

Pemodelan *Generalized Poisson Regression* (GPR) terhadap Jumlah Kasus Penyakit *Tuberculosis* di Provinsi Nusa Tenggara Timur

Joseph F. Seran Berek¹, Robertus Dole Guntur²

^{1,2} Program Studi Matematika, Universitas Nusa Cendana

Jl. Adisucipto Penfui, PO BOX 104, Kupang, 85001, NTT

Email: fberek08@gmail.com; robertus_guntur@staf.undana.ac.id

Korespondensi penulis : robertus_guntur@staf.undana.ac.id

Abstrak

Provinsi Nusa Tenggara Timur (NTT) merupakan salah satu wilayah dengan jumlah kasus tuberkulosis (TBC) yang cukup tinggi pada tahun 2022. Jumlah kasus TBC merupakan contoh variabel diskrit yang umumnya dianalisis menggunakan Regresi Poisson. Model ini mengasumsikan bahwa nilai mean dan variansi dari variabel respon adalah sama. Namun, dalam banyak kasus, asumsi ini tidak terpenuhi karena variansi variabel respon sering kali lebih besar daripada mean (overdispersi). Untuk mengatasi masalah ini, digunakan model *Generalized Poisson Regression* (GPR) yang lebih fleksibel dalam menangani overdispersi. Penelitian ini bertujuan untuk memodelkan jumlah kasus tuberkulosis di Provinsi Nusa Tenggara Timur (NTT) dengan menggunakan model *Generalized Poisson Regression* (GPR). Variabel dependen dalam penelitian ini adalah jumlah kasus TBC Paru per kabupaten/kota pada tahun 2023. Sementara itu, variabel independen terdiri dari beberapa faktor yang diduga berpengaruh, yaitu: jumlah penduduk miskin, persentase perokok aktif, jumlah balita kurang gizi, jumlah puskesmas, persentase rumah tangga dengan akses sanitasi layak, Persentase penggunaan air bersih dan jumlah kasus HIV AIDS. Hasil analisis menunjukkan bahwa terdapat 31 model yang mungkin terbentuk, dan model optimal dengan 5 variabel prediktor menghasilkan nilai *Akaike Information Criterion* (AIC) sebesar 1413.285. Variabel yang berpengaruh signifikan terhadap jumlah kasus tuberkulosis adalah jumlah penduduk miskin, jumlah puskesmas, persentase rumah tangga dengan akses sanitasi layak, persentase air bersih, dan jumlah penderita HIV/AIDS.

Kata Kunci: *Generalized Poisson Regression*, kriteria informasi Akaike, overdispersi, Provinsi NTT, tuberkulosis.

Abstract

East Nusa Tenggara Province (NTT) is one of the provinces with a fairly high number of tuberculosis cases in 2022. The number of the cases is an example of a discrete variable which is usually analyzed using Poisson regression model assuming that the mean and variance of the response variable must be the same. However, this condition is often violated where the variance of the response variable is greater than the mean (overdispersion). To overcome this, the Generalized Poisson Regression (GPR) model is applied. This study aims to model the number of tuberculosis cases in ENTP and investigate the variables influencing it using the GPR model. The dependent variable is the number of pulmonary tuberculosis cases, with eight independent variables. The study shows that data on the distribution of the number of tuberculosis cases in ENTP is over-dispersed. The results of the analysis using the GPR model show that there are 31 possible models to be formed and modeling involving 5 predictor variables produces an optimal GPR model with an Akaike Information Criterion value is 1413.285. The five independent variables that have a significant influence on the number of tuberculosis cases in ENTP in 2023 are the number of poor people, the number of community health centers, the percentage of households with access to proper sanitation, the percentage of clean water and the number of HIV and AIDS sufferers.

Keywords: Akaike Information Criterion, East Nusa Tenggara Province, Generalized Poisson Regression Model, Overdispersion, tuberculosis.

