

# PEMODELAN JUMLAH KASUS PENYAKIT KUSTA DI PROVINSI SULAWESI TENGGARA MENGGUNAKAN METODE REGRESI POISSON INVERSE GAUSSIAN

Muhammad Nurul Ramadhan<sup>1</sup>, Indah Manfaati Nur<sup>2</sup>, Tiani Wahyu Utami<sup>3</sup>

<sup>123</sup>Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Muhammadiyah Semarang

Alamat e-mail : [ramadhanstk96@gmail.com](mailto:ramadhanstk96@gmail.com)

## ABSTRAK

Data penyakit kusta merupakan data cacahan sehingga dalam pemodelannya bisa menggunakan regresi Poisson. Data penyakit juga berpotensi mengalami *overdispersi* sehingga dalam mengatasi kasus *overdispersi*, dibentuk beberapa pemodelan yang merupakan perpaduan antara distribusi Poisson dengan beberapa distribusi baik diskrit maupun kontinu (*mixed poisson distribution*). Salah satunya adalah distribusi Poisson Invers Gaussian (PIG) yang merupakan *mixed poisson distribution* dengan random efek yang memiliki distribusi Invers Gaussian. Pada penerapan metode regresi Poisson Inverse Gaussian terdapat faktor-faktor yang mempengaruhi terjadinya penyakit kusta yang ada di Sulawesi Tenggara. Adapun faktor-faktornya yaitu Kepadatan Penduduk, Penduduk yang Mengalami Keluhan Kesehatan dan Pernah Berobat Jalan, Penggunaan Air Bersih, Rumah Tangga BER-PHBS, Fasilitas Kesehatan, Tenaga Kesehatan, dan Pertumbuhan Ekonomi dari website BPS dan Profil Dinas Kesehatan Provinsi Sulawesi Tenggara yang ada pada lampiran pdf dengan hasil analisis dan pembahasan yang dilakukan maka diperoleh model regresi Poisson Inverse Gaussian dengan variabel yang signifikan terhadap model yaitu persentase kepadatan penduduk dan persentase penggunaan air bersih. Dengan penambahan 1 rasio pada persentase kepadatan penduduk akan sebanding dengan kenaikan rata-rata jumlah penyakit kusta di Provinsi Sulawesi Tenggara sebesar 1.00177. Dan dengan penambahan 1 rasio pada persentase penggunaan air bersih maka akan sebanding dengan kenaikan rata-rata jumlah penyakit kusta di Provinsi Sulawesi Tenggara sebesar 1.01205.

**Kata Kunci:** Penyakit Kusta, Regresi Poisson Inverse Gaussian, Overdispersi

## ABSTRACT

Leprosy data is census data so in its modeling, it can use Poisson regression. Leprosy data also has the potential to experience *overdispersion* so that in overcoming the case of *overdispersion*, several models are formed which are a combination of the Poisson distribution with several distributions both discrete and continuous (*mixed Poisson distribution*). One of them is the distribution of Poisson Inverse Gaussian (PIG) which is a *mixed Poisson distribution* with random effects that have the Gaussian inverse distribution. In the application of a Poisson Inverse Gaussian regression method, there are factors affecting the occurrence of leprosy in Southeast Sulawesi. As for the factors namely the population density, the residents who experience health complaints and ever ambulatory care, the use of clean water, the Clean and Healthy Behavior (PHBS) Households, the health facilities, the health workers, and the economic growth from the BPS website and the Profile of the Health Office of Southeast Sulawesi Province in the PDF appendix. The results of data analysis and discussion, the Poisson Inverse Gaussian regression model is obtained with a significant variable to the model that is the percentage of population density and the percentage of clean water use. With the addition of 1 ratio to the percentage of population density will be proportional to the increase in the average number of leprosy in Southeast Sulawesi Province by 1.00177. And with the addition of 1 ratio on the percentage of the use of clean water, it would be comparable to the average increase in the number of leprosy in southeast Sulawesi province by 1.01205.

**Keywords:** Leprosy, Poisson Inverse Gaussian Regression, Overdispersion

## PENDAHULUAN

Data cacahan (count data) adalah data yang menggambarkan sejumlah kejadian yang terjadi pada suatu kurun waktu tertentu. Suatu data cacahan hanya dapat berupa bilangan bulat positif karena suatu kejadian tidak mungkin terjadi dalam sejumlah bilangan negatif. Pemodelan data cacahan banyak digunakan dalam berbagai ilmu seperti sosiologi, kedokteran, aktuaria dan demografi. 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). Distribusi bersyarat dari data cacahan cenderung menceng kanan dengan tidak ada kejadian yang di bawah nol dan jumlah kejadian dengan frekuensi tinggi hanya sedikit sehingga melanggar asumsi kenormalan. 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).

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). Dalam mengatasi kasus overdispersi, dibentuk beberapa pemodelan yang merupakan perpaduan antara distribusi Poisson dengan beberapa distribusi baik diskrit maupun kontinu (*mixed poisson distribution*). *Mixed poisson distribution* tersebut merupakan solusi alternatif untuk kasus overdispersi, tetapi hanya beberapa distribusi yang sering digunakan dalam penelitian dikarenakan penghitungannya yang rumit. Salah satunya adalah distribusi Poisson Invers Gaussian (PIG) yang merupakan *mixed poisson distribution* dengan random efek yang memiliki distribusi Invers Gaussian. Distribusi ini pertama kali diperkenalkan oleh Holla pada tahun 1966 (Karlis dan Nikoloulopoulos, 2005). Distribusi PIG sendiri merupakan bentuk dari distribusi *Siche* (SI) dengan dua parameter. SI disebut 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).

Willmot (1987) menunjukkan potensi dari pemodelan dengan regresi poisson inverse gaussian sebagai alternatif dari regresi binomial negatif pada data klaim asuransi mobil. Enam set data klaim asuransi mobil disajikan dengan karakteristik hampir 80 persen data yang mengandung nol, dan menghasilkan kesimpulan bahwa pemodelan dengan regresi PIG merupakan model yang lebih baik dibandingkan model regresi binomial negatif. Penelitian lainnya yang menggunakan model regresi PIG adalah penelitian dari Shoukri, Selanjutnya

dalam beberapa penelitian keselamatan di jalan raya seperti pemodelan data kecelakaan dan penelitian asuransi motor, model regresi PIG sering digunakan sebagai alternatif dari model regresi binomial negatif (Denis, Ghitany dan Gupta, 2014). Pemodelan dengan regresi PIG juga digunakan pada penelitian Zha, Lord dan Zou (2014) pada kasus jumlah kecelakaan motor yang terjadi di dua tempat berbeda yaitu di Texas dan Washington. Dalam penelitian tersebut data jumlah kecelakaan motor menceng kanan dengan ekor yang sedikit panjang serta 37 persen data mengandung nilai nol. Dengan nilai *Akaike Information Criterion* (AIC) dan *Bayesian Information Criterion* (BIC) yang lebih kecil, menunjukkan bahwa pemodelan dengan regresi PIG lebih baik untuk memodelkan kasus jumlah kecelakaan motor di dua tempat tersebut.

