

PEMODELAN JUMLAH KASUS HIV DI PROVINSI JAWA TIMUR MENGGUNAKAN REGRESI BINOMIAL NEGATIF

Mohammad Ichsan Maulana

Program Studi Matematika, Universitas Negeri Surabaya
e-mail: mohammadmaulana.22085@mhs.unesa.ac.id

Danang Ariyanto

Program Studi Matematika, Universitas Negeri Surabaya
e-mail: danangariyanto@unesa.ac.id

Abstrak

Kasus Orang dengan HIV (ODHIV) di Jawa Timur tetap tinggi dan belum menunjukkan penurunan signifikan, sehingga diperlukan pemodelan statistik yang akurat untuk langkah preventif. Penelitian ini bertujuan membangun model regresi binomial negatif dan menganalisis faktor-faktor yang memengaruhi jumlah kasus ODHIV. Pendekatan penelitian yang digunakan adalah kuantitatif dengan desain analisis regresi binomial negatif untuk mengatasi masalah overdispersi pada data cacah. Lokasi penelitian mencakup 38 kabupaten/kota di Provinsi Jawa Timur menggunakan data tahun 2024. Sumber data diperoleh secara sekunder dari Dinas Kesehatan Provinsi Jawa Timur dan Badan Pusat Statistik (BPS). Teknik analisis data melibatkan pengujian asumsi overdispersi, estimasi parameter model, dan uji signifikansi. Hasil penelitian menunjukkan bahwa terdapat empat faktor yang berpengaruh signifikan terhadap jumlah kasus ODHIV, yaitu jumlah pernikahan, rata-rata lama sekolah, tingkat pengangguran terbuka, dan kepadatan penduduk. Simpulan penelitian menegaskan bahwa faktor demografi dan sosial-ekonomi berperan penting dalam penyebaran kasus di Jawa Timur. Disarankan bagi penelitian selanjutnya untuk mempertimbangkan variabel yang lebih spesifik pada level individu, seperti penggunaan NAPZA, orientasi seksual, dan perilaku berisiko lainnya guna memperdalam akurasi model prediktif.

Kata kunci: ODHIV, Regresi Binomial Negatif, Overdispersi, Jawa Timur

Abstract

The prevalence of People Living with HIV (PLHIV) in East Java remains high and has not shown a significant decline, necessitating accurate statistical modeling for preventive measures. This study aims to construct a negative binomial regression model and analyze the factors influencing the number of PLHIV cases. A quantitative approach was employed, utilizing a negative binomial regression design to address the issue of overdispersion commonly found in count data. The research scope covers 38 regencies and cities in East Java Province using 2024 data. Secondary data were sourced from the East Java Provincial Health Office and the Central Bureau of Statistics (BPS). Data analysis involved testing for overdispersion, parameter estimation, and significance testing. The results indicate that four factors significantly affect PLHIV cases: the number of marriages, average years of schooling, the open unemployment rate, and population density. The study concludes that demographic and socio-economic factors play a crucial role in the distribution of cases across East Java. It is suggested that future research incorporate variables that more closely reflect individual behaviors, such as illicit drug use (NAPZA), sexual orientation, and other high-risk behaviors, to enhance the accuracy of predictive models.

Keywords: PLHIV, Negative Binomial Regression, Overdispersion, East Java.

PENDAHULUAN

Human Immunodeficiency Virus (HIV) masih menjadi tantangan serius dalam kesehatan global. Menurut UNAIDS (2024), sekitar 40 juta orang hidup dengan HIV pada akhir tahun 2024. Di Indonesia, beban kasus HIV juga masih tinggi dengan 63.707 kasus baru yang ditemukan sepanjang tahun 2024 (Rizal, 2025). Provinsi Jawa Timur menjadi wilayah dengan jumlah kasus tertinggi di Indonesia, yaitu

sebanyak 10.556 kasus Orang dengan HIV (ODHIV) berdasarkan data Dinas Kesehatan dan Badan Pusat Statistik tahun 2024. Tingginya jumlah kasus menunjukkan bahwa HIV masih menjadi permasalahan kesehatan masyarakat yang serius, terutama karena distribusi kasus antar wilayah di Jawa Timur menunjukkan variasi yang besar.

Faktor seperti kepadatan penduduk, tingkat kemiskinan, pengangguran, pendidikan, dan literasi

masyarakat diketahui berkaitan dengan perilaku berisiko, akses terhadap layanan kesehatan, serta efektivitas penyampaian informasi kesehatan (Cust dkk, 2021). Wilayah dengan kepadatan penduduk tinggi cenderung memiliki intensitas interaksi sosial yang lebih besar sehingga meningkatkan risiko transmisi HIV, sedangkan kondisi ekonomi dan pendidikan yang rendah dapat memengaruhi akses terhadap layanan tes HIV, pengobatan antiretroviral (ARV), serta pemahaman mengenai perilaku pencegahan.

Indikator seperti Indeks Pembangunan Literasi Masyarakat (IPLM), rerata lama sekolah (RLS), indeks pendidikan dan pengetahuan, serta tingkat kegemaran membaca dapat mencerminkan kemampuan masyarakat dalam menerima dan memahami informasi kesehatan. Rendahnya literasi kesehatan dapat menyebabkan kurang optimalnya pemahaman masyarakat mengenai pencegahan HIV dan penggunaan alat proteksi seperti kondom (Etowa dkk, 2021). Oleh karena itu, integrasi faktor literasi dan pendidikan menjadi penting dalam menjelaskan variasi jumlah kasus ODHIV antarwilayah di Provinsi Jawa Timur.

