

# Pemodelan Jumlah Kasus HIV/AIDS di Provinsi Lampung Menggunakan Regresi Binomial Negatif

Cici Tria Fatmala<sup>1</sup>, Ma'rufah Hayati<sup>2\*</sup>, Reni Permata Sari<sup>3</sup>, Mahfuz Hudori<sup>4</sup>

<sup>1,2,3,4</sup> Program Studi Statistika, Universitas Nahdlatul Ulama Lampung, Purbolinggo, Indonesia

Corresponding Email\*: [marufahhayatim1@gmail.com](mailto:marufahhayatim1@gmail.com)

## Abstrak

Penelitian ini bertujuan untuk menentukan metode terbaik dalam memodelkan jumlah kasus HIV/AIDS di Provinsi Lampung dengan membandingkan regresi Poisson dan regresi Binomial Negatif. Data penelitian diperoleh dari Badan Pusat Statistik (BPS) dan Dinas Kesehatan Provinsi Lampung. Regresi Poisson umum digunakan untuk data berjenis count, namun memiliki keterbatasan dalam menghadapi overdispersion, di mana varians lebih besar dari rata-rata. Sebagai alternatif, regresi Binomial Negatif mampu mengatasi masalah overdispersion ini dengan lebih baik. Hasil analisis menunjukkan bahwa model terjadi overdispersi dalam pemodelan regresi Poisson dan regresi Binomial Negatif memiliki nilai AIC yang lebih kecil, sehingga lebih efektif dalam memodelkan jumlah kasus HIV/AIDS di Provinsi Lampung dibandingkan dengan regresi Poisson. Selain itu, analisis menunjukkan bahwa variabel pengangguran memiliki pengaruh yang signifikan terhadap jumlah kasus HIV/AIDS.

**Kata Kunci:** Jumlah Kasus HIV/AIDS, Regresi Poisson, Regresi Binomial Negatif, *Generalized Linear Model*, *Overdispersi*.

## Abstract

This study aims to determine the best method for modeling the number of HIV/AIDS cases in Lampung Province by comparing Poisson regression and Negative Binomial regression. The research data were obtained from the Central Statistics Agency and the Lampung Provincial Health Office. Poisson regression is commonly used for count data but has limitations in dealing with overdispersion, where the variance is greater than the mean. As an alternative, Negative Binomial regression can better address this overdispersion issue. The analysis results show that there is overdispersion in the Poisson regression model, and the Negative Binomial regression has a smaller AIC value, making it more effective in modeling the number of HIV/AIDS cases in Lampung Province compared to Poisson regression. Additionally, the analysis indicates that the unemployment variable has a significant influence on the number of HIV/AIDS cases.

**Keywords:** Number of HIV/AIDS Cases, Poisson Regression, Negative Binomial Regression, *Generalized Linear Model*, *Overdispersion*

Received :25-07-2024, Revised :10-09-2024, Accepted :23-09-2024

## 1. Pendahuluan

Kesehatan adalah hak asasi manusia yang mendasar dan merupakan salah satu aspek penting dalam kualitas sumber daya manusia yang harus diperhatikan. Karena perannya yang krusial dalam investasi sumber daya manusia, peran pemerintah dalam pembangunan sektor kesehatan secara berkelanjutan menjadi sangat penting [1]. Salah satu tantangan terbesar dalam sektor kesehatan saat ini adalah penyebaran HIV (*Human Immunodeficiency Virus*), virus yang menyerang sel darah putih, khususnya limfosit T CD4, dan menyebabkan penurunan kekebalan tubuh yang parah pada individu yang terinfeksi. Infeksi ini dapat berkembang menjadi AIDS (*Acquired Immune Deficiency Syndrome*), yang ditandai dengan munculnya berbagai gejala penyakit akibat penurunan kekebalan tubuh [2]. HIV dapat ditularkan melalui hubungan seksual, transfusi darah, berbagi jarum suntik, serta dari ibu ke anak selama proses kelahiran dan menyusui, menjadikannya ancaman serius bagi kesehatan masyarakat [3].

Masalah HIV tidak hanya terbatas pada satu wilayah tetapi telah menjadi salah satu masalah kesehatan global yang mendesak. Menurut laporan epidemi HIV global dari United Nations Programme on HIV and AIDS (UNAIDS), pada tahun 2019 terdapat 38 juta orang di seluruh dunia yang hidup dengan HIV. Yang lebih mengkhawatirkan, sekitar 7,1 juta orang di antaranya tidak menyadari bahwa mereka telah terinfeksi [4]. Indonesia termasuk dalam negara dengan laju peningkatan kasus HIV/AIDS tercepat di Asia Tenggara, dengan perkiraan peningkatan infeksi HIV lebih dari 36% [5]. Bahkan, Indonesia menempati peringkat pertama di kawasan ini dalam hal jumlah kumulatif kasus HIV/AIDS tertinggi, dengan total 549.291 kasus. Dalam dekade terakhir, kasus baru HIV/AIDS di Indonesia terus mengalami peningkatan, dengan jumlah tertinggi mencapai 50.282 kasus baru pada tahun 2019 [6].