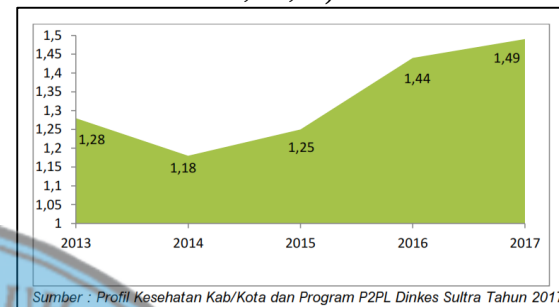
Widiari (2016) menggunakan model regresi Poisson Inverse Gaussian pada data jumlah kasus baru HIV di Jawa Timur Tahun 2013 dan menghasilkan variabel yang signifikan berpengaruh adalah persentase penduduk dengan pendidikan terakhir SLTA, persentase PUS yang menggunakan kondom, dan rasio fasilitas kesehatan. Kemudian Herindrawati, Latra dan Purhadi (2017) juga melakukan penelitian dengan menggunakan regresi Poisson Inverse Gaussian pada data kasus baru penyakit HIV di Provinsi Jawa Tengah Tahun 2015. Dalam penelitiannya diperoleh model regresi Poisson Invers Gaussian dan faktor-faktor yang berpengaruh secara signifikan terhadap kasus baru penyakit HIV di Provinsi Jawa Tengah yaitu persentase PUS yang menggunakan kondom, rasio fasilitas kesehatan, persentase daerah perkotaan, dan persentase penduduk usia 25-34 tahun.

Pada penelitian sebelumnya juga yang mengambil kasus penyakit kusta dengan beberapa judul yaitu faktor-faktor yang berhubungan dengan derajat kecacatan pasien *morbus Hansen* (Kusta) di kota Bandar Lampung, faktor-faktor yang berhubungan dengan kejadian kusta di wilayah kerja puskesmas kabunan kabupaten pematang tahun 2014, analisis faktor risiko tingkat kecacatan pada penderita kusta di puskesmas Padas kabupaten Ngawi pada tahun 2010, sebaran kasus penyakit kusta berdasarkan wilayah kecamatan dikota Bekasi tahun 2006-2008.

Selain itu penyakit Kusta dikenal juga sebagai penyakit Lepra atau penyakit Hansen, disebabkan oleh bakteri *Mycobacterium leprae*.

Bakteri ini mengalami proses pembelahan cukup lama antara 2-3 minggu. Daya tahan hidup kuman kusta mencapai 9 hari di luar tubuh manusia. Kuman kusta memiliki masa inkubasi 2-5 tahun bahkan lebih.

Pada Tahun 2017, jumlah kasus kusta baru di Sulawesi Tenggara sebesar 327 kasus dengan angka penemuan kasus baru (*New Case Detection Rate/NCDR*) per 100.000 penduduk sebesar 12,40, lebih rendah dari tahun 2016 (342 kasus baru / NCDR; 13,20).



Gambar 1.1 Angka Prevalensi Kusta Per 10.000 Penduduk Provinsi Sulawesi Tenggara Tahun 2013-2017

Berdasarkan grafik di atas bahwa pada tahun 2013-2017, angka prevalensi kusta Sulawesi Tenggara sempat turun mencapai angka terendah pada tahun 2014, tetapi kembali naik pada tiga tahun terakhir, bila merujuk pada target nasional yang ditetapkan mengenai prevalensi kusta sebesar kurang dari 1/10.000 penduduk (lebih dari 10/100.000), maka prevalensi kusta Sulawesi Tenggara sebesar 1.49/10.000 penduduk makin menjauh dari target, setelah hampir mencapai target pada tahun 2014. Tampaknya upaya penanggulangan kusta masih membutuhkan upaya yang lebih keras beberapa tahun ke depan agar angka prevalensi kusta Sulawesi Tenggara bisa mencapai target nasional.

Pada penelitian ini yang akan diteliti yaitu jumlah penyakit kusta di Sulawesi Tenggara. Dimana jumlah penyakit kusta merupakan salah satu indikator penting bagi pemerintah dalam mengevaluasi dibidang kesehatan. Dengan mengetahui faktor-faktor yang mempengaruhi, tentu saja pemerintah memiliki gambaran langkah apa saja yang harus dilakukan dalam menekan jumlah penyakit kusta. Mengingat dalam tahun terakhir yaitu Pada Tahun 2017, jumlah kasus kusta di Sulawesi Tenggara sebesar 327 kasus dengan angka penemuan kasus per 100.000 penduduk sebesar 12,40, lebih rendah dari tahun 2016 (342 kasus baru / NCDR; 13,20). Berdasarkan prevalensi,



kasus kusta di Sulawesi Tenggara menunjukkan kurva menaik pada rentang tahun 2013-2017, dengan prevalensi tertinggi terjadi pada tahun 2017 sebesar 1,49.

Data penyakit kusta merupakan data cacahan sehingga dalam pemodelannya bisa menggunakan regresi Poisson. Data penyakit kusta 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).

Distribusi Poisson memiliki ciri-ciri sebagai berikut:

1. Banyaknya percobaan yang terjadi dalam suatu selang waktu atau suatu daerah tertentu, tidak tergantung pada banyaknya hasil percobaan yang terjadi pada selang waktu atau daerah lain yang terpisah.
2. Peluang terjadinya satu hasil percobaan selama suatu selang waktu yang singkat sekali atau dalam suatu daerah yang kecil. Sebanding dengan panjang selang waktu tersebut atau besarnya daerah tersebut dan tidak bergantung pada banyak hasil percobaan yang terjadi diluar selang waktu dan daerah tertentu.
3. Peluang bahwa lebih dari satu hasil percobaan akan terjadi dalam selang waktu yang singkat tersebut atau dalam daerah yang kecil tersebut dapat diabaikan.

Fungsi peluang untuk data berdistribusi Poisson bergantung pada parameter tunggal, yaitu rata-rata  $\mu$ . Fungsi peluangnya adalah sebagai

berikut:

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

dan  $\mu > 0$

Dalam distribusi Poisson, rata-rata dan variansi adalah bernilai sama dan dapat dituliskan sebagai berikut:

$$E(Y) = Var(Y) = \mu$$

### 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 memiliki dua parameter dan fungsi kepadatan peluang yang dapat ditulis sebagai berikut:

$$f(y) = (2\pi y^3 \sigma)^{-0,5} e^{-a(y-\mu)^2 / 2y\mu^2 \sigma^2}, y > 0$$

Dengan rata-rata dan varians:

$$E(Y) = \mu \text{ dan } Var(Y) = \sigma^2 \mu^2.$$

dan  $\sigma^2$  adalah parameter dispersi. 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

Dalam 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 rata-ratanya. 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 dapat ditulis:

$$Var(Y) > E(Y)$$

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).