Data jumlah kasus HIV merupakan data cacah yang umumnya dimodelkan menggunakan Regresi Poisson (RP). Namun, model RP memiliki asumsi equidispersi, yaitu nilai rata-rata dan varians data respon relatif sama. Namun, asumsi tersebut sering kali tidak terpenuhi akibat adanya heterogenitas karakteristik wilayah sehingga menyebabkan overdispersi, yaitu kondisi ketika varians melebihi nilai rata-rata (Tohari dkk, 2021). Apabila overdispersi diabaikan, maka estimasi parameter dan simpangan baku dapat menjadi bias sehingga menghasilkan kesimpulan yang kurang akurat (Rahayu, 2020). Untuk mengatasi kondisi tersebut, Regresi Binomial Negatif digunakan karena mampu mengakomodasi overdispersi melalui parameter dispersi tambahan sehingga lebih sesuai untuk data cacah dengan variasi tinggi.

Beberapa penelitian sebelumnya telah menggunakan pendekatan regresi dalam pemodelan kasus HIV. Azizah dkk (2025) menggunakan regresi linier berganda untuk memprediksi penyebaran HIV di Kabupaten Sidoarjo, namun metode tersebut kurang sesuai untuk data cacah yang berpotensi mengalami overdispersi. Sementara itu, Fatmala dkk (2024) menerapkan Regresi Binomial Negatif pada

kasus HIV di Provinsi Lampung dan memperoleh pendekatan yang lebih sesuai secara metodologis, tetapi penelitian tersebut belum mempertimbangkan faktor literasi dan pendidikan masyarakat. Padahal, faktor pendidikan, literasi, dan perilaku kesehatan memiliki peran penting dalam membentuk kesadaran serta perilaku pencegahan HIV.

Berdasarkan uraian tersebut, penelitian ini bertujuan memodelkan jumlah ODHIV di Provinsi Jawa Timur menggunakan Regresi Binomial Negatif dengan mempertimbangkan faktor sosiodemografi, pendidikan, literasi, ekonomi, dan perilaku kesehatan. Pendekatan ini diharapkan mampu menghasilkan model yang lebih sesuai untuk data kasus HIV yang mengalami overdispersi serta memberikan informasi mengenai faktor-faktor yang memengaruhi jumlah ODHIV di Jawa Timur.

KAJIAN TEORI

GENERALIZED LINEAR MODEL

Generalized Linear Model (GLM) merupakan perluasan regresi linear klasik yang digunakan untuk dilakukan pemodelan pada variabel respon non-normal, dengan menghubungkan nilai rata-ratanya ke kombinasi linear prediktor melalui fungsi link tertentu (Santi & Rahayuningsih, 2024). Persamaan GLM memiliki bentuk umum sebagai berikut:

$$g(\mu_i) = \beta_0 + \sum_{j=1}^p \beta_j x_{ij} = \mathbf{x}_i^T \boldsymbol{\beta} \quad (1)$$

dengan $g(\mu_i)$ merupakan fungsi penghubung dan bagian $\beta_0 + \sum_{j=1}^p \beta_j x_{ij} = \mathbf{x}_i^T \boldsymbol{\beta}$ merupakan kombinasi linear variabel prediktor x_i dengan koefisien parameter regresi $\boldsymbol{\beta}$. Komponen-komponen tersebut menjelaskan bahwa pembentukan model dengan GLM perlu diperhatikan bahwa variabel prediktor bersifat linear, variabel respon terkategori ke dalam keluarga eksponensial, serta harus terdapat fungsi penghubung (Armadi Mori dkk, 2024).

REGRESI POISSON

Regresi Poisson (RP) adalah salah satu regresi yang digunakan untuk menganalisis hubungan antara variabel respon yang berupa bilangan cacah dengan satu atau lebih variabel prediktor Y dengan asumsi bahwa Y merupakan variabel acak yang mengikuti distribusi Poisson (Cameron & Trivedi, 2013). RP berasumsi bahwa nilai harapan dan

variannya adalah sama. Jika asumsi equidisersi tidak dipenuhi, model regresi Poisson berpotensi menghasilkan estimasi parameter yang tidak efisien (Rahayu, 2020).

Regresi Poisson menggunakan fungsi link logaritmik (log-link) untuk menghubungkan nilai ekspektasi variabel respon dengan kombinasi linear dari variabel prediktor dengan tujuan untuk memastikan bahwa nilai prediksi selalu bernilai positif (Sundari, 2023). Bentuk umum RP dapat ditulis sebagai berikut:

$$\ln(\mu_i) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_p X_{pi} \quad (2)$$

$$\mu_i = \exp(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_p X_{pi}) \quad (3)$$

Parameter $\beta_0, \beta_1, \beta_2, \dots, \beta_p$ diestimasi menggunakan metode Maximum Likelihood Estimation (MLE). Proses estimasi MLE dimulai dari membangun fungsi likelihood dari fungsi distribusi Poisson sebagai berikut:

$$\begin{aligned} L(\beta) &= \prod_{i=0}^{\infty} P(Y = y_i | \mu_i) \\ &= \prod_{i=0}^{\infty} \frac{e^{-\mu_i} \mu_i^{y_i}}{y_i!} \end{aligned} \quad (4)$$

Fungsi *likelihood* diatas ditransformasikan ke dalam bentuk *log-likelihood* untuk memudahkan proses estimasi.

OVERDISPERSI

Overdispersi adalah kondisi saat varians variabel respon pada data cacah lebih besar daripada nilai rata-ratanya (Saudidin dkk., 2020). Apabila asumsi ini tidak dipenuhi, maka model RP menjadi kurang sesuai untuk digunakan. Pengujian overdispersi diawali dengan melihat nilai devians (D) sebagai berikut:

$$D = 2 \sum_{i=1}^n \left[y_i \ln \left(\frac{y_i}{\hat{\mu}_i} \right) - (y_i - \hat{\mu}_i) \right] \quad (5)$$

Kemudian dilanjutkan untuk keputusan dengan rumus sebagai berikut:

$$\hat{\phi}_D = \frac{D}{n-p} \quad (6)$$

Kriteria pengambilan keputusan didasarkan pada nilai $\hat{\phi}_D$, yaitu:

- $\hat{\phi}_D \approx 1$: Terindikasi equidisersi
- $\hat{\phi}_D > 1$: Terindikasi overdispersi
- $\hat{\phi}_D < 1$: Terindikasi underdispersi

Bila hasil uji overdispersi menunjukkan adanya overdispersi, maka model RP dianggap kurang memadai, yang berarti diperlukan model alternatif. Pada penelitian ini, akan dipilih Regresi Binomial Negatif untuk mengakomodasi varians yang lebih besar dari nilai harapannya.

