

## ANALISIS FAKTOR LINGKUNGAN PENYEBAB KASUS DIARE DI JAWA BARAT TAHUN 2016 MENGGUNAKAN REGRESI BINOMIAL NEGATIF

Ninik Kardinah Lestari<sup>1</sup>, Dimas Wahyu Hikmawan<sup>2</sup>,  
Ilham Safitrah<sup>3</sup>, Edy Widodo<sup>4</sup>

<sup>1)2)3)4)</sup>Program Studi Statistika, FMIPA, Universitas Islam Indonesia  
15611052@students.uui.ac.id<sup>1</sup>, 15611058@students.uui.ac.id<sup>2</sup>,  
15611051@students.uui.ac.id<sup>3</sup>, edywidodo@uui.ac.id<sup>4</sup>

### Abstrak

*Analisis regresi binomial negatif merupakan alternatif analisis yang digunakan ketika asumsi equidispersion pada analisis regresi poisson tidak terpenuhi atau disebut dengan overdispersion. Gangguan kesehatan merupakan penghambat bagi manusia untuk melaksanakan aktifitas sehari-hari. Diare merupakan gangguan kesehatan terbanyak yang ditemukan di Indonesia mencapai 2.544.084 kasus, 930.176 kasus diantaranya ditemukan di Jawa Barat dan gangguan kesehatan ini berbahaya karena menyebabkan kekurangan gizi dan kematian. Lingkungan yang diperhatikan berkaitan dengan terciptanya kesehatan yang baik sehingga gangguan kesehatan seperti penyakit diare menjadi terkendali. Oleh karena itu penelitian ini bertujuan untuk mengetahui faktor lingkungan terdiri dari 5 variabel independen yang berpengaruh terhadap kasus diare serta model terbaik yang dihasilkan dengan membandingkan hasil model analisis regresi poisson dan analisis regresi binomial negatif. Dengan melihat kriteria pemilihan model terbaik AIC diperoleh hasil bahwa model terbaik adalah model dari hasil analisis regresi binomial negatif dan hasil menunjukkan juga bahwa dari 5 variabel independen dari faktor lingkungan, terdapat satu faktor yang berpengaruh terhadap kasus diare di Jawa Barat tahun 2016 yaitu persentase penduduk dengan akses air minum berkualitas (layak).*

**Kata Kunci:** Diare; Faktor Lingkungan; Overdispersion; Regresi Binomial Negatif; Regresi Poisson

### 1. PENDAHULUAN

Penyakit diare adalah penyakit atau gangguan kesehatan menular langsung dan menjadi penyakit potensial Kejadian Luar Biasa (KLB) yang sering disertai dengan kematian dan malnutrisi atau kekurangan gizi. Penyakit diare merupakan penyakit menular langsung yang memiliki kasus terbanyak dibandingkan penyakit menular langsung lainnya. Diare merupakan penyakit terbanyak yang ditemukan di Indonesia mencapai 2.544.084 kasus, 930.176 kasus diantaranya ditemukan di Jawa Barat (Kementrian Kesehatan RI, 2017).

Diare merupakan salah satu penyakit berbasis lingkungan. Dua faktor lingkungan yang dominan yaitu, sarana air bersih dan pembuangan tinja. Kedua faktor ini akan berinteraksi bersama dengan perilaku manusia. Apabila faktor lingkungan tidak sehat karena tercemar bakteri diare dan berakumulasi dengan perilaku manusia yang tidak sehat juga yaitu, melalui makanan dan minuman, maka dapat menimbulkan diare (Wijoyo, 2013).

Untuk mengetahui pengaruh faktor lingkungan terhadap kasus diare menggunakan metode regresi poisson, dan binomial negatif. Model regresi poisson merupakan model regresi nonlinear yang digunakan untuk menganalisis data diskrit (*count*). Dalam regresi poisson terdapat asumsi equidispersi, namun sering terjadi pelanggaran seperti kasus overdispersi. penanganan overdispersi pada regresi Poisson juga dapat dilakukan menggunakan pendekatan model binomial negatif.