1. Pendahuluan

Tuberkulosis (TBC) merupakan masalah kesehatan masyarakat yang menyerang paru-paru dan *Mycobacterium Tuberculosis* adalah patogennya[1]. Kemungkinan terjangkit penyakit ini biasanya tersebar melalui air liur atau dahak penderita yang mengandung bakteri *Mycobacterium Tuberculosis*. Penyakit ini ditemukan oleh Robert Koch pada tanggal 24 Maret 1882 di Berlin, Jerman [2].

Tuberkulosis masih menjadi masalah kesehatan di berbagai negara. Penyakit ini merupakan salah satu penyebab utama kematian di dunia. Menurut *World Health Organization* ditahun 2020, sekitar 10 juta orang di seluruh dunia terindikasi terkena penyakit tuberkulosis [3]. Berdasarkan data pada *Global Tuberculosis Report*, Indonesia menduduki peringkat kedua tertinggi jumlah kasus TBC pada tahun 2022 setelah India dengan 969.000 kasus, setara dengan 11 kematian per jam [4]. Jumlah ini terus tumbuh sebesar 17% pada tahun 2020, dengan jumlah kasus sebanyak 824.000 kasus, dimana 430.667 kasus diantaranya tidak terdeteksi [5]. Persebaran jumlah kasus TBC ditemukan pada semua provinsi di Indonesia dengan penyumbang terbanyak berasal dari Provinsi Jawa Barat dengan 184.406 kasus dan terendah dari Provinsi Kalimantan Utara dengan 1738 kasus pada tahun 2022 [6].

Provinsi Nusa Tenggara Timur (NTT) juga berkontribusi pada tingginya kasus TBC di Indonesia dengan jumlah kasus sebanyak 8.802 di tahun 2022 [6]. Jumlah ini menunjukkan kenaikan yang signifikan yang hanya mencapai 5.361 kasus di tahun 2020. Persebaran jumlah kasus ini tidak sama disetiap kabupaten dengan jumlah yang terbanyak di Kota Kupang sebanyak 806 kasus dan jumlah yang terendah di Kabupaten Sabu Raijua sebanyak 82 kasus [7].

Jumlah kasus tuberkulosis di Provinsi NTT ini adalah jenis data hitungan (*count data*) yang merupakan karakteristik distribusi Poisson [8]. Apabila rata-rata dan variansnya sama (distribusi yang sama), regresi Poisson sangat cocok untuk digunakan untuk

menganalisis data hitungan. Namun, situasi ini sebenarnya jarang terjadi. Data *count* biasanya memiliki *overdispersion* atau nilai varians lebih besar dari *mean*. Dalam kondisi ini model regresi Poisson tidak bisa digunakan lagi. Salah satu metode yang dapat digunakan untuk mengatasi masalah ini adalah dengan menggunakan metode *Generalized Poisson Regression (GPR)* [9]. Penggunaan model GPR sudah digunakan sebelumnya untuk mengatasi masalah overdispersi pada jumlah kasus stunting tahun 2022 di Provinsi NTT [9], jumlah kematian balita [10], jumlah kasus AIDS [11], jumlah kasus pneumonia [12], dan jumlah kasus berat badan lahir rendah pada balita di Indonesia tahun 2021 [13]. Penerapan model GPR pada penelitian ini digunakan untuk mengatasi masalah overdispersi pada data jumlah kasus *tuberculosis* di Provinsi NTT pada tahun 2022. Penelitian ini bertujuan untuk memberikan wawasan baru tentang faktor-faktor yang mempengaruhi jumlah kasus tuberkulosis di NTT, serta memberikan rekomendasi kebijakan yang dapat membantu pemerintah dalam merumuskan strategi pengendalian penyakit ini.