REGRESI BINOMIAL NEGATIF

Regresi Binomial Negatif (RBN) adalah regresi yang diaplikasikan untuk menganalisis hubungan variabel respon dengan bentuk data yang mengalami overdispersi. Model RBN dikonstruksi berdasarkan asumsi bahwa variabel respon berdistribusi Binomial Negatif bahwa kesamaan antara nilai varians dan rata-rata pada data tidak lagi menjadi suatu keharusan (Winata, 2023).

Model ini mengandung parameter dispersi ψ yang memungkinkan varians pada RBN bervariasi lebih fleksibel dari rata-rata, sehingga memiliki fleksibilitas lebih banyak baik pada data dengan variasi yang tinggi (Hilbe, 2011). Fungsi distribusi Binomial Negatif dapat dinyatakan sebagai berikut:

$$\begin{aligned} P(Y = y) &= \frac{\Gamma\left(y + \frac{1}{\psi}\right)}{\Gamma\left(\frac{1}{\psi}\right) y!} \left(\frac{\exp(\mathbf{X}\beta)\psi}{1 + \exp(\mathbf{X}\beta)\psi} \right)^y \left(\frac{1}{1 + \exp(\mathbf{X}\beta)\psi} \right)^{\frac{1}{\psi}} \end{aligned} \quad (7)$$

Estimasi parameter RBN β dan ψ menggunakan *Maximum Likelihood Estimation* (MLE). Proses estimasi parameter β dan ψ RBN dimulai dari membangun fungsi *likelihood* untuk setiap observasi:

$$L(\beta, \alpha) = \prod_{i=0}^n P(Y = y) \quad (8)$$

RBN menggunakan fungsi link-log dalam kerangka GLM. Penggunaan link-log memastikan bahwa nilai rata-rata μ_i tetap bernilai positif. Apabila dalam proses iterasi telah memenuhi kriteria konvergen dengan ketentuan memenuhi (2.16) atau (2.17), maka dapat disusun persamaan RBN sebagai berikut:

$$\ln(\mu_i) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_p X_{pi} \quad (9)$$

$$\mu_i = \exp(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_p X_{pi}) \quad (10)$$

UJI MULTIKOLINIARITAS

Uji Multikolinearitas dilakukan untuk mengetahui apakah terdapat variabel bebas yang korelasi kuat antara variabel bebas lain model regresi. Model regresi yang baik seharusnya tidak menunjukkan adanya korelasi antara variabel

prediktor, sehingga gejala multikolinearitas dapat dihindari (Effiyaldi dkk, 2022). Multikolinearitas dalam suatu model regresi dapat dideteksi dari *Variance Inflation Factor* (VIF) dengan rumus:

$$VIF = \frac{1}{(1 - R^2)} \quad (11)$$

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \quad (12)$$

Pada persamaan (11) *VIF* merupakan *Variance Inflation Factor* pada variabel prediktor, sedangkan pada persamaan (12), R^2 merupakan koefisien determinasi, SS_{res} jumlah kuadrat dari selisih antara nilai yang diamati dan nilai prediksi, dan SS_{tot} Jumlah kuadrat dari selisih nilai yang diamati dari rata-rata data. Indikasi gejala multikolinearitas adalah sebagai berikut:

$VIF_j > 10$: Variabel j multikolinearitas

$VIF_j \leq 10$: Variabel j tidak multikolinearitas

UJI SIMULTAN

Uji signifikansi secara simultan dilakukan untuk melihat apakah setidaknya satu variabel prediktor pada model regresi signifikan. Uji G dilakukan untuk uji signifikansi variabel prediktor secara simultan pada model RP dan RBN dilakukan dengan hipotesis (Mustafidah & Sofro, 2024):

$$H_0 : \beta_0 = \beta_1 = \beta_2 = \dots = \beta_j = 0$$

$$H_1 : \text{Setidaknya satu } \beta_j \neq 0, \quad j = 0, 1, 2, 3, \dots, k$$

Rumus uji G adalah:

$$G = -2(\log L_0 - \log L_1) \quad (13)$$

dengan G merupakan *Likelihood Ratio Test* (LRT), L_0 adalah *likelihood* ketika tidak ada variabel prediktor, dan L_1 adalah *likelihood* ketika variabel prediktor disertakan. Bandingkan statistik G dengan $\chi^2_{\alpha, df}$ dari tabel *chi-square* dengan α merupakan tingkat signifikansi dan df merupakan derajat kebebasan, sehingga kriteria keputusan pada uji G adalah sebagai berikut:

$$G > \chi^2_{\alpha, df} : \text{Tolak } H_0$$

$$G \leq \chi^2_{\alpha, df} : \text{Gagal tolak } H_0$$

Uji parsial adalah uji untuk melihat pengaruh masing-masing variabel prediktor pada model RP dan RBN. Peneliti ini menggunakan Uji Wald untuk melihat signifikansi setiap variabel prediktor dengan hipotesis (Ariyanto dkk, 2024):

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0$$

dengan statistik uji:

$$W = \left[\frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \right]^2 \quad (14)$$

Pada persamaan (14) ditunjukkan bahwa $\hat{\beta}_j$ merupakan taksiran parameter dan $SE(\hat{\beta}_j)$ merupakan taksiran *standard error* dari $\hat{\beta}_j$. Apabila didapati $W > \chi^2_{(\alpha, 1)}$, maka $\hat{\beta}_j$ tolak H_0 . Namun apabila didapati $W \leq \chi^2_{(\alpha, 1)}$, maka $\hat{\beta}_j$ gagal tolak H_0 . dengan α adalah tingkat signifikansi dan derajat bebas 1. Apabila H_0 ditolak, itu mengindikasikan parameter yang diuji signifikan terhadap Y .