Penelitian tentang faktor penyebab diare ini dilakukan oleh Fauziah Paiman tahun 2013 adalah Penelitian yang bersifat *teoritis* yaitu menguji autokorelasi pada Model *Regresi Spasial Lag* dengan *Lagrange Multiplier* dengan studi kasus diare di Jawa Timur tahun 2010. Delapan variabel independen yang menggambarkan faktor penyebab diare yaitu persentase penduduk terhadap akses air bersih, persentase penduduk terhadap akses sanitasi, persentase penduduk terhadap ketersediaan jamban, persentase penduduk terhadap pengelolaan sampah yang baik, persentase rumah tangga dengan kategori rumah sehat, jumlah penduduk setiap kota/kabupaten, kepadatan penduduk setiap kota/kabupaten, dan jumlah kematian balita. Dari delapan variabel independen yang diteliti terhadap kasus diare diperoleh bahwa variabel prediktor yang signifikan terhadap kasus diare adalah persentase penduduk terhadap ketersediaan jamban dan jumlah penduduk setiap kota/kabupaten.

Penelitian oleh Nurvita Arumsari dan Sutikno tahun 2010 mengenai pemodelan regresi penyakit diare dengan membandingkan 3 metode regresi yaitu regresi OLS (Ordinary Least Square), poisson dan spasial GWPR (Geographically Weighted Poisson Regression). Hasil analisis dan pembahasan yang ada pemodelan kejadian diare di Kabupaten Tuban baik dengan model regresi OLS, regresi poisson, dan spasial GWPR diperoleh variabel yang berpengaruh terhadap kejadian diare yaitu jarak dan Fasilitas air minum dan model terbaik yang diperoleh merupakan model dari hasil analisis regresi spasial GWPR dengan melihat kriteria kebaikan model R-square dan AIC.

Penelitian ini bertujuan untuk mengetahui faktor lingkungan terdiri dari 5 variabel independen yang berpengaruh terhadap kasus diare serta model terbaik yang dihasilkan dengan membandingkan hasil model analisis regresi poisson dan analisis regresi binomial negatif.

Dari uraian masalah serta penelitian sebelumnya mengenai analisis faktor yang berhubungan dengan kasus diare, penulis melakukan penelitian ini dengan menambahkan dua variabel independen baru sebagai faktor lingkungan yaitu persentase tempat pengelolaan makanan yang memenuhi syarat dan persentase tempat-tempat umum yang memenuhi syarat karena diare adalah penyakit yang diderita seseorang akibat makanan yang terkontaminasi mikroba penyebab penyakit. Pengelolaan makanan yang higienis serta prosesnya dapat mematikan penyebab penyakit ini. Pencegahan penyakit diare juga dilakukan dengan melakukan penyajian makanan tidak

terkena lalat, debu, udara kotor, peralatan makan yang higienis dan mendapat surat keterangan sehat (terutama di tempat-tempat umum) (Soemirat, 2011).

Kajian teori yang digunakan untuk mendukung penelitian ini adalah sebagai berikut:

1. Statistika deskriptif adalah analisis statistika yang membahas tentang metode-metode untuk menyajikan data sehingga menarik dan informatif. Secara umum statistika deskriptif dapat diartikan sebagai metode-metode yang berkaitan dengan pengumpulan dan penyajian suatu gugus data sehingga memberikan informasi yang berguna. Mean (Rataan) didefinisikan sebagai jumlah seluruh data dibagi dengan banyaknya data (Walpole, 1995). Varians adalah suatu nilai yang menunjukkan ukuran variabilitas yang dihitung dengan cara mengkuadratkan standar deviasi (Walpole, 1995)
2. Asumsi multikolinearitas merupakan bentuk pengujian yang menyatakan bahwa variabel independen harus terbebas dari gejala multikolinearitas. Gejala multikolinearitas adalah gejala korelasi antar variabel independen (Nugraha, 2016). Untuk menemukan ada atau tidak adanya korelasi antar variabel independen dengan mendeteksi nilai *Tolerance* dan *Variance Inflation Factor (VIF)*. *Tolerance* diperoleh dari 1 dibagi dengan nilai *VIF*. Para ahli mengatakan bahwa jika nilai *tolerance* = 1, berarti tidak terjadi adanya korelasi antar variabel independen atau jika *VIF* lebih dari 10 dikatakan ada kolinearitas yang tinggi. Model regresi yang baik jika tidak terjadi korelasi antar variabel independen (Subiyakto, 2001).
3. Dalam inferensi statistik, dikenal nilai  $\alpha$  yang merupakan probabilitas kesalahan Tipe I yang biasa disebut dengan tingkat sinifikansi. Nilai  $\alpha$  terkecil sedemikian sehingga dapat menolak  $H_0$  disebut *p-value*. *P-value* dihitung berdasarkan distribusi statistiknya (Nugraha, 2016).
4. Model Regresi Poisson merupakan model regresi nonlinear yang digunakan untuk menganalisis data diskrit (*count*). Beberapa karakteristik dari percobaan yang mengikuti sebaran distribusi poisson (Cameron & Trivedi, 1998).

