

ANALISIS FAKTOR PENYEBAB PENYAKIT DBD DI JAWA TENGAH MENGGUNAKAN REGRESI BINOMIAL NEGATIF

Edy Widodo, Putri Meliana Ariani

Universitas Islam Indonesia
edywidodo.uui.ac.id, 14611257@students.uui.ac.id

ABSTRAK

Latar Belakang: Dalam analisis regresi Poisson, variabel terikat harus memenuhi asumsi yaitu nilai variansi sama dengan rata-ratanya. Pada kenyataannya yang terjadi ketika melakukan analisis adalah variansi dari variabel terikatnya lebih besar daripada rata-ratanya yang disebut dengan terjadinya kasus overdispersi.

Tujuan: Untuk memilih model terbaik yang digunakan untuk memodelkan kasus terjadinya penderita penyakit DBD di Jawa Tengah tahun 2016.

Metode: Metode yang digunakan yaitu Analisis Deskriptif untuk mengetahui gambaran umum data, regresi Poisson yang digunakan untuk data diskrit berupa bilangan cacah dan regresi Binomial negatif sebagai alternatif dari data overdispersi.

Hasil: Hasil analisis menunjukkan bahwa model regresi Binomial Negatif yang terbaik digunakan untuk mengatasi kasus overdispersi pada model regresi Poisson.

Kesimpulan: Penerapan regresi Poisson dan regresi Binomial Negatif pada data jumlah kasus penyakit DBD di Jawa Tengah tahun 2016 yang mengandung overdispersi terjadi pada penarikan kesimpulan uji signifikansi parameter.

Kata kunci: Overdispersi, Regresi Poisson, Regresi Binomial Negatif

ABSTRACT

Background : In the Poisson regression analysis, the dependent variable must satisfy the assumption that the variance value is equal to the mean. In fact, what happens when doing the analysis is the variance of the dependent variable is greater than the average called the occurrence of cases of overdispersion.

Objective : To choose the best model the case of dengue fever case in Central Java Year 2016.

Methods : The method used is Descriptive Analysis to know the general description of data, Poisson regression used for discrete data in the form of count and Binomial negative regression as an alternative to overdispersion data.

Results : The results of the analysis show that the best binomial negative regression model is used to overcome the case of overdispersion in the Poisson regression model.

Conclusion : The application of Poisson regression and Binomial Negative regression to data on the number of DHF cases in Central Java in 2016 containing overdispersion occurred in the withdrawal of conclusions of the parameter significance test.

Keywords : Overdispersion, Poisson Regression, Negative Binomial Regression

PENDAHULUAN

Tingkat curah hujan dan kelembaban yang tinggi di Indonesia merupakan salah satu faktor yang mengakibatkan cepatnya perkembangan dari sumber penyakit. Banyak penyakit yang dapat terjadi pada saat musim penghujan salah satunya penyakit *DBD*. Gejala yang di derita penyakit *DBD* yaitu dapat membuat suhu tubuh penderita menjadi sangat tinggi dan pada umumnya di sertai sakit kepala, nyeri sendi, otot dan tulang serta nyeri di bagian belakang mata (Basuki, 2016). Berdasarkan hasil penelitian yang dilakukan oleh Kementerian Kesehatan RI, tercatat bahwa jumlah penderita *DBD* di Indonesia pada bulan Januari-Februari 2016 sebanyak 8.487 orang dengan jumlah kematian 108 orang.

Berdasarkan keadaan-keadaan di atas, maka peneliti bermaksud untuk melakukan penelitian terhadap faktor-faktor yang mempengaruhi penyakit *DBD* di Provinsi Jawa Tengah untuk mendapatkan model terbaik dalam menggambarkan penyakit *DBD*. Manfaat yang bisa didapatkan yaitu masyarakat bisa melakukan pencegahan awal dengan cepat karena faktor utama penyebab *DBD* telah diketahui dari model terbaik yang didapatkan. Penelitian yang dilakukan oleh Fatmasari (2011) mengenai Regresi Poisson dan oleh Sartika (2012) mengenai Regresi Binomial Negatif sebagai alternatif dari data yang mengalami overdispersi. Dari kajian pustaka tersebut, maka peneliti ingin mencari model terbaik dari regresi Poisson dan regresi Binomial Negatif dalam mengatasi data yang mengalami kasus overdispersi di Jawa Tengah. Sepengetahuan peneliti belum pernah ada penelitian tentang Analisis faktor penyebab penyakit *DBD* di Jawa Tengah tahun 2016 yang menggunakan regresi Binomial Negatif dan alasan tersebut yang menjadikan penelitian ini berbeda dari penelitian sebelumnya.