METODE

DATA PENELITIAN

Data jumlah kasus Orang dengan HIV (ODHIV) diakses dari Dinas Kesehatan (Dinkes) Provinsi Jawa Timur tahun 2024, sedangkan seluruh variabel prediktor diakses dari Badan Pusat Statistik (BPS) Provinsi Jawa Timur tahun 2024. Variabel respon pada penelitian ini adalah jumlah kasus ODHIV, sedangkan variabel prediktor meliputi Kepadatan penduduk, Indeks Pembangunan Literasi Masyarakat (IPLM), tingkat kemiskinan, Indeks Pendidikan, Jumlah pernikahan, Rerata lama sekolah (RLS), tingkat pengangguran terbuka (TPT), tingkat kegemaran membaca (TGB), dan peserta keluarga berencana (KB) pasca persalinan dengan kondom.

TAHAPAN PENELITIAN

Adapun tahapan-tahapan penelitian sebagai berikut:

1. Menentukan rumusan masalah, tujuan, dan batasan masalah, hingga studi literatur ruang lingkup penelitian.
2. Pengumpulan data yang bersifat sekunder, mengakses data dari Dinkes dan BPS.
3. Membangun model RP dengan menaksir parameter RP dengan MLE.
4. Melakukan pemeriksaan multikolinearitas pada model RP dengan mencari nilai VIF untuk setiap variabel prediktor.
5. Melakukan uji asumsi equidispersi pada model RP. Apabila model RP memenuhi asumsi, maka dilanjutkan uji signifikansi parameter secara

simultan dan parsial pada model RP dan melakukan interpretasi. Apabila asumsi tidak terpenuhi, maka dilanjutkan dengan membangun model RBN.

6. Membangun model RBN dengan menaksir parameter RBN dengan MLE, dilanjutkan dengan uji signifikansi parameter model RBN.
7. Melakukan interpretasi model RBN

HASIL DAN PEMBAHASAN

ANALISIS DESKRIPTIF

Dalam penelitian yang penulis lakukan, diidentifikasi jumlah kasus ODHIV sebagai variabel respon (y), kemudian variabel prediktor (x) diantaranya kepadatan penduduk (x_1), IPLM (x_2), tingkat kemiskinan (x_3), Indeks Pendidikan (x_4), jumlah pernikahan (x_5), RLS (x_6), TPT (x_7), TGB (x_8), dan peserta KB pasca persalinan dengan kondom (x_9). Berikut adalah ringkasan analisis deskriptif:

Tabel 1. Analisis Deskriptif

Var	Mean	Varians	sd.	Max	Min
y	277,79	55.606	235.81	1.261	34
x_1	65,82	275,89	16,61	100	29
x_2	0,66	0,0064	0,08	0,81	0,52
x_3	6.504	18.950.960	4.353	19.506	803
x_4	6,36	8,88	2,98	9,94	0,03
x_5	0,04	0,0001	0,01	0,06	0,02
x_6	77,14	74,65	8,64	100	64,62
x_7	1.928	5.274.142	2.296	8.698	111
x_8	0,10	0,0016	0,04	0,21	0,03
x_9	408,9	434.768,8	659,37	3.713	7,00

Setelah diperoleh analisis deskriptif, pada variabel respon dapat diketahui bahwa nilai varians 55.606 lebih besar dari nilai rata-ratanya yaitu 277,79 yang diduga mengalami overdispersi.

REGRESI POISSON

Setelah dilakukan eksplorasi data dengan analisis deskriptif, selanjutnya akan dilakukan analisis untuk melihat hubungan antar variabel prediktor terhadap jumlah kasus ODHIV dengan ringkasan analisis Regresi Poisson (RP) sebagai berikut:

Tabel 2. Analisis Regresi Poisson

Prediktor	Estimasi	std. error	p-value
Intercept	3,759	1,902e-1	2e-16
x_1	-4,844e-3	9,538e-4	38e-7
x_2	-0,387	2,657e-1	0,202
x_3	9,5173e-5	2,566e-6	2e-16
x_4	0,1551	5,170e-3	2e-16
x_5	0,1842	1,226	2e-16
x_6	-8,549e-3	1,622e-3	136e-6
x_7	2,49e-4	9,068e-6	2e-16
x_8	0,4028	3,592e-1	0,262
x_9	1,525e-5	1,973e-5	0,440

Berdasarkan ringkasan analisis RP diatas, dapat dibentuk model RP sebagai berikut:

$$\ln(\mu) = 3,759 - 4,844 \times 10^{-3}x_1 + 9,517 \times 10^{-5}x_3 + 0,1551x_4 + 0,1842x_5 - 8,549 \times 10^{-3}x_6 + 2,49 \times 10^{-4}x_7 \quad (15)$$

Ringkasan tabel (2) tersebut menunjukkan hanya terdapat 3 variabel yang tidak signifikan secara parsial pada p -value 0,05, diantaranya x_2 , x_8 , dan x_9 . Selain daripada ketiga variabel tersebut, tolak H_0 yang berarti signifikan terhadap variabel respon.

PEMERIKSAAN MULTIKOLINEARITAS

Sebelum dilakukan uji asumsi equidispersi model RP, perlu dilakukan pemeriksaan multikolinearitas terhadap model RP untuk mengetahui apakah terdapat hubungan yang kuat antar variabel prediktor, dengan ringkasan sebagai berikut:

Tabel 3. Pemeriksaan Multikolinearitas

Prediktor	VIF
x_1	2,773345
x_2	4,156217
x_3	1,942119
x_4	1,783547
x_5	1,815131
x_6	1,945213
x_7	4,098107
x_8	1,532502
x_9	1,322526

Hasil pemeriksaan multikolinearitas yang ditunjukkan pada tabel diatas menginformasikan bahwa seluruh variabel prediktor terbebas dari masalah multikolinearitas.

