



Negative Binomial and Generalized Poisson Regression Model for Death Due to Dengue Hemorrhagic Fever Data

Risnawati Ibtnas^a, Satriani^b, Khalilah Nurfadilah^{c*}

^{a,b,c}Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Alauddin Makassar. Email: khalilah@uin-alauddin.ac.id

ABSTRACT

Data on the number of deaths due to Dengue Fever in statistics is count data often approximated by a Poisson distribution. However, if overdispersion occurs, Poisson regression is no longer sufficient, so the Negative Binomial and Generalized Poisson Regression approaches are used. From the two models, the best model was chosen based on the smallest AIC value, 66.50, namely the Negative Binomial Regression model. From this model, factors that have a significant effect are determined based on the p-value, and the factor ratio of health facilities per 100,000 population (X_4) is obtained.

Keywords: Number of deaths due to DHF, overdispersion, Negative Binomial regression, Generalized Poisson regression

ABSTRAK

Data jumlah kematian akibat Demam Berdarah (DBD) dalam statistik merupakan data *count* yang seringkali didekati dengan sebaran Poisson. Namun, jika terjadi overdispersi, maka regresi Poisson tidak lagi memadai, sehingga digunakan pendekatan Regresi Binomial Negatif dan *Generalized Poisson*. Dari kedua model, selanjutnya dipilih model terbaik berdasarkan nilai AIC terkecil, sebesar 66.50, yaitu model Regresi Binomial Negatif. Dari model tersebut kemudian ditentukan faktor yang berpengaruh signifikan berdasarkan nilai p-value, diperoleh faktor rasio sarana kesehatan per 100.000 penduduk (X_4).

Keywords: jumlah kematian akibat DBD, overdispersi, regresi Binomial Negatif, regresi Generalized Poisson.

Diserahkan: 08-12-2023; Diterima: 26-06-2023;

Doi: <https://doi.org/10.29303/emj.v6i1.153>

1. PENDAHULUAN

Penyakit Demam Berdarah *Dengue* (DBD) terjadi karena gigitan dari nyamuk *Aedes Aegypti*. Penyakit tersebut sudah menyebar di Provinsi Sulawesi Selatan. Hal tersebut dijelaskan oleh

dinas kesehatan Provinsi Sulawesi Selatan yang menunjukkan angka jumlah kematian akibat DBD setiap tahunnya. Pada tahun 2013 sebanyak 48 jumlah kematian akibat DBD, kemudian tahun 2015 menurun menjadi 30 jumlah kematian,

* Corresponding author.

Alamat e-mail: khalilah@uin-alauddin.ac.id

tahun 2016 bertambah menjadi 48 orang, serta kembali menurun di tahun 2017 sebanyak 17 orang, tahun 2018 meningkat menjadi 19 orang, dan tahun 2019 meningkat menjadi 25 orang serta terdapat 19 orang yang meninggal dunia pada tahun 2020.

Jumlah kematian akibat DBD merupakan data cacahan dengan peluang kejadiannya kecil pada interval waktu tertentu. Pemodelan yang tepat digunakan untuk mengkaji hubungan antara jumlah kematian DBD dengan faktor yang mempengaruhinya adalah model regresi Poisson. Analisis data dengan regresi Poisson wajibenuhi asumsi yaitu nilai varians serta rata-rata dari peubah respons harus sama atau *equidispersi*, akan tetapi tidak sepenuhnya asumsi tersebut dapat terpenuhi. Beberapa penelitian yang sudah dilakukan ditemukan pelanggaran asumsi regresi Poisson yaitu nilai varians lebih besar dari nilai rata-ratanya yang disebut sebagai *overdispersi*.

Overdispersi pada data bisa terjadi sebab proporsi nilai nol yang berlebih pada peubah terikat (*excess zeros*). Akibatnya data yang mengandung *overdispersi* menciptakan galat baku yang lebih kecil dari nilai sesungguhnya (*underestimate*). Perihal tersebut menimbulkan kesimpulan yang diperoleh jadi tidak valid. Sehingga digunakan pendekatan alternatif untuk menangani permasalahan *overdispersi*, antara lain metode regresi *Generalized Poisson* (GPR) dan metode regresi Binomial Negatif. Kedua metode tersebut memungkinkan untuk mengatasi data *overdispersi* karena nilai varians dan rata-ratanya tidak harus sama sehingga dapat menyesuaikan kehadiran *overdispersi*.

Listiani (2010) menggunakan GPR untuk memodelkan faktor-faktor yang mempengaruhi angka kematian bayi di Jawa timur. Selanjutnya, oleh Brianika dan Purhadi (2012) membandingkan antara model GPR dan Binomial Negatif untuk memodelkan jumlah kasus penderita kanker serviks.

Berdasarkan uraian di atas, penelitian ini bertujuan untuk memilih model terbaik antara model regresi *Generalized Poisson* (GPR) dan Binomial Negatif pada kasus kematian akibat DBD di Provinsi Sulawesi Selatan tahun 2020 berdasarkan nilai AIC terkecil.