METODE

Populasi yang diambil dalam penelitian ini adalah semua penduduk Provinsi Jawa Tengah yang pernah

mengalami penyakit *DBD*. Sampel yang digunakan dalam penelitian ini adalah sebagian jumlah penderita *DBD* yang ada di Provinsi Jawa Tengah selama tahun 2016. Jenis data yang digunakan dalam penelitian ini adalah data sekunder dalam bentuk tahunan selama tahun 2016. Sumber data di ambil dari Dinas Kesehatan Provinsi Jawa Tengah dan Badan Pusat Statistik Provinsi Jawa Tengah. Penelitian ini menggunakan dua metode pengumpulan data yaitu dari studi pustaka yaitu mengumpulkan data dan teori yang relevan terhadap permasalahan yang diteliti dan studi documenter yaitu mengumpulkan data dari Dinas Kesehatan dan Badan Pusat Statistik Jawa Tengah.

Metode analisis yang digunakan adalah Analisis regresi Binomial Negatif. Regresi Binomial neaktif adalah model *non linier* yang berasal dari distribusi *poisson-gamma mixture* yang merupakan penerapan dari *Generalized Linear Model (GLM)* yang menggambarkan hubungan antara variabel dependen dengan variabel independent. Regresi Binomial Negatif digunakan untuk memodelkan data dengan variabel respon berupa data *count*. Regresi Binomial Negatif digunakan sebagai alternatif dari model regresi Poisson yang mengalami overdispersi yaitu nilai variansi lebih besar dari *mean* (Pingit, 2009).

HASIL DAN PEMBAHASAN

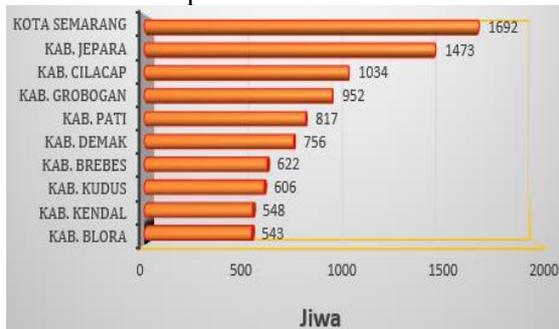
A. Analisis Deskriptif

Hal pertama yang di lakukan dalam penelitian yaitu melakukan analisis deskriptif dari variabel yang digunakan.

Tabel 1. Statistika Deskriptif

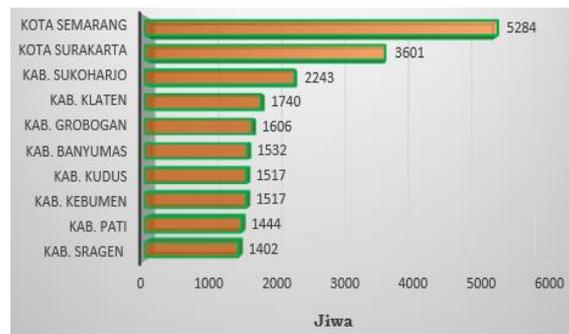
| Variabel | Rata-rata | Variansi |
|---|-----------|----------|
| Jumlah Kasus <i>DBD</i> (Y) | 462.3 | 139113 |
| Kepadatan Penduduk (X ₁) | 2015.4 | 5734197 |
| Ketinggian Wilayah (X ₂) | 775.6 | 693336.5 |
| Jumlah Tenaga Kesehatan (X ₃) | 1290 | 769480.6 |
| Jumlah Sarana Kesehatan (X ₄) | 66.1 | 2512.387 |
| Jumlah Curah Hujan (X ₅) | 1737.8 | 330395.4 |

Terlihat bahwa nilai variansi dari variabel dependen lebih besar dari rata-rata, sehingga bisa dikatakan data Jumlah Kasus *DBD* di Jawa Tengah mengalami kasus overdispersi.



Gambar 1. Kabupaten dengan penderita penyakit *DBD* terbanyak berdasarkan jumlah penyakit *DBD*

Gambar 1 menunjukkan sepuluh kabupaten/kota dengan jumlah penderita *DBD* paling banyak di Provinsi Jawa Tengah Tahun 2016. Jumlah penderita penyakit *DBD* paling banyak terjadi pada Kota Semarang sebanyak 1692 jiwa, hal tersebut dikarenakan masih banyak masyarakat yang tidak sadar dan tidak peduli terhadap lingkungan disekitar. Masih banyaknya genangan-genangan air yang menyebabkan jentik nyamuk sangat cepat berkembang (Prayoga, 2016).