PENGUJIAN ASUMSI EQUIDISPERSI

Model RP membawa asumsi dasar bahwa varians dan *mean* haruslah sama. Namun, pada suatu fenomena, asumsi ini tidak selalu dapat dipenuhi sehingga perlu dilakukan pengujian asumsi agar dapat diketahui keandalan hasil uji signifikansi parameter model RP. Pengujian asumsi equidispersi diringkas dalam tabel dibawah ini:

Tabel 4. Uji Asumsi Equidispersi

Model	df	D	$\hat{\phi}_D$
RP	32	4,9493	136,50625

Tabel diatas membuktikan dugaan kuat overdispersi pada model RP (15). Berdasarkan pengujian asumsi ini, dapat disimpulkan model RP mengalami overdispersi, sehingga perlu model alternatif Regresi Binomial Negatif.

REGRESI BINOMIAL NEGATIF

Regresi Binomial Negatif (RBN) pada penelitian ini adalah sebagai alternatif dari model RP yang terkonfirmasi mengalami overdispersi yang ditunjukkan pada Tabel (4) diatas. Estimasi parameter model RBN dilakukan menggunakan MLE, dengan hasil yang disajikan pada tabel berikut:

Tabel 5. Analisis Regresi Binomial Negatif

Prediktor	Estimasi	std. error	p-value
Intercept	2,044	1,302	0,11652
x_1	-7,442e-3	0,007	0,28860
x_2	1,367	1,825	0,45385
x_3	1,205e-4	2,234e-5	6,93e-8
x_4	0,1274	0,03153	5,33e-5
x_5	21,91	8,068	0,00662
x_6	4,116e-3	0,01133	0,71643
x_7	1,594e-4	6,22e-5	0,01038
x_8	-1,248	2,078	0,54813
x_9	-3,099e-5	1,228e-4	0,80073

Tabel (5) hanya menunjukkan 4 variabel prediktor yang signifikan. Variabel-variabel yang signifikan tersebut kemudian digunakan dalam pembentukan model RBN dengan model sebagai berikut:

$$\ln(\mu) = 1,205 \times 10^{-4}x_3 + 0,1274x_4 + 21,91x_5 + 1,594 \times 10^{-4}x_7 \quad (16)$$

Model diatas dapat digunakan sebagai model untuk melakukan intervensi terhadap jumlah kasus ODHIV.

UJI SIMULTAN

Setelah diperoleh hasil penaksiran parameter pada model RBN, akan dilakukukan uji signifikansi parameter secara simultan dengan Uji G atau *Likelihood Ratio Test* (LRT) dengan ringkasan uji G sebagai berikut:

Tabel 6. Uji Simultan

	df	G-Statistik	p-value
Model penuh	9	46,23	5,453e-7

Berdasarkan hasil uji G diatas, diperoleh *p-value* lebih kecil dari $\alpha = 0.05$. Nilai *p* yang sangat kecil tersebut menunjukkan bahwa pada uji ini tolak H_0 yang berarti setidaknya satu variabel prediktor berpengaruh signifikan terhadap variabel respon secara simultan.

UJI PARSIAL

Mengacu dari Tabel (5), ditunjukkan bahwa hanya terdapat 4 variabel prediktor yang berpengaruh signifikan terhadap jumlah ODHIV. Variabel tersebut antara lain Jumlah pernikahan (x_3), RLS (x_4), TPT (x_5), dan Kepadatan penduduk (x_7). Karena seluruh *p-value* lebih kecil dari taraf signifikansi $\alpha = 0,05$, maka dapat disimpulkan bahwa secara parsial keempat variabel tersebut tolak H_0 yang berarti memiliki pengaruh yang signifikan terhadap jumlah ODHIV, sementara variabel lainnya tidak menunjukkan pengaruh yang signifikan.

Namun, secara teoritis IPLM (x_1) dan tingkat kegembiraan membaca (x_6) diharapkan berkorelasi negatif dengan kasus HIV melalui peningkatan pengetahuan pencegahan dan reduksi stigma (Amry dkk, 2025). Data empiris di Jawa Timur menunjukkan pengaruh literasi sering kali tidak signifikan atau lemah karena kalah dominan oleh faktor demografis dan perilaku berisiko langsung seperti seks bebas dan narkoba suntik (Solihah dkk, 2023). *Health Belief Model* (HBM) mendukung literasi sebagai penguat persepsi kerentanan, kenyataan di lapangan menunjukkan literasi umum tidak selalu linier dengan literasi kesehatan spesifik HIV (Eny dkk, 2024).

Serupa dengan literasi, Indeks Pendidikan (x_2) diasumsikan berbanding terbalik dengan transmisi HIV karena kapasitas kognitif yang lebih baik (Wardhani dkk, 2023). Akan tetapi, fenomena di wilayah urban seperti Surabaya menunjukkan lonjakan kasus tetap terjadi di kalangan penduduk

terdidik akibat resistensi terhadap *Voluntary Counseling and Testing* (VCT) (Martini, 2024). Ketidaksignifikanan x_2 utama menunjukkan bahwa intervensi harus bergeser dari pendidikan umum ke edukasi perilaku spesifik (Viontika & Yasmin, 2024).

Tingkat kemiskinan (x_3) yang semula diprediksi meningkatkan risiko melalui keterbatasan akses kesehatan, ternyata tidak menjadi faktor determinan tunggal. Data di Jawa Timur menunjukkan kasus HIV justru terkonsentrasi di wilayah urban dengan angka kemiskinan rendah, yang mengindikasikan bahwa perilaku berisiko dan stigma berperan daripada faktor struktural kemiskinan (Lusiana, 2025).