Model Regresi Poisson merupakan *Generalized Linear Model (GLM)* yang data responnya diasumsikan berdistribusi Poisson. Model regresi Poisson diberikan sebagai berikut:

$$y_i \sim \text{Poisson}(\mu_i)$$

$$\mu_i = \exp(\mathbf{x}_i^T \boldsymbol{\beta})$$

maka,

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

Dan jika persamaan (1) diturunkan menjadi seperti sebagai berikut:

$$\mu = \exp(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki}) \quad (2)$$

Sehingga persamaan (2) ini adalah model yang digunakan untuk interpretasi model.

5. Nilai  $D(\boldsymbol{\beta})$  disebut juga nilai devians dari model, semakin kecil nilai devians artinya semakin bagus model yang dihasilkan. Hal ini dikarenakan

nilai devians model regresi poisson mengikuti distribusi  $\chi^2$  sehingga tolak  $H_0$  jika  $D(\beta) > \chi^2(\alpha, k)$ .

6. Regresi poisson dikatakan mengandung overdispersi apabila nilai variansnya lebih besar dari nilai meannya. Overdispersi memiliki dampak yang sama dengan pelang-garan asumsi jika pada data diskrit terjadi overdispersi namun tetap digunakan regresi poisson, anak dugaan dari parameter koefisien regresinya tetap konsisten namun tidak efisien. Hal ini berdampak pada nilai *standar error* yang menjadi *under estimate*, sehingga kesimpulannya menjadi tidak valid. Fenomena overdispersi (McCullagh dan Nelder, 1983) dapat dituliskan  $var(Y) > E(Y)$ . Penanganan overdispersi pada regresi poisson juga dapat dilakukan menggunakan pendekatan model binomial negatif. Dalam regresi binomial negatif, jika  $\theta$  menuju nol maka  $var(Y_i)$  menuju  $\mu_i$  sehingga binomial negatif akan konvergen menuju poisson.
7. Model regresi binomial negatif memiliki bentuk yang sama dengan model regresi poisson yaitu pada persamaan (1) dan persamaan (2).

**2. METODE PENELITIAN**

Metode penelitian yang digunakan penulis adalah analisis regresi poisson dan analisis regresi binomial negatif. Jenis penelitian yang dilakukan adalah penerapan metode penelitian terhadap studi kasus.

Populasi dalam penelitian ini adalah seluruh penderita diare yang ditemukan di Indonesia tahun 2016. Sampel dalam penelitian ini adalah sebagian penderita diare yang ditemukan di Indonesia tahun 2016. Variabel dan definisi operasional variabel penelitian agar lebih mudah dipahami tampak seperti pada tabel 1.

Tabel 1. Variabel dan Definisi Operasional

Variabel	Definisi Operasional Variabel	Rumus	Satuan	Skala
Y	Banyaknya penderita diare yang ditangani di sarana kesehatan di Jawa Barat tahun 2016	-	Jiwa	Rasio
X <sub>1</sub>	Persentase Penduduk Dengan Akses terhadap Fasilitas Sanitasi Yang Layak (Jamban Sehat)	$X_1 = \frac{g}{d} \times 100\%$ g = jumlah penduduk dengan akses terhadap fasilitas sanitasi yang layak d = jumlah penduduk setiap kab/kota	Persen	Rasio
X <sub>2</sub>	Persentase penduduk dengan akses air minum berkualitas (layak)	$X_2 = \frac{c}{d} \times 100\%$ c = jumlah penduduk dengan akses air minum berkualitas d = jumlah penduduk setiap kab/kota	Persen	Rasio