2. TINJAUAN PUSTAKA

2.1 Generalized Linear Model (GLM)

GLM merupakan perluasan dari model regresi ketika peubah respon memiliki model distribusi eror selain distribusi normal. Seluruh model dalam GLM memiliki tiga komponen yaitu:

1) Komponen acak

Komponen acak diidentifikasi oleh peubah respon (y_i) dengan $Y_i(Y_1, Y_2, \dots, Y_n)$ saling bebas dan diasumsikan memiliki distribusi peluang yang bergantung pada parameter disperse (θ). Fungsi kepadatan peluangnya merupakan keluarga eksponensial. Bentuk keluarga eksponensial adalah:

$$f(y) = \exp[a(y)b(\theta) + c(\theta) + d(y)] \quad (1)$$

Untuk beberapa fungsi yang diketahui $a(y)$, $b(\theta)$, $c(\theta)$, $d(y)$. θ adalah parameter disperse.

2) Komponen sistematis

Meliputi peubah-peubah prediktor dari model. Dalam notasi matriks ditulis $\eta = X\beta$ dengan η merupakan vector dari n predictor linear $\eta = (\eta_1, \eta_2, \dots, \eta_n)^t$ dan β adalah vector dari p parameter $\beta = (\beta_0, \beta_1, \beta_2, \dots, \beta_p)^t$ dan X adalah matriks berukuran $(n \times p)$

3) Fungsi Link

Fungsi *link* menghubungkan ekspektasi dari peubah respon dengan peubah-peubah predictor melalui persamaan linear.

$$\mu_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_j x_{ij} + \dots + \beta_p x_{ip} \quad (2)$$

2.2 Distribusi Poisson

Distribusi Poisson merupakan suatu percobaan menghasilkan peubah random Y yang menyatakan banyaknya sukses dalam daerah tertentu atau selama interval waktu tertentu. Jumlah keluaran Y sepanjang satu percobaan Poisson disebut peubah random Poisson, dan distribusi peluangnya disebut distribusi Poisson. Distribusi Poisson memiliki ciri-ciri yaitu:

1. Banyaknya sukses yang terjadi dalam suatu selang waktu ataupun daerah tertentu tidak tergantung pada banyaknya hasil uji

percobaan yang terjadi pada selang waktu ataupun daerah lain yang terpisah.

2. Probabilitas terjadinya suatu sukses dalam interval waktu yang singkat atau daerah yang kecil tidak tergantung banyaknya sukses yang terjadi di luar interval waktu atau daerah lain yang terpisah.

Probabilitas terjadinya lebih dari satu sukses dalam interval waktu yang singkat atau daerah yang kecil dapat diabaikan.

Suatu peubah acak diskret Y dikatakan memiliki sebaran Poisson apabila memiliki fungsi kepekatan peluang seperti berikut

$$f(y; \mu) = \frac{e^{-\mu} \mu^y}{y!} \quad y = 0, 1, 2, \dots \quad (3)$$

dengan parameter $\mu > 0$.

Dimana: μ = mean keberhasilan (banyak sukses yang terjadi) dalam interval waktu atau daerah tertentu.

y = banyaknya unsur berhasil dalam sampel atau peubah random diskret
 $e = 2,71828$.

2.3 Model Regresi Poisson

Model regresi Poisson yaitu model regresi nonlinear yang berasal dari sebaran Poisson. Distribusi Poisson sering digunakan untuk memodelkan data cacah. Peubah y adalah peubah acak yang menyebar Poisson dengan fungsi massa peluang pada persamaan (3).

Model regresi poisson dapat ditulis sebagai berikut:

$$\ln(\mu_i) = \beta_0 + \beta_1 X_{i1} + \dots + \beta_k X_{ik}$$

$$\mu_i = \exp(\beta_0 + \beta_1 X_{i1} + \dots + \beta_k X_{ik}) \quad (4)$$

Atau dapat ditulis dalam bentuk matriks

$$\ln(\mu_i) = X_i^T \beta$$

$$\mu_i = \exp(X_i^T \beta)$$

$$y_i = \mu_i + \varepsilon_i = \exp(X_i^T \beta) + \varepsilon_i$$

$$i = 1, 2, \dots, n \quad (5)$$

dimana μ_i nilai ekspektasi y_i berdistribusi Poisson. Dalam model regresi Poisson koefisien regresi β_j menyatakan perubahan yang diharapkan terhadap logaritma natural *mean* per unit perubahan pada prediktor X_j . Dengan

demikian regresi Poisson memenuhi 3 komponen GLM sehingga terbukti bahwa regresi Poisson merupakan salah satu penerapan GLM.

