

# Model Regresi Multilevel Negative Binomial Pada Kasus Kronis Filariasis di Indonesia

Rizal Usman, Salmun K. Nasib, Djihad Wungguli, and Siti Nurmardia Abdussamad



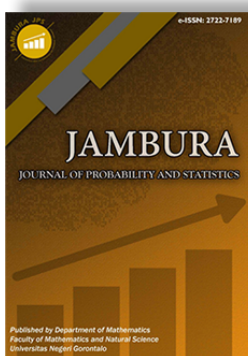
Volume 6, Issue 2, Pages 78–84, November 2025

Received 15 May 2025, Revised 17 May 2025, Accepted 31 July 2025, Published Online 30 November 2025

To Cite this Article : R. Usman, S. K. Nasib, D. Wungguli, and S. N. Abdussamad, " Model Regresi Multilevel Negative Binomial Pada Kasus Kronis Filariasis di Indonesia ", *Jambura J. Probab. Stat.*, vol. 6, no. 2, pp. 78–84, 2025, <https://doi.org/10.34312/jjps.v6i2.31648>

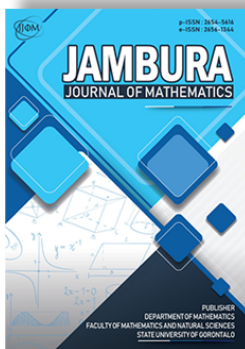
© 2025 by author(s)

## JOURNAL INFO • JAMBURA JOURNAL OF PROBABILITY AND STATISTICS



	Homepage	: <a href="https://ejournal.ung.ac.id/index.php/jps/index">https://ejournal.ung.ac.id/index.php/jps/index</a>
	Journal Abbreviation	: Jambura J. Probab. Stat.
	Frequency	: Biannual (May and November)
	Publication Language	: English (preferable), Indonesia
	DOI	: <a href="https://doi.org/10.34312/jjps">https://doi.org/10.34312/jjps</a>
	Online ISSN	: 2722-7189
	Editor-in-Chief	: Ismail Djakaria
	Publisher	: Department of Mathematics, Universitas Negeri Gorontalo
	Country	: Indonesia
	OAI Address	: <a href="http://ejournal.ung.ac.id/index.php/jps/oai">http://ejournal.ung.ac.id/index.php/jps/oai</a>
	Google Scholar ID	: kWdujzMAAAJ
	Email	: <a href="mailto:redaksi.jjps@ung.ac.id">redaksi.jjps@ung.ac.id</a>

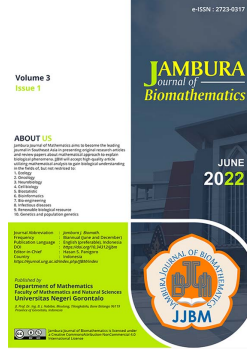
## JAMBURA JOURNAL • FIND OUR OTHER JOURNALS



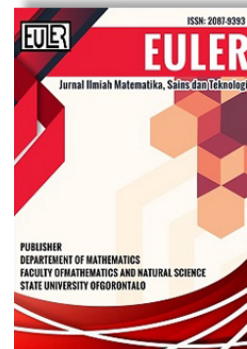
Jambura Journal of Mathematics



Jambura Journal of Mathematics Education



Jambura Journal of Biomathematics



EULER : Jurnal Ilmiah Matematika, Sains, dan Teknologi



# Model Regresi Multilevel Negative Binomial Pada Kasus Kronis Filariasis di Indonesia

Rizal Usman<sup>1</sup>, Salmun K. Nasib<sup>1\*</sup>, Djihad Wungguli<sup>2</sup>, Siti Nurmardia Abdussamad<sup>1</sup>

<sup>1</sup>Program Studi Statistika, Fakultas MIPA, Universitas Negeri Gorontalo

<sup>2</sup>Program Studi Matematika, Fakultas MIPA, Universitas Negeri Gorontalo

## ARTICLE HISTORY

Received 15 May 2025

Revised 17 May 2025

Accepted 31 July 2025

Published 30 November 2025

## KATA KUNCI

Filariasis  
Multilevel Negative Binomial  
Regression  
overdispersi.

## KEYWORDS

Filariasis  
Multilevel Negative Binomial  
Regression  
overdispersi.

**ABSTRAK.** Filariasis merupakan penyakit menular yang disebabkan oleh infeksi cacing parasit *Filaria* dan ditularkan melalui gigitan nyamuk yang terinfeksi. Analisis jumlah kasus kronis filariasis di Indonesia sering menghadapi permasalahan statistik berupa overdispersi dan excess zero. Untuk mengatasi hal tersebut, digunakan model Multilevel Negative Binomial Regression yang mampu menangani varians data yang lebih besar dari rata-rata serta banyaknya nilai nol dalam data. Hasil penelitian menunjukkan bahwa model tersebut efektif dalam mengatasi permasalahan overdispersi dan excess zero. Berdasarkan uji signifikansi parameter menggunakan uji Wald, variabel lingkungan seperti keberadaan sumur tak terlindung ( $X_4$ ) dan kedekatan rumah tangga dengan penampungan limbah ( $X_5$ ) berpengaruh signifikan terhadap jumlah kasus kronis filariasis. Sebaliknya, variabel sosial ekonomi seperti persentase penduduk laki-laki ( $X_1$ ), penduduk usia produktif ( $X_2$ ), sanitasi layak ( $X_3$ ), persentase penduduk miskin ( $X_6$ ), dan Indeks Pembangunan Manusia ( $X_7$ ) tidak menunjukkan pengaruh signifikan. Temuan ini menegaskan bahwa faktor lingkungan memiliki peran penting dalam penyebaran kasus kronis filariasis di Indonesia.