Gambar 2. Kabupaten dengan penderita penyakit *DBD* terbanyak berdasarkan jumlah tenaga kesehatan

Diketahui bahwa jumlah tenaga kesehatan paling banyak terjadi di Kota Semarang sebanyak 5284 jiwa berdasarkan Gambar 2 di atas. Hal tersebut dapat menjadi indikasi karena kota Semarang juga mempunyai jumlah kasus *DBD* paling banyak pula seperti Gambar 1.

B. Uji Kolmogorov-Smirnov

Tujuan dari Uji Kolmogorov-Smirnov adalah untuk menentukan apakah data jumlah kasus *DBD* di Provinsi Jawa Tengah tahun 2016 pada 35 kabupaten/kota mengikuti distribusi Poisson atau tidak.

Tabel 2. Tampilan Uji Kolmogorov-Smirnov

| Jumlah Sampel | Nilai Asymp. Sig |
|---------------|------------------|
| 35 | 0.203 |

Analisis *output* untuk uji Kolmogorov-Smirnov adalah sebagai berikut:

- (i) Hipotesis
 - $H_0 : F(X) = F_0(X)$ (Data berasal dari populasi berdistribusi Poisson)
 - $H_0 : F(X) \neq F_0(X)$ (Data bukan berasal dari populasi berdistribusi Poisson)
- (ii) Tingkat Signifikansi
 - $\alpha = 0.05$
- (iii) Keputusan
 - Berdasarkan hasil *output* dari Tabel 2, diperoleh nilai *Asymp. Sig* = 0.203 lebih besar dari Nilai $\alpha = 0.05$, yang berarti gagal tolak H_0 .
- (iv) Kesimpulan
 - Sehingga dapat disimpulkan bahwa variabel terikat adalah berasal dari populasi berdistribusi Poisson.

C. Uji Multikolinieritas

Uji Multikolinieritas diperlukan untuk mengetahui ada tidaknya variabel independent yang lain dalam satu model. Pengujian multikolinieritas menggunakan nilai VIF dimana nilai VIF > 10 maka dikatakan mengalami multikolinieritas. Uji hipotesisnya adalah sebagai berikut:

- (i) Hipotesis
 H_0 : Tidak terdapat hubungan antar variabel independen
 H_1 : Terdapat hubungan antar variabel independen
- (ii) Tingkat Signifikansi
 $\alpha = 0.05$
- (iii) Daerah Kritis
Tolak H_0 jika nilai VIF > 10
- (iv) Keputusan

Tabel 3. Keputusan Hasil Pengujian Multikolinieritas

| Variabel | VIF | Kriteria |
|-------------------|-----------|----------|
| (X ₁) | 1.609 (*) | 10 |
| (X ₂) | 1.147 (*) | 10 |
| (X ₃) | 1.664 (*) | 10 |
| (X ₄) | 1.288 (*) | 10 |
| (X ₅) | 1.076 (*) | 10 |

(*) Tidak terjadi multikolinieritas

- (v) Kesimpulan
Dapat di lihat bahwa nilai VIF dari masing-masing variabel bebas tidak ada yang lebih dari 10. Hal tersebut menunjukkan bahwa antar variabel bebas tidak terjadi kasus multikolinieritas, sehingga layak di iikutsertakan dalam pembentukan model regresi Poisson dan regresi Binomial Negatif.

D. Model Regresi Poisson

Hasil pendugaan parameter untuk model regresi Poisson dapat di lihat pada Tabel 4. Hasil ini di peroleh menggunakan perangkat lunak R 3.4.2.

Tabel 4. Nilai dugaan parameter model regresi Poisson

| P | Estimate | Pr (> z) |
|-----------|-------------------------|------------|
| β_0 | 6.739 | <2e-16 *** |
| β_1 | 1.874x10 ⁻⁴ | <2e-16 *** |
| β_2 | 1.038x10 ⁻⁴ | <2e-16 *** |
| β_3 | 4.113x10 ⁻⁴ | <2e-16 *** |
| β_4 | 1.474x10 ⁻⁴ | <2e-16 *** |
| β_5 | -5.180x10 ⁻⁴ | <2e-16 *** |

Deviance : 9590.8; derajat bebas: 34; Rasio disperse: 5221,6; AIC: 493

Maka diperoleh model regresi Poisson sebagai berikut:

$$\ln(\mu_i) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5$$

$$(\mu_i) = \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5)$$

$$(\mu_i) = \exp(6.739 - 1.874 \times 10^{-4} - 1.038 \times 10^{-4} + 4.113 \times 10^{-4} + 1.474 \times 10^{-3} - 5.180 \times 10^{-4})$$

Model tersebut menunjukkan bahwa setiap penambahan satu orang di suatu kabupaten/kota akan menyebabkan nilai harapan penderita DBD meningkat sebesar $\exp(6.739) = 844.71559$ kali dengan asumsi peubah lain dianggap tetap. Artinya, setiap penambahan 10000 penduduk akan meningkatkan nilai harapan jumlah penderita DBD sebanyak 8447156 orang dengan asumsi peubah lain di anggap tetap.

