mu_{\mathrm{i}}{}^{y_i}e^{\frac{1}{\tau}}}{y_i!}\left(\frac{2}{\pi\tau}\right)^{\frac{1}{2}}(2\mu_i\tau+1)^{-\frac{\left(y_i-\frac{1}{2}\right)}{2}}K_{si}(\mathrm{Z}_i)\right\}$$

Fungsi likelihood tersebut diubah bentuk logaritma natural (ln) sehingga persamaannya menjadi sebagai berikut:

$$L(\beta;\tau)=\ln L(\beta;\tau)$$

$$=\sum_{i=1}^{n}y_ix_i^T\beta+\frac{n}{\tau}-\ln\left(\sum_{i=1}^{n}y_i!\right)+\frac{n}{2}\ln\left(\frac{2}{\pi}\right)-\frac{n}{2}\ln\tau$$

$$-\sum_{i=1}^{n}\left(\frac{2y_i-1}{4}\right)\ln(2x_i^T\beta+1)+\sum_{i=1}^{n}\ln K_{si}(Z_i)$$

Apabila persamaan di atas implisit dan nonlinear dalam parameter β dan τ sehingga untuk mendapatkan taksiran dari $\theta=\left[\beta^T\tau\right]^T$ fungsi dimaksimumkan dengan menggunakan *Fisher Scoring Algorithm* dengan persamaan berikut:

$$\hat{\theta}_{(r+1)}=\hat{\theta}_{(r)}+\mathrm{I}^{-1}(\hat{\theta}_{(m)})D(\hat{\theta}_{(m)})$$

Dimana:

$$\hat{\theta}=\left[\hat{\beta}^T\hat{\tau}\right]^T$$

$$D(\hat{\theta})=\left(\frac{\partial l}{\partial\hat{\tau}}\cdot\frac{\partial l}{\partial\hat{\beta}^T}\right)$$

$$\mathrm{I}(\hat{\theta}_{(m)})=-E\left[H(\hat{\theta}_{(m)})\right]$$

$$\mathrm{I}(\hat{\theta}_{(m)})_{(k+1)(k+1)}=\begin{bmatrix}\frac{\partial^2 l}{\partial\hat{\tau}^2} & \frac{\partial^2 l}{\partial\hat{\tau}\partial\hat{\beta}^T}\\ \frac{\partial^2 l}{\partial\hat{\beta}^T\partial\hat{\tau}} & \frac{\partial^2 l}{\partial\hat{\beta}\partial\hat{\beta}^T}\end{bmatrix}$$

Sehingga

$$\mathrm{I}(\hat{\theta}_{(m)})=-E\begin{bmatrix}\frac{\partial^2 l}{\partial\hat{\tau}^2} & \frac{\partial^2 l}{\partial\hat{\tau}\partial\hat{\beta}^T}\\ \frac{\partial^2 l}{\partial\hat{\beta}^T\partial\hat{\tau}} & \frac{\partial^2 l}{\partial\hat{\beta}\partial\hat{\beta}^T}\end{bmatrix}$$

Matriks hessian merupan matriks yang berisi turunan kedua dari fungsi likelihood terhadap parameter β dan τ.

7. Pengujian Parameter

Pengujian parameter pada model Poisson Inverse Gaussian dilakukan dengan menggunakan pengujian hipotesis secara serentak pada parameter β serta pengujian parsial pada parameter β dan τ. Pengujian parameter dilakukan untuk mengetahui ada atau tidaknya pengaruh variabel prediktor terhadap variabel repon.

1. Uji secara serentak (Simultan)

   Langkah-langkah pengujian yang dilakukan secara serentak yaitu:

   H_0 : $\beta_1=\beta_2=\ldots=\beta_k$ (variabel prediktor secara simultan tidak berpengaruh terhadap variabel respon)

H$_1$: minimal ada satu $\beta_i \neq 0$ dengan i = 1, 2, …, k. (minimal ada satu variabel prediktor yang berpengaruh terhadap variabel respon)

Statistik uji yang digunakan adalah ukuran statistik *likelihood ratio* yang dibentuk dengan menentukan himpunan parameter dibawah populasi (Ω) yaitu $\Omega = (\beta,\tau)$ dan himpunan parameter dibawah H$_0$ benar (ω) yaitu $\omega = (\beta_0,\tau_0)$. Pada himpunan parameter dibawah populasi, dibentuk fungsi *likelihood* untuk model penuh (*saturated*) yang melibatkan seluruh variabel prediktor $L(\Omega)$. Sedangkan pada himpunan parameter dibawah H$_0$ benar, dibentuk fungsi *likelihood* untuk model yang tidak melibatkan variabel prediktor $L(\omega)$.