**ABSTRACT.** Filariasis is a contagious disease caused by infection with the parasitic worm *Filaria* and transmitted through the bite of an infected mosquito. Analysis of the number of chronic filariasis cases in Indonesia often faces statistical problems in the form of overdispersion and excess zero. To overcome this, a Multilevel Negative Binomial Regression model is used which is able to handle data variance that is greater than the average as well as the number of zero values in the data. The results showed that the model was effective in overcoming overdispersion and excess zero problems. Based on the parameter significance test using the Wald test, environmental variables such as the presence of unprotected wells ( $X_4$ ) and household proximity to waste storage ( $X_5$ ) have a significant effect on the number of chronic filariasis cases. In contrast, socioeconomic variables such as percentage of male population ( $X_1$ ), productive age population ( $X_2$ ), proper sanitation ( $X_3$ ), percentage of poor population ( $X_6$ ), and Human Development Index ( $X_7$ ) did not show a significant effect. These findings confirm that environmental factors play an important role in the spread of chronic filariasis cases in Indonesia.



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 JJPS: Department of Statistics, Universitas Negeri Gorontalo, Jln. Prof. Dr. Ing. B. J. Habibie, Bone Bolango 96554, Indonesia.

## 1. Pendahuluan

Filariasis merupakan penyakit menular yang disebabkan oleh infeksi cacing parasit yang ditularkan melalui gigitan nyamuk [1]. Pada tahun 2020, *World Health Organization* (WHO) membuat sebuah program *Global Programme to Eliminate Lymphatic Filariasis* (GPELF) untuk mengeliminasi filariasis limfatik sebagai masalah kesehatan masyarakat yang signifikan. Namun, hal ini belum sepenuhnya tercapai mengurangi penyebaran penyakit ini secara global. Lebih lanjut, Kasus filariasis di Indonesia tetap menjadi masalah kesehatan masyarakat yang serius, terutama di wilayah dengan sanitasi buruk dan kepadatan penduduk tinggi [2]. Meskipun upaya pengendalian terus dilakukan, berdasarkan publikasi dinas kesehatan pada tahun 2022 tercatat 8.742 kasus kronis filariasis yang tersebar di 34 provinsi di Indonesia, menandakan tantangan besar dalam upaya eradikasi penyakit ini.

Analisis kasus kronis filariasis seringkali dihadapkan pada permasalahan data, yaitu adanya overdispersi dan excess zero pada jumlah kasus filariasis. Hal ini disebabkan oleh karakteristik data filariasis kronis di tingkat kabupaten/kota, di mana sebagian besar kasus memiliki nilai nol.

Menurut [3] excess zero merujuk pada fenomena di mana terdapat jumlah pengamatan dengan nilai nol yang jauh lebih tinggi daripada yang diharapkan berdasarkan model distribusi tertentu, seperti distribusi poisson, yang memicu terjadinya overdispersion. Kondisi ini perlu diperhatikan karena dapat menyebabkan bias dalam estimasi parameter dan inferensi yang tidak akurat [4].

Overdispersi terjadi ketika varians data lebih besar daripada rata-rata, sehingga bisa menyebabkan bias pada estimasi jika tidak ditangani dengan tepat [5]. Model regresi Poisson yang biasanya digunakan untuk memodelkan data cacah memiliki keterbatasan dalam menangani dua permasalahan ini, sehingga sering kali menghasilkan estimasi yang kurang akurat [6].

\*Corresponding Author.

Untuk mengatasi hal ini, model regresi *negative binomial* sering digunakan karena mampu mengakomodasi banyaknya nilai nol dalam data dan data dengan varians yang lebih besar dari rata-rata (*overdispersi*) [7]. Namun, dalam kasus data filariasis yang bersifat hierarki, model *multilevel negative binomial* menjadi pilihan yang lebih tepat karena mampu mempertimbangkan perbedaan karakteristik antar wilayah.

Berbagai studi sebelumnya telah mengkaji penyebaran kasus filariasis di Indonesia dengan pendekatan metodologis yang beragam. Secara umum, pendekatan tersebut dapat diklasifikasikan ke dalam beberapa kategori utama, yakni pendekatan spasial, studi berbasis wilayah tertentu, dan model regresi multilevel. Pendekatan spasial, seperti yang digunakan oleh [8] dengan *Geographically Weighted Negative Binomial Regression* (GWNBR), mampu menangkap variasi geografis antar wilayah dalam menganalisis kasus kronis filariasis. Namun demikian, model ini belum mempertimbangkan struktur hierarkis dalam data, di mana unit pengamatan (kabupaten/kota) dikelompokkan dalam wilayah administratif yang lebih tinggi (provinsi).

Sementara itu, studi lain seperti yang dilakukan oleh [2] berfokus pada faktor-faktor lingkungan dan sosial ekonomi yang memengaruhi penyebaran kasus filariasis, namun terbatas pada wilayah administratif tertentu, yakni Kabupaten Asahan. Keterbatasan cakupan wilayah dalam studi tersebut menyulitkan generalisasi hasil temuan ke tingkat nasional dan mengabaikan kemungkinan adanya variasi antar wilayah yang lebih luas.

Pendekatan regresi multilevel telah digunakan dalam beberapa studi untuk menganalisis data yang memiliki struktur hierarkis. Seperti penelitian [9] dan [10] menggunakan regresi logistik dan regresi multilevel untuk mengevaluasi pengaruh karakteristik individu, rumah tangga, dan wilayah terhadap status kemiskinan atau capaian pendidikan. Namun, sebagian besar studi ini tidak memperhatikan karakteristik distribusi data yang khas dalam konteks epidemiologi penyakit, seperti *overdispersion* dan *excess zero*. Meskipun studi oleh [11] telah menerapkan regresi multilevel negatif binomial, fokus penelitian tersebut tidak secara eksplisit ditujukan pada penanganan masalah *overdispersion* dan *excess zero*. Selain itu, sebagian penelitian multilevel lainnya seperti penelitian [12], tidak secara khusus menyoroti kasus filariasis sebagai objek kajian. Berdasarkan tinjauan tersebut, dapat disimpulkan bahwa masih terdapat keterbatasan dalam studi-studi sebelumnya, terutama dalam mempertimbangkan struktur hierarkis data serta karakteristik *overdispersion* dan *excess zero*. Kondisi ini menegaskan perlunya pengembangan model analisis yang lebih sesuai untuk menggambarkan kompleksitas data jumlah kasus filariasis kronis di Indonesia secara lebih akurat. Oleh karena itu, model regresi multilevel negatif binomial dianggap sebagai pendekatan yang lebih tepat dalam menganalisis kasus kronis filariasis. Model ini tidak hanya mengatasi masalah *overdispersion* dan *excess zero*, tetapi juga mempertimbangkan variasi antar kelompok dalam data.