Provinsi Lampung, seperti banyak daerah lain di Indonesia, tidak luput dari dampak epidemi ini. Penyakit HIV/AIDS pertama kali dilaporkan di Provinsi Lampung pada tahun 2002, dan sejak saat itu, insidensi penyakit ini terus meningkat setiap tahunnya. Prevalensi kasus HIV di provinsi ini pada periode 2010-2015 berkisar antara 0,03% hingga 0,04% [7]. Data dari Dinas Kesehatan Kota Bandar Lampung pada Triwulan ke-III Tahun 2017 menunjukkan bahwa kasus HIV/AIDS lebih banyak terjadi pada laki-laki, dengan 248 kasus, dibandingkan dengan 73 kasus pada perempuan [8]. Peningkatan yang terus-menerus ini menegaskan perlunya pendekatan yang tepat dalam memodelkan HIV/AIDS agar faktor-faktor yang berkontribusi terhadap peningkatan kasus dapat diidentifikasi dan ditangani.

Salah satu metode statistik yang umum digunakan untuk memodelkan data count seperti kasus HIV/AIDS adalah regresi Poisson yang berbasis *Generalized Linear Model* (GLM) [4]. Namun, dalam banyak kasus, regresi Poisson menghadapi masalah overdispersi, di mana variansi data lebih besar dari rata-rata, yang dapat menyebabkan kesalahan interpretasi hasil [9]. Untuk mengatasi overdispersi ini, regresi binomial negatif dapat digunakan sebagai alternatif yang lebih sesuai, memberikan estimasi yang lebih akurat dalam kondisi di mana data menunjukkan variabilitas yang tinggi. Pendekatan ini diharapkan dapat memberikan pemahaman yang lebih baik tentang faktor-faktor yang mempengaruhi peningkatan jumlah kasus HIV/AIDS di Provinsi Lampung.

Salah satu metode yang dapat digunakan untuk mengatasi overdispersi adalah regresi binomial negatif. Regresi binomial negatif adalah teknik statistik yang khusus digunakan untuk memodelkan data count dengan overdispersi, di mana variansi data lebih besar dari rata-ratanya. Metode ini sering diandalkan dalam analisis epidemiologi untuk mengatasi masalah overdispersi yang kerap muncul dalam data kasus penyakit menular, termasuk HIV/AIDS [10]. Penelitian sebelumnya mengenai HIV di Provinsi Lampung telah menerapkan berbagai metode statistik, termasuk regresi Poisson, namun sering kali mengalami kendala overdispersi, yang mengurangi keakuratan hasil model [3].

Penelitian ini bertujuan untuk membandingkan efektivitas metode regresi Poisson dengan regresi binomial negatif dalam memodelkan data HIV/AIDS di Provinsi Lampung. Pemilihan metode terbaik akan didasarkan pada nilai Akaike Information Criterion (AIC) terkecil, yang menunjukkan model dengan kecocokan terbaik dan kompleksitas terendah [6]. Hasil dari penelitian ini diharapkan dapat memberikan kontribusi signifikan dalam pemahaman dan penanganan epidemi HIV/AIDS di wilayah tersebut.

## 2. Landasan Teori

### 1.1. *Generalized Linier Model* (GLM)

[11] menyatakan bahwa GLM adalah pengembangan dari model linier di mana variabel responsnya tidak harus mengikuti distribusi normal. Variabel respons biasanya mengikuti salah satu distribusi dalam keluarga eksponensial, seperti Normal, Poisson, Binomial, Gamma, dan Inverse Gaussian [12]. Salah satu bentuk GLM yang cukup populer, selain model regresi logistik, adalah model regresi Poisson. Model regresi Poisson digunakan untuk menganalisis hubungan antara beberapa variabel, di mana variabel respons berbentuk diskrit tetapi tidak biner [10]. Komponen *Generalized Linier Model* (GLM) ada 3 yaitu:

#### a. Fungsi *Link* atau fungsi hubung (*Link Function* $g(\cdot)$ )

Fungsi link digunakan untuk menghubungkan nilai ekspektasi dari respons  $\mu = E(Y)$  dengan prediktor linear  $\eta$ . Fungsi link mengubah rata-rata respons menjadi kombinasi linear dari parameter model. Fungsi link yang sering digunakan meliputi logit untuk regresi logistik dan log untuk regresi

Poisson. Model regresi untuk GLM dengan menggunakan fungsi hubung memiliki bentuk sebagai berikut:

$$g(\mu_i) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{2i} + \dots + \beta_k X_{ik}$$

b. Prediktor linier (*Linear Predictor*)