2.4 Estimasi Parameter Regresi Poisson

Metode untuk menduga koefisien parameter regresi Poisson adalah metode *Maximum Likelihood Estimation* (MLE). Fungsi *likelihood* dari regresi Poisson adalah:

$$L(y_i, \beta) = \prod_{i=1}^n f(y_i, \beta) = \frac{e^{-\sum_{i=1}^n [\exp(X_i^T \beta)]} \prod_{i=1}^n [\exp(X_i^T \beta)]^{y_i}}{\prod_{i=1}^n y_i!} \quad (6)$$

Fungsi *log-likelihood* digunakan untuk memudahkan perhitungan dalam memperkirakan koefisien parameter regresi *Poisson*. Logaritma natural dari fungsi *likelihood* adalah:

$$\ln L(y_i, \beta) = \sum_{i=1}^n y_i \ln[\exp(X_i^T \beta)] - \sum_{i=1}^n \exp(X_i^T \beta) - \sum_{i=1}^n \ln(y_i!) \quad (7)$$

2.5 Multikolinearitas

Multikolinearitas yaitu suatu kondisi dimana terjadi korelasi antara peubah bebas atau antar peubah bebas tidak bersifat saling bebas. Besaran (*quality*) yang dapat digunakan untuk mendeteksi adanya multikolinearitas adalah faktor inflasi ragam VIF (*Variance Inflation Factor*). Nilai VIF lebih besar dari 10 mengidentifikasi adanya masalah multikolinearitas (Ryan, 1997).

Langkahnya sebagai berikut:

Hipotesis:

H_0 : Model regresi memiliki masalah multikolinearitas

H_1 : Model regresi tidak memiliki masalah multikolinearitas

Taraf Signifikansi, $\alpha = 0,05$

Statistik Uji

$$VIF = \frac{1}{1-r_{ij}^2} \quad (8)$$

dimana VIF = *Variance Inflation Factor*

r_{ij}^2 = koefisien korelasi antara X_i dan X_j

Kriteria Uji:

Tolak H_0 jika seluruh peubah prediktor memiliki nilai VIF kurang dari 10 dan nilai *Tolerance* lebih dari 0,1. Sebaliknya jika seluruh peubah prediktor

memiliki nilai VIF lebih besar 10 dan nilai *Tolerance* kurang dari 0,1 maka H_0 gagal tolak.

2.6 Overdispersi

Overdispersi bisa diidentifikasi dengan teknik mengukur rasio antara deviance dan derajat bebas. Jika nilai rasio antara deviance dengan derajat bebas lebih besar dari satu maka model dapat dikatakan mengalami overdispersi (Numra 2009).

Parameter disperse (ϕ) diperoleh dari rumus:

$$\phi = \frac{\text{nilai deviance}}{df} \quad (9)$$

dimana; $df = \text{degree of freedom}$.

Nilai deviance didefinisikan

$$G^2 = 2 \sum_{i=1}^n y_i \ln \left(\frac{y_i}{\mu_i} \right) \quad (10)$$

Dimana G = nilai deviance

y_i = nilai sesungguhnya amatan ke- i dari peubah respon

μ_i = nilai dugaan peubah respon untuk amatan ke- i

Jika nilai $\phi > 1$ maka terjadi overdispersi dan jika $\phi < 1$ maka akan terjadi underdispersi.

2.7 Model Regresi Generalized Poisson (GPR)

Fungsi distribusi *Generalized Poisson* adalah

$f(y; \mu; \theta) =$

$$\left(\frac{\mu}{1+\theta\mu} \right)^y \left(\frac{(1+\theta y)}{y!} \right)^{y-1} \exp \left(\frac{-\mu(1+\theta y)}{1+\theta\mu} \right) \quad (11)$$

dimana y = nilai dari data cacah (0,1,2, ...)

θ = parameter disperse

Model GPR adalah

$$\ln(\mu_i) = X_i^T \beta = \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip}$$

$$\mu_i = \exp(X_i^T \beta), i = 1, 2, \dots, n \quad (12)$$

dengan, x_i =vektor berdimensi $k - 1$

β =vektor berdimensi k dari parameter regresi

Estimasi Parameter Model Regresi Generalized Poisson (GPR)

Metode pendugaan *maximum likelihood* digunakan untuk menduga parameter koefisien

regresi μ dan dispersi θ pada GPR. Fungsi \ln kemungkinan dari sebaran GPR adalah sebagai berikut:

$$\ln L(y|\beta, \theta) = \sum_{i=1}^n y_i \ln \left(\frac{\mu}{1+\theta\mu} \right) + (y_i - 1) \ln(1 + \theta y_i) - \left(\frac{-\mu(1+\theta y)}{1+\theta\mu} \right) - \ln(y!) \quad (13)$$

Fungsi $\ln L(y|\beta, \theta)$ dapat dimaksimumkan dengan mencari turunan terhadap masing-masing parameter β dan θ setelah itu disamakan dengan nol.