Dengan demikian, penelitian ini bertujuan untuk mengembangkan model regresi multilevel negatif binomial guna mengatasi masalah *overdispersion* dan *excess zero* pada data jumlah kasus filariasis kronis di Indonesia, khususnya pada tingkat kabupaten/kota. Model ini diharapkan dapat memberikan informasi yang lebih akurat mengenai faktor-faktor yang signifikan dalam penyebaran filariasis, sehingga dapat mendukung upaya pencegahan dan pengendalian penyakit tersebut di Indonesia.

gahan dan pengendalian penyakit tersebut di Indonesia.

## 2. Metode Penelitian

Penelitian ini dilakukan pada bulan September hingga Desember 2024. Penelitian ini menggunakan data sekunder yang diperoleh dari jumlah kasus kronis filariasis tahun 2022 dalam publikasi profil Kementerian Kesehatan di Indonesia pada laman kemkes.go.id, publikasi Badan Pusat Statistik berjudul Publikasi Statistik Kesejahteraan Rakyat Kabupaten/Kota di Indonesia, dan publikasi Provinsi dalam Angka tahun 2023 pada laman bps.go.id. Populasi yang digunakan dalam penelitian ini adalah jumlah kasus kronis filariasis tahun 2022. Teknik penarikan sampel yang digunakan adalah teknik sampling jenuh, sehingga sampel yang digunakan pada penelitian ini berjumlah 514 Kabupaten/kota di Indonesia. Analisis data pada penelitian ini menggunakan perangkat lunak R Studio dengan memanfaatkan package *gmmTMB*. Adapun variabel yang digunakan dalam penelitian ini dapat dilihat pada tabel 1.

Tabel 1. Variabel Penelitian

Simbol	Variabel	Skala
Y	Jumlah Kronis Filariasis	Rasio
X1	Persentase Penduduk Laki-Laki	Rasio
X2	Persentase Penduduk Umur 15–64 Tahun	Rasio
X3	Persentase Sanitasi Layak	Rasio
X4	Persentase Sumur Tak Terlindung	Rasio
X5	Persentase Rumah Tangga Dekat Limbah	Rasio
X6	Jumlah Penduduk Miskin	Rasio
X7	Indeks Pembangunan Manusia	Rasio

Adapun langkah-langkah yang dilakukan dalam penelitian ini sebagai berikut:

1. Mendeteksi multikolinearitas menggunakan *Variance Inflation Factor* (VIF). Jika nilai VIF yang diperoleh  $> 10$ , maka mengindikasikan adanya multikolinearitas. Sebaliknya, jika nilai VIF  $< 10$ , maka data tidak menunjukkan adanya multikolinearitas.
2. Membentuk model regresi Poisson.
3. Mendeteksi *overdispersi*, di mana jika nilai *dispersion ratio* lebih besar dari 1 maka terjadi *overdispersi*, sebaliknya jika lebih kecil dari 1 maka tidak terjadi *overdispersi*.
4. Mendeteksi *excess zero*, di mana jika nilai 0 lebih dari 50%, maka terindikasi *excess zero*.
5. Membentuk model regresi *Negative Binomial*.
6. Membentuk model regresi *Multilevel Negative Binomial*.
7. Melakukan pengujian *Likelihood Ratio Test*.
8. Uji signifikansi parameter, yaitu uji simultan dan parsial.

## 3. Hasil dan Pembahasan

### 3.1. Uji Multikolinearitas

Uji multikolinearitas dilakukan untuk mendeteksi adanya hubungan linier yang sangat tinggi antara variabel bebas dalam model regresi [12]. Tabel 2 menunjukkan hasil Pengujian Multikolinearitas.

Hasil pengujian multikolinearitas di atas menunjukkan bahwa seluruh variabel independen memiliki nilai VIF kurang dari 10. Oleh karena itu, dapat disimpulkan bahwa tidak mengalami multikolinearitas. Dengan demikian, keputusan yang di-

Tabel 2. Hasil Pengujian Multikolinearitas

Var	X1	X2	X3	X4	X5	X6	X7
X1	–	1,027	1,187	1,009	1,026	1,079	1,185
X2	1,027	–	1,018	1,024	1,009	1,044	1,040
X3	1,187	1,018	–	1,018	1,021	1,002	1,274
X4	1,009	1,024	1,018	–	1,002	1,026	1,006
X5	1,026	1,009	1,021	1,002	–	1,001	1,022
X6	1,031	1,033	1,000	1,007	1,000	–	1,010
X7	1,217	1,002	1,954	1,026	1,023	1,010	–

ambil adalah menolak H0, yang berarti tidak terdapat indikasi adanya multikolinearitas.

### 3.2. Regresi Poisson

Analisis regresi Poisson adalah teknik statistika yang digunakan untuk memodelkan hubungan antara variabel respon Y yang berbentuk data cacah (count data) yang mengikuti distribusi Poisson, dengan satu atau lebih variabel prediktor X. Menurut [13], regresi Poisson digunakan untuk menganalisis hubungan antara variabel respon berupa data diskrit dengan satu atau lebih variabel prediktor, di mana rata-rata (mean) dan variansi diasumsikan sama (equidispersi). Menurut [14] model regresi Poisson berdistribusi Poisson dengan fungsi densitas ditulis sebagai berikut:

$$f(y_i | \mathbf{x}_i) = \frac{e^{-\mu_i} \mu_i^{y_i}}{y_i!}, \quad y_i = 0, 1, 2, \dots \quad (1)$$

dan parameter nilai harapan,

$$\mathbb{E}[y_i | \mathbf{x}_i] = \mu_i = \exp(\mathbf{x}_i^\top \boldsymbol{\beta}) \quad (2)$$

Dengan,

- $y_i$ : variabel respons (jumlah kejadian) untuk pengamatan ke- $i$ ,
- $\mathbf{x}_i$ : vektor kovariat atau variabel prediktor untuk pengamatan ke- $i$ ,
- $\mu_i$ : rata-rata atau parameter Poisson yang merupakan nilai harapan dari  $y_i$ ,
- $\boldsymbol{\beta}$ : vektor parameter regresi,
- $\exp(\mathbf{x}_i^\top \boldsymbol{\beta})$ : fungsi *link* log-linear yang memastikan bahwa  $\mu_i > 0$ .