Variabel peserta KB pasca persalinan dengan kondom (x_5) terbukti tidak signifikan dalam model RBN. Meski penggunaan efektif secara medis, cakupan program KB hanya menasar populasi monogami berisiko rendah (Dinas Kesehatan, 2025). Transmisi HIV di Jawa Timur diduga didominasi oleh kelompok heteroseksual berisiko tinggi dan homoseksual tersembunyi dalam pernikahan yang disebut *Mixed Orientation Marriage* atau MOM, di mana laki-laki yang masih menyimpan gairah terhadap sesama jenis dan melakukan kegiatan berisiko terhadap sesama jenis. Penggunaan pengaman juga sering kali diabaikan karena stigma atau ketidaktahuan pasangan (Sekarjati dkk, 2025). Hal ini menegaskan perlunya perluasan skema intervensi yang melampaui program KB konvensional dengan kondom (Adelliani dkk, 2023).

INTERPRETASI

Interpretasi hanya dilakukan untuk model RBN yang ditunjukkan pada persamaan (16) karena model RP telah terbukti mengalami overdispersi.

1. Jumlah pernikahan

Variabel jumlah pernikahan memiliki koefisien sebesar 0,0001205. Nilai koefisien $e^{0,0001205} = 1,00012$ bermakna bahwa akan adanya peningkatan 1,00012 kali kasus HIV apabila terjadinya satu pernikahan, dengan asumsi variabel lain konstan. Hal ini relevan dengan penelitian Widia & Rayasari (2017) yang menginformasikan sebesar 25% kasus HIV terjadi akibat penularan dari pasangan (suami atau istri), yang diduga dipengaruhi oleh rendahnya pemahaman terkait upaya pencegahan penularan HIV. Ini terjadi karena

pasangan serodiskordan memiliki sikap pasrah menghadapi risiko infeksi, demi keinginan memiliki anak dan keinginan merasakan seks tanpa pembatas.

2. Rerata lama sekolah

Variabel Rerata Lama Sekolah (RLS) memiliki koefisien sebesar 0,1274. Nilai $e^{0,1274} = 1,136$ mengindikasikan bahwa peningkatan satu tahun RLS, meningkatkan 1,136 kali ODHIV, dengan asumsi variabel lain konstan. Hasil empiris ini sejalan dengan teori bahwa orang yang memiliki durasi sekolah yang lebih lama juga memiliki tingkat kesadaran HIV yang lebih baik, sehingga pada suatu wilayah dengan durasi sekolah yang lebih tinggi akan terdapat aktivitas pencatatan kasus HIV yang lebih aktif (Sabo dkk, 2024).

3. Tingkat pengangguran terbuka

Variabel Tingkat Pengangguran Terbuka (TPT) menghasilkan koefisien bernilai 21,91. Mengingat variabel ini diekspresikan dalam bentuk proporsi dengan interval nilai 0 sampai dengan 1, maka kenaikan sebesar 1 satuan merepresentasikan perubahan sebesar 100%. Maka dari itu, interpretasi koefisien dilakukan pada perubahan yang lebih kecil, yaitu sebesar 1% atau 0,01 satuan, sehingga diperoleh nilai $e^{21,91 \times 0,01} = 1,245$ dengan makna, apabila TPT meningkat 1%, meningkatkan jumlah ODHIV 1,245 kali, dalam asumsi variabel lain konstan.

Hasil empiris ini relevan dengan bahasan teoritis sebelumnya, yang menyatakan bahwa pengangguran dapat meningkatkan kerentanan terhadap berbagai masalah kesehatan, termasuk penyakit menular, melalui mekanisme tekanan ekonomi, stres psikososial, serta potensi keterlibatan dalam perilaku berisiko akibat keterbatasan sumber daya (Kasim & Hendra, 2023).

4. Kepadatan penduduk

Variabel Kepadatan Penduduk menghasilkan koefisien dengan nilai 0,0001594. Nilai $e^{0,0001594} = 1,00016$ menunjukkan kenaikan jumlah ODHIV sebesar 1,00016 kali setiap kepadatan penduduk per km² bertambah satu jiwa, dimana variabel lain dalam asumsi konstan. Hasil ini mendukung kajian teoritis sebelumnya, apabila terdapat kepadatan penduduk yang tinggi, maka intensitas interaksi sosial dan

mobilitas penduduk akan meningkat, yang pada akhirnya memperbesar risiko penularan penyakit menular, termasuk HIV (Taniwan dkk., 2024).

PENUTUP

KESIMPULAN

Model RBN yang digunakan dalam penelitian ini dibangun melalui tahapan awal dengan membentuk model RP. Namun, berdasarkan hasil pengujian overdispersi, ditemukan bahwa model RP tidak memenuhi asumsi kesamaan antara nilai *mean* dan *varians*, sehingga terjadi overdispersi pada Tabel (5). Oleh karena itu, digunakan model RBN sebagai alternatif yang lebih sesuai, karena mampu mengakomodasi adanya overdispersi dalam data. Model RBN yang dihasilkan adalah:

$$\ln(\mu) = 1,205 \times 10^{-4}x_3 + 0,1274x_4 + 21,91x_5 + 1,594 \times 10^{-4}x_7 \quad (17)$$

Berdasarkan uji signifikansi parameter simultan yang ditunjukkan Tabel 4.8, dihasilkan statistik uji *G* bernilai 46,23 dengan derajat kebebasan 9 dan *p-value* bernilai $5,45 \times 10^{-7}$. Itu berarti setidaknya satu prediktor signifikan terhadap variabel respon secara simultan. Sedangkan faktor yang signifikan secara parsial hanya Jumlah pernikahan (x_3), RLS (x_4), TPT x_5 , dan Kepadatan penduduk (x_7).

SARAN

Mengacu dari hasil empiris yang telah didapatkan, selanjutnya disarankan untuk mempertimbangkan penggunaan variabel yang lebih merefleksikan karakteristik, kecenderungan praktik perilaku berisiko, orientasi seksual, serta pengguna napza suntik, dibandingkan variabel faktor eksternal. Karena itu, diharapkan gambaran akan lebih spesifik terhadap faktor-faktor yang memengaruhi jumlah ODHIV. Penelitian selanjutnya juga dapat mempertimbangkan variabel tambahan yang berkaitan dengan praktik prostitusi, yang berpotensi meningkatkan risiko penyebaran HIV.