Prediktor linear  $\eta_i$  merupakan kombinasi linear dari parameter model ( $\beta$ ) dan variabel bebas ( $X$ ). Prediktor linear ini membentuk dasar dari fungsi link dan menggambarkan hubungan antara variabel bebas dan ekspektasi dari variabel respons yang memiliki bentuk sebagai berikut:

$$\eta_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{2i} + \dots + \beta_k X_{ik}$$

c. Distribusi Eksponensial (*Exponential Family Distribution*)

GLM mengasumsikan bahwa variabel respons  $Y$  mengikuti distribusi dari keluarga eksponensial. Distribusi ini menentukan bentuk fungsi *log-likelihood* yang digunakan dalam estimasi parameter. Distribusi yang termasuk dalam keluarga eksponensial meliputi distribusi normal, binomial, Poisson, gamma, dan inverse Gaussian. Dikatakan keluarga eksponensial jika distribusi tertentu bisa dibentuk dalam fungsi berikut:

$$f(y; \theta, \phi) = \exp\left(\frac{y\theta - b(\theta)}{\phi} + c(y, \phi)\right)$$

Di mana  $\theta$  adalah parameter kanonik,  $\phi$  adalah parameter dispersi, dan  $b(\theta)$  serta  $c(y, \phi)$  adalah fungsi spesifik distribusi [13].

### 1.2. Regresi Poisson

Distribusi Poisson adalah suatu distribusi yang digunakan untuk peristiwa dengan probabilitas kejadian kecil, dimana kejadian tersebut bergantung pada interval waktu tertentu atau di suatu daerah tertentu dengan hasil pengamatan berupa variabel diskrit [14]. Regresi Poisson adalah pendekatan regresi yang populer untuk memodelkan data cacah. Model regresi Poisson dikembangkan melalui Generalized Linear Models (GLM). Dalam pemodelan regresi Poisson, asumsi ekuidispersi harus dipenuhi. Asumsi ini menunjukkan bahwa varians dan rata-rata dari variabel dependen memiliki nilai yang sama [15]. Model regresi Poisson digunakan untuk memodelkan data count atau hitungan, khususnya ketika data tersebut mengikuti distribusi Poisson. Secara matematis, model regresi Poisson dinyatakan sebagai berikut:

$$\log(\lambda_i) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{2i} + \dots + \beta_k X_{ik}$$

Di mana:

- $\lambda_i$  adalah rata-rata (mean) jumlah kejadian untuk pengamatan ke-i.
- $\beta_0$  adalah intersep (titik potong dengan sumbu Y).
- $\beta_1, \beta_2, \dots, \beta_k$  adalah koefisien regresi yang mengukur pengaruh masing-masing variabel bebas  $X_{i1}, X_{2i}, \dots, X_{ik}$  terhadap jumlah kejadian.
- $X_{i1}, X_{2i}, \dots, X_{ik}$  adalah variabel bebas (independen) untuk pengamatan ke-i.

Model ini dapat ditulis ulang dalam bentuk eksponensial untuk menunjukkan bahwa  $\lambda_i$  mengikuti distribusi Poisson:

$$\lambda_i = e^{\beta_0 + \beta_1 X_{i1} + \beta_2 X_{2i} + \dots + \beta_k X_{ik}}$$

Fungsi eksponensial memastikan bahwa prediksi  $\lambda_i$  selalu positif, sesuai dengan sifat data count yang tidak bisa negatif. Model regresi Poisson membantu dalam memahami bagaimana variabel bebas mempengaruhi rata-rata jumlah kejadian dari data count yang diberikan [16][17][18].

### 1.3. Regresi Binomial Negatif

Regresi binomial negatif adalah model yang mengasumsikan bahwa peubah responnya mengikuti distribusi binomial negatif. Salah satu keunggulan dari regresi binomial negatif adalah dapat digunakan dalam kondisi equidispersion maupun overdispersion. Selain itu, karena peubah respon pada regresi binomial negatif diasumsikan mengikuti distribusi binomial negatif, maka tidak mengharuskan nilai variansnya sama dengan rata-ratanya [19]. Regresi binomial negatif digunakan sebagai alternatif dari model regresi Poisson yang mengalami overdispersi (varian > rata-rata) [20]. Berikut adalah model dari regresi binomial negatif:

$$g(\mu_i) = \eta_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} \text{ dengan } i = 1, 2, \dots, n$$

## 3. Metode

### 3.1. Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari profil dinas kesehatan provinsi Lampung dan publikasi Badan Pusat Statistika (BPS) provinsi Lampung. Jumlah sampel yang digunakan, yaitu sebanyak 15 kabupaten/kota di provinsi Lampung.