Hasil analisis regresi poisson disajikan dalam tabel estimasi parameter regresi poisson dapat dilihat pada tabel 3:

Tabel 3. Estimasi Parameter Regresi Poisson

Variabel	Estimate
Intercept	-7,522
Presentase Penduduk (X1)	2,198
Presentase Penduduk Umur 15–64 Tahun (X2)	-0,007
Presentase Sanitasi Layak (X3)	-0,006
Presentase Sumur Tak Terlindung (X4)	-0,029
Presentase Rumah Tangga tidak Limbah (X5)	-0,022
Jumlah Penduduk Miskin (X6)	0,000
Indeks Pembangunan Manusia (X7)	0,058

Berdasarkan hasil regresi poisson di dapatkan persamaan model regresi sebagai berikut:

$$\log(\hat{Y}) = -7,522 + 2,198X_1 - 0,007X_2 + 0,006X_3 - 0,029X_4 - 0,022X_5 + 0,000X_6 - 0,058X_7$$

### 3.3. Uji Overdispersi dan Excess Zero

Overdispersi merujuk pada suatu kondisi di mana variabilitas data lebih besar dari yang diharapkan berdasarkan distribusi teoritis yang diasumsikan, seperti distribusi Poisson. Dalam konteks regresi Poisson, overdispersi terjadi ketika variansi dari variabel respon lebih besar daripada rata-rata hitungannya [5]. Dispersi ratio ini dinyatakan dalam bentuk persamaan berikut [15]:

$$\Phi = \frac{D}{n - p} \quad (3)$$

Nilai *deviance* diungkapkan melalui persamaan sebagai berikut:

$$D = 2 \sum_{i=1}^n y_i \ln \left( \frac{y_i}{\hat{\mu}_i} \right), \quad \text{dengan } i = 1, 2, \dots, n \quad (4)$$

Berdasarkan hasil analisis regresi *poisson*, nilai *deviance* dan derajat bebas dari Tabel 2 berturut-turut adalah 587,659 dan 506. Sehingga menghasilkan nilai dispersi ratio sebesar 1,161. Berdasarkan [16], jika dispersi ratio besar dari 1 maka overdispersi, sebaliknya, tidak terindikasi overdispersi. Berdasarkan hasil nilai dispersi ratio adalah sebesar 1,161 maka nilai tersebut lebih besar dari 1. Jadi terdapat indikasi overdispersi pada data.

Menurut [17], *Excess zero* merujuk pada situasi dimana lebih dari 50% nilai dalam variabel respon adalah nol. Dalam suatu penelitian, sering kali menghadapi situasi di mana terdapat banyak nilai nol, kondisi ini lebih sering muncul dan dapat menyebabkan overdispersi. Proporsi data yang memiliki nilai nol dapat memengaruhi akurasi estimasi yang dilakukan.

Menurut [18] nilai nol dalam penelitian ini memiliki arti penting dan tidak dapat diabaikan, karena akan memengaruhi hasil analisis. Oleh karena itu, nilai nol harus tetap dimasukkan dalam analisis. besarnya proporsi nilai nol dapat berdampak pada ketepatan inferensi yang dilakukan. Berdasarkan hasil analisis regresi *poisson*, presentase observasi dengan nilai nol dari Tabel 2 mencapai 90,07%. Hal ini menunjukkan bahwa lebih dari sembilan dari sepuluh observasi memiliki nilai nol, yang menandakan adanya *excess zero* dalam data.

### 3.4. Regresi Negatif Binomial

Menurut [19] regresi negatif binomial merupakan salah satu metode regresi yang digunakan untuk memodelkan data cacah yang mengalami overdispersi, yaitu situasi di mana varians data lebih besar dari rata-ratanya. Model ini merupakan perluasan dari regresi poisson yang mengasumsikan bahwa varians dan rata-rata data *count* adalah sama. Namun, ketika asumsi ini tidak terpenuhi karena varians yang lebih besar (overdispersi), regresi poisson menjadi tidak cocok. Sehingga regresi negatif binomial digunakan sebagai alternatif. Model regresi negatif binomial dapat dinyatakan sebagai berikut [20]:

$$\hat{\mu}_i = \exp(x_i\beta + \varepsilon_i) \tag{5}$$

Dimana:

- $\hat{\mu}_i$ : nilai harapan acak dari variabel respons  $y_i$  untuk observasi ke- $i$ . Berbeda dengan model *Poisson*, nilai harapan di sini dipengaruhi oleh komponen acak  $\varepsilon_i$ .
- $x_i$ : vektor variabel independen (prediktor) untuk observasi ke- $i$ .
- $\beta$ : vektor parameter regresi yang menunjukkan pengaruh masing-masing variabel independen terhadap nilai log-harapan.
- $\varepsilon_i$ : galat acak yang mencerminkan heterogenitas tak teramati antar unit observasi.