Kedua fungsi diatas dibandingkan dalam bentuk devians berikut:

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

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

Statistik G adalah pendekatan dari distribusi *chi square* dengan derajat bebas v sehingga kriteria pengujiannya adalah tolak H$_0$ apabila $G_{hit} > x^2_{(\alpha,v)}$ dimana v adalah derajat bebas yang diperoleh dari jumlah parameter dibawah populasi dikurangi jummlah parameter dibawah H$_0$

2. Uji secara individu (Parsial)

   Pengujian hipotesis secara parsial (Individu) melalui kriteria keputusan penolakan H$_0$. Hipotesis yang digunakan adalah:

   Hipotesis pengujian parameter β:

   H$_0$: $\beta_j = 0$, j = 1, 2, …,k (variabel prediktor ke-j tidak berpengaruh terhadap variabel respon)

   H$_1$: $\beta_j \neq 0$, j = 1, 2, …,k (variabel prediktor ke-j berpengaruh terhadap variabel respon)

Uji statistik yang digunakan dalam pengujian signifikan parameter β adalah:

$$Z = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)}$$

Kriteria penolakan H$_0$ apabila $|Z_{hit}| > Z_{\frac{\alpha}{2}}$ atau

p-value < α dimana α adalah tingkat signifikan yang digunakan dan $SE(\hat{\beta}_j)$ merupakan elemen diagonal yang diperoleh dari elemen diagonal utama ke-(m+20) dari matrik varians dan covarians yang diperoleh dari:

$$\widehat{Cov}(\hat{\theta}) = -\left(H^{-1}(\hat{\theta})\right)$$

Hipotesis pengujian parameter τ:

H$_0$: τ = 0 (variabel prediktor tidak berpengaruh terhadap variabel respon)

H$_1$: τ ≠ 0 (variabel prediktor berpengaruh terhadap variabel respon)

Statistik uji yang digunakan:

$$Z = \frac{\hat{\tau}}{SE(\hat{\tau})}$$

Kriteria penolakan H0 apabila $|Z_{hit}| > Z_{\frac{\alpha}{2}}$ atau

p-value < α dimana α adalah tingkat signifikan.

8. Akaike Informantion Criteria (AIC)

Metode AIC adalah satu metode yang dapat digunakan untuk memilih model regresi terbaik yang ditemukan oleh Akaike afn Schwarz. Metode tersebut didasarkan pada metode *maximum likelihood estimation* (MLE) (Fathurahman, 2009). Untuk menghitung nilai AIC dapat digunakan rumus berikut:

$$AIC = e^{\frac{2k}{n}}\frac{\sum_{i=1}^{n}\hat{u}_i^2}{n} \text{ atau } \ln AIC = \frac{2k}{n} + \left(\frac{\sum_{i=1}^{n}\hat{u}_i^2}{n}\right)$$

Dengan:

k = jumlah parameter yang di estimasi dalam model regresi

n = jumlah observasi

e = sisa (residual)

METODE PENELITIAN

Data yang digunakan dalam penelitian ini adalah data sekunder tentang jumlah angka kematian bayi (AKB) dan faktor-faktor yang mempengaruhi jumlah angka kematian pada bayi. Data yang diambil dari website BPS Provinsi Jawa Barat tahun 2017 dan Profil Kesehatan Provinsi Jawa Barat tahun 2017 dari hasil publikasi dari Dinas Kesehatan Jawa Barat Tahun 2017 dengan unit pengamatan yang diambil yaitu pada tingkat Kabupaten / Kota di Provinsi Jawa Barat dengan 27 Kabupaten / Kota. Adapun variabel yang digunakan pada penelitian ini terdapat beberapa variabel yang diambil pada penelitiaan sebelumnya sebagai berikut:

Tabel 1 Keterangan Variabel Penelitian Dan Tipe Data

VARIABEL	KETERANGAN
Y	Jumlah Angka Kematian Bayi 2017.
X1	Persentase Berat Badan Bayi Lahir Rendah Tahun 2017
X2	Persentase Persalinan dalam kurun waktu tertentu yang ditolong oleh tenaga kesehatan Tahun 2017

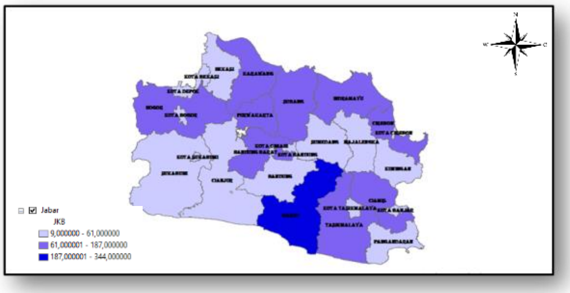
X3	Persentase Penduduk Miskin menurut Kabupaten/Kota Pada Tahun 2017
X4	Persentase Rumah tangga yang berprilaku bersih dan sehat menurut Kabupaten/Kota Pada Tahun 2017
X5	Persentase Kehamilan Resiko Tinggi pada Tahun 2017
X6	Persentase Pelayanan Kesehatan Bayi Yang Ada Di Kabupaten/Kota Pada Tahun 2017

Langkah-langkah analisis yang digunakan dalam penelitian ini adalah sebagai berikut.

1. Membuat gambaran pola sebaran jumlah angka kemaatian bayi
2. Menguji korelasi antara variabel respon dengan variabel predictor
3. Melakukan pemeriksaan multikolinearitas dengan menggunakan kriteria VIF
4. Melakukan uji overdispersi
5. Menentukan nilai penaksiran model Regresi PIG
6. Melakukan pengujian hipotesis untuk Regresi PIG
7. Membandingkan nilai AIC untuk mencari model terbaik
8. Melakukan interpretasi model PIGR yang didapatkan
9. Membuat kesimpulan.

HASIL DAN PEMBAHASAN

1. Gambaran Pola Sebaran Angka Kematian Bayi di Provinsi Jawa Barat



Gambar 1 Pola Sebaran Angka Kematian Bayi di Provinsi Jawa Barat Tahun 2017

Berdasarkan gambar 1 menunjukan persebaran angka kematian bayi di Provinsi Jawa Barat tahun 2017. Dapat diketahui bahwa wilayah dengan warna semakin gelap menunjukan angka kematian bayi semakin tinggi. Terlihat bahwa kabupaten/kota yang memiliki angka kematian bayi tinggi berkisar antara 187,001 sampai 344,000. Berdasarkan hasil analisis terdapa 1 Kabupaten/kota yang memiliki jumlah angka kematian bayi tinggi. Adapun Kabupaten/Kota tersebut adalah Garut. Kabupaten/kota yang memiliki angka kematian bayi sedang berkisar 61,0001 sampai 187,000. Berdasarkan hasil analisis terdapat 15 Kabupaten/kota yang memiliki angka kematian bayi sedang. Adapun Kabupaten/Kota tersebut adalah Kota Tasikmaya, Ciamis, Tasikmaya, Kota Banjar, Kota Cirebon, Cirebon, Indramayu, Kota Bandung, Bnadung Barat, Kota Cimahi, Purwakarta, Subang, Karawang, Kota Bogor, Bogor. Kabupaten/kota yang memiliki angka kematian bayi rendah berkisar antara 9,00000 samai 61,0000. Berdasarkan hasil analisis terdapat 11 Kabupaten/kota yang memiliki angka kematian bayi rendah. Adapun Kabupaten/Kota tersebut adalah Pangandaran, Kuningan, Majalengka, Semedang, Bandung, Cianjur, Sukabumi, Kota Sukabumi, Kota Depok, Kota Bekasi, Bekasi.