**Tabel 1.** Variabel penelitian

Variabel	Keterangan	Satuan
Y	Jumlah kasus AIDS	Kasus
X <sub>1</sub>	Kepadatan penduduk	Jiwa
X <sub>2</sub>	Penduduk miskin	Persen
X <sub>3</sub>	Pengangguran	Persen

Sumber: [21]

### 3.2. Langkah-Langkah Penelitian

Data yang digunakan dalam penelitian ini diolah menggunakan software R-Studio, berikut Langkah-langkah dalam melakukan penelitian:

- 1) Kumpulkan data yang relevan, termasuk variabel dependen (y) yang merupakan jumlah kasus HIV/AIDS dan variabel independen (x) seperti kepadatan penduduk, persentase penduduk miskin, dan persentase pengangguran
- 2) Statistika deskriptif menggunakan ringkasan numerik dan visualisasi data
- 3) Terapkan regresi Poisson menggunakan fungsi glm() dari package MASS di R-Studio, dengan model link log dan distribusi Poisson.
- 4) Uji apakah data mengalami overdispersi atau underdispersi dengan membandingkan nilai *deviance* dan *degrees of freedom*.
- 5) Jika ditemukan overdispersi, lakukan analisis regresi binomial negatif menggunakan fungsi glm.nb() dari package MASS.
- 6) Bandingkan kedua model dengan melihat nilai *Akaike Information Criterion* (AIC) dan pilih model dengan nilai AIC terkecil.
- 7) Gunakan model yang terpilih untuk interpretasi lebih lanjut mengenai jumlah kasus HIV/AIDS.
- 8) Menjelaskan hasil dari model terpilih, termasuk hubungan antara variabel independen dan variabel dependen, serta implikasinya dalam konteks epidemiologi HIV/AIDS.

## 4. Hasil dan Pembahasan

Bagian ini menyajikan hasil analisis data kasus HIV/AIDS di Provinsi Lampung menggunakan metode regresi Poisson dan regresi binomial negatif. Analisis bertujuan untuk menemukan model terbaik dengan nilai AIC terkecil. Hasilnya ditampilkan dalam tabel dan grafik, dengan pembahasan mengenai faktor

signifikan yang mempengaruhi peningkatan kasus HIV/AIDS, serta kelebihan dan kekurangan masing-masing metode. Implikasi praktis dari hasil ini akan turut dibahas.

4.1 Analisis statistika Deskriptif

Bagian ini menyajikan ringkasan numerik dari variabel-variabel yang dianalisis dalam penelitian ini, Data tersebut mencakup nilai minimum, maksimum, rata-rata, dan varians, yang memberikan gambaran umum tentang distribusi dan penyebaran masing-masing variabel. Ringkasan ini membantu dalam memahami karakteristik dasar data sebelum analisis lebih lanjut.:

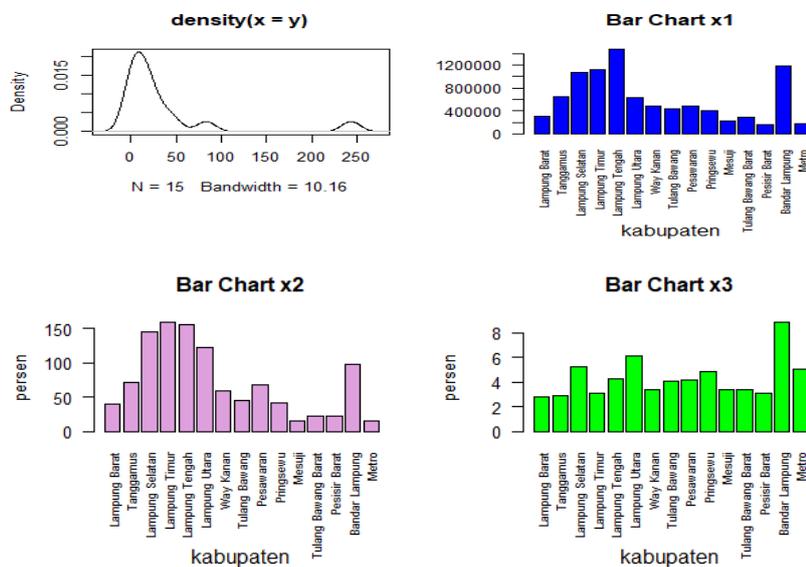
Tabel 2. Ringkasan numerik seluruh variabel

variabel	Minimum	Maksimum	Rata-rata	Varians
Y	0.0	243	34.5	8766.9
X <sub>1</sub>	163641	1477395	605456	227257424021.7
X <sub>2</sub>	15.2	159.8	72.2	3807.3
X <sub>3</sub>	2.8	8.9	4.3	29.7