Hasil analisis regresi negatif binomial disajikan dalam tabel estimasi parameter regresi negatif binomial sebagai berikut:

**Tabel 4.** Estimasi Parameter Regresi Negatif Binomial

Variabel	Estimate
Intercept	1,576
Presentase Penduduk (X1)	0,050
Presentase Penduduk Umur 15–64 Tahun (X2)	-0,033
Presentase Sanitasi Layak (X3)	-0,040
Presentase Sumur Tak Terlindung (X4)	-0,016
Presentase Rumah Tangga dekat Limbah (X5)	-0,048
Jumlah Penduduk Miskin (X6)	0,000
Indeks Pembangunan Manusia (X7)	0,011
Deviance	488,2

Berdasarkan hasil tersebut maka didapatkan model regresi negative binomial sebagai berikut:

$$\log(\hat{Y}) = -1,576 + 0,050X_1 - 0,033X_2 - 0,040X_3 - 0,016X_4 - 0,048X_5 + 0,000X_6 + 0,011X_7$$

### 3.5. Regresi Multilevel Negative Binomial

Multilevel Negatif Binomial adalah metode regresi yang digunakan untuk menganalisis data cacah yang memiliki struktur hirarkis atau berlapis, dengan mempertimbangkan variasi antar kelompok atau *cluster*. Model ini merupakan perluasan dari regresi negatif binomial, yang mampu menangani data dengan overdispersi (di mana varians lebih besar dari rata-rata) dan *excess zeros* (keberadaan jumlah nol yang berlebihan) [11].

Menurut [21] model multilevel yang paling dasar adalah model 2 level, dimana level 1 merupakan unit pengamatan (misalnya individu, waktu, atau pengamatan berulang) dan level 2

merupakan grup atau klaster yang lebih besar (misalnya kabupaten, sekolah, atau rumah sakit). Dalam penelitian ini, level-1 terdiri atas 514 kabupaten/kota, yang dikelompokkan ke dalam 34 provinsi sebagai level-2. Provinsi dimodelkan sebagai *random intercept*, yang berarti perbedaan rata-rata jumlah kasus kronis filariasis antar provinsi diperhitungkan dalam model. Dengan pendekatan ini, diasumsikan bahwa pengaruh variabel-variabel prediktor level-1 bersifat tetap (*fixed*) di seluruh provinsi, namun *baseline* kasus kronis filariasis dapat bervariasi antar provinsi.

Menurut [22], pemodelan regresi multilevel dengan dua level pada level-1 dapat dijelaskan sebagai berikut:

$$Y_{ij} = \beta_{0j} + \beta_{1j}x_{ij} + R_{ij} \tag{6}$$

Bentuk pemodelan pada level-2 dapat dinyatakan sebagai berikut:

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + \gamma_{01}z_j + U_{0j} \\ \beta_{1j} &= \gamma_{10} + \gamma_{11}z_j + U_{1j} \end{aligned} \tag{7}$$

Jika persamaan (6) dimasukkan ke dalam persamaan (7), maka akan diperoleh model multilevel dua level sebagai berikut:

$$Y_{ij} = \gamma_{00} + \gamma_{10}x_{ij} + U_{0j} + U_{1j}x_{ij} + R_{ij} \tag{8}$$

Dimana:

- $Y_{ij}$ : respons individu ke- $i$  dalam kelompok ke- $j$ .
- $x_{ij}$ : prediktor level-1 (individu).
- $\beta_{0j}$ : intercept khusus kelompok ke- $j$ .
- $\beta_{1j}$ : slope khusus kelompok ke- $j$ .
- $R_{ij}$ : error individual (level-1).
- $z_j$ : prediktor pada level kelompok.
- $\gamma_{00}, \gamma_{10}$ : nilai rata-rata intercept dan slope keseluruhan.
- $\gamma_{01}, \gamma_{11}$ : pengaruh prediktor level-2 terhadap intercept dan slope.
- $U_{0j}, U_{1j}$ : error acak pada level kelompok.

### 3.5. Regresi Multilevel (2 level) tanpa Variabel Z

**Tabel 5.** Estimasi Regresi Multilevel (2 level) tanpa Variabel Z

Variabel	Estimate
Intercept	-0,007
Presentase Penduduk (X1)	-0,074
Presentase Penduduk Umur 15–64 Tahun (X2)	0,008
Presentase Sanitasi Layak (X3)	0,006
Presentase Sumur Tak Terlindung (X4)	-0,080
Presentase Rumah Tangga dekat Limbah (X5)	-0,027
Deviance	467,0

Berdasarkan Tabel 5 diperoleh nilai *p-value* untuk variabel  $X_1$  sampai dengan  $X_5$  lebih besar dari  $\alpha = 0,05$ , sehingga dapat disimpulkan bahwa tidak terdapat variabel yang signifikan memengaruhi jumlah kasus kronis filariasis pada level kabupaten/kota.

Berdasarkan hasil tersebut didapatkan model dengan peubah penjelas level 1 sebagai berikut:

$$\log(\hat{Y}) = -0,007 - 0,074X_1 + 0,008X_2 + 0,006X_3 - 0,080X_4 - 0,027X_5$$

Berdasarkan model (3), Setiap peningkatan satu unit pada variabel X1 maka akan menurunkan jumlah kejadian filariasis sebesar  $\exp(\beta_1) = \exp(-0,074) = 0.930$  kali dari rata rata jumlah kejadian filariasis.

### 3.5. Regresi Multilevel (2 level) dengan Variabel Z

**Tabel 6.** Estimasi Regresi Multilevel (2 level) dengan Variabel Z

Variabel	Estimate
Intercept	1,260
Presentase Penduduk (X1)	-0,162
Presentase Penduduk Umur 15–64 Tahun (X2)	0,009
Presentase Sanitasi Layak (X3)	-0,003
Presentase Sumur Tak Terlindung (X4)	-0,111
Presentase Rumah Tangga dekat Limbah (X5)	-0,039
Jumlah Penduduk Miskin (Z1)	0,000
Indeks Pembangunan Manusia (Z2)	0,077
Deviance	465,9

Berdasarkan Tabel 6 diperoleh nilai *p-value* untuk variabel  $X_1, X_2, X_3, Z_1,$  dan  $Z_2$  lebih besar dari  $\alpha = 0,05$ , sehingga dapat disimpulkan bahwa tidak terdapat variabel yang signifikan memengaruhi jumlah kasus kronis filariasis pada level kabupaten/kota. Sedangkan pada variabel  $X_4$  dan  $X_5$  nilai *p-value* lebih kecil dari  $\alpha = 0,05$ , sehingga dinyatakan bahwa variabel  $X_4$  dan  $X_5$  berpengaruh signifikan terhadap  $Y$ .