2.8 Model Regresi Binomial Negatif

Misalkan r adalah banyaknya kejadian berhasil dan peubah acak Y adalah banyaknya percobaan sampai terjadinya kejadian keberhasilan ke- r , bentuk sebaran Binomial Negatif dengan peluang berhasil sebesar P dan peluang gagal sebesar $(1 - P)$ adalah sebagai berikut:

$$P(Y = y|r, p) = \binom{y-1}{r-1} p^r (1-p)^{y-r};$$

$$Y = r, r+1, r+2, \dots \quad (14)$$

dimana; P = peluang sukses pada percobaan tertentu

Y = banyaknya percobaan sampai terjadinya kejadian keberhasilan ke- r

y = percobaan ke-

p = peluang sukses

r = sukses ke-

$1 - p$ = peluang gagal

Model Binomial Negatif adalah:

$$\ln(\mu_i) = X_i^T \beta = \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip}$$

$$\mu_i = \exp(X_i^T \beta), i = 1, 2, \dots, n \quad (15)$$

Estimasi Parameter Regresi Binomial Negatif

Metode *Maksimum Likelihood Estimation* (MLE) merupakan metode yang juga umum digunakan dalam pendekatan regresi Binomial Negatif untuk mempermudah pengestimasian parameternya. Fungsi kemungkinan dari sebaran Binomial Negatif adalah:

$L(\beta, \phi|y, x) =$

$$\prod_{i=1}^n \left\{ \frac{\Gamma(y_i + \phi^{-1})}{\Gamma(\phi^{-1}) y_i!} \left(\frac{\phi \mu_i}{1 + \phi \mu_i} \right)^{y_i} \left(\frac{1}{1 + \phi \mu_i} \right)^{\phi^{-1}} \right\} \quad (16)$$

Fungsi log *likelihood* Binomial Negatif adalah:

$$\ln L(\beta, \phi | y, x) = \sum_{i=1}^n \left\{ \sum_{r=1}^{y_i-1} \ln(1 + \phi r) \right\} - y_i \ln(\phi) - \ln(y_i!) + y_i \ln(\phi \mu_i) - (y_i + \phi^{-1}) \ln(1 + \phi \mu_i) \quad (17)$$

2.9 Statistik Uji

Uji Serentak

Uji serentak dalam regresi Poisson bertujuan untuk melihat secara serentak pengaruh peubah independen terhadap peubah dependen. Selain itu dapat pula digunakan sebagai uji kelayakan model.

a) Hipotesis

$H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$ (tidak terdapat pengaruh peubah independen terhadap peubah dependen)

H_1 : paling tidak ada satu $\beta_j \neq 0$; $j = 1, 2, 3, \dots, k$ (paling sedikit terdapat satu peubah independen yang berpengaruh terhadap peubah dependen)

b) Taraf Signifikan, $\alpha = 0,05$

c) Statistik Uji

Dalam menentukan nilai statistik uji digunakan

$$\begin{aligned} D(\hat{\beta}) &= -2 \ln \frac{L(\hat{\omega})}{L(\hat{\Omega})} \\ &= 2(\ln(\hat{\Omega}) - \ln(\hat{\omega})) \\ &= 2 \sum_{i=1}^n \left(y_i X_i^T \hat{\beta} - \exp(X_i^T \hat{\beta}) - (y_i \hat{\beta}_0 - \exp(\hat{\beta}_0)) \right) \end{aligned} \quad (18)$$

dengan D = Statistik uji rasio *Likelihood*

$\ln(\hat{\Omega})$ = *likelihood* tanpa peubah *independent*

$\ln(\hat{\omega})$ = *likelihood* dengan peubah *independent*.

$\hat{\beta}_0$ = dugaan awal parameter

Kriteria uji yaitu tolak H_0 pada taraf signifikan $\alpha < 0.05$.

Uji Parsial

Setelah dilakukan pengujian signifikansi model, selanjutnya dilakukan pengujian signifikansi

masing-masing parameter dari model. Pengujian yang dilakukan adalah uji *Wald*.

a) Hipotesis

$H_0: \beta_j = 0$ (peubah independen tidak berpengaruh terhadap peubah dependen)

$H_1: \beta_j \neq 0$ (peubah independen berpengaruh terhadap peubah dependen)

b) Taraf Signifikan

$$\alpha = 0,05$$

c) Statistik Uji

$$W = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \quad (19)$$

dengan W = statistik uji wald

$\hat{\beta}_j$ = nilai dugaan untuk parameter β_j

$SE(\hat{\beta}_j)$ = taksiran standar error $\hat{\beta}_j$

Nilai uji W mengikuti distribusi *chi-square* sehingga dibandingkan dengan *chi-square* tabel $\chi^2_{(\alpha, ab=1)}$, maka kriteria uji untuk pengambilan keputusan dengan taraf nyata adalah tolak H_0 jika nilai $W > \chi^2_{(\alpha, ab=1)}$ dengan α adalah tingkat signifikansi dan derajat bebas satu.

2.10 Pemilihan Model Terbaik

Dalam analisis regresi, terdapat beberapa metode yang dapat digunakan untuk memilih model terbaik antara lain *Akaike Information Criterion* (AIC), *Bayes Information Criterion* (BIC), *cross validation*, dan *bootstrap*. Namun, kriteria yang paling banyak digunakan adalah AIC dimana model terbaik yaitu model yang memiliki nilai AIC terkecil. AIC dapat didefinisikan sebagai berikut:

$$AIC = -2 \ln L(y|\hat{\mu}) + 2p \quad (20)$$

dimana AIC = *Akaike Information Criterion*

$\ln L(y|\hat{\mu})$ = merupakan \ln (logaritma natural) dari model yang melibatkan semua peubah penjelas.

p = jumlah parameter dalam model.