Jumlah kasus HIV/AIDS di wilayah studi sangat bervariasi, mulai dari 0 hingga 243 kasus. Rata-rata kasus per wilayah adalah sekitar 34,5 kasus, tetapi varians yang cukup besar (8766,9) menunjukkan bahwa ada ketidakhomogenan yang cukup tinggi antar wilayah. Ini mengindikasikan adanya daerah dengan jumlah kasus yang jauh lebih tinggi dibandingkan daerah lainnya, yang mungkin memerlukan perhatian khusus dalam upaya pencegahan dan penanganan. Kepadatan penduduk juga menunjukkan variasi yang signifikan antara wilayah studi. Kepadatan penduduk bervariasi dari 163.641 hingga 1.477.395 orang per km<sup>2</sup> dengan rata-rata 605.456 orang per km<sup>2</sup>. Varians yang sangat besar (227.257.424.021,7) menunjukkan perbedaan yang ekstrem antara wilayah dengan kepadatan penduduk rendah dan tinggi. Hal ini penting karena kepadatan penduduk dapat mempengaruhi penyebaran HIV/AIDS, dengan daerah yang lebih padat mungkin memiliki risiko lebih tinggi.

Persentase penduduk miskin di wilayah studi memiliki variasi yang luas, mulai dari 15,2% hingga 159,8%, dengan rata-rata 72,2%. Varians sebesar 3807,3 mengindikasikan adanya variasi yang signifikan dalam tingkat kemiskinan di berbagai wilayah. Persentase penduduk miskin yang tinggi dapat menjadi faktor risiko tambahan untuk penyebaran HIV/AIDS, karena kemiskinan sering kali berkaitan dengan akses yang lebih rendah terhadap layanan kesehatan. Persentase pengangguran di wilayah studi lebih konsisten dengan minimum 2,8% dan maksimum 8,9%, serta rata-rata 4,3%. Varians yang relatif kecil (29,7) menunjukkan bahwa tingkat pengangguran cenderung lebih seragam di seluruh wilayah studi. Meskipun demikian, pengangguran masih dapat mempengaruhi penyebaran HIV/AIDS, terutama jika dikaitkan dengan faktor-faktor lain seperti kemiskinan dan kepadatan penduduk.

Berikut adalah grafik dari setiap variabel yang digunakan dalam penelitian:



**Gambar 1.** Density plot variabel Y dan *bar-chart* variabel  $X_1$ ,  $X_2$  dan  $X_3$

Berdasarkan **Gambar 1** terlihat bahwa plot density Y menunjukkan bahwa plot kepadatan kernel tidak menggambarkan distribusi normal. Namun, plot ini menampilkan kurva yang tidak simetris, dengan ekor panjang di salah satu sisi dan beberapa puncak kecil, mengindikasikan distribusi yang lebih kompleks dan adanya outlier. Secara keseluruhan, plot ini menunjukkan bahwa distribusi jumlah kasus penderita HIV/AIDS memiliki puncak di sekitar nilai yang lebih rendah dengan beberapa kasus ekstrem di kedua sisi.

Barchart untuk variabel  $X_1$  menjelaskan bahwa kepadatan penduduk tertinggi adalah Lampung Selatan. Barchart variabel  $X_2$  yaitu variabel kemiskinan terlihat bahwa kabupaten termiskin adalah Lampung Timur. Barchatr variabel  $X_3$  yaitu variabel pengangguran menunjukkan bahwa angka pengangguran tertinggi adalah Bandar Lampung

**4.2 Pemodelan Jumlah Kasus HIV/AIDS Menggunakan Regresi Poisson**

Data jumlah kasus HIV/AIDS dimodelkan menggunakan regresi Poisson. Hasil estimasi model regresi poisson dengan menggunakan bantuan dari *software* R-Studio dan dibantu *packages* AER disajikan dalam tabel sebagai berikut:

**Tabel 3.** Hasil Estimasi Prameter Regresi Poisson

Parameter	Estimasi	z-Value	Pr(> z )
<b>Intercept</b>	5.512e-01	3.302	0.00096
$\beta_1$	-4.835e-07	-1.621	0.10510
$\beta_2$	-3.140e-03	-1.501	0.13347
$\beta_3$	6.640e-01	17.315	<2e-16

Dari pemodelan regresi Poisson dapat dilihat bahwa tingkat pengangguran berpengaruh terhadap penyebaran virus HIV/AIDS. Sehingga diperoleh persamaan regresi sebagai berikut:

$$\mu = \exp(5.512e-01 - 4.835e-07X_1 - 3.140e-03X_2 + 6640e-01X_3)$$

Dalam metode regresi Poisson, diasumsikan bahwa rata-rata dan varians dari variabel respon adalah sama. Akan tetapi ada kalanya varians lebih besar dari rata-rata (*overdispersi*) atau varians lebih kecil dari rata-rata respon (*underdispersi*).