Uji statistik yang bisa juga digunakan untuk mendeteksi *overdispersi* pada suatu data adalah uji *overdispersi* yang dapat menggunakan *package* AER dari *software* R-*Consule* (Herindrawati, Latra, Purhadi, 2017).

Keputusan hipotesis:

$H_0$  : Tidak terjadi Overdispersi

$H_1$  : terjadi Overdispersi

Keputusan yang diambil untuk uji *overdispersi* menggunakan *software* R-*Consule* yaitu jika nilai  $p$  - value  $< \alpha$  maka  $H_0$  ditolak yang berarti terjadi *overdispersi*. Tetapi jika nilai  $p$  - value  $> \alpha$  maka  $H_0$  diterima yang berarti tidak terjadi *overdispersi*.

#### 4. Distribusi Poisson Inverse Gaussian

Distribusi Poisson Inverse Gaussian merupakan salah satu distribusi mixed poisson yang ditentukan oleh dua parameter yaitu rata-rata ( $\mu$ ) sebagai parameter lokasi dan parameter disperse ( $\tau$ ) sebagai parameter bentuk (Herindrawati, Latra, Purhadi, 2017). Dimana kedua parameter tersebut dapat dituliskan sebagai berikut:

$$P(Y = y | \mu) = \frac{\mu^y e^{\frac{1}{\tau}}}{y!} \left( \frac{2}{\pi\tau} \right)^{\frac{1}{2}} (2\mu\tau + 1)^{\frac{1}{2}} K_{\frac{y-1}{2}} \left( \frac{1}{\tau} \sqrt{2\mu\tau + 1} \right)$$

Rata-rata untuk distribusi Poisson Inverse Gaussian yaitu:

$$E(Y) = E\{E(Y | \mu\tau)\} = E(\mu\tau) = \mu$$

Variansi untuk distribusi Poisson Inverse Gaussian yaitu:

$$Var(Y) = Var\{E(Y | \mu\tau)\} + E\{Var(Y | \mu\tau)\} = \mu + \tau\mu^2$$

#### 5. Regresi Poisson Inverse Gaussian

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

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

#### 6. Estimasi Parameter Regresi Poisson Inverse Gaussian

Persamaan 2.4 merupakan model regresi PIG dan parameter  $\beta$  pada regresi PIG ditaksir dengan metode *maximum likelihood*. Langkahnya adalah dengan menentukan fungsi likelihood dari distribusi PIG sebagai berikut:

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

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

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

$$\begin{aligned} L(\beta; \tau) &= \ln L(\beta; \tau) \\ &= \sum_{i=1}^n y_i x_i^T \beta + \frac{n}{\tau} - \ln \left( \sum_{i=1}^n y_i! \right) + \frac{n}{2} \ln \left( \frac{2}{\pi} \right) - \frac{n}{2} \ln \tau \\ &\quad - \sum_{i=1}^n \left( \frac{2y_i - 1}{4} \right) \ln(2x_i^T \beta + 1) + \sum_{i=1}^n \ln K_{\frac{y_i-1}{2}}(Z_i) \end{aligned}$$

Selanjutnya ditentukan turunan pertama dan turunan kedua terhadap  $\beta$  dan  $\tau$ :

$$\begin{aligned} \frac{\partial L}{\partial \beta} &= \frac{\partial}{\partial \beta} \left[ \sum_{i=1}^n y_i x_i^T \beta + \frac{n}{\tau} - \ln \left( \sum_{i=1}^n y_i! \right) + \frac{n}{2} \ln \left( \frac{2}{\pi} \right) - \frac{n}{2} \ln \tau \right. \\ &\quad \left. - \sum_{i=1}^n \left( \frac{2y_i - 1}{4} \right) \ln(2x_i^T \beta + 1) + \sum_{i=1}^n \ln K_{\frac{y_i-1}{2}}(Z_i) \right] \\ &= \sum_{i=1}^n y_i x_i^T - \sum_{i=1}^n \left( \frac{2y_i - 1}{4} \right) \frac{1}{(2x_i^T \beta + 1)} 2x_i^T \\ &= \frac{\partial L}{\partial \tau} \\ &= \frac{\partial}{\partial \tau} \left[ \sum_{i=1}^n y_i x_i^T \beta + \frac{n}{\tau} - \ln \left( \sum_{i=1}^n y_i! \right) + \frac{n}{2} \ln \left( \frac{2}{\pi} \right) - \frac{n}{2} \ln \tau \right. \\ &\quad \left. - \sum_{i=1}^n \left( \frac{2y_i - 1}{4} \right) \ln(2x_i^T \beta + 1) + \sum_{i=1}^n \ln K_{\frac{y_i-1}{2}}(Z_i) \right] \\ &= -\frac{n}{\tau^2} - \frac{n}{2\tau} = 0 \end{aligned}$$

Berikutnya dari persamaan (2.6) akan ditentukan turunan kedua dari parameter  $\beta$  dan  $\tau$  dan turunan  $\tau$  dari persamaan (2.7)

Apabila persamaan diatas implisit dan nonlinear dalam parameter  $\beta$  dan  $\tau$  sehingga untuk mendapatkan taksiran dari  $\theta = [\beta^T \tau]^T$

fungsi dimaksimumkan dengan menggunakan *Fisher Scoring Algorithm* dengan persamaan berikut:

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

Dimana:

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

$$D(\hat{\theta}) = \begin{pmatrix} \frac{\partial l}{\partial \hat{\tau}} & \frac{\partial l}{\partial \hat{\beta}^T} \end{pmatrix}$$

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

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

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

Matriks hessian merupakan matriks yang berisi turunan kedua dari fungsi *likelihood* terhadap parameter  $\beta$  dan  $\tau$ . langkah-langkah pengerjaan dengan menggunakan *Fisher Scoring Algorithm* (Ummah, Sulyanto dan Sediono, 2013).yaitu:

1. Menentukan vektor awal parameter  $\hat{\theta}_0$  dengan mengasumsikan data memenuhi model regresi linear berganda
2. Membentuk vektor gradien  $D(\hat{\theta}_0)$
3. Membentuk matriks hessian  $H(\hat{\theta}_0)$
4. Membentuk matrik informasi *Fisher*  $I(\hat{\theta}_0)$
5. Memasukkan nilai  $\hat{\theta}_{(0)}$  sehingga diperoleh vector gradient  $D(\hat{\theta}_0)$  dan matriks hessian  $H(\hat{\theta}_0)$
6. Memulai dari  $m = 0$  dilakukan iterasi pada, nilai  $\hat{\theta}_{(m)}$  merupakan sekumpulan penaksir parameter yang konvergen saat iterasi ke-m
7. Jika belum diperoleh penaksiran parameter yang konvergen saat iterasi ke-m, maka dilanjutkan kembali kelangkah 5 sampai iterasi ke-m+1. Iterasi akan berhenti apabila nilai dari  $\|\hat{\theta}_{(m+1)} - \hat{\theta}_{(m)}\| \leq \varepsilon$  dan  $\varepsilon > 0$ .