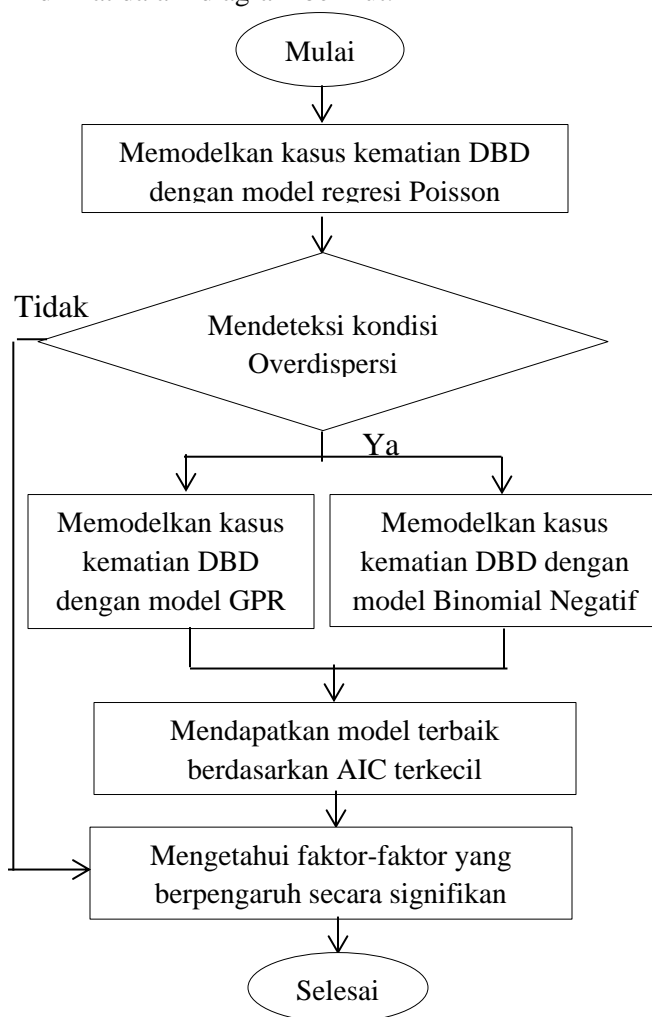
3. METODOLOGI

Data yang digunakan adalah data sekunder yang diperoleh dari Dinas Kesehatan dan Badan Pusat

Statistik (BPS) Provinsi Sulawesi Selatan tahun 2020. Peubah yang digunakan pada penelitian ini adalah jumlah kasus kematian akibat DBD di Provinsi Sulawesi Selatan tahun 2020 (Y), tingkat kepadatan penduduk (X_1), persentase penduduk miskin (X_2), persentase rumah tangga berperilaku hidup bersih dan sehat (PHBS) (X_3), dan rasio sarana kesehatan (X_4).

Prosedur Analisis

Adapun analisis data dalam penelitian ini dapat dilihat dalam diagram berikut:



Gambar 1 – Diagram Alir Penelitian

4. HASIL DAN PEMBAHASAN

Statistik Deskriptif

Tabel 1. Analisis Statistika Deskriptif Data Penelitian untuk Peubah Y dan Peubah X

Peubah	Min	Max	Mean	Var
Y	0,00	7,00	1,13	4,38
X_1	42,73	8100,8	640,88	2619362
X_2	4,54	14,58	9,42	8,52
X_3	30,3	77,1	54,5	203,41
X_4	3,81	10,96	6,73	3,91

Pada Tabel 1 statistik deskriptif terdapat nilai mean dan variance yang dapat digunakan untuk mengetahui terjadinya pelanggaran asumsi pada regresi Poisson.

Uji Multikolinearitas

Tabel 2. Nilai VIF Peubah Prediktor

Peubah	VIF
X_1	1,230931
X_2	1,309108
X_3	1,139538
X_4	1,039090

Berdasarkan Tabel 2 terdapat nilai VIF dari setiap peubah prediktor. Pada pengujian multikolinearitas dikatakan nilai tersebut mempunyai hubungan korelasi yang kuat jika nilai VIF dari seluruh peubah prediktor lebih besar dari 10, dalam pengujian pada Tabel 2 tidak terjadi multikolinearitas karena tidak didapatkan data yang mempunyai nilai VIF lebih dari 10 sehingga data yang ada pada tabel tersebut tidak mengalami multikolinearitas dan dapat dilanjutkan dengan analisis lebih lanjut.

Model Regresi Poisson

Model estimasi parameter regresi Poisson didapatkan dengan menggunakan metode *Maximum Likelihood Estimation* (MLE).