Variabel	Definisi Operasional Variabel	Rumus	Satuan	Skala
$X_3$	Persentase Desa/Kelurahan yang melaksanakan STBM	$X_3 = \frac{a}{b} \times 100\%$ a = jumlah desa/kel yang melaksanakan STBM b = jumlah desa/kel setiap kab/kota	Persen	Rasio
$X_4$	Persentase Tempat - tempat Umum (TTU) yang memenuhi syarat kesehatan	$X_4 = \frac{h}{i} \times 100\%$ h = jumlah TTU yang memenuhi syarat kesehatan i = jumlah TTU setiap kab/kota	Persen	Rasio
$X_5$	Persentase Tempat Pengelolaan Makanan (TPM) yang memenuhi syarat kesehatan	$X_5 = \frac{j}{k} \times 100\%$ j = jumlah TPM yang memenuhi syarat kesehatan k = jumlah TPM setiap kab/kota	Persen	Rasio

Aturan dalam penelitian ini meliputi langkah-langkah analisis regresi binomial negatif. Berikut adalah langkah-langkah analisis regresi binomial negative sebagai berikut:

1. Menganalisis statistika deskriptif untuk variabel dependen dan variabel independen.
2. Menganalisis korelasi antar variabel-variabel independen untuk mendeteksi adanya gejala multikolinearitas.
3. Mendapatkan model terbaik untuk regresi poisson.
4. Mendeteksi adanya kasus overdispersi pada data dengan melihat nilai *Pearson Chi-squares* dan *Deviance* yang dibagi derajat bebasnya.
5. Mendapatkan model terbaik untuk regresi binomial negatif dengan menaksir parameter, menguji signifikansi parameter secara serentak dan parsial serta menghitung nilai AIC (Akaike Information Criteration).
6. Membandingkan model terbaik hasil regresi poisson dan regresi binomial negatif menggunakan nilai AIC (Qudratullah, 2013).

### 3. HASIL PENELITIAN DAN PEMBAHASAN

Jumlah kasus diare di Jawa Barat mencapai 1.032.284 dengan rata-rata 38.233 kasus. Kabupate Bogor merupakan kabupaten yang memiliki jumlah kasus diare tertinggi yaitu 159.405 kasus pada tahun 2016, sedangkan Kabupaten Subang merupakan kabupaten yang memiliki jumlah kasus diare terendah yaitu 25 kasus. Nilai variansi dari variabel dependen yaitu 1125033191,15. Hal ini menunjukkan varians kasus diare di Jawa Barat sangat besar karena terdapat kabupaten dengan jumlah kasus diare sampai jutaan kasus namun terdapat kabupaten dengan jumlah kasus diare yang hanya 25 kasus.

Nilai varians tertinggi dari beberapa variabel independen yang diduga mempengaruhi jumlah kasus kanker diare di Jawa Barat terdapat pada variabel  $X_2$  (Persentase penduduk dengan akses air minum berkualitas (layak)) yaitu

sebesar 712,07 dengan nilai minimum sebesar 40,98 dan nilai maksimum sebesar 144,48. Artinya bahwa setiap kabupaten/kota di Jawa Barat memiliki akses air minum berkualitas yang berbeda-beda. Semakin besar range data, maka semakin besar nilai varians.

Salah satu syarat dalam regresi yang melibatkan beberapa variabel independen adalah antara variabel independennya saling bebas. Jika terdapat adanya hubungan antara variabel bebas maka terjadi adanya kasus multikolinieritas. Identifikasi kasus multikolinieritas dalam penelitian ini digunakan 2 kriteria yaitu nilai korelasi dan VIF. Berdasarkan hasil analisis diperoleh nilai korelasi lebih kecil dari 0,95 dan nilai VIF lebih kecil dari 10. Dengan demikian pada variabel independen untuk setiap kabupaten/kota tidak terdapat adanya kasus multikolinieritas. Analisis selanjutnya yaitu pemodelan dengan regresi poisson, GPR, dan regresi binomial negatif.

Pemodelan menggunakan analisis regresi poisson ini dilakukan untuk mengetahui faktor-faktor yang mempengaruhi jumlah kasus diare. Diperoleh estimasi parameter model regresi poisson yang disajikan dalam Tabel 2.