## 7. Pengujian Parameter

Pengujian parameter pada model Poisson Inverse Gaussian dilakukan dengan menggunakan pengujian hipotesis secara serentak pada parameter  $\beta$  serta pengujian parsial pada parameter  $\beta$  dan  $\tau$ . Pengujian parameter dilakukan untuk mengetahui ada atau tidaknya pengaruh variabel prediktor terhadap variabel respon.

### ➤ Uji secara serentak (Simultan)

Langkah-langkah pengujian yang dilakukan secara serentak yaitu:

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

(variabel prediktor secara simultan tidak berpengaruh terhadap variabel respon)

$$H_1 : \text{minimal ada satu } \beta_i \neq 0$$

dengan  $i = 1, 2, \dots, k$ .

(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 ( $\Omega$ ) yaitu  $\Omega = (\beta, \tau)$  dan himpunan parameter dibawah  $H_0$  benar ( $\omega$ ) 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  $L(\omega)$ .

$$L(\hat{\Omega}) = \prod_{i=1}^n \left\{ \frac{e^{x_i^T \hat{\beta}_i} e^{\frac{1}{\tau}}}{y_i!} \left( \frac{2}{\pi \tau} \right)^{\frac{1}{2}} (2e^{x_i^T \hat{\beta}_i} + 1)^{-\frac{(y_i - \frac{1}{2})}{2}} K_{si}(Z_i) \right\}$$

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

$$L(\hat{\omega}) = \prod_{i=1}^n \left\{ \frac{(e^{\hat{\beta}_0})^{y_i} e^{\frac{1}{\tau}}}{y_i!} \left( \frac{2}{\pi \hat{\tau}_\omega} \right)^{\frac{1}{2}} (2(e^{\hat{\beta}_0}) \hat{\tau}_\omega + 1)^{-\frac{(y_i - \frac{1}{2})}{2}} K_{si}(Z_i) \right\}$$

$$\ln(L(\hat{\omega})) = \sum_{i=1}^n y_i x_i^T \hat{\beta} + \frac{n}{\tau} - \ln(\sum_{i=1}^n y_i!) + \frac{n}{2} \ln\left(\frac{2}{\pi}\right) - \frac{n}{2} \ln \tau - \sum_{i=1}^n \left( \frac{2y_i - 1}{4} \right) \ln(2x_i^T \hat{\beta}_0 \hat{\tau}_\omega + 1) + \sum_{i=1}^n K_{si}(Z_i)$$

Kedua fungsi diatas dibandingkan dalam bentuk devians berikut:

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

Statistik G adalah pendekatan dari distribusi *chi square* dengan derajat bebas  $v$  sehingga kriteria pengujiannya adalah tolak  $H_0$  apabila  $G_{hit} > \chi^2_{(\alpha, v)}$  dimana  $v$  adalah derajat bebas yang diperoleh dari jumlah parameter dibawah populasi dikurangi jumlah parameter dibawah  $H_0$

➤ Uji secara individu (Parsial)

Pengujian hipotesis secara parsial (Individu) melalui kriteria keputusan penolakan  $H_0$ . Hipotesis yang digunakan adalah:

Hipotesis pengujian parameter  $\beta$ :

$$H_0 : \beta_j = 0, j = 1, 2, \dots, k$$

(variabel prediktor ke- $j$  tidak berpengaruh terhadap variabel respon)

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

(variabel prediktor ke- $j$  berpengaruh terhadap variabel respon)

Uji statistik yang digunakan dalam pengujian signifikan parameter  $\beta$  adalah:

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

Kriteria penolakan  $H_0$  apabila  $|Z_{hit}| > Z_{\frac{\alpha}{2}}$  atau  $p\text{-value} < \alpha$  dimana  $\alpha$  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 kovarians yang diperoleh dari:

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

Hipotesis pengujian parameter  $\tau$ :

$$H_0 : \tau = 0$$

(variabel prediktor tidak berpengaruh terhadap variabel respon)

$$H_1 : \tau \neq 0$$

(variabel prediktor berpengaruh terhadap variabel respon)

Statistik uji yang digunakan:

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

Kriteria penolakan  $H_0$  apabila  $|Z_{hit}| > Z_{\frac{\alpha}{2}}$  atau  $p\text{-value} < \alpha$  dimana  $\alpha$  adalah tingkat signifikan.

## 8. Uji Korelasi

Uji korelasi merupakan bagian dari ilmu statistika yang digunakan untuk menentukan

hubungan keeratan antara dua variabel atau lebih dengan menggunakan analisis koefisien korelasi. Koefisien korelasi digunakan untuk mengukur derajat erat tidaknya hubungan antara satu variabel terhadap variabel lainnya dimana pengamatan pada masing-masing variabel tersebut pada pemberian peringkat tertentu serta pasangannya (Pradeka, 2012).

Untuk pengambilan keputusan dalam analisis korelasi yaitu:

$H_0$  : Tidak terdapat korelasi antarvariabel

$H_1$  : terdapat korelasi antarvariabel

Jika  $\text{sig} < \alpha$  maka  $H_0$  ditolak. Berarti terdapat korelasi antarvariabel

Jika  $\text{sig} > \alpha$  maka  $H_0$  diterima. Berarti tidak terdapat korelasi antarvariabel.

## 9. Uji Multikolinearitas

Menurut Montgomery dan Peck (1999), untuk mendeteksi atau mengetahui ada atau tidaknya multikolinearitas didalam model regresi dapat dilihat pada nilai *variance inflation factors* (VIF) (Candraningtyas, Safitri, Ispriyanti, 2013) yaitu:

$$VIF_j = \frac{1}{(1 - r_j^2)}$$

Dengan  $r_j^2$  adalah nilai koefisien determinasi yang diperoleh dari meregresikan antar variabel bebas lainnya. Jika nilai  $VIF > 10$  maka menunjukkan multikolinearitas yang kuat. Jika nilai  $VIF < 10$  maka tidak terjadi multikolinearitas (Candraningtyas, Safitri, Ispriyanti, 2013).

## 10. 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:

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

Dengan:

$k$  = jumlah parameter yang di estimasi dalam model regresi

$n$  = jumlah observasi

$e$  = sisa (residual)



### 11. Penyakit Kusta Di Sulawesi Tenggara

Penyakit Kusta dikenal juga sebagai penyakit Lepra atau penyakit Hansen, disebabkan oleh bakteri *Mycobacterium leprae*. Bakteri ini mengalami proses pembelahan cukup lama antara 2-3 minggu. Daya tahan hidup kuman kusta mencapai 9 hari di luar tubuh manusia. Kuman kusta memiliki masa inkubasi 2-5 tahun bahkan lebih. Penatalaksanaan kasus yang buruk dapat menyebabkan kusta menjadi progresif, sehingga menyebabkan kerusakan permanen pada kulit, saraf, anggota gerak dan mata.