Tabel 3. Hasil Estimasi Parameter Model Regresi Poisson

Parameter	Estimasi	Standard Error	z.value	p.value
β_0	6,684059	1,658669	4,030	0.0000558***
β_1	-0,002626	0,001801	-1,458	0.144812
β_2	-0,087900	0,088202	-0,997	0,318968
β_3	-0,018688	0,015775	-1,185	0,236174
β_4	-0,699132	0,204139	-3,425	0.000615***

Berdasarkan tabel 3 Estimasi Regresi Poisson yang telah dilakukan, dapat dibentuk model sebagai berikut:

$$(\hat{\mu}_i) = \exp(6,684059 - 0,002626X_1 - 0,087900X_2 - 0,018688X_3 - 0,699132X_4)$$

Pada model yang terbentuk semua variabel respon berpengaruh negatif yang menjelaskan bahwa setiap penambahan satu persen atau satu satuan pada masing-masing variabel akan menyebabkan nilai harapan jumlah kematian DBD mengalami penurunan. X_1 mengalami penurunan sebesar $\exp(0,002626) = 1,00263$ kali, X_2 yaitu $\exp(0,087900) = 1,09188$ kali, X_3 yaitu $\exp(0,018688) = 1,01886$ kali, dan X_4 sebesar $\exp(0,699132) = 2,01201$ kali dari jumlah kasus semula. Berdasarkan model keseluruhan didapatkan nilai AIC sebesar 72,74. Untuk dapat menggunakan model Poisson, dibutuhkan asumsi *equidispersi* harus terpenuhi. Pada Tabel 1 terlihat bahwa asumsi tersebut terlanggar, sehingga model Poisson tidak digunakan.

Uji Overdispersi

Tabel 4. Hasil Uji Overdispersi

Kriteria	Derajat bebas	Nilai	Nilai/Derajat Bebas
Deviance	23	73,929	3,21
Khi Kuadrat	19	41,135	2,17
Pearson			

Berdasarkan Tabel 4 rasio deviance dan khi kuadrat pearson sebesar 3,21 dan 2,17 yang lebih dari 1 menandakan pemodelan kasus kematian akibat DBD Provinsi Sulawesi Selatan tahun 2020 mengalami overdispersi.

Model Regresi Binomial Negatif (BN)

Salah satu model yang dapat digunakan dalam mengatasi overdispersi adalah model Binomial Negatif. Adapun estimasi parameter model Binomial Negatif yaitu dengan menggunakan metode *Maximum Likelihood Estimation* (MLE).

Tabel 5. Hasil Estimasi Parameter Regresi Binomial Negatif

Parameter	Estimasi	Standard Error
β_0	6,736515	3,408672
β_1	-0,004189	0,003471
β_2	-0,063911	0,163954
β_3	-0,020798	0,032616
β_4	-0,657145	0,337912

Berdasarkan Tabel 5 dapat dibentuk model Regresi Binomial Negatif sebagai berikut:

$$(\hat{\mu}_i) = \exp(6,736515 - 0,004189 - 0,063911X_2 - 0,020798X_3 - 0,657145X_4)$$

Pada model yang terbentuk semua variabel berpengaruh negatif yang menjelaskan bahwa setiap penambahan satu persen atau satu satuan pada masing-masing variabel akan menyebabkan nilai harapan jumlah kematian akibat DBD mengalami penurunan. X_1 mengalami penurunan sebesar $\exp(0,004189) = 1,0042$ kali, X_2 yaitu $\exp(0,063911) = 1,066$ kali, X_3 yaitu $\exp(0,020798) = 1,02102$ kali, dan X_4 sebesar $\exp(0,657145) = 1,92928$ kali dari jumlah kasus semula. Berdasarkan model keseluruhan didapatkan nilai AIC sebesar 66,50

Model Regresi Binomial Negatif Terbaik

StepAIC digunakan untuk mengetahui model terbaik pada Binomial Negatif.

Tabel 6. Model Regresi Binomial Negatif Terbaik

Parameter	Estimasi	Standar Error	z-value	p-value
β_0	6,80384	1,97644	3,442	0,00258*
β_3	-0,04151	0,02545	-1,631	0,11847
β_4	-0,50791	0,18348	-2,768	0,01186*

Didapatkan model terbaik yang terbentuk berdasarkan *stepAIC* dengan nilai AIC sebesar 62,82 pada peubah X_1 dan X_4 .

Model Regresi *Generalized Poisson* (GPR)

Model lain yang dapat digunakan untuk mengatasi overdispersi yaitu model GPR. Adapun estimasi parameter model GPR yaitu dengan menggunakan metode *Maximum Likelihood Estimation* (MLE).