E. Uji Parsial Model Regresi Poisson

Langkah berikutnya adalah uji parsial, uji ini digunakan untuk menentukan apakah secara parsial variabel-variabel independent berpengaruh terhadap variabel dependen. Sehingga digunakan uji hipotesis sebagai berikut:

- (i) Hipotesis
 H_0 : $\beta_j = 0$ (Variabel independent tidak berpengaruh terhadap variabel dependen)
 H_1 : Paling sedikit ada satu j dengan $\beta_j \neq 0$ (Variabel independent berpengaruh terhadap variabel dependen)
- (ii) Keputusan

Tabel 5. Uji Parsial Model Regresi Poisson

| Variabel | $\chi^2_{0.05,1}$ | Nilai W |
|----------------|-------------------|---------------------------|
| Konstanta | 3.841 | 85897.78 (*) |
| X ₁ | 3.841 | 3.40x10 ¹⁰ (*) |
| X ₂ | 3.841 | 8.95x10 ¹¹ (*) |
| X ₃ | 3.841 | 1.46x10 ¹⁰ (*) |
| X ₄ | 3.841 | 3.80x10 ⁶ (*) |
| X ₅ | 3.841 | 3.40x10 ⁹ (*) |

(*) Signifikan karena tolak H_0 (Nilai W lebih besar dari $\chi^2_{0.05,1}$)

(iii) Kesimpulan

Secara parsial variabel independen berpengaruh terhadap variabel dependen sehingga model layak digunakan.

F. Overdispersi

Overdispersi pada data kasus DBD di Jawa Tengah tahun 2016 ditunjukkan pada Tabel 1 dimana variansi Y lebih besar dari rataan Y. Selain itu, fenomena overdispersi pada data kasus DBD di Jawa Tengah dapat di lihat berdasarkan nilai *Pearson Chi-Squares* dan *Deviance* yang dibagi dengan derajat bebas bernilai lebih dari 1.

Tabel 6. Hasil Uji Overdispersi

| Nilai Deviance | Db |
|----------------|----|
| 9591 | 34 |

Jika nilai *deviance* dibagi dengan derajat bebas maka didapatkan nilai 282. Nilai yang dihasilkan melebihi angka 1, hal tersebut menunjukkan terjadi kasus overdispersi pada model regresi Poisson. Adanya overdispersi menyebabkan model regresi Poisson menjadi kurang baik, karena memiliki tingkat kesalahan yang tinggi. Salah satu cara untuk mengatasi adanya kasus overdispersi dalam regresi Poisson adalah dengan mengganti asumsi distribusi Poisson dengan distribusi Binomial Negatif.

G. Model Regresi Binomial Negatif

Hasil pendugaan parameter untuk model regresi Binomial Negatif dapat dilihat pada Tabel 7 berikut.

Tabel 7. Nilai dugaan parameter model regresi Binomial Negatif

| P | Estimate | Pr (> z) |
|-----------|-------------------------|------------|
| β_0 | 6.557 | <2e-16 *** |
| β_1 | -2.003x10 ⁻⁴ | <2e-16 *** |
| β_2 | -9.408x10 ⁻⁵ | <2e-16 *** |
| β_3 | 4.548x10 ⁻⁴ | <2e-16 *** |
| β_4 | 3.920x10 ⁻³ | <2e-16 *** |
| β_5 | -5.392x10 ⁻⁴ | <2e-16 *** |

Deviance: 60.869; derajat bebas: 34; Rasio disperse: 37.549; AIC: 5499,2

Maka diperoleh model regresi Binomial Negatif sebagai berikut:

$$\ln(\mu_i) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5$$

$$(\mu_i) = \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5)$$

$$(\mu_i) = \exp(6.557 - 2.003 \times 10^{-4} - 9.408 \times 10^{-5} + 4.548 \times 10^{-4} + 3.920 \times 10^{-3} - 5.392 \times 10^{-4})$$

Model tersebut menunjukkan bahwa setiap penambahan satu orang di suatu kabupaten/kota akan menyebabkan nilai harapan penderita DBD meningkat sebesar $\exp(6.557) = 704.15605$ kali dengan asumsi peubah lain di anggap tetap. Artinya, setiap penambahan 10000 penduduk akan meningkatkan nilai harapan jumlah penderita DBD sebanyak 7041561 orang dengan asumsi peubah lain di anggap tetap.