**4.3 Deteksi Overdispersi Dan Underdispersi**

*Overdispersi* dapat diperiksa menggunakan statistik uji serentak yang dibagi dengan derajat bebas dan disimbolkan dengan  $\theta$  Jika nilai  $\theta$  lebih besar dari nol, maka terdapat kondisi *overdispersi* [22]. *Overdispersi* dapat dideteksi dengan membagi nilai Pearson chi-square atau deviansi dengan derajat bebasnya. karena *overdispersi* terjadi pada regresi Poisson, maka salah satu alternatif yang dapat digunakan untuk menanganinya adalah regresi binomial *negative* [23]

**Tabel 4.** Hasil uji overdispersi

Nilai Deviance	Df	Nilai Deviance/Df
132.66	11	12.1

Dapat dilihat bahwa hasil dari uji overdispersi adalah  $12.06 > 1$ , yang menunjukkan bahwa adanya *overdispersi* pada model yang dibuat. Adanya *overdispersi* menyebabkan kesimpulan yang tidak valid dalam analisis statistika.

#### 4.4. Pemodelan Regresi Binomial Negatif

Salah satu alternatif untuk menangani overdispersi yang terjadi pada model regresi poisson adalah regresi binomial negatif. Berikut ini adalah tabel hasil estimasi parameter dari model regresi binomial negatif :

**Tabel 5.** Hasil Estimasi Parameter Regresi Binomial Negatif

Parameter	Estimasi	z-value	Pr(> z )
Intercept	-4.035e-01	-0.555	0.579
$\beta_1$	2.418e-08	0.016	0.988
$\beta_2$	-6.927e-04	-0.058	0.953
$\beta_3$	7.534e-01	4.434	9.26e-06

Dari pemodelan regresi binomial negative dapat dilihat bahwa tingkat pengangguran berpengaruh signifikan terhadap penyebaran jumlah kasus HIV/AIDS dengan p-value = 9.26e-06.

#### 4.5. Pemilihan Model Terbaik

Setelah dilakukan pengujian regresi binomial negatif , Langkah selanjutnya adalah pemilihan model terbaik antara model regresi poisson dan model regresi binomial negatif dengan melihat nilai AIC yang lebih rendah. Berikut adalah hasil dari perhitungan nilai AIC dari model regresi poisson dan model regresi binomial negatif :

**Tabel 6.** Nilai AIC

Model	Nilai AIC
Regresi Poisson	201.2244
Regresi Binomial Negatif	125.0593

Berdasarkan tabel diatas nilai AIC terkecil diperoleh pada model regresi binomial negatif. Maka dapat disimpulkan bahwa model terbaik yang di gunakan untuk jumlah kasus penderita HIV/AIDS di provinsi Lampung adalah model regresi binomial negative

#### 4.6. Model Terbaik yang Terbentuk

Model yang terpilih dalam penelitian ini adalah model regresi binomial negatif dengan nilai AIC lebih rendah dibandingkan dengan nilai AIC model regresi poisson. Maka model regresi binomial negatif adalah model terbaik untuk digunakan dalam penelitian ini. berikut adalah model regresi binomial yang terbentuk:

$$\mu_i = \exp(-4.035e-01 + 2.418e-08X_1 - 6.927e-04X_2 + 7.534e-01X_3)$$

Dalam model regresi binomial negatif ini, koefisien variabel prediktor memberikan gambaran tentang pengaruh masing-masing faktor terhadap ekspektasi jumlah kasus HIV/AIDS:

- Kepadatan penduduk ( $X_1$ ) berkontribusi positif terhadap ekspektasi jumlah kasu HIV/AIDS, dengan setiap kenaikan satu satuan meningkatkan ekspektasi sebesar  $\exp(2.418e-08)$ . Hal ini berarti bahwa nilai eksponensial dari 2.418e-08 adalah sekitar 1.0000002418. Ini menunjukkan bahwa setiap peningkatan 1 jiwa per kilometer persegi dalam kepadatan penduduk akan meningkatkan jumlah kasus AIDS sekitar 0.000002418% (dalam angka relatif dan sangat kecil).
- Penduduk miskin ( $X_2$ ) memiliki pengaruh negatif, di mana setiap kenaikan satu satuan mengurangi ekspektasi sebesar  $\exp(-6.927e-04)$ . Hal ini berarti bahwa nilai eksponensial dari -6.927e-04 adalah sekitar 0.9993077. Ini berarti bahwa setiap peningkatan 1% dalam persentase penduduk miskin akan mengurangi jumlah kasus AIDS sekitar 0.06923%. Jadi, persentase penduduk miskin memiliki pengaruh negatif yang sangat kecil terhadap jumlah kasus AIDS.
- pengangguran ( $X_3$ ) memiliki pengaruh positif di mana setiap kenaikan satu satuan meningkatkan ekspektasi sebesar  $\exp(7.534e-01)$ . Hal ini berarti bahwa Nilai eksponensial dari 7.534e-01 adalah sekitar 2.109. Ini menunjukkan bahwa setiap peningkatan 1% dalam tingkat pengangguran akan

meningkatkan jumlah kasus AIDS sekitar 110.9%. Ini merupakan efek yang cukup besar, menunjukkan bahwa pengangguran memiliki pengaruh signifikan terhadap jumlah kasus AIDS.