2. Uji Overdispersi dan Contoh Terapan

Sebagai contoh penerapan regresi poisson *inverse gaussian* di gunakan data suatu penelitian di Provinsi Jawa Barat yang diambil dari website BPS Provinsi Jawa Barat tahun 2017 dan Profil Kesehatan Provinsi Jawa Barat tahun 2017 dari hasil publikasi dari Dinas Kesehatan Jawa Barat Tahun 2017 tentang angka kematian bayi (AKB) dan faktor-faktor yang mempengaruhi jumlah angka kematian pada bayi dengan unit pengamatan yang diambil yaitu pada tingkat Kabupaten / Kota di Provinsi Jawa Barat dengan 27 Kabupaten / Kota. dengan unit pengamatan yang diambil yaitu pada tingkat Kabupaten / Kota di Provinsi Jawa Barat dengan 27 Kabupaten / Kota. Untuk menganalisis model regresi poisson *inverse gaussian* maka terlebih dahulu dilakukan pengecekan asumsi bahwa variabel respon Y_i berdistribusi poisson menggunakan uji *kolmogorov-semirnov.*

Tabel 2 *One-Sampel Kolmogorov-Semirnov Test*

		AKB (Y)
N		27
Poisson Prameter (a,b)	Mean	80.93
Most	Absolute	0.202
Extream	Positive	0.202
Differences	Negatif	-0.159
Kolmogorov-Semirnov Z		0.202
Asymp. Sig. (2.tailed)		0.126

Analisis Output untuk uji *komogorov-semirnov* adalah sebagai berikut:
Hipotesis:
H_0 : F(X) = F_0(X) (Data berdistribusi Poisson)

H_1 : F(X) $\neq$ F_0(X) (Data tidak berdistribusi Poisson)
Tingkat signifikansi α = 1%

Statistik Uji:
$D = max|F(x) - F_0(X)|$

Tolak H_0 JIKA D $\geq$ D*(α), dimana D*(α) merupakan nilai krisis yang diperoleh dari tabel "Klmogorov-Semirnov", atau menggunakan tolak H_0 jika *p-value* $\leq$ α.

Berdasarkan hasil output pada Tabel 4.4 diperoleh nilai *asymp. Sig (2-tailed)* = 0.126, jika dibandingkan dengan α = 0.01 maka nilai *p-value* = 0.126 > α = 0.126 yang berarti H_0 diterima. Sehingga dapat sisimpulkan bahwa variabel respon AKB (Y) adalah berdistribusi Poisson.

Tabel 3 Overdispersi Model Regresi Poisson

Goodness of Fit			
Criterion	DF	Value	Value/DF
Deviance	27	1383.833	53.244
Scaled Deviance	27	1383.833	
Pearson Chi-Square	27	1663	63.962
Scaled Pearson X2	27	1663	
Log Likelihood		771.401	

Berdasarkan hasil analisis overdispersi pada model regresi poisson nilai *Pearson Chi-Square* model dibagi derajat bebasnya adalah 63,962. Nilai tersebut lebih besar dari satu yang berarti dapat disimpulkan model mengalami overdispersi. Oleh karena itu pada kasus angka kematian bayi di Provinsi Jawa Barat pada tahun 2017 telah memenuhi asusmsi sehingga dapat dilajutkan untuk pengujian overdispersi menggunakan metode regresi poisson *inverse gaussian.*

Tabel 4 Overdispersi Model Regresi Poisson *Inverse Gaussian (PIG)*

Overdispersion	
Data	
z = 3.9429	p-value = 4.025e-05
Sampel estimates:	
dispersion : 26.55784	

Pada hasil uji overdispersi menggunakan *package* AER dari *software* R- *Studio* (dapat dilihat pada lampiran 1) juga menunjukkan jika terjadi overdispersi karena nilai p – *value* = 0,00004025 dimana nilai α sebesar 1% sehingga data penelitian mengalami overdispersi akan dilanjutkan menggunakan langkah meode berikutnya yaitu menggunakan Regresi Poisson Inverse Gaussian.