Penelitian selanjutnya juga disarankan untuk melakukan eksplorasi model yang lebih beragam serta membandingkan kinerja model RBN dengan model alternatif lainnya, seperti *Generalized Poisson Regression* yang lebih sensitif terhadap variasi data, serta Conway-Maxwell-Poisson fleksibel dalam menangani distribusi dengan karakteristik ekstrem

(Rahayu, 2020). Hal ini bertujuan untuk memperoleh model yang paling sesuai dalam memodelkan dengan mempertimbangkan skor *Akaike Information Criterion* (AIC) untuk mengonfirmasi model yang lebih baik.

DAFTAR PUSTAKA

- Adelliani, N., Najmah, Ariesandi, Z., Aprina, F., Setiawan, Y., Misnaniarti, Dewi, Y. K., Vanessa, A. S., & Andajani, S. (2023). Non-Disclosure of Sexual Orientation of Men Having Sex with Men (MSM): a Hidden Threat of HIV Transission in Palembang, Indonesia. *JIKM: Jurnal Ilmu Kesehatan Masyarakat*, 14(2). <https://doi.org/https://doi.org/10.26553/jikm.2023.14.2.147-161>
- Amry, R. Y., Maria, D. Y., & Rahayu, B. A. (2025). Hubungan Health Literacy dengan Pengetahuan Pencegahan HIV/AIDS pada Remaja. *Jurnal Pembaruan Kesehatan Indonesia*, 1(1), 40-48.
- Ariyanto, D., Sofro, A., Hanifah, A. N., Prihanto, J. B., Maulana, D. A., & Romadhonia, R. W. (2024). Logistic And Probit Regression Modeling to Predict The Opportunities of Diabetes ini Prospective Athletes. *BAKERENG: Journal of Mathematics and Its Applications*, 18(3), 1391-1402. <https://doi.org/https://doi.org/10.30598/barekengvol18iss3pp1391-1402>
- Armadi Mori, S., Martha, S., & Aprizkiyandari, S. (2024). Pemodelan Tingkat Pengangguran Terbuka Menggunakan Generalized Linear Model. *Buletin Ilmiah Math. Stat. Dan Terapanya (Bimaster)*, 13(1), 137-146.
- Astuti. AW, W., & Rayasari, F. (2017). Pengalaman Seksual Pasangan Menderita HIV dalam Mempertahankan Status HIV Negatif di RSPI Prof. Dr. Sulianti Saroso. *Indonesian Journal of Nursing Practices*, 1(2), 32-43. <https://doi.org/10.18196/ijnp.1259>
- Azizah, R. N., Nisak, U. K., & Indahyanti, U. (2025). Analisis Jumlah Prediksi Penyebaran HIV / AIDS di Kabupaten Sidoarjo menggunakan Metode Multiple Linier Regression. *Physical Sciences, Life Science and Engineering*, 1(1), 1-11.
- Cameron, A. C., & Trivedi, P. K. (2013). Regression Analysis for Count Data. In *Econometric Society Monographs* (Issue 30).
- Cust, H., Jones, H., Powell-Jackson, T., Lepine, A., & Rosabla, R. (2021). Economic Shocks and Risky Sexual Behaviours A Systematic Review of the Literature. *Journal of Development Effectiveness*, 13, 166-203. <https://doi.org/10.1080/19439342.2021.19287>