Pada Tahun 2017, jumlah kasus kusta baru di Sulawesi Tenggara sebesar 327 kasus dengan angka penemuan kasus baru (*New Case Detection Rate/NCDR*) per 100.000 penduduk sebesar 12,40, lebih rendah dari tahun 2016 (342 kasus baru / NCDR; 13,20). Berdasarkan prevalensi, kasus kusta di Sulawesi Tenggara menunjukkan kurva menaik pada rentang tahun 2013-2017, dengan prevalensi tertinggi terjadi pada tahun 2017 sebesar 1.49.

Berdasarkan data sebelum tahun 2017, angka prevalensi kusta Sulawesi Tenggara sempat turun mencapai angka terendah pada tahun 2014, tetapi kembali naik pada tiga tahun terakhir, bila merujuk pada target nasional yang ditetapkan mengenai prevalensi kusta sebesar kurang dari 1/10.000 penduduk (lebih dari 10/100.000), maka prevalensi kusta Sulawesi Tenggara sebesar 1.49/10.000 penduduk makin menjauh dari target, setelah hampir mencapai target pada tahun 2014. Tampaknya upaya penanggulangan kusta masih membutuhkan upaya yang lebih keras beberapa tahun ke depan agar angka prevalensi kusta Sulawesi Tenggara bisa mencapai target nasional, selanjutnya agar dapat mengeliminasi penyakit kusta dari Sulawesi Tenggara.

Berdasarkan bebannya, kusta dibagi menjadi 2 kategori yaitu beban kusta tinggi (*high burden*) dan beban kusta rendah (*low burden*). Suatu daerah disebut 'beban tinggi' jika angka penemuan kasus baru kusta (NCDR) mencapai lebih dari 10 per 100.000 penduduk sedangkan 'beban rendah' jika NCDR kurang dari 10 per 100.000 penduduk. Dengan NCDR mencapai 12.40 maka Sulawesi Tenggara termasuk ke dalam kategori provinsi dengan beban kusta tinggi (*high burden*). Kabupaten/kota dengan angka kasus kusta tertinggi masih didominasi daerah-daerah yang sejak beberapa tahun terakhir

merupakan daerah endemik kusta, namun daerah lain juga harus tetap mewaspadai penyebaran penyakit ini, karena angka yang dilaporkan mungkin hanya yang dapat ditemukan dan dicatat oleh petugas kesehatan, sedangkan kasus lain yang tidak terekspos masih banyak di tengah-tengah masyarakat, apalagi penyakit kusta adalah salah satu penyakit dengan masa inkubasi yang relatif lama, sehingga sering tidak disadari dengan segera oleh si penderita.

### METODE PENELITIAN

#### 1. Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder tentang jumlah penyakit kusta dan faktor-faktor yang mempengaruhi jumlah penyakit kusta. Data yang diambil dari website BPS Provinsi Sulawesi Tenggara tahun 2017 dan Profil Kesehatan Sulawesi Tenggara tahun 2017 dari hasil publikasi dari Dinas Kesehatan Sulawesi Tenggara Tahun 2017 dengan unit pengamatan yang diambil yaitu pada tingkat Kabupaten / Kota di Sulawesi Tenggara dengan 17 Kabupaten / Kota.

#### 2. Variabel Penelitian

Adapun variabel yang digunakan pada penelitian ini adalah sebagai berikut.

Tabel 3. 1 Keterangan Variabel Penelitian Dan Tipe Data

| VARIABEL  | KETERANGAN                                                                    | TYPE DATA |
|-----------|-------------------------------------------------------------------------------|-----------|
| <b>Y</b>  | Jumlah Penyakit Kusta                                                         | Diskrit   |
| <b>X1</b> | Persentase Kepadatan Penduduk                                                 | Kontinu   |
| <b>X2</b> | Persentase Penduduk Yang Mengalami Keluhan Kesehatan Dan Pernah Berobat Jalan | Kontinu   |
| <b>X3</b> | Persentase Penggunaan Air Bersih                                              | Kontinu   |
| <b>X4</b> | Persentase Rumah Tangga Berperilaku Hidup Bersih dan Sehat                    | Kontinu   |



|           |                                |         |
|-----------|--------------------------------|---------|
| <b>X5</b> | Persentase Fasilitas Kesehatan | Kontinu |
| <b>X6</b> | Persentase Tenaga Kesehatan    | Kontinu |
| <b>X7</b> | Persentase Pertumbuhan Ekonomi | Kontinu |

### 3. Langkah-Langkah Penelitian

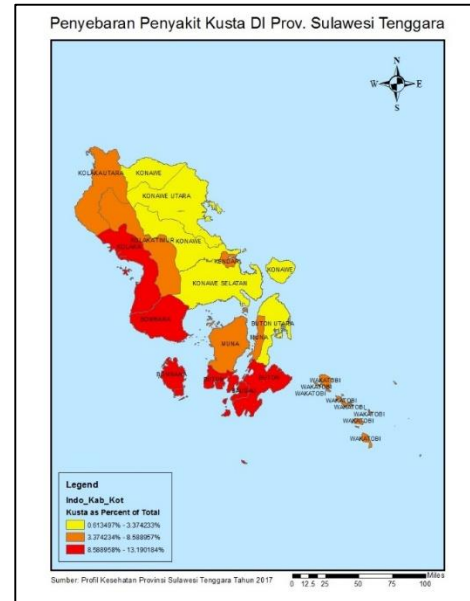
Langkah-langkah analisis yang digunakan dalam penelitian ini adalah sebagai berikut.

1. Membuat gambaran pola sebaran penyakit kusta
2. Menguji korelasi antara variabel respon dengan variabel prediktor
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 Penyakit Kusta di Provinsi Sulawesi Tenggara

Berdasarkan data yang dikumpulkan dari Website Badan Pusat Statistik tahun 2017 dan Profil Dinas Kesehatan Provinsi Sulawesi Tenggara tahun 2017, terdapat data tentang kasus penyakit Kusta dan faktor faktor yang mempengaruhi kasus penyakit Kusta di Provinsi Sulawesi Tenggara dengan pola penyebaran sebagai berikut:



Gambar 4.1 Pola Sebaran Penyakit Kusta di Provinsi Sulawesi Tenggara Tahun 2017

Berdasarkan gambar 4.1 menunjukkan persebaran penyakit kusta di Provinsi Sulawesi Tenggara tahun 2017. Dapat diketahui bahwa wilayah dengan warna semakin gelap menunjukkan penyakit kusta semakin tinggi berkisar antara 8,5889 sampai 13,1902. Berdasarkan hasil analisis terdapat 5 Kabupaten/kota yang memiliki penyakit kusta tinggi. Adapun Kabupaten/Kota tersebut adalah Kota Kolaka, Kabupaten Bombana, Kabupaten Buton, Kabupaten Buton Tengah, dan Kota Bau-Bau.

Kabupaten/kota yang memiliki penyakit kusta rendah berkisar antara 0,6134 sampai 3,3742. Berdasarkan hasil analisis terdapat 6 Kabupaten/kota yang memiliki penyakit kusta rendah. Adapun Kabupaten/Kota tersebut adalah Kabupaten Konawe, Kabupaten Konawe Timur, Kabupaten Konawe Selatan, Kabupaten Konawe Utara, Kabupaten Konawe Kepulauan, dan Kabupaten Buton Utara.