Berdasarkan hasil tersebut diperoleh model dengan peubah penjelas level-1 dan level-2 sebagai berikut:

$$\log(\hat{Y}) = 1,260 + 0,000Z_{1j} + 0,077Z_{2j} - 0,162X_{1ij} + 0,009X_{2ij} - 0,003X_{3ij} - 0,111X_{4ij} - 0,039X_{5ij}$$

Berdasarkan model tersebut, setiap peningkatan satu unit pada variabel  $X_4$  akan meningkatkan jumlah kejadian filariasis sebesar  $\exp(\beta_4) = \exp(-0,111) = 0,895$  kali dari rata-rata jumlah kejadian filariasis. Setiap peningkatan satu unit pada variabel  $X_5$  akan meningkatkan jumlah kejadian filariasis sebesar  $\exp(\beta_5) = \exp(-0,039) = 0,962$  kali dari rata-rata jumlah kejadian filariasis.

### 3.5. Pemilihan Model Terbaik

Pemilihan model multilevel negatif binomial terbaik dilakukan dengan membandingkan dua model, yaitu model tanpa variabel Z (model 1) dan model dengan variabel Z (model 2). Menurut [23] Semakin kecil nilai *deviance*, semakin baik model menjelaskan variabilitas dalam data. Selain itu, pemilihan model juga mempertimbangkan nilai *Akaike Information Criterion* (AIC), di mana model dengan nilai AIC yang lebih rendah dianggap memiliki keseimbangan yang lebih baik antara kompleksitas model dan kecocokan terhadap data [24]. Dengan demikian, model terbaik dipilih berdasarkan nilai *deviance* dan AIC terkecil.

**Tabel 7.** Pemilihan Model Terbaik

Model	Deviance
Regresi Multilevel (2 level) tanpa Variabel Z (Model 1)	467,0
Regresi Multilevel (2 level) dengan Variabel Z (Model 2)	465,9

Berdasarkan Tabel 7, diketahui bahwa nilai *deviance* model 2 (465,9) lebih kecil dibandingkan model 1 (467,0), yang mengindikasikan bahwa penambahan variabel level-2 meningkatkan kemampuan model dalam menjelaskan variasi data. Namun demikian, nilai AIC pada model 2 (485,9) sedikit lebih tinggi dibandingkan model 1 (483,0), yang menunjukkan adanya peningkatan kompleksitas model akibat penambahan parameter. Oleh karena itu, pemilihan model terbaik tidak hanya mempertimbangkan *deviance*, tetapi juga perlu memperhatikan keseimbangan antara kecocokan model dan kompleksitasnya. Meskipun model 2 menunjukkan perbaikan dari sisi *deviance*, perbedaan nilai AIC yang relatif kecil menunjukkan bahwa peningkatan kompleksitas masih dalam batas yang wajar. Dengan demikian, model multilevel dengan variabel Z atau variabel level-2 (model 2) tetap dapat dipertimbangkan sebagai model yang lebih tepat untuk konteks ini, terutama jika dilihat dari sudut pandang substantif dan kebutuhan untuk menangkap variasi antar provinsi.

### 3.6. Likelihood Ratio Test

**Tabel 8.** Likelihood Ratio Test

Model	Loglik	Pr(> $\chi^2$ )
Negative Binomial	-244,08	
Multilevel Negative Binomial	-232,94	2,36E-06

Berdasarkan Tabel 8 dapat dilihat bahwa nilai *p-value* yang sangat kecil menunjukkan bahwa model multilevel negatif binomial secara signifikan lebih baik dibandingkan dengan model negatif binomial. Oleh karena itu, penggunaan model multilevel negatif binomial diperlukan untuk mempertimbangkan perbedaan karakteristik antar provinsi sehingga menghasilkan estimasi yang lebih akurat serta inferensi yang lebih tepat.

### 3.7. Uji Hipotesis

Uji signifikansi parameter digunakan untuk menentukan apakah setiap variabel independen dalam model memiliki pengaruh yang bermakna secara statistik terhadap variabel dependen. Uji signifikansi parameter terdiri dari uji simultan dan uji parsial, yang bertujuan untuk mengevaluasi kebermaknaan model secara keseluruhan serta kontribusi masing-masing variabel independen dalam menjelaskan variasi variabel dependen.

Pada uji simultan menggunakan uji F, uji F digunakan untuk mengevaluasi apakah model regresi secara keseluruhan signifikan, yaitu untuk mengetahui apakah terdapat hubungan yang signifikan antara variabel dependen dan set variabel independen secara bersama-sama. Berdasarkan hasil analisis didapatkan bahwa  $F_{hitung}$  (5,610) lebih besar dari  $F_{tabel}$  (3,014) pada taraf signifikansi 5% ( $\alpha = 0,05$ ). Hal ini menunjukkan bahwa secara keseluruhan, terdapat setidaknya satu variabel independen yang berpengaruh signifikan terhadap jumlah kasus filariasis.

Adapun uji parsial menggunakan uji Wald, uji Wald digunakan untuk menguji signifikansi koefisien dalam model regresi, khususnya untuk menentukan apakah variabel independen tertentu berpengaruh signifikan terhadap variabel dependen. Hasil analisis uji Wald pada tabel 9:

Tabel 9. Hasil Uji Wald

Var	$W_{hitung}$	$W_{tabel}$	p-value	$\alpha$	Keputusan
X1	0,371	3,841	0,543	0,05	Gagal tolak $H_0$
X2	0,050	3,841	0,823	0,05	Gagal tolak $H_0$
X3	0,061	3,841	0,804	0,05	Gagal tolak $H_0$
X4	4,703	3,841	0,030	0,05	Tolak $H_0$
X5	7,710	3,841	0,005	0,05	Tolak $H_0$
X6	2,487	3,841	0,115	0,05	Gagal tolak $H_0$
X7	0,864	3,841	0,352	0,05	Gagal tolak $H_0$