Tabel 2. Nilai Dugaan Parameter Model Regresi Poisson

Parameter	Estimasi	p-value
$\beta_0$	1,217E+01	< 2e-16
$\beta_1$	3,45E-03	< 2e-17
$\beta_2$	-1,69E-02	< 2e-18
$\beta_3$	-5,63E-03	< 2e-19
$\beta_4$	-4,61E-03	< 2e-20
$\beta_5$	-3,84E-04	5,01E-12

Deviance: 656872; derajat bebas: 26; Rasio disperse: 399546; AIC: 399879

Berikut merupakan hipotesis untuk uji *overall* pada jumlah kasus diare, kemudian dianalisis menggunakan *output Null deviance* pada analisis regresi Poisson.

$$H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$$

$$H_1: \text{paling sedikit ada satu } \beta_j \neq 0; j = 1, 2, \dots, k$$

Tingkat sigifikansi yang digunakan yaitu 5% dan hasil perhitungan akan tolak  $H_0$  jika nilai  $D(\beta) > \chi^2_{(5,0.05)}$ .

Dari analisis kasus diare didapatkan hasil perhitungan *Null Deviance* sebesar 656.872 dan  $\chi^2_{(5,0.05)} = 11,070$ . Jika dibandingkan maka:  
 $D(\beta) = 656.872 > \chi^2_{(5,0.05)} = 11,070$ .

Keputusan yang dihasilkan yaitu tolak  $H_0$ , sehingga kesimpulan yang didapatkan yaitu paling sedikit ada satu  $\beta_j \neq 0; j = 1, 2, \dots, k$ . Artinya bahwa terdapat variabel prediktor yang berpengaruh terhadap jumlah kasus diare.

Setelah diketahui bahwa terdapat variabel independen yang berpengaruh terhadap jumlah kasus diare, maka selanjutnya melakukan uji parsial di tiap variabel dependen. Hipotesis yang digunakan yaitu

$$H_0: \beta_j = 0$$

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

Tingkat signifikansi yang digunakan yaitu 5% dan hasil perhitungan akan tolak  $H_0$  apabila  $p\text{-value} < \alpha$ . Hasil analisis uji parsial dapat dilihat seperti pada tabel 3.

Tabel 3. Hasil Uji Parsial pada Model Regresi Poisson

Parameter	Estimasi	p-value	$\alpha$	Keputusan
$\beta_0$	1.217E+01	$< 2e-16$	0,05	Tolak $H_0$
$\beta_1$	3.45E-03	$< 2e-17$	0,05	Tolak $H_0$
$\beta_2$	-1.69E-02	$< 2e-18$	0,05	Tolak $H_0$
$\beta_3$	-5.63E-03	$< 2e-19$	0,05	Tolak $H_0$
$\beta_4$	-4.61E-03	$< 2e-20$	0,05	Tolak $H_0$
$\beta_5$	-3.84E-04	$5.01E-12$	0,05	Tolak $H_0$

Dari tabel 3 dapat dilihat bahwa semua variabel independen berpengaruh terhadap jumlah kasus diare.

Model yang terbentuk dari hasil estimasi parameter yaitu:

$$\mu = \exp (0,1217 + 0,003454 X_1 - 0.0169 X_2 - 0,005628 X_3 - 0,004612 X_4 - 0,0003843 X_5)$$

Model tersebut menunjukkan bahwa dengan menganggap variabel lain konstan maka setiap penambahan 1% Penduduk Dengan Akses terhadap Fasilitas Sanitasi Yang Layak (Jamban Sehat) akan mempengaruhi kasus diare sebesar  $\exp(0,003454) = 1,00346 \approx 1$  jiwa dari rata – rata kasus diare semula.

Kemudian dilakukan pengujian overdispersi pada data kasus diare dimana variansi Y yang bernilai 1125033191,15 lebih besar dari rata-rata Y yang bernilai 38.233. Selain itu, keadaan overdispersi pada data kasus diare dapat di lihat berdasarkan nilai Deviance yang tertera pada tabel 4 yang dibagi dengan derajat bebas bernilai lebih dari 1.

Tabel 4. Hasil Uji Overdispersi

Kriteria	Nilai	db	Nilai/db
Deviance	656872	26	25264.31

Berdasarkan Tabel 4 dapat diketahui bahwa nilai  $deviance/db$  lebih besar dari 1 sehingga dapat disimpulkan pada model regresi poisson jumlah kasus diare di Jawa Barat terjadi overdispersi.