### 2. Pemeriksaan Korelasi

Koefisien korelasi antar variabel respon dengan variabel prediktor dapat dilihat pada tabel 4.1.

Pengujian hipotesis sebagai berikut:

$H_0$ : Tidak terdapat hubungan antar kedua variabel

$H_1$ : Terdapat hubungan antar kedua variabel

Jika nilai sig. lebih kecil dari  $\alpha$ , dengan  $\alpha$  bernilai 0,1 maka  $H_0$  ditolak.

Tabel 4. 1 Koefisien Korelasi Variabel Respon dan Variabel Prediktor

|    | Korelasi        | Y            |
|----|-----------------|--------------|
| X1 | Pearson Cor.    | 0.403        |
|    | Sig. (2-tailed) | <b>0.109</b> |
| X2 | Pearson Cor.    | 0.171        |
|    | Sig. (2-tailed) | 0.512        |
| X3 | Pearson Cor.    | 0.384        |
|    | Sig. (2-tailed) | 0.128        |
| X4 | Pearson Cor.    | -0.127       |
|    | Sig. (2-tailed) | 0.626        |
| X5 | Pearson Cor.    | 0.050        |
|    | Sig. (2-tailed) | 0.848        |
| X6 | Pearson Cor.    | 0.462        |
|    | Sig. (2-tailed) | <b>0.062</b> |
| X7 | Pearson Cor.    | 0.391        |
|    | Sig. (2-tailed) | 0.121        |

Berdasarkan tabel 4.1 diperoleh ada variabel prediktor yang memiliki hubungan (korelasi) dengan variabel respon yaitu variabel X1 (persentase kepadatan penduduk) dan X6 (persentase tenaga kesehatan yang ada di kabupaten/kota).

### 3. Pemeriksaan Multikolinearitas

Pengujian uji multikolinearitas antar variabel respon dengan variabel prediktor dapat dilihat pada tabel 4.2.

Tabel 4. 2 Nilai VIF Variabel Prediktor

| Variabel Prediktor | Nilai VIF |
|--------------------|-----------|
| <b>X1</b>          | 1.211     |
| <b>X2</b>          | 1.354     |
| <b>X3</b>          | 1.697     |
| <b>X4</b>          | 1.300     |
| <b>X5</b>          | 1.191     |
| <b>X6</b>          | 1.402     |
| <b>X7</b>          | 1.306     |

Pada tabel 4.2 disajikan nilai VIF dari masing-masing prediktor. Pada pengujian multikolinearitas dengan menggunakan nilai VIF, variabel dikatakan bermultikolinearitas jika nilai VIF lebih dari 10. Karena nilai VIF dari setiap variabel prediktor yang disajikan pada tabel bernilai dibawah 10, maka dapat disimpulkan bahwa tidak ada kasus multikolinearitas antara variabel. Sehingga semua variabel prediktor memenuhi asumsi nonmultikolinearitas.

### 4. Uji Overdispersi

Pada hasil uji overdispersi menggunakan package AER dari software R-Console (dapat dilihat pada lampiran 2) juga menunjukkan jika terjadi overdispersi karena nilai P-value =  $0.0001435 < \alpha$ , dimana nilai  $\alpha$  sebesar 1 %, sehingga data penelitian yang mengalami overdispersi.

### 5. Menentukan Penaksiran Model Regresi Poisson Inverse Gaussian

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 tujuh variabel yang digunakan yaitu persentase kepadatan penduduk (X1), persentase penduduk yang mengalami keluhan kesehatan dan pernah berobat jalan (X2), persentase penggunaan air bersih (X3), persentase rumah tangga berperilaku hidup bersih dan sehat (X4), persentase fasilitas kesehatan yang ada di kabupaten/kota (X5), persentase tenaga kesehatan yang ada di kabupaten/kota (X6), dan persentase pertumbuhan ekonomi (X7) menghasilkan tiga kombinasi model yang sudah konvergen. Ketiga kombinasi kemungkinan model regresi Poisson Inverse Gaussian yang sudah konvergen, kemudian dicari model terbaiknya. Berikut merupakan ketiga kemungkinan model Poisson Inverse Gaussian yang sudah konvergen:

$$\mu = \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \beta_7 x_7)$$

$$\mu = \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6)$$

$$\mu = \exp(\beta_0 + \beta_1 x_1 + \beta_3 x_3)$$

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 4.3

Tabel 4. 3 Estimasi Parameter Kemungkinan Model Regresi Poisson Inverse Gaussias

| Variabel Dari Model  | $\beta_0$ | $\beta_1$ | $\beta_2$ | $\beta_3$ | $\beta_4$ | $\beta_5$ | $\beta_6$ | $\beta_7$ | $\tau$  |
|----------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|---------|
| X1,X2,X3,X4,X5,X6,X7 | 1.21788   | 0.00118   | 0.02462   | 0.01276   | -         | 0.10862   | 0.02463   | 0.03378   | -1.182  |
|                      |           |           |           |           | 0.01545   |           |           |           |         |
| X1,X2,X3,X4,X5,X6    | 1.28454   | 0.00166   | 0.03868   | 0.01275   | -         | 0.10567   | 0.03047   |           | -1.396  |
|                      |           |           |           |           | 0.01552   |           |           |           |         |
| X1,X3                | 2.05129   | 0.00177   |           | 0.01198   |           |           |           |           | -0.8842 |

Keterangan:  $\tau$  = nilai Sigma Coefficients berdasarkan nilai estimasi

Berdasarkan tabel 4.3 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 ( $\tau$ ) yang lebih kecil sehingga untuk mengestimasi model yang didapat. Setelah memperoleh estimasi parameternya, selanjutnya adalah pengujian hipotesis untuk regresi Poisson Inverse Gaussian.

## 6. Pengujian Hipotesis

### 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 statistik 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 variabel respon)

$$H_1 : \text{paling sedikit ada } \beta_j \neq 0 \text{ dengan } j = 1, 2, 3, 4, 5, 6$$

(Variabel prediktor secara simultan berpengaruh terhadap variabel respon)

$$\alpha = 0,1$$

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 4.4 berikut:

Tabel 4. 4 Pengujian Parameter Regresi Poisson Inverse Gaussian Secara Serentak

| Variabel dari Model    | $G_{\text{Hitung}}$ | $\nu$ | $\chi^2_{(\alpha, \nu)}$ | Keputusan   |
|------------------------|---------------------|-------|--------------------------|-------------|
| X1,X2,X3, X4,X5,X6, X7 | 121.6573            | 10    | 15.9871                  | Tolak $H_0$ |
| X1,X2,X3, X4,X5,X6     | 121.7312            | 11    | 17.2750                  | Tolak $H_0$ |
| X1,X3                  | 128.2424            | 15    | 22.3071                  | Tolak $H_0$ |

Dari tabel 4.4 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  $G_{\text{Hitung}}$  dan chi-square berdasarkan derajat bebas ( $\nu$ ).