H. Uji Parsial Model Regresi Binomial Negatif

Digunakan uji hipotesis sebagai berikut:

(i) Hipotesis

$H_0 : \beta_j = 0$ (Variabel independent tidak berpengaruh terhadap variabel dependen)

$H_1 : \text{Paling sedikit ada satu } j \text{ dengan } \beta_j \neq 0$ (Variabel independent berpengaruh terhadap variabel dependen)

(ii) Keputusan

Tabel 8. Uji Parsial Model Regresi Binomial Negatif

| Variabel | $\chi^2_{0.05,1}$ | Nilai W |
|-----------|-------------------|--------------------------|
| Konstanta | 3.841 | 4.864 (*) |
| X_1 | 3.841 | 2.81x10 ⁸ (*) |
| X_2 | 3.841 | 0.480 |
| X_3 | 3.841 | 3.68x10 ⁷ (*) |
| X_4 | 3.841 | 0.155 |
| X_5 | 3.841 | 2.42x10 ⁷ (*) |

(*) Signifikan karena tolak H_0 (Nilai W lebih besar dari $\chi^2_{0.05,1}$)

(iii) Kesimpulan

Secara parsial hanya variabel Kepadatan penduduk, Jumlah tenaga kesehatan dan Jumlah curah hujan

yang berpengaruh dan memiliki kontribusi terhadap variabel dependen.

I. Kesesuaian Model Regresi Binomial Negatif

Pemilihan model terbaik dilakukan untuk mendapatkan model terbaik dari kedua model regresi yang dibandingkan untuk mengatasi kasus overdispersi pada model regresi Poisson.

Tabel 9. Pemilihan Model Terbaik

| Model | Nilai AIC | Deviance/db |
|-------|-----------|-------------|
| RP | 5499 | 330.7 |
| RBN | 493 | 1.7 |

RP : Regresi Poisson

RBN : Regresi Binomial Negatif

Nilai AIC untuk model regresi Binomial Negatif lebih kecil dibandingkan dengan model regresi Poisson sehingga model terbaik untuk data kasus DBD di Jawa Tengah yaitu dengan menggunakan model regresi Binomial Negatif.

KESIMPULAN

Model terbaik untuk menggambarkan penyebab penyakit DBD di Provinsi Jawa Tengah tahun 2016 yaitu dengan menggunakan regresi Binomial Negatif karena memiliki nilai AIC yang paling kecil dibandingkan regresi Poisson.

DAFTAR PUSTAKA

Agresti, A. 2002. *Categorical Data Analysis. Second Edition*. New York: John Wiley and Sons, Inc.

Algifari. 2000. *Penyajian Data Statistik*. Kudus: Tidak Dipublikasikan.

Aradea. 2011. *Demam Berdarah Dengue*. (Online)
<http://id.shvoong.com/medisine-andhealth/alternative-medicine/1869006-anti-demam-berdarah-dengue-bukan/.html>
(diakses 07 Oktober 2017 pukul 10.58 WIB)

BPS. 2014. *Provinsi Jawa Tengah Dalam Angka*. Jawa Tengah: BPS

BPS. 2015. *Provinsi Jawa Tengah Dalam Angka*. Jawa Tengah: BPS

BPS. 2016. *Provinsi Jawa Tengah Dalam Angka*. Jawa Tengah: BPS

Pingit. 2009. *Analisis Data Kategorik*. Surabaya: Jurusan Statistika ITS.

Prayoga, Nyoman. 2016. *Kota Semarang Bergerak Bersama Cegah DBD, 100% Bebas Jentik* (Online)
http://www.kompasiana.com/nyomanprayoga/kota-semarang-bergerak-bersama-cegah-dbd-100-bebas-jentik_573a8f28ba93731f05c055a5.html (diakses pada 8 February 2018 pukul 21.49 WIB).

Walpole, R, Myers, R. 1995. *Ilmu Peluang Dan Statistik Untuk Insinyur dan Ilmuwan*. Bandung: ITB.

Wikipedia. 2017. *Sensus Provinsi Jawa Tengah*. (Online).
<https://id.wikipedia.org/wiki/SensusJawaTengah>, diakses 28 Desember 2017 pukul 14.00 WIB).