Berdasarkan kriteria uji Wald, yaitu tolak  $H_0$  jika nilai statistik Wald ( $W_i$ ) >  $W_{\alpha}$ , df = 1 (nilai kritis Chi-Square pada tingkat signifikansi  $\alpha$  dan derajat kebebasan 1) atau p-value <  $\alpha$  (0,05). Maka didapatkan variabel X4 (*Persentase Sumur Tak Terlindung*) dan variabel X5 (*Persentase Rumah Tangga dekat Limbah*) menunjukkan nilai  $W_i$  yang lebih besar dari nilai kritis  $W_{\alpha}$ , df = 1 sebesar 3,84, atau memiliki nilai p-value yang lebih kecil dari tingkat signifikansi  $\alpha = 0,05$ . Oleh karena itu, hipotesis nol ( $H_0$ ) untuk variabel-variabel ini ditolak, yang berarti variabel-variabel tersebut berpengaruh secara signifikan terhadap model. Hasil ini konsisten dengan literatur seperti dalam penelitian [25] yang menunjukkan bahwa keberadaan sumur dapat menjadi tempat berkembang biak bagi nyamuk vektor filariasis yang cenderung memilih genangan air untuk bertelur hingga menjadi nyamuk dewasa serta jarak terbang nyamuk yang mencapai 1–2 km, sehingga keberadaan limbah dalam jarak tersebut membuat manusia lebih sering bersentuhan dengan nyamuk.

Sebaliknya, variabel-variabel sosial-ekonomi seperti X1 (persentase penduduk laki-laki), X2 (persentase penduduk usia 15–64 tahun), X3 (persentase sanitasi layak), X6 (jumlah penduduk miskin), dan X7 (*Indeks Pembangunan Manusia*) tidak menunjukkan pengaruh yang signifikan secara statistik terhadap jumlah kasus kronis filariasis pada tingkat signifikansi 5%. Padahal, dalam sejumlah literatur, indikator sosial-ekonomi telah dikaitkan dengan kondisi kesehatan masyarakat, termasuk penyebaran penyakit berbasis lingkungan seperti filariasis. Misalnya, penelitian [26] menunjukkan bahwa prevalensi filariasis lebih tinggi pada laki-laki dibandingkan perempuan karena laki-laki cenderung memiliki aktivitas luar ruangan yang lebih tinggi sehingga lebih berisiko terpapar vektor filariasis. Selain itu, kelompok usia produktif (15–64 tahun), khususnya individu berusia 26 hingga 55 tahun, dilaporkan memiliki risiko 4,6 kali lebih tinggi untuk terkena filariasis dibandingkan kelompok usia lainnya. Ketidaksignifikanan ini kemungkinan disebabkan oleh beberapa faktor, seperti sifat data agregat yang dapat menyembunyikan variasi individu, adanya variabel perantara yang tidak dimasukkan dalam model, serta heterogenitas wilayah yang tinggi di Indonesia. Selain itu, interaksi antarvariabel dan keterbatasan kualitas data juga dapat memengaruhi signifikansi statistik.

#### 4. Kesimpulan

Penelitian ini bertujuan untuk merancang model regresi multilevel negatif binomial dua-level pada kasus kronis filariasis di Indonesia tahun 2022, serta mengidentifikasi faktor-faktor signifikan yang memengaruhi jumlah kasus tersebut. Model regresi multilevel negatif binomial dipilih karena mampu menangani struktur data hierarkis (kabupaten/kota dalam provinsi), mengkomodasi adanya variasi antar-provinsi, serta mengatasi permasalahan overdispersi yang umum terjadi pada data jumlah kasus penyakit. Hasil analisis menunjukkan bahwa faktor lingkungan, yaitu persentase sumur tak terlindung (X4) dan persentase rumah tangga yang tinggal dekat penampungan limbah (X5), berpengaruh signifikan terhadap jumlah kasus kronis filariasis. Sebaliknya, variabel sosial ekonomi seperti persentase penduduk laki-laki (X1), penduduk usia 15–64 tahun (X2), sanitasi layak (X3), jumlah penduduk miskin (X6), dan Indeks Pembangunan Manusia (X7) tidak menunjukkan pengaruh signifikan. Ketidaksesuaian ini dapat disebabkan oleh keterbatasan data agregat yang digunakan, potensi variabel pengganggu yang tidak terukur, atau faktor lokal yang tidak tercakup dalam model. Penelitian selanjutnya disarankan menggunakan model regresi multilevel *Zero-Inflated Negative Binomial* (ZINB). Selain itu, penelitian mendatang juga disarankan untuk menggunakan data yang lebih besar dan mencakup lebih banyak variabel yang mungkin berpengaruh terhadap penyebaran filariasis, seperti faktor lingkungan atau intervensi kesehatan. Penelitian yang lebih luas dengan periode waktu yang lebih panjang juga akan memberikan pemahaman yang lebih mendalam dan komprehensif mengenai faktor-faktor yang mempengaruhi jumlah kasus filariasis di Indonesia. Dengan demikian, diharapkan kebijakan pencegahan dan penanggulangan filariasis dapat lebih tepat sasaran dan efektif dalam mengurangi prevalensi penyakit filariasis di masa depan.