3. Menentukan Penaksiran Model Regresi Poisson Inverse Gaussian

Regresi Poisson Inverse Gaussian (PIG) merupakan regresi yang diaplikasikan pada data yang mengalami overdispersi. Pemodelan regresi poisson inverse gaussian dapat dilakukan setelah pengujian asumsi multikolinearitas. Data yang digunakan untuk penerapan Regresi Poisson Inverse Gaussian adalah data jumlah angka kematian bayi di Provinsi Jawa Barat pada tahun 2017. Sesuai dengan hasil pengujian yang sudah dilakukan, didapatkan bahwa data mengalami overdispersi dan tidak terdapat multikolinearitas, sehingga asumsi untuk melakukan pemodelan dengan regresi Poisson Inverse Gaussian (PIG) sudah terpenuhi.

Berdasarkan enam variabel yang digunakan yaitu persentase berat badan bayi lahir rendah (X1), persentase persalinan dalam kurun waktu tertentu di suatu wilayah yang ditolong oleh tenaga kesehatan (X2), persentase penduduk miskin (X3), persentase rumah tangga yang berprilaku bersih dan sehat menurut kabupaten/kota (X4), persentase kehamilan resiko tinggi (X5), persentase pelayanan kesehatan bayi yang ada di kabupaten/kota (X6), menghasilkan lima kombinasi model yang sudah konvergen. Kelima kombinasi kemungkinan model regresi Poisson Inverse Gaussian yang sudah konvergen, kemudian dicari model terbaiknya (kelima model dapat dilihat pada lampiran 2). Berikut merupakan keempat kemungkinan model Poisson Inverse Gaussian yang sudah konvergen:

$\mu = \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 6_6)$
$\mu = \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 6_6)$
$\mu = \exp(\beta_0 + \beta_1 x_1 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 6_6)$
$\mu = \exp(\beta_0 + \beta_1 x_1 + \beta_4 x_4 + \beta_5 x_5)$
$\mu = \exp(\beta_0 + \beta_1 x_1 + \beta_5 x_5)$

Sebelum menentukan model terbaiknya, maka perlu dilakukan penaksiran parameter, dan pengujian parameter. Pengujian parameter dilakukan secara serentak dan secara individu. Berikut merupakan estimasi parameter dari model-model yang mungkin menjadi model terbaik dalam regresi Poisson Inverse Gaussian yang ditunjukkan pada Tabel 4.4 (lebih lanjut dapat dilihat pada lampiran 2).

Tabel 5 Estimasi Parameter Kemungkinan Model Regresi Poisson Inverse Gaussias

Variabel Dari Model	β_0	β_1	β_2	β_3	β_4	β_5	β_6	τ
X1,X2,X3,X4,X5,X6	2.561313	0.173363	-0.003626	-0.005963	0.009189	0.254777	0.001147	-0.8616
X1,X2,X4,X5,X6	2.515175	0.162223	-0.003710		0.008549	0.251869	0.001868	-0.8551
X1,X4,X5,X6	2.418475	0.160077			0.009109	0.254088	-0.001027	-0.8631
X1,X4,X5	2.626118	0.129735			0.005139	0.250409		-0.8399
X1,X5	2.92283	0.14140				0.24485		-0.8496

Keterangan: τ = nilai Sigma Coefficients berdasarkan nilai estimasi

Berdasarkan tabel 4.7 diatas maka didapat estimasi parameter sementara untuk memodelkan pemodelan regresi Poisson Inverse Gaussian yang signifikan terhadap variabel respon dengan melihat nilai P-value yang lebih kecil pada setiap iterasi yang dilakukan dan menampilkan parameter disperse (τ) yang lebih kecil sehingga untuk mengestimasikan model yang didapat (dapat dilihat pada lampiran 2). Setelah memperoleh estimasi parameternya, selanjutnya adalah pengujian hipotesis untuk regresi Poisson Inverse Gaussian.

Pengujian Hipotesis

Pengujian parameter dilakukan secara serentak (simultan) dan individu (parsial) untuk mengetahui signifikansi dari masing-masing parameter yang diperoleh.