### b. Pengujian Parameter Secara Parsial

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

#### ➤ Parameter $\beta$

$$H_0 : \beta_i = 0$$

$$H_1 : \beta_i \neq 0$$

$$\alpha = 0,1$$

#### ➤ Parameter $\tau$

$$H_0 : \tau = 0$$

$$H_1 : \tau \neq 0$$

$$\alpha = 0,1$$

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 4.5

Tabel 4. 5 Pengujian Parameter Regresi Poisson Inverse Gaussian Secara Parsial

| Variabel dari Model   | Parameter Signifikan | Nilai AIC |
|-----------------------|----------------------|-----------|
| X1,X2,X3,X4, X5,X6,X7 | -                    | 139.6573  |
| X1,X2,X3,X4, X5,X6    | $\beta_1, \beta_3$   | 137.7312  |
| X1,X3                 | $\beta_1, \beta_3$   | 136.2424  |

Berdasarkan tabel 4.5 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 = 136.2424 dengan variabel prediktor persentase kepadatan penduduk (X1) dan



persentase penggunaan air bersih (X3) yang signifikan terhadap model.

## 7. Pemilihan Model Terbaik

Pada penelitian ini pemodelan dilakukan dengan menggunakan *package* *glmss* 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, didapatkan hasil estimasi parameter model Regresi Poisson Invers Gaussian yang dapat dilihat pada tabel. 4.6

Tabel 4. 6 Penaksiran Parameter Model Regresi Poisson Inverse Gaussian Pada Jumlah Penyakit Kusta di Provinsi Sulawesi Tenggara

| Parameter | Estimasi | Std. Error | Zhitung | P-value        |
|-----------|----------|------------|---------|----------------|
| $\beta_0$ | 2.05129  | 0.41753    | 4.913   | 0.000284 (***) |
| $\beta_1$ | 0.00177  | 0.00090    | 1.963   | 0.071470 (.)   |
| $\beta_3$ | 0.01198  | 0.00655    | 1.829   | 0.071470 (.)   |
| $\tau$    | -0.8842  | 0.5119     | -1.727  | 0.108          |

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

Nilai pada tabel di atas menunjukkan nilai P-value pada uji parameter dispersi  $\tau$  yang lebih kecil dari nilai  $\alpha$  sehingga tolak  $H_0$  yang berarti nilai parameter dispersi tidak sama dengan nol dan dapat disimpulkan bahwa terjadi fenomena overdispersi pada jumlah kasus penyakit kusta di Provinsi Sulawesi Tenggara tahun 2017.

Berdasarkan tabel. 4.6 maka didapat nilai P-value dan nilai AIC yang signifikan untuk membentuk model. Sehingga diperoleh model Regresi Poisson Invers Gaussian berikut:

$$\hat{\mu} = \exp(\beta_0 + \beta_1 x_1 + \beta_3 x_3)$$

$$\hat{\mu} = \exp(2.05129 + 0.00177 X_1 + 0.01198 X_3)$$

Berdasarkan model tersebut maka dapat diinterpretasikan bahwa setiap penambahan 1 persen variabel  $X_1$  yang signifikan akan melipatgandakan rata-rata variabel respon  $Y$  sebesar  $\exp(0,00177) = 1.00177$ . Dengan kata lain, penambahan 1 rasio persentase kepadatan penduduk akan sebanding dengan kenaikan rata-rata jumlah penyakit kusta di Provinsi Sulawesi Tenggara sebesar 1.00177.

Setiap penambahan 1 persen variabel  $X_3$  yang signifikan akan melipatgandakan rata-rata

variabel respon  $Y$  sebesar  $\exp(0.01198) = 1.01205$ . Dengan kata lain, penambahan 1 rasio persentase penggunaan air bersih maka akan sebanding dengan kenaikan rata-rata jumlah penyakit kusta di Provinsi Sulawesi Tenggara sebesar 1.01205.

## KESIMPULAN

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

1. Berdasarkan pola persebaran penyakit kusta di Provinsi Sulawesi Tenggara tahun 2017. Dapat diketahui bahwa wilayah dengan memiliki penyakit kusta tinggi berkisar antara 8,5889 sampai 13,1902. Dimana terdapat 5 Kabupaten/kota yang memiliki penyakit kusta tinggi. Adapun Kabupaten/Kota tersebut adalah Kota Kolaka, Kabupaten Bombana, Kabupaten Buton, Kabupaten Buton Tengah, dan Kota Bau-Bau. Dan adapun Kabupaten/kota yang memiliki penyakit kusta rendah berkisar antara 0,6134 sampai 3,3742. Dimana terdapat 6 Kabupaten/kota yang memiliki penyakit kusta rendah. Adapun Kabupaten/Kota tersebut adalah Kabupaten Konawe, Kabupaten Konawe Timur, Kabupaten Konawe Selatan, Kabupaten Konawe Utara, Kabupaten Konawe Kepulauan, dan Kabupaten Buton Utara
2. Setelah dilakukan uji overdispersi pada data jumlah penyakit kusta di Provinsi Sulawesi Tenggara pada tahun 2017 mengalami overdispersi sehingga diuji lanjut menggunakan regresi Poisson Inverse Gaussian (PIG).

Model regresi PIG yang terbentuk adalah:

$$\hat{\mu} = \exp(2.05129 + 0.00177 X_1 + 0.01198 X_3)$$

Berdasarkan hasil analisis maka didapat model regresi PIG yang terbentuk dari beberapa variabel yang berpengaruh secara signifikan terhadap jumlah penyakit kusta di Provinsi Sulawesi Tenggara adalah persentase kepadatan penduduk dan persentase penggunaan air bersih. Bahwa setiap penambahan 1 persen variabel  $X_1$  yang signifikan akan melipatgandakan rata-rata variabel respon  $Y$  sebesar  $\exp(0,00177) = 1.00177$ . Dengan kata lain, penambahan 1 rasio persentase kepadatan penduduk akan sebanding dengan kenaikan rata-rata jumlah penyakit kusta di Provinsi Sulawesi

Tenggara sebesar 1.00177. Dan setiap penambahan 1 persen variabel X3 yang signifikan akan melipatgandakan rata-rata variabel respon Y sebesar  $\exp(0.01198) = 1.01205$ . Dengan kata lain, penambahan 1 rasio persentase penggunaan air bersih maka akan sebanding dengan kenaikan rata-rata jumlah penyakit kusta di Provinsi Sulawesi Tenggara sebesar 1.01205.