## 5. Kesimpulan

Penelitian ini menemukan bahwa model regresi Binomial Negatif adalah metode terbaik untuk memodelkan jumlah kasus HIV/AIDS di Provinsi Lampung, terutama karena mampu menangani masalah overdispersi yang ada dalam data. Model Binomial Negatif dipilih karena memiliki nilai Akaike Information Criterion (AIC) yang lebih rendah dibandingkan dengan model regresi Poisson, menandakan kecocokan model yang lebih baik dengan data yang tersedia.

Hasil model menunjukkan bahwa tingkat pengangguran memiliki pengaruh paling signifikan terhadap peningkatan jumlah kasus HIV/AIDS, dengan setiap peningkatan 1% dalam tingkat pengangguran berhubungan dengan peningkatan yang substansial dalam jumlah kasus. Sementara itu, kepadatan penduduk dan persentase penduduk miskin menunjukkan pengaruh yang jauh lebih kecil terhadap jumlah kasus HIV/AIDS, sehingga model regresi binomial negatif yang diperoleh adalah:

$$\mu_i = \exp(-4.035e-01 + 2.418e-08X_1 - 6.927e-04X_2 + 7.534e-01X_3)$$

Seluruh variabel tetap dituliskan dalam model walau sebenarnya hanya variable pengangguran yang berpengaruh. Karena hal ini tergantung dari urgensi setiap variable dalam model menurut bidangnya. Jika variabel tersebut harus ada dalam model maka variable yang tidak signifikan tidak perlu dihilangkan dari modelnya.

## Ucapan Terima Kasih

Penulis ingin mengucapkan terimakasih kepada Dr. Ma'rufah Hayati MT, S.Si., M.Sc selaku dosen mata kuliah analisis data kategorik yang telah meluangkan banyak waktunya untuk memberikan ilmu, bimbingan serta arahnya kepada penulis dalam menyelesaikan penelitian ini.

## Referensi

- [1] N. T. Ratnasari and Purhadi, "Pemodelan Faktor yang Mempengaruhi Jumlah Hivdan Aids Provinsi Jawa Timur Menggunakan RegresiPoisson Bivariat," *J. Sains Dan Seni Pomits*, vol. 2, no. 2, pp. 2337–3520, 2013.
- [2] S. T. Wahyuni, T. W. Utami, and M. Y. Darsyah, "Pemodelan Generalized Additive Model For Location, Scale, and Shape (Gamlss) Dengan Pemulusan Locally Estimated Scatterplot Smoothing (Loess) pada Kasus Hiv/Aids Di Jawa Timur," *J. Litbang Edusaintech*, vol. 2, no. 1, pp. 18–26, 2021, doi: 10.51402/jle.v2i1.7.
- [3] N. N. S. T. Ulandari, I. Wahina, G. A. M. Adhi, and F. Astuti, "Hubungan Pengetahuan Tentang Infeksi Hiv/Aids Dengan Perilaku Pencegahan Hiv/Aids Pada Remaja Smkn 2 Mataram," *JISIP (Jurnal Ilmu Sos. dan Pendidikan)*, vol. 7, no. 1, pp. 804–809, 2023, doi: 10.58258/jisip.v7i1.4586.
- [4] D. Rohmatullailah and D. Fikriyah, "Hiv Faktor Resiko," *J. Biostat. Kependudukan, dan Inform. Kesehat.*, vol. 2, no. 1, p. 45, 2021.
- [5] M. Zunaidi, A. H. Nasyuha, and S. M. Sinaga, "Penerapan Data Mining Untuk Memprediksi Pertumbuhan Jumlah Penderita Human Immunodeficiency Virus (HIV) Menggunakan Metode Multiple Linier Regression (Studi Kasus Dinas Kesehatan Provinsi Sumatera Utara)," *J-SISKO TECH (Jurnal Teknol. Sist. Inf. dan Sist. Komput. TGD)*, vol. 3, no. 1, p. 137, 2020, doi: 10.53513/jsk.v3i1.205.
- [6] S. F. Thahar and T. Sirait, "Analisis Spasial Variabel-Variabel yang Memengaruhi Jumlah Kasus

Baru HIV/AIDS di Provinsi Jawa Timur Tahun 2021,” *Semin. Nas. Off. Stat.*, vol. 2023, no. 1, pp. 287–296, 2023, doi: 10.34123/semnasoffstat.v2023i1.1602.