A. Pengujian Parameter Secara Serentak

Pengujian parameter secara serentak dilakukan pada kemungkinan model yang sesuai dengan model regresi poisson Inverse Gaussian. Pada uji hipotesis ini dapat dilihat dari nilai statistic G dengan hipotesis berikut:

H_0 : $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = 0$ (variabel prediktor secara simultan tidak berpengaruh terhadap varibel respon)

H_1: paling sedikit ada $\beta_j \neq 0$ dengan j = 1, 2, 3, 4, 5, 6 (Variabel prediktor secara simultan berpengaruh terhadap variabel respon)

α = 0,1

Nilai statistik uji G yang diperoleh dari hasil regresi poisson Inverse Gaussian akan dibandingkan dengan nilai chi-kuadrat. Jika nilai statistik uji G > nilai chi kuadrat maka H_0 ditolak. Hasil pengujian hipotesis secara serentak dapat dilihat pada tabel 4.5 berikut:

Tabel 6 Pengujian Parameter Regresi Poisson Inverse Gaussian Secara Serentak

Variabel dari Model	G_{Hitung}	N	$X^2_{(\alpha,v)}$	Keputusan
X1,X2,X3,X4,X5,X6	269.6477	21	38.93	Tolak H_0
X1,X2,X4,X5,X6	269.6538	22	40.29	Tolak H_0
X1,X4,X5,X6	269.6903	23	41.64	Tolak H_0
X1,X4,X5	269.7296	24	42.98	Tolak H_0
X1,X5	270.0029	25	44.31	Tolak H_0

Dari tabel 6 ditunjukkan jika nilai statistik uji G semua kemungkinan model lebih besar dari nilai chi kuadrat maka keputusannya adalah tolak H_0. Sehingga dalam pengujian parameter secara serentak memenuhi syarat dari hasil perbandingan G_{Hitung} dan chi-square berdasarkan derajat bebas (ν). (Nilai pada tabel 4.5. dapat dilihat pada lampiran 2).

B. Pengujian Parameter Secara Parsial

Pengujian parameter secara individu digunakan untuk mencari variabel prediktor yang berpengaruh signifikan terhadap Jumlah Angka Kematian Bayi di Provinsi Jawa Barat dengan melihat nilai P - value dan Z_{hitung}. Jika P - value < α atau $Z_{hitung} \geq Z_{tabel} = Z_{(\alpha/2)}$ = 0,512 maka akan ditolak H_0 parameter secara individu sebagai berikut.

➢ Parameter β	➢ Parameter τ
H_0 : $\beta_i = 0$	H_0 : $\tau = 0$
H_1 : $\beta_i \neq 0$	H_1 : $\tau \neq 0$
α = 0,1	α = 0,1

Dari hasil uraian regresi poisson invers Gaussian diperoleh hasil parameter secara parsial diperoleh parameter yang signifikan terhadap semua kemungkinan model. Parameter yang signifikan untuk setiap kemungkinan model dapat dilihat pada Tabel 4.6 (data dapat lihat pada lampiran 2).

Tabel 7 Pengujian Parameter Regresi Poisson Inverse Gaussian Secara Parsial

Variabel dari Model	Parameter Signifikan	Nilai AIC
X1,X2,X3,X4,X5,X6	-	285.6477
X1,X2,X4,X5,X6		283.6538
X1,X4,X5,X6		281.6903
X1,X4,X5,		279.7296
X1,X5	β_1, β_5	278.0029

Berdasarkan tabel 4.9 maka dapat disajikan nilai AIC untuk setiap kemungkinan model. Dan dari tabel tersebut dapat diketahui model regresi poisson Inverse Gaussian yang memiliki nilai AIC terkecil dan terdapat variabel yang signifikan terhadap model yaitu pada model ketiga dengan nilai AIC = 278.0029 dengan variabel prediktor persentase berat badan bayi lahir rendah (X1) dan persentase kehamilan resiko tinggi (X5) yang signifikan terhadap model.

a. Pemilihan Model Terbaik

Pada penelitian ini pemodelan dilakukan dengan menggunakan *package* gamlss yang tersedia pada *software* R-Studio. Pemilihan model terbaik yang digunakan adalah dengan menggunakan metode *backward* yang bertujuan untuk mendapatkan model poisson Inverse Gaussian dengan variabel prediktor yang signifikan berdasarkan nilai AIC terkecil.