#### References

- [1] R. Rahanyamtel, N. Nurjazuli, and S. Sulistiyani, "Faktor lingkungan dan praktik masyarakat berkaitan dengan kejadian filariasis di kabupaten semarang," *Jurnal Kesehatan Lingkungan Indonesia*, vol. 18, no. 1, pp. 8–11, 2019.
- [2] U. Usalma, A. Abdullah, and A. Aletta, "Faktor penyebaran kasus filariasis di kabupaten asahan," *Jukema (Jurnal Kesehatan Masyarakat Aceh)*, vol. 6, no. 2, pp. 134–140, 2020.
- [3] M. Adarabioyo and R. Ipinoyomi, "Comparing zero-inflated poisson, zero-inflated negative binomial and zero-inflated geometric in count data with excess zero," *Asian Journal of Probability and Statistics*, vol. 4, no. 2, pp. 1–10, 2019.
- [4] N. Achmad, M. R. F. Payu, and Y. Rahim, "Pemodelan pneumonia berat menggunakan regresi zero inflated negative binomial di gorontalo," *Euler: Jurnal Ilmiah Matematika, Sains dan Teknologi*, vol. 10, no. 1, pp. 45–53, 2022.
- [5] Y. Setyawan, K. Suryowati, and D. Octaviana, "Applying negative binomial regression analysis to overcome the overdispersion of poisson regression model for malnutrition cases in indonesia," *Bulletin of Applied Mathematics and Mathematics Education*, vol. 2, no. 2, pp. 87–96, 2022.
- [6] L. E. Afri et al., "Pemodelan regresi hurdle negative binomial tersensor pada kasus penyakit difteri," *Jurnal Absis: Jurnal Pendidikan Matematika dan Matematika*, vol. 2, no. 1, pp. 148–158, 2019.
- [7] L. P. L. Fávero, P. Belfiore, M. A. dos Santos, and R. F. Souza, "Overdisp: A stata (and mata) package for direct detection of overdispersion in poisson and negative binomial regression models," *Statistics, Optimization & Information Computing*, vol. 8, no. 3, pp. 773–789, 2020.
- [8] S. R. Y. Sinurat and E. Pasaribu, "Pemodelan kasus kronis filariasis di indonesia tahun 2019 menggunakan geographically weighted negative binomial regression (gwnbr)," *Indonesian Journal of Applied Statistics*, vol. 5, no. 1, pp. 19–30, 2022.
- [9] O. V. Amida and J. R. H. Sitorus, "Penerapan regresi logistik biner multilevel dalam analisis pengaruh karakteristik individu, rumah tangga, dan wilayah

- terhadap status kemiskinan balita di kepulauan maluku dan pulau papua,” in *Seminar Nasional Official Statistics*, vol. 2020, no. 1, 2020, pp. 967–977.
- [10] A. E. S. H. Maharani, D. A. Azka, A. Rahim, and L. G. Dwikasari, “Analisis regresi multilevel pada nilai rata-rata sma/ma pelaksana unbk,” *Jurnal Riset dan Aplikasi Matematika (JRAM)*, vol. 7, no. 2, pp. 158–169, 2023.
- [11] A. Z. Alem, B. A. Tegegne, F. M. Aragaw, R. E. Teklu, and T. A. Baykeda, “Multilevel negative binomial analysis of factors associated with numbers of antenatal care contacts in low and middle income countries: Findings from 59 nationally representative datasets,” *Plos one*, vol. 19, no. 4, p. e0301542, 2024.
- [12] I. N. Azizah, P. R. Arum, and R. Wasono, “Model terbaik uji multikolinearitas untuk analisis faktor-faktor yang mempengaruhi produksi padi di kabupaten blora tahun 2020,” in *Prosiding Seminar Nasional UNIMUS*, vol. 4, 2021.
- [13] L. Zuhrot, D. Devianto, and I. R. HG, “Pemodelan jumlah kasus dbd yang meninggal di kota padang dengan menggunakan regresi poisson,” *Jurnal Matematika UNAND*, vol. 4, no. 4, pp. 57–64, 2015.
- [14] A. C. Cameron and P. K. Trivedi, *Regression analysis of count data*. Cambridge university press, 2013, no. 53.
- [15] J. Hinde and C. G. Demétrio, “Overdispersion: models and estimation,” *Computational statistics & data analysis*, vol. 27, no. 2, pp. 151–170, 1998.
- [16] M. Sundari and P. R. Sihombing, “Penanganan overdispersi pada regresi poisson:(studi kasus: Pengaruh faktor iklim terhadap jumlah penderita penyakit demam berdarah di kota bogor),” *Jurnal Lebesgue: Jurnal Ilmiah Pendidikan Matematika, Matematika Dan Statistika*, vol. 2, no. 1, pp. 1–9, 2021.
- [17] F. Famoye and K. P. Singh, “Zero-inflated generalized poisson regression model with an application to domestic violence data,” *Journal of Data Science*, vol. 4, no. 1, pp. 117–130, 2006.
- [18] Q. W. Prianggada, “Pemilihan regresi zero-inflated poisson (zip) dan regresi zero-inflated binomial negatif (zibn) pada angka kematian ibu kabupaten bojonegoro tahun 2016 (data mengalami overdispersi),” Ph.D. dissertation, Universitas Brawijaya, 2017.
- [19] T. W. Utami, “Analisis regresi binomial negatif untuk mengatasi overdispersion regresi poisson pada kasus demam berdarah dengue,” *Jurnal Statistika Universitas Muhammadiyah Semarang*, vol. 1, no. 2, 2013.
- [20] J. Scott Long, “Regression models for categorical and limited dependent variables,” *Advanced quantitative techniques in the social sciences*, vol. 7, 1997.
- [21] H. Goldstein, *Multilevel statistical models*. John Wiley & Sons, 2011.
- [22] T. A. Snijders and R. Bosker, “Multilevel analysis: An introduction to basic and advanced multilevel modeling,” 2011.
- [23] B. Tantular, H. Wijayanto *et al.*, “Pemilihan model regresi linier multilevel terbaik,” in *Forum Statistika dan Komputasi*, vol. 14, no. 2, 2009.
- [24] B. L. Manurung, T. Priharyanto, and U. Alifah, “Pemodelan jumlah kekerasan terhadap perempuan di jawa timur dengan regresi poisson dan binomial negatif,” *JSN: Jurnal Sains Natural*, vol. 2, no. 3, pp. 68–77, 2024.
- [25] I. R. Rahmi, D. Sutiningsih, R. Hestningsih, and L. D. Saraswati, “Faktor-faktor yang berhubungan dengan kasus filariasis di indonesia: Sistematis review,” *Jurnal Epidemiologi Kesehatan Komunitas*, vol. 7, no. 2, pp. 501–521, 2022.
- [26] I. K. Faizin, M. Nasir *et al.*, “Karakteristik penderita filariasis di wilayah kerja puskesmas kaleke,” *Jurnal Medical Profession (Medpro)*, vol. 1, no. 1, pp. 43–49, 2019.