Pemodelan menggunakan analisis regresi binomial negatif ini dilakukan untuk mengetahui faktor-faktor yang mempengaruhi jumlah kasus diare. Diperoleh estimasi parameter model regresi poisson yang disajikan dalam Tabel 5 dan nilai AIC yang dihasilkan sebesar 399879.

Tabel 5. Nilai Dugaan Parameter Model Regresi Binomial Negatif

Parameter	Estimasi	p-value
$\beta_0$	12.410793	$< 2e-16$

Parameter	Estimasi	p-value
$\beta_1$	-0.002832	0.826589
$\beta_2$	-0.022230	0.000656
$\beta_3$	-0.003359	0.649210
$\beta_4$	-0.001575	0.888021
$\beta_5$	0.003532	0.694543

Deviance: 41839; derajat bebas: 26; Rasio disperse: 29,900; AIC: 627,4

Berikut merupakan hipotesis untuk uji *overall* pada jumlah kasus diare, kemudian dianalisis menggunakan *output Null deviance* pada analisis regresi Poisson.

$$H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$$

$$H_1: \text{paling sedikit ada satu } \beta_j \neq 0; j = 1, 2, \dots, k$$

Tingkat sigifikansi yang digunakan yaitu 5% dan hasil perhitungan akan tolak  $H_0$  jika nilai  $D(\beta) > \chi^2_{(5,0.05)}$ .

Dari analisis kasus diare didapatkan hasil perhitungan *Null Deviance* sebesar 41,839 dan  $\chi^2_{(5,0.05)} = 11,070$ . Jika dibandingkan maka:

$$D(\beta) = 41,839 > \chi^2_{(5,0.05)} = 11,070.$$

Keputusan yang dihasilkan yaitu tolak  $H_0$ , sehingga kesimpulan yang didapatkan yaitu paling sedikit ada satu  $\beta_j \neq 0; j = 1, 2, \dots, k$ . Artinya bahwa terdapat variabel prediktor yang berpengaruh terhadap jumlah kasus diare.

Setelah diketahui bahwa terdapat variabel independen yang berpengaruh terhadap jumlah kasus diare, maka selanjutnya melakukan uji parsial di tiap variabel dependen. Hipotesis yang digunakan yaitu

$$H_0: \beta_j = 0$$

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

Tingkat signifikansi yang digunakan yaitu 5% dan hasil perhitungan akan tolak  $H_0$  apabila  $p\text{-value} < \alpha$ . Hasil analisis uji parsial dapat dilihat seperti pada tabel 6.

Tabel 6. Hasil Uji Parsial pada Model Regresi Binomial Negatif

Parameter	Estimasi	p-value	$\alpha$	Keputusan
$\beta_0$	12,410793	$< 2e-16$	0,05	Gagal Tolak $H_0$
$\beta_1$	-0,002832	0,826589	0,05	Gagal Tolak $H_0$
$\beta_2$	-0,022230	0,000656	0,05	Tolak $H_0$
$\beta_3$	-0,003359	0,649210	0,05	Gagal Tolak $H_0$
$\beta_4$	-0,001575	0,888021	0,05	Gagal Tolak $H_0$
$\beta_5$	0,003532	0,694543	0,05	Gagal Tolak $H_0$

Model yang terbentuk dari hasil estimasi parameter yaitu:

$$\mu = \exp(12,4107 - 0,022230X_2)$$

Model tersebut menunjukkan bahwa dengan menganggap variabel lain konstan. Setiap penambahan 1% Persentase penduduk dengan akses air



minum berkualitas (layak) akan mempengaruhi kasus diare sebesar  $\exp(-0,022230) = 0,978 \approx 1$  jiwa dari rata – rata kasus diare semula.

Model yang memiliki nilai AIC terkecil adalah model regresi binomial negatif dengan nilai sebesar 627,4 yang dapat dilihat perbandingannya nilai AIC pada tabel 7.