Berdasarkan model tersebut, dengan menggunakan *package* gamlss yang tersedia pada *software* R-Studio didapatkan hasil estimasi parameter model Regresi Poisson Invers Gaussian yang dapat dilihat pada tabel. 4.7 (nilai dari tabel dapat dilihat pada lampiran 2).

Tabel 8 Penaksiran Parameter Model Regresi Poisson Inverse Gaussian Pada Jumlah Angka Kematian Bayi di Provinsi Jawa Barat

Parameter	Estimasi	Std. Error	Z_{hitung}	P-value
β_0	2.92283	0.33630	8.691	0.00000101
β_1	0.14140	0.7592	1.862	0.00000754
β_5	0.24485	0.05080	4.820	0.00000729
τ	-0.8496	0.3265	-2.602	0.0159*

Keterangan: *** = 5 %, . = 1 %

Nilai pada tabel di atas menunjukkan nilai P-value pada uji parameter dispersi τ yang lebih kecil dari nilai α sehingga tolak H_0 yang berarti nilai parameter dispersi tidak sama dengan nol dan dapat disimpulkan bahwa terjadi fenomena overdispersi pada jumlah angka kematian bayi di Provinsi Jawa Barat Tahun 2017.

Berdasarkan tabel. 8 maka didapat nilai P-value dan nilai AIC yang signifikan untuk membentuk model. Sehingga diperoleh model Regresi Poisson Invers Gaussian berikut:

$$\mu = \exp(\beta_0 + \beta_1 x_1 + \beta_5 x_5)$$

$$\mu = \exp(2.92283 + 0.14140 + 0.24485)$$

Berdasarkan model tersebut dengan melihat nilai AIC terkecil bahwa setiap penambahan 1 persen variabel X1 yang signifikan akan melipat gandakan rata-rata variabel respon Y sebesar exp (0,14140) = 1.14140. Dengan kata lain, penambahan 1 rasio persentase berat badan bayi lahir rendah akan sebanding dengan kenaikan rata-rata jumlah angka kematian bayi di Provinsi Jawa Barat pada Tahun 2017.

Setiap penambahan 1 persen variabel X5 yang signifikan akan melipat gandakan rata-rata variabel respon Y sebesar exp (0.24485) = 1.24485. Dengan kata lain, penambahan 1 rasio persentase kehamilan resiko tinggi maka akan sebanding dengan kenaikan rata-rata jumlah angka kematian bayi di Provinsi Jawa Barat pada Tahun 2017.

KESIMPULAN

Berdasarkan hasil analisis dan pembahasan yang dilakukan oleh peneliti adalah sebagai berikut:

1. Berdasarkan pola persebaran jumlah angka kematian bayi di Provinsi Jawa Barat pada Tahun 2017. Dapat diketahui bahwa wilayah dengan memiliki angka kematian bayi tinggi berkisar antara 187,001 sampai 344,000 dimana terdapat 1 Kabupaten/kota yang memiliki angk a kematian bayi tinggi. adapun Kabupaten/Kota tersebut adalah Garut. Kabupaten/Kota yang memiliki angka kematian bayi sedang berkisar antara 61,0001 sampai 187,000 dimana terdapat 15 Kabupaten/kota yang memiliki angka kematian bayi sedang adapun Kabupaten/Kota tersebut adalah Kota Tasikmaya, Ciamis, Tasikmaya, Kota Banjar, Kota Cirebon, Cirebon, Indramayu, Kota. Kabupaten/kota yang memiliki angka kematian bayi rendah berkisar antara 9,00000 samai 61,0000 dimana terdapat 11 Kabupaten/kota yang memiliki angka kematian bayi rendah adapun Kabupaten/Kota tersebut adalah Pangandaran, Kuningan, Majalengka, Semedang, Bandung, Cianjur, Sukabumi, Kota Sukabumi, Kota Depok, Kota Bekasi, Bekasi.
2. Setelah dilakukan uji overdispersi pada data jumlah angka kematian bayi di Provinsi Jawa Barat pada tahun 2017 mengalami overdispersi sehingga diuji lanjut menggunakan regresi Poisson Inverse Gaussian (PIG). Model regresi PIG yang terbentuk adalah:
$$\mu = \exp(2.92283 + 0.14140 + 0.24485)$$
3. Berdasarkan hasil analisis maka didapat model regresi PIG yang terbentuk dari