Tabel 7. Hasil Estimasi Parameter Regresi *Generalized Poisson*

Parameter	Estimasi	Standard Error
β_0	8,0065517	2,4699993
β_1	-0,0002402	0,0002415
β_2	-0,1120160	0,1380978
β_3	-0,0420522	0,0263674
β_4	-0,5026212	0,1815310

Berdasarkan Tabel 7 dapat dibentuk model Regresi *Generalized Poisson* sebagai berikut:

$$(\hat{\mu}_i) = \exp(8,0065517 - 0,0002402X_1 - 0,1120160X_2 - 0,0420522X_3 - 0,5026212X_4)$$

Pada model yang terbentuk semua variabel berpengaruh negatif yang menjelaskan bahwa setiap penambahan satu persen atau satu satuan pada masing-masing variabel akan menyebabkan nilai harapan jumlah kematian akibat DBD mengalami penurunan. X_1 mengalami penurunan sebesar $\exp(0,0002402) = 1,00024$ kali,

X_2 yaitu $\exp(0,1120160) = 1,11853$ kali, X_3 yaitu $\exp(0,0420522) = 1,04295$ kali, dan X_4 sebesar $\exp(0,5026212) = 1,65305$ kali dari jumlah kasus semula. Berdasarkan model keseluruhan didapatkan nilai AIC sebesar 105,28.

Model Regresi *Generalized Poisson* Terbaik

StepAIC digunakan untuk mengetahui model terbaik pada GPR.

Tabel 8. Model Regresi *Generalized Poisson* Terbaik.

Parameter	Estimasi	Standard Error	z-value	p-value
β_0	5,435952	2,467303	2,203	0,0276*
β_1	-0,003995	0,003288	-1,215	0,2244
β_4	-0,732456	0,347446	-2,108	0,0350*

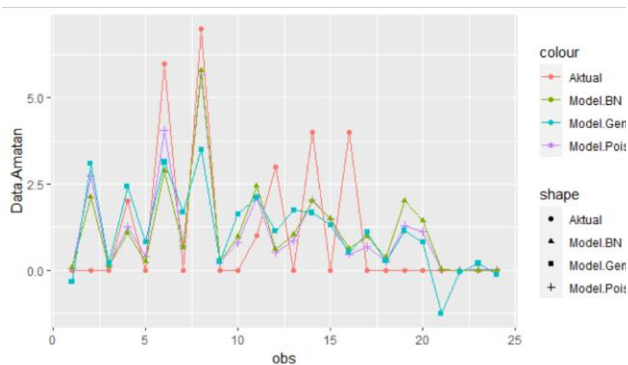
Didapatkan model terbaik yang terbentuk berdasarkan *stepAIC* dengan nilai AIC sebesar 102,48 pada peubah X_3 dan X_4 .

Pemilihan Model Terbaik

Tabel 9. Pemilihan Model Terbaik

Model	AIC model keseluruhan	AIC berdasarkan <i>stepAIC</i>
Regresi Binomial Negatif	66,50	62,82
Regresi <i>Generalized Poisson</i>	105,28	102,48

Tabel 9 Menunjukkan bahwa regresi Binomial Negatif memiliki nilai AIC terkecil sebesar 66,50 jika dibandingkan dengan model yang GPR. Sehingga model regresi Binomial Negatif dipilih untuk memodelkan jumlah kematian akibat DBD di Provinsi Sulawesi Selatan.



Gambar 1. Plot Nilai Aktual dan Dugaan Jumlah Kematian Akibat DBD Keseluruhan Model.

Gambar 1 menunjukkan nilai dugaan terhadap kematian DBD menggunakan model Binomial Negatif lebih mendekati nilai aktual jika dibandingkan dengan nilai dugaan model lainnya.

Uji Signifikansi Model Terbaik

Uji Kesesuaian Model Binomial Negatif

Secara Simultan (*Overall test*)

Tabel 10. Uji Simultan

Likelihood Ratio Chi-Square	Derajat bebas	<i>p-value</i>
7,2218	2	0,02702711

Nilai $p - value < \alpha$ yaitu $0,02702711 < 0,05$ maka 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 peubah prediktor yang berpengaruh terhadap jumlah kasus kematian akibat DBD.

Uji Signifikansi Parameter Binomial Negatif

Secara Individu (*Partial test*)

Untuk uji signifikansi individu dapat dilihat nilai $p-value$ pada tabel 6. Berdasarkan tingkat signifikansi $\alpha = 0,05$ disimpulkan bahwa pada penduga X_1 dengan $p-value 0,2244 > 0,05$ tidak signifikan atau tidak terdapat pengaruh terhadap kematian DBD. Sedangkan pada penduga X_4

dengan $p-value 0,0350 < 0,05$ signifikan atau terdapat pengaruh terhadap kasus kematian DBD. Hal ini sejalan dengan penelitian M. Al Haris (2017) Berdasarkan nilai AIC yang terkecil, model regresi Binomial Negatif dipilih untuk memodelkan jumlah kematian akibat DBD di Jawa Tengah tahun 2012 dengan melibatkan peubah persentase penduduk yang berusia dibawah 15 tahun, rasio sarana kesehatan masyarakat per 100,000 penduduk dan persentase balita usia 0 - 4 tahun yg diberi imunisasi lengkap.