- [7] E. Krisdayanti and J. I. Hutasoit, “Pengaruh Coping Strategies terhadap Kesehatan Mental dan Kualitas Hidup Penderita HIV/AIDS positif,” *J. Ilmu Keperawatan Jiwa*, vol. 2, no. 3, p. 179, 2019, doi: 10.32584/jikj.v2i3.440.
- [8] F. Listina and S. nandar Baharza, “Hubungan Pengetahuan Dan Sikap Lsl Terhadap Upaya Pencegahan HIV & AID Di Puskesmas Simpur Kota Bandar Lampung,” *Manuju Malahayati Nurs. J.*, vol. 2, no. 1, pp. 151–159, 2020.
- [9] D. Ratnasari and Purhadi, “Model Regresi Poisson untuk Data yang mengalami Overdispersi,” *J. Mat. dan Stat.*, vol. 9, no. 2, pp. 129–136, 2013.
- [10] M. B. R. U. Utama and N. Hajarisman, “Metode Pemilihan Variabel pada Model Regresi Poisson Menggunakan Metode Nordberg,” *J. Ris. Stat.*, vol. 1, no. 1, pp. 35–42, 2021, doi: 10.29313/jrs.v1i1.24.
- [11] P. McCullagh and J. A. Nelder, *Generalized Linear Models (2nd ed.)*. Chapman & Hall/CRC, 1989.
- [12] Y. Wilandari, S. H. Kartiko, and A. R. Effendie, “Estimasi Cadangan Klaim Menggunakan Generalized Linear Model (Glm) Dan Copula,” *J. Gaussian*, vol. 9, no. 4, pp. 411–420, 2020, doi: 10.14710/j.gauss.v9i4.29260.
- [13] J. W. Hardin and J. M. Hilbe, *Generalized Linear Models and Extensions*. Stata Press, 2018.
- [14] N. A. H. Salamah *et al.*, “Pemodelan Faktor-Faktor Yang Mempengaruhi Jumlah Penderita Demam Berdarah Dengue (Dbd) Di Kota Kendari Menggunakan Regresi Poisson Inverse Gaussian,” *Pemodelan Fakt. Yang Mempengaruhi Jumlah Penderita Demam Berdarah Dengue Di*, vol. 2, no. 2, pp. 1–8, 2022.
- [15] M. Fathurahman, “Regresi Binomial Negatif untuk Memodelkan Kematian Bayi di Kalimantan Timur,” *Ekspensial*, vol. 13, no. 1, p. 79, 2022, doi: 10.30872/ekspensial.v13i1.888.
- [16] A. F. Zuur, E. N. Ieno, and G. M. Smith, “Analyzing Count Data with Poisson Regression,” *Methods Ecol. Evol.*, vol. 11, no. 8, pp. 976–987, 2020.
- [17] G. J. Ragland and J. G. Kingsolver, “Using Poisson Regression Models to Analyze Experimental Count Data.,” *J. Exp. Biol.*, vol. 224, no. 3, p. jeb232967, 2021.
- [18] S. Lee and J. A. Nelder, “Applications, Generalized Linear Models for Count Data: Recent Advances and Applications,” *Stat. Sci.*, vol. 37, no. 1, pp. 52–70, 2022.
- [19] H. M. Winata, “Mengatasi Overdispersi Dengan Regresi Binomial Negatif Pada Angka Kematian Ibu Di Kota Bandung,” *J. Gaussian*, vol. 11, no. 4, pp. 616–622, 2023, doi: 10.14710/j.gauss.11.4.616-622.
- [20] A. Sauddin, N. I. Auliah, and W. Alwi, “Pemodelan Jumlah Kematian Ibu di Provinsi Sulawesi Selatan Menggunakan Regresi Binomial Negatif,” *J. MSA ( Mat. dan Stat. serta Apl. )*, vol. 8, no. 2, p. 42, 2020, doi: 10.24252/msa.v8i2.17409.
- [21] A. Safitri, I. Rahml, and D. Devianto, “Penerapan Regresi Poisson Dan Binomial Negatif Dalam Memodelkan Jumlah Kasus Penderita Aids Di Indonesia Berdasarkan Faktor Sosiodemografi,” *J. Mat. UNAND*, vol. 3, no. 4, pp. 58–65, 2014.
- [22] A. D. Chaniago and S. P. Wulandari, “Pemodelan Generalized Poisson Regression (GPR) dan Negative Binomial Regression (NBR) untuk Mengatasi Overdispersi pada Jumlah Kematian Bayi di Kabupaten Probolinggo,” *J. Sains dan Seni ITS*, vol. 11, no. 6, 2023, doi: 10.12962/j23373520.v11i6.93240.
- [23] N. Delvia, M. Mustafid, and H. Yasin, “Geographically Weighted Negative Binomial Regression Untuk Menangani Overdispersi Pada Jumlah Penduduk Miskin,” *J. Gaussian*, vol. 10, no. 4, pp.

*Cici Tria Fatmala, dkk (2024)*

532–543, 2021, doi: 10.14710/j.gauss.v10i4.33106.



© **The Author(s) 2024.** This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial 4.0 International License. Editorial of Journal of Mathematics: Theory and Applications, Department of Mathematics, Universitas Sulawesi Barat, Jalan Prof. Dr. Baharuddin Lopa, S.H., Talumung, Majene 91412, Sulawesi Barat.