beberapa variabel yang berpengaruh secara signifikan terhadap jumlah angka kematian bayi di Provinsi Jawa Barat pada tahun 2017 adalah persentase berat badan bayi lahir rendah (X1) dan persentase kehamilan resiko tinggi (X5). Bahwa setiap penambahan 1 persen variabel X1 yang signifikan akan melipat gandakan rata-rata variabel respon Y sebesar exp (0,14140) = 1.14140. Dengan kata lain, penambahan 1 rasio persentase berat badan bayi lahir rendah akan sebanding dengan kenaikan rata-rata jumlah angka kematian bayi di Provinsi Jawa Barat pada tahun 2017 sebesar 1.14140. dan setiap penambahan 1 persen variabel X3 yang signifikan akan melipat gandakan rata-rata variabel respon Y sebesar exp (0.24485) = 1.22485 dengan kata lain, penambahan 1 rasio persentase kehamilan resiko tinggi maka akan sebanding dengan kenaikan rata-rata jumlah angka kematian bayi di Provinsi Jawa Barat pada tahuh 2017 sebesar 1.22485.

4. Berdasarkan model regresi PIG yang terbentuk maka didapatkan bahwa faktor yang berpengaruh secara signifikan terhadap jumlah angka kematian bayi di Provinsi Jawa Barat adalah presentase berat badan lahir rendah (yaitu berat badan bayi yang kurang 2500 gram) dan presentase kehamilan resiko tinggi (dimana ibu hamil resiko tinggi adalah ibu hamil dengan keadaan penyimpangan dari normal yang secara langsung menyebabkan kesakitan dan kematian bagi ibu maupun bayinya).

DAFTAR PUSTAKA

Cahyandari. 2014. *Pengujian Overdispersi pada Model Regresi Poisson*. UIN Sunan Gunung Djati.

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

Coxe, S., West, S. G. dan Aiken, L.S. 2009. *The Analysis of Count Data: A Gentle Introduction to Poisson Regression and Its Alternatives*. Journal of Personality Assesment, Vol. 91, No. 2, hal. 121-136.

Darnah, 2011. Mengatasi Overdispersi pada Model Regresi Poisson dengan Generalized Poisson I. *Skripsi*. Universitas Mulawarman.

De Jong, P. dan Heller, G.Z. 2008. *Generalized Linear Models for Insurance Data*.1st edition, Cambridge University, Press., New York.Darnah, 2011. *Mengatasi Overdispersi pada Model Regresi Poisson dengan Generalized Poisson I*. Universitas Mulawarman.

Fathurahman. 2009. Pemilihan Model Regresi Terbaik Menggunakan Metode Akaike's Information Criteria dan Schwarz Information Criterion. *Skripsi*. Universitas Mulawarman.

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

Herindrawati, Latra, Purhadi. 2017. *Pemodelan Regresi Poisson Inverse Gaussian*. Institut Teknologi Sepuluh Nopember.

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

Kusuma, Komalasari, Hadijati. 2013. *Model Regresi Zero Inflated Poisson pada Data Overdispersion.*

Keswari, N. M. R. Sumarjaya, I W. Suciptawati, N. L. P. 2014. *Perbandingan Regresi Binomial Negatif dan Regresi Generalisasi Poisson dalam Mengatasi Overdispersi.* Universitas Udayana.

Nur, I.M. 2016. Pemodelan Regresi *Zero Inflated Poisson* pada Kasus Angka Kematian Bayi (AKB) di Provinsi Jawa Tengah. Prosiding Seminar Nasional Sains dan Entrepreneurship III Tahun 2016.

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.