Tabel 7. Pemilihan Model Terbaik

Model	Variabel Signifikan	AIC
Regresi Poisson	$X_1 X_2 X_3 X_4 X_5$	399879
Regresi Binomial Negatif	$X_2$	627,4

#### 4. SIMPULAN

Model regresi binomial negatif merupakan model terbaik untuk menggambarkan penyebab diare di Provinsi Jawa Barat pada tahun 2016 karena memiliki nilai AIC yang paling kecil dibandingkan regresi poisson dan faktor lingkungan yang mempengaruhi kasus diare adalah persentase penduduk dengan akses air minum berkualitas (layak).

#### 5. DAFTAR PUSTAKA

- Agresti, (2002). *Categorical Data Analysis Second Edition*. New York: John Wiley & Sons.
- Anies. (2015). *Penyakit Berbasis Lingkungan*. Yogyakarta: Ar-Ruzz Media.
- Ariani, P. M. (2018). Analisis Faktor Yang Berpengaruh Terhadap Pencegahan Penyakit DBD di Provinsi Jawa Tengah Menggunakan Regresi Binomial Negatif. *Jurnal Kesehatan Vokasional* Vol 3 No 1, ISSN 2599-3275 (Online). Diakses dari <https://jurnal.ugm.ac.id/jkesvo/article/view/33870/20994>
- Armawati, Welly, Nur. (2018). Metode Regresi Poisson Bivariat dalam Pemodelan Faktor Jumlah Kasus HIV dan AIDS di Jawa Tengah Tahun 2016. *Prosiding SENDIKA*, Diselenggarakan oleh Universitas Muhammadiyah Purworejo, 12 Mei 2018 (Vol 4 No 1, ISSN: 2459-962X). Diakses dari <http://e proceedings.umpwr.ac.id/index.php/sendika/article/view/297/273>
- Arumsari, N & Sutikno. (2010). *Permodelan Kejadian Diare Dengan Pendekatan Regresi Spasial Studi Kasus: Kabupaten Tuban Jawa Timur*. Diakses dari <https://anzdoc.com/permodelan-kejadian-diare-dengan-pendekatan-regresi-spasial-.html>
- Cameron, A.C, & Trivedi, P.K.,. (1998). *Regression Analysis of Count Data*. Cambridge University Press.
- Dinas Kesehatan Provinsi. (2017). *Profil Kesehatan Provinsi Jawa Barat 2016*. Bandung: Dinas Kesehatan Provinsi Jawa Barat.
- Kementerian Kesehatan RI. (2017). *Profil Kesehatan Republik Indonesia 2016*. Jakarta: Kementerian Kesehatan Republik Indonesia.

- McCullagh P & Nelder JA. (1983). *Generalized Linier Models*. London: Chapman and Hall.
- Myers, R. H. (1990). *Classical and Modern Regression with Applications, second edition*. Boston : PWS KENT Publishing Company.
- Nugraha, J. (2016). *Pengantar Analisis Data Kategorik Edisi 1 Cetakan ke 3*. Yogyakarta: Deepublish.
- Paiman, F. (2013). *Pengujian Autokorelasi Pada Model Regresi Spasial Lag Dengan Lagrange Multiplier*. Diakses dari <https://docplayer.info/52013502-Pengujian-autokorelasi-pada-model-regresi-spasial-lag-dengan-lagrange-multiplier-studi-kasus-penyakit-diare-di-jawa-timur-tahun-2010-skripsi.html>
- Quadratullah, M. F. (2013). *Analisis Regresi Terapan: Teori, Contoh Kasus dan Aplikasi dengan SPSS*. Yogyakarta: Andi Offset.
- Rahayu, Dian. Wiwiek, Setya. Winahju, Adatul Mukarromah. (2012). *Pemodelan Pengaruh Iklim Terhadap Angka Kejadian Demam Berdarah Dengue di Surabaya*. Jurnal Sains dan Seni ITS Vol. 1, ISSN: 2301-928X. doi: 10.12962/j23373520.v1i1.594.
- Soemirat, J. (2011). *Kesehatan Lingkungan Cetakan Kedelapan*. Yogyakarta: Gajah Mada University Press.
- Walpole, R. E., & Raymond H Myers. (1995). *Ilmu Peluang dan Statistika untuk Insinyur dan Ilmuwan*. Edisi ke-4. Bandung: Penerbit ITB.
- Wijoyo, Y. (2013). *Diare: Pahami Penyakit dan Obatnya*. Yogyakarta: PT Citra Aji Parama.