Nilai estimasi parameter X_4 sebesar $\exp(0,732456)$. Artinya, setiap penambahan satu satuan sarana kesehatan masyarakat per 100.000 penduduk (X_4) menyebabkan nilai harapan jumlah kematian akibat DBD mengalami penurunan sebesar $\exp(0,732456) = 1,99$ kali dari jumlah kasus semula.

5. KESIMPULAN

Adapun kesimpulan dalam penelitian ini adalah:

- 1) Model yang terbentuk yaitu:
 - a. Model regresi Binomial Negatif

$$(\hat{\mu}_i) = \exp(6,736515 - 0,004189X_1 - 0,063911X_2 - 0,020798X_3 - 0,657145X_4)$$
 - b. Model regresi *Generalized Poisson*

$$(\hat{\mu}_i) = \exp(8,0065517 - 0,0002402X_1 - 0,1120160X_2 - 0,0420522X_3 - 0,5026212X_4)$$
- 2) Model terbaik yang didapatkan berdasarkan nilai AIC terkecil yaitu model regresi Binomial Negatif. Nilai dari keseluruhan model memiliki nilai AIC sebesar 66,50 sedangkan nilai AIC dari model yang terpilih sebesar 62,82.
- 3) Berdasarkan hasil analisis regresi Binomial Negatif, faktor-faktor yang berpengaruh

secara signifikan terhadap kasus kematian akibat DBD di Provinsi Sulawesi Selatan tahun 2020 yakni rasio sarana kesehatan per 100.000 penduduk (X_4).

DAFTAR PUSTAKA

- Ashwan, N., & Kamel, M. (2011). Using Generalized Poisson Log Linear Regression Models in Analyzing Two-Way Contingency Tables. *Journal Applied Mathematical Science*, 5(5), 213-222.
- Darnah. (2011). Mengatasi Overdispersi pada Model Regresi Poisson dengan Generalized Poisson Regression I. *Jurnal Eksponensial*, 2(2), 6.
- Dobson, A. J. (2002). *AN Introduction to Generalized Linear Models, Secon Edition*. New York: A CRC Press Company.
- Haris, A. (2017). *Regresi Binomial Negatif dan General Poisson pada Data DBD di Jawa Tengah*. Bogor: Institut Pertanian Bogor.
- Irawati, B. & Purhadi. (2012). Perbandingan Analisis Generalized Poisson Regression (GPR) dan Regresi Binomial Negatif untuk Mengatasi Overdispersi Studi Kasus: Pemodelan Jumlah Kasus Kanker Serviks di Jawa Timur. *Jurnal Matematika*, 2(2), 2.
- Ismah, Kurnia, A., & Sadik, K. (2020). Analysis Of Overdispersed Count Data By Poisson Model. *European Journal of Moleculer & Clinical Medicine*, 54.
- Ismail, N., & Jemain, A. (2007). *Handling Overdispersion with Negative Binomial and Generalized Poisson Regression Models*. Virginia (US): Casualty Actuarial Society Forum.
- Jansakul, N., & Hinde, J. (2002). Score Test For Zero Inflated Poisson Models. *Computational Statistics and Data Analysis*, 40.
- Listiani, Y. (2010). *Pemodelan Regresi Generalized Poisson pada Faktor-Faktor yang Mempengaruhi Angka Kematian Bayi di Jawa Timur Tahun 2007*. Surabaya: Institut Teknologi Sepuluh Nopember.
- McCullagh, P., & FRS, J. N. (1989). *Generalized Linear Models Second Edition*. London: Chapman and Hall.
- Nugraha, J. (2013). *Pengantar Analisis Data Kategorik*. Yogyakarta: Deepublish.
- Nugroho, S. (2008). *Pengantar Statistika Matematika*. Bengkulu: UNIB Press.
- Numna, S. (2009). *Analysis of Extra Zero Counts Using Zero-Inflated Poisson*. Songkhla: Prince of Songkhla University.
- Putri, S. (2013). *Perbandingan Analisis Generlized Poisson Regression (GPR) dan Regresi Binomial Negatif (Studi Kasus: Pemodelan Jumlah Penderita Kusta di Jawa Timur Tahun 2012)*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Soepardi, J. (2010). Demam Berdarah Dengue. *Buletin Jendela Epidemiologi*, 13.
- Sriningsih, M., Hatidja, D., & Prang, J. D. (2018). Penanganan Multikolinearitas dengan Menggunakan Analisis Regresi Komponen Utama Pada Kasus Impor Beras di Provinsi Sulut. *Jurnal Ilmiah Sains*, 18(1), 19.
- Tiro, M. A., Sukarna, & Aswi. (2008). *Pengantar Teori Peluang*. Makassar: Andira Publisher.
- Yulianingsih, K. A., Sukarsa, K. G., & Suciptawati, L. P. (2012). Penerapan Regresi Poisson untuk Mengetahui Faktor-Faktor yang Mempengaruhi Jumlah Siswa SMA/SMK yang Tidak Lulus UN di Bali. *Jurnal Matematika*, 1(1), 61.