- Dhahari, N. M., & Sofro, A. (2024). Analisis Regresi Binomial Negatif Untuk Pemodelan Angka Positif Penyakit Kusta di Jawa Timur. *Jurnal Ilmiah Matematika*, 12(01), 49–56.
- Dinas Kesehatan. (2025). *Profil Kesehatan Provinsi Jawa Timur 2024*.
- Effiyaldi, Pasaribu, J. P. K., Suratno, E., Kadar, M., Gunardi, Naibaho, R., Hati, S. K., & Aryati, V. (2022). Penerapan Uji Multikolinearitas Penelitian Manajemen Sumber Daya Manusia. *Jumanage: Jurnal Ilmiah Manajemen Dan Kewirausahaan*, 1(2), 94–102.
- Eny, Dwiyantri, D., & Kurnayanti. (2024). Pengaruh Literasi Kesehatan Menggunakan Ular Tangga tentang HIV/AIDS Terhadap Perilaku Berisiko Remaja. *Miracle Journal of Public Health (MJPH)*, 7(2).
<https://doi.org/10.36566/mjph/Vol7.Iss2/382>
- Etowa, J., Ghose, B., Loemba, H., Etowa, E. B., Husbands, W., Omorodion, F., Luginaah, I., & Wong, J. P. (2021). Factors Associated with Condom Knowledge, Attitude, and Use among Black Heterosexual Men in Ontario, Canada. *The Scientific World Journal*, 2021.
- Fadil, M., Raupong, & Ilyas, N. (2024). Mengatasi Overdispersi Menggunakan Regresi Binomial Negatif dengan Penaksir Maksimum Likelihood pada Kasus Demam Berdarah di Kota Makassar. *Journal of Statistics and Its Application*, 5(1), 37–51.
- Fatmala, C. T., Hayati, M., Sari, R. P., & Hudori, M. (2024). Pemodelan Jumlah Kasus HIV / AIDS di Provinsi Lampung Menggunakan Regresi Binomial Negatif. *Journal of Mathematics: Theory and Applications*, 6(2), 168–177.
<https://doi.org/10.31605/jomta.v6i2.4069>
- Febriani, R. E., & Nugraha, J. (2025). Analisis Regresi Poisson untuk Mengetahui Variabel Berpengaruh Pada Kasus Gizi Buruk di Kabupaten Bangka. *Seminar Nasional Matematika Dan Pendidikan Matematika UNY*, 85–94.
- Hilbe, J. (2011). Negative Binomial Regression. In *Cambridge University Press*. Cambridge University Press.
- Jusmansyah, M. (2020). Analisis pengaruh current ratio, debt to equity ratio, total asset turn over, dan return on equity terhadap harga saham. *Jurnal Ekonomika Dan Manajemen*, 9(2), 178–198.
- Kasim, F., & Hendra. (2023). Pengaruh Pengangguran dan Kemiskinan Terhadap Tindakan Kriminal di Kabupaten Tolitoli Periode 2021. *Jurnal Ilmu Ekonomi (JIE)*, 5(2), 247–261.
- Lusiana, E. F. (2025). *Kabupaten dan Kota di Jawa Timur dengan Kasus Baru HIV/AIDS Tertinggi*. DetikJatim.
- Martini, S. (2024). *Tinjauan Sistematis dan Meta-Analisis Perilaku Risiko Penularan HIV, Variasi Genetik, dan Resistensi Antiretroviral (ARV) pada Populasi LGBT*. Universitas Airlangga.
<https://unair.ac.id/tinjauan-sistematis-dan-meta-analisis-perilaku-risiko-penularan-hiv-variasi-genetik-dan-resistensi-antiretroviral-arv-pada-populasi-lgbt/>
- Mustafidah, M. E., & Sofro, A. (2024). Pemodelan Jumlah Kematian Akibat Filariasis di Jawa Timur dengan Menggunakan Regresi Hurdle Negative Binomial. *Jurnal Ilmiah Matematika*, 12(2), 198–199.
<https://ejournal.unesa.ac.id/index.php/mathunesa/article/view/64649>
- Rahayu, A. (2020). Model-Model Regresi untuk Mengatasi Masalah Overdispersi Pada Regresi Poisson. *Journal Pegguruang: Conference Series*, 2(May), 1–5.
- Rizal, F. (2025). *Tahun 2025: Pengidap HIV di Indonesia Tembus 564 Ribu*. Warta Pontianak.
<https://wartapontianak.pikiran-rakyat.com/nasional/pr-1179438390/tahun-2025-pengidap-hiv-di-indonesia-tembus-564-ribu/>
- Sabo, K. G., Seifu, B. L., Kase, B. F., Asebe, H. A., Asmare, Z. A., Asgedom, Y. S., Shibeshi, A. H., Tebeje, T. M., Lombebo, A. A., Fente, B. M., & Mare, K. U. (2024). Factors influencing HIV testing uptake in Sub-Saharan Africa: a comprehensive multi-level analysis using demographic and health survey data (2015–2022). *BMC Infectious Diseases*, 24(1), 1–12.
<https://doi.org/10.1186/s12879-024-09695-1>
- Santi, V. M., & Rahayuningsih, Y. (2024). Negative Binomial Regression in Overcoming Overdispersion Poverty Data in Extreme Poverty Data in Indonesia. *Pattimura International Journal of Mathematics (PIJMath)*, 4(1), 1–10.
<https://doi.org/10.57059/formasi.v4i1.67>
- Sauddin, A., Auliah, N. I., & Alwi, W. (2020). Pemodelan Jumlah Kematian Ibu di Provinsi Sulawesi Selatan Menggunakan Regresi Binomial Negatif. *Jurnal Matematika Dan Statistika Serta Aplikasinya*, 8(2).
- Sekarjati, A., Widiarini, R., & Bachrun, E. (2025). Experiences of Sexuality of Bisexual HIV/AIDS Patients (a Phenomenological Study in Madiun Regency). *Asian Journal Healthy and Science*, 4(9), 332–344.
- Soliha, N. L., Lestari, D., & Widyaningsih, Y. (2023). Analisis Faktor-Faktor yang Menjelaskan

- Kasus AIDS Provinsi Jawa Timur Menggunakan Model Geographically Weighted Logistic Regression (GWLR). *Jurnal Statistika Dan Aplikasinya*, 7(1), 37–48. <https://journal.unj.ac.id/unj/index.php/statistika/article/download/33353/15356>
- Sundari, I. (2023). Regresi Poisson dan Penerapannya Untuk Memodelkan Hubungan Usia dan Perilaku Merokok Terhadap Jumlah Kematian Penderita Penyakit Kanker Paru-Paru. *Jurnal Matematika UNAND*, 1(1).
- Taniwan, P., Bilbina, A. F., Ganap, C. R. S., & Faidah, D. Y. (2024). Pemodelan Faktor-Faktor yang Mempengaruhi Kejadian HIV AIDS di Provinsi Jawa Barat Pemodelan Faktor-Faktor yang Memengaruhi Kejadian HIV/AIDS di Provinsi Jawa Barat. *Indo-MathEdu Intellectuals Journal*, 5(3), 3298–3308. <https://doi.org/10.54373/imej.v5i3.1278>
- Tohari, A., Chamidah, N., Fatmawati, & Lestari, B. (2021). Modelling the number of HIV and AIDS Cases in East Java Using Biresponse Multipredictor Negative Binomial Regression Based On Local Linear Estimator. *Communications in Mathematical Biology and Neuroscience*, 2021. <https://doi.org/https://doi.org/10.28919/cm bn/5652>
- Viontika, V., & Yasmin, I. F. (2024). Korelasi Tingkat Pendidikan dan Nilai Akademik dengan Pengetahuan dan Stigma HIV/AIDS pada Mahasiswa/i Fakultas Kedokteran Universitas YARSI Angkatan 2022 dan Siswa/i MAN 13 Jakarta. *Journal Medical Journal*, 2(9), 1123–1132. <https://doi.org/https://doi.org/10.33476/jmj.v2i9.4169>
- Wardhani, A. C., Haryanto, S., Fahmi, T., & Saud, M. (2023). HIV/AIDS-related knowledge among Indonesian women: the role of media exposure and socio-demographic factors. *International Journal of HIV-Related Problems*, 22(1), 43–52. <https://doi.org/https://doi.org/10.5114/hiv ar.2023.124548>
- Winata, H. M. (2023). Mengatasi Overdispersi Dengan Regresi Binomial Negatif Pada Angka Kematian Ibu di Kota Bandung. *Jurnal Gaussian*, 11(4), 616–622. <https://doi.org/10.14710/j.gauss.11.4.616-622>