## DAFTAR PUSTAKA

- Berk, McDonald, J. M. 2008. *Overdispersion and Poisson Regression*. Philadelphia: Springer.
- Cahyandari. 2014. *Pengujian Overdispersi pada Model Regresi Poisson*. UIN Sunan Gunung Djati.
- Cameron, A.C. dan Trivedi, P.K. 1990. *Regression-Based Test For Overdispersion In The Poisson Model*. Journal of Econometrics. Vol. 46, No. 1, hal. 347-364.
- Cameron, A.C. dan Trivedi, P.K. 1998. *Regression Analysis of Data Count*. 1 edition, Cambridge University, Press., New York.
- Cole, T.J. dan Green, P.J. 1992. *Smoother Reference Centile Curves: The LMS Method and Penalized Likelihood*. Statistics in Medicine. Vol. 11, hal. 1305-1319.
- 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 Assessment, Vol. 91, No. 2, hal. 121-136.
- Darnah, 2011. Mengatasi Overdispersi pada Model Regresi Poisson dengan Generalized Poisson I. *Skripsi*. Universitas Mulawarman.
- Dean, C., Lawless, J. F. dan Willmot, G.E. 1989a. *A Mixed Poisson-inverse Gaussian Regression Model*. The Canadian Journal of Statistics, Vol. 17, No. 2, hal. 171-181.
- Dean, C., Lawless, J. F. dan Willmot, G.E. 1989b. *Tests for Detecting Overdispersion in Poisson Regression Model*. Journal of the American Statistical Association, Vol. 84, No. 406, hal. 467-472.
- Déniz, E. G., Ghitany, M. E. dan Gupta, R. C. 2014. *Poisson-Mixed Inverse Gaussian Regression Model And Its Application*. Communications in Statistics - Simulation and Computation, DOI: 10.1080/03610918.2014.925924
- De Jong, P. dan Heller, G.Z. 2008. *Generalized Linear Models for Insurance Data*. 1<sup>st</sup> edition, Cambridge University, Press., New York.
- Darnah, 2011. Mengatasi Overdispersi pada Model Regresi Poisson dengan Generalized Poisson I. Universitas Mulawarman.
- Dinas Kesehatan Provinsi Sulawesi Tenggara. 2018. *Profil Kesehatan Provinsi Sulawesi Tenggara 2017*. Kendari: DinKes Sultra
- Ernawati, 2016. *Analisis Faktor-Faktor yang Mempengaruhi Angka Prevalensi Penyakit Kusta di Jawa Timur dengan Pendekatan Spatial Durbin Model*. Jurnal Sains dan Seni ITS. Vol. 5, No. 2
- Fathurahman. 2009. *Pemilihan Model Regresi Terbaik Menggunakan Metode Akaike's Information Criteria dan Schwarz Information Criterion*. *Skripsi*. Universitas Mulawarman.
- Filliben, J.J. 1975. *The Probability Plot Correlation Coefficient Test for Normality*. Technometrics. Vol. 17, No. 1, hal. 111-117.
- Herindrawati, Latra, Purhadi. 2017. *Pemodelan Regresi Poisson Inverse Gaussian*. Institut Teknologi Sepuluh Nopember.

- Herindrawati. 2017. *Pemodelan Regresi Poisson Inverse Gaussian*. Skripsi. Institut Teknologi Sepuluh Nopember.
- Hilbe, J.M. 2007. *Negative Binomial Regression*. 1<sup>st</sup> edition, Cambridge University, Press., New York.
- Karlis, D. dan Nikoloulopoulos, E. 2005. *Mixed Poisson Distribution*. International Statistical Review. Vol. 73, No. 1, hal. 35-58.
- Karlis, D. dan Xekalaki, E. 2000. *A Simulation Comparison of Several Procedures for Testing the Poisson Assumption*. The Statistician. Vol. 49, No. 3, hal. 355-382.
- 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.
- Kismiantini. 2008. *Perbandingan Model Regresi Poisson Dan Model Regresi Binomial Negatif*. Skripsi. Universitas Negeri Yogyakarta
- Muharry, A. 2014. *Faktor Resiko Kejadian Kusta*. Jurnal Kesehatan Masyarakat. Vol. 9, hal. 174-182
- Nuraeni, 2018. *Pemodelan Jumlah Kematian Bayi Di Provinsi Sulawesi Selatan Menggunakan Regresi Poisson Inverse Gaussian*. Skripsi. Program Studi Matematika. Jurusan Matematika. Fakultas Matematika dan Ilmu Pengetahuan Alam. Universitas Negeri Makassar. Makassar
- Pradeka, Rangga. 2012. *Uji Koefisien Korelasi Spearman dan Kendall Menggunakan Metode Bootstrap dan sampel yang Dibandingkan Berdasarkan Estimasi Densitas Kernel Multivariat*. Skripsi. Universitas Kristen Satya Wacana.
- Pujiastuti, A.G. 2009. *Sebaran Kasus Penyakit Kusta Berdasarkan Wilayah Kecamatan di Kota Bekasi Tahun 2006-2008*. Skripsi. Program Studi Kesehatan Masyarakat, Fakultas Kesehatan Masyarakat, Universitas Indonesia. Depok
- Rigby, R.A. dan Stasinopoulos, D.M. 1996 *A Semi-parametric Additive Model for variance Heterogeneity*. Statistics and Computing. Vol. 6, hal. 57-65.
- Shoukri, M.M., Asyali, M.H., Vandorp, R. dan Kelton, R. 2004. *The Poisson Inverse Gaussian Regression Model in the Analysis of Clustered Counts Data*. Journal of Data Science. Vol. 2, No. 2, hal. 17-32.
- Stasinopoulos, D.M. dan Rigby, R.A. 2007. *Generalized Additive Models for Location Scale and Shape*. Journal of Statistical Software. Vol. 23, hal. 1-46.
- Stasinopoulos, D.M. dan Rigby, R.A. 2007. *Generalized Additive Models for Location Scale and Shape (GAMLSS) in R*. Journal of Statistical Software. Vol. 20, hal. 1-64.
- Ummah, Z., Suliyanto dan Sediono 2013. *Estimasi Model Linier Tergeneralisasi Gaussian Berdasarkan Maksimum Likelihood Estimator dengan Menggunakan Algoritma Fisher Scoring*. Jurnal Matematika, Vol. 1, No.1, hal. 110-120.
- Walpole, R.E. 1995. *Pengantar Metode Statistika*. Ahli Bahasa: Ir. Bambang Sumantri, Jakarta: PT Gramedia Pustaka Utama.
- Widiari, S. M. 2013. *Penaksiran Parameter Dan Statistik Uji Dalam Model Regresi Poisson Inverse Gaussian (PIG)*. Tesis. Jurusan Statistika. Fakultas Matematika dan Ilmu Pengetahuan Alam. Institut Teknologi Sepuluh Nopember. Surabaya
- Willmot, G.E. 1987. *The Poisson-Inverse Gaussian Distribution as An Alternative to the Negative Binomial*. Scandinavian Actuarial Journal. Vol. 3, No. 4, hal. 113 127.



- Widiari, S, M. 2016. *Penaksiran Parameter dan Statistik Uji dalam Model Regresi Poisson Inverse Gaussian (PIG)*. Tesis. Institut Teknologi Sepuluh Nopember.
- Zha, L., Lord, D. dan Zou, Y. 2014. *The Poisson Inverse Gaussian (PIG) Generalized Linear Regression Model for Analyzing Motor Vehicle Crash Data*. Journal of Transportation Safety and Security. DOI: 20.2080/19439962.2014.977502.