2. Metode Penelitian

Penelitian ini menggunakan metode kuantitatif dengan analisis *Generalized Poisson Regression (GPR)* untuk mengatasi masalah overdispersion. Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari website resmi Badan Pusat Statistik Provinsi NTT dengan jumlah kasus *tuberculosis* pada setiap kabupaten di Provinsi NTT pada tahun 2023 sebagai variable respon(Y). Kemudian variable bebasnya adalah jumlah penduduk miskin(X_1), persentase perokok aktif(X_2), jumlah balita kurang gizi(X_3), jumlah puskesmas(X_4), persentase rumah tangga dengan akses sanitasi layak (X_5), persentase air bersih (mata air terlindungi)(X_6) dan jumlah kasus HIV/AIDS(X_7). Analisis dilakukan dengan menggunakan *software* statistic SPSS untuk menguji hipotesis dan menentukan signifikansi hubungan antar variabel. Tahapan Analisis pada penelitian ini adalah sebagai berikut:

1. Pengumpulan Data
Mengumpulkan data terkait jumlah kasus TBC Paru serta faktor-faktor yang berpengaruh dari sumber sekunder.
2. Visualisasi Persebaran
Menampilkan distribusi geografis kasus TBC Paru dalam bentuk peta.
3. Uji Multikolinearitas
Menganalisis hubungan antar variabel independen menggunakan *Variance Inflation Factor (VIF)*.
4. Estimasi Model Regresi Poisson
Menghitung parameter model dengan metode *Maximum Likelihood Estimation (MLE)*.
5. Deteksi Overdispersi
Mengevaluasi kemungkinan overdispersi pada data dengan uji *Pearson chi-square/df* atau *likelihood ratio test*.
6. Pemodelan dengan *Generalized Poisson Regression (GPR)*
Jika ditemukan overdispersi, maka digunakan model GPR sebagai solusi alternatif.
7. Pemilihan Model Terbaik
Menentukan model optimal berdasarkan nilai *Akaike Information Criterion (AIC)*.
8. Interpretasi Model

Menafsirkan hasil estimasi parameter untuk memahami pengaruh variabel independen terhadap jumlah kasus TBC.

2.1 Statistika Deskriptif

Dengan menggunakan metode statistik sederhana seperti minimum, rata-rata, maksimum, dan standar deviasi, analisis deskriptif bertujuan untuk memberikan gambaran tentang karakteristik data. Analisis juga dilakukan dengan menggunakan visualisasi data, seperti peta dan grafik.

2.2 Multikolinearitas

Multikolinearitas merupakan suatu kondisi dimana dua atau lebih variabel bebas dalam model regresi berkorelasi tinggi satu sama lain. Saat membentuk model regresi dengan beberapa variabel independen (bebas), penting untuk memastikan bahwa tidak adanya multikolinearitas antar variabel-variabel tersebut. Untuk mendeteksi multikolinearitas, digunakan kriteria nilai VIF (*Variance Inflation Factor*). Jika nilai VIF lebih besar 10, artinya terdapat multikolinearitas [14].

Nilai VIF dihitung dengan rumus:

$$VIF = \frac{1}{1-R_j^2}; j = 1, 2, \dots, k \quad (1)$$

Dimana R_j^2 adalah koefisien determinasi antara variabel-variabel bebas lainnya.

2.3 Poisson Regression

Poisson Regression (PR) adalah suatu model regresi non-linear yang digunakan untuk menganalisis data berdistribusi Poisson dimana variabel respon berbentuk data *count* dan nilainya adalah *integer* tidak negative [13]. Regresi Poisson merupakan salah satu aplikasi dari *Generalized Linear Model* (GLM), yang merupakan perluasan dari model regresi umum untuk variabel respon yang memiliki sebaran eksponensial. Dalam regresi Poisson, masing-masing variabel prediktor tidak memiliki multikolinearitas dan variabel respon Y diasumsikan berdistribusi Poisson. Secara sistematis distribusi Poisson memiliki fungsi peluang sebagai berikut:

$$P(Y = y) = \frac{\lambda^y e^{-\lambda}}{y!}, \quad y = 0, 1, 2, \dots$$

Keterangan:

Y : Variabel acak dengan distribusi Poisson

y : Nilai relisasi dari variabel acak Y

λ : Parameter rata-rata

e : Bilangan euler ($\approx 2,718$)

$y!$: Faktorial dari y

2.4 Penaksiran Parameter Model Regresi Poisson

Saat menaksir parameter model regresi Poisson kita dapat menggunakan metode *maximum likelihood estimator* (MLE) dengan melibatkan asumsi bahwa variabel terikat dengan mengikuti distribusi Poisson dan bahwa hubungan antara variabel bebas dan variabel terikat bersifat linear dalam logaritma.

Fungsi *log-likelihood* pada model regresi Poisson dirumuskan dengan persamaan, sebagai berikut:

$$\ln \mathcal{L}(\beta) = \sum_{i=1}^n Y_i \ln[\lambda(x, \beta)] - \sum_{i=1}^n \lambda(x, \beta) - \sum_{i=1}^n \ln(Y_i) \quad (2)$$

Selanjutnya, memaksimumkan fungsi dalam Persamaan (2) untuk mendapatkan estimasi parameter yang optimal dengan menggunakan algoritma *Fisher Scoring* [15]. Untuk menguji parameter model regresi Poisson dilakukan dengan pengujian sebagai berikut

1. Uji Simultan

Hipotesis yang digunakan adalah:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0,$$

Artinya tidak ada variabel bebas yang berpengaruh pada variabel terikat.

$$H_1 : \beta_1 \neq \beta_2 \neq \dots \neq \beta_p \neq 0,$$

Artinya setidaknya ada satu variabel bebas yang mempengaruhi variabel terikat.

Statistik uji:

$$D(\hat{\beta}) = -2 \ln \quad (3)$$

Kriteria uji adalah tolak H_0 apabila $D(\hat{\beta}) > \chi^2_{(i, \alpha)}$ dengan i adalah banyaknya variabel bebas atau nilai $p - value < \alpha$.

2. Uji Parsial

Hipotesis Uji Wald yaitu:

$$H_0 : \beta_i = 0,$$

variabel bebas tidak mempengaruhi variabel terikat.

$$H_1 : \beta_i \neq 0,$$

variabel bebas yang mempengaruhi variabel terikat.

Statistik Uji Wald yaitu:

$$Z_{Hitung} = \frac{\hat{\beta}_i}{SE(\hat{\beta}_i)} \quad (4)$$

Kriteria uji adalah tolak H_0 apabila $|Z_{Hit}| > Z_{\frac{\alpha}{2}}$ atau nilai $p - value < \alpha$.

2.5 Pengujian Overdispersi

Secara umum, regresi Poisson memiliki varians dan rata-rata yang sama, tetapi terkadang terjadi overdispersi, di mana variansnya lebih besar dari rata-rata. Untuk memeriksa overdispersi digunakan rasio dispersi (ϕ) yang didefinisikan sebagai ratio antara nilai *Chi-squared* (χ^2) dan derajat bebasnya (df), yaitu:

$$\phi = \frac{\chi^2}{df} \quad (5)$$

$$\phi = \frac{\sum_{i=1}^N \frac{(y_i - E(Y))^2}{Var(Y)}}{N - p} \quad (6)$$

Dimana, y_i merupakan nilai pengamatan ke- i dari variabel terikat, jumlah pengamatan (N) dan jumlah parameter (p). Jika nilai $\phi > 1$ menunjukkan adanya overdispersi dalam model [16].

2.6 Pemodelan dengan *Generalized Poisson Regression*

Untuk menghitung jumlah data saat terjadi overdispersi dan underdispersi, model GPR mengambil distribusi Generalized Poisson (GP) untuk setiap variabel random. Berikut adalah model distribusi model GPR [17].

$$f(y; \mu; \theta) = \left(\frac{\mu}{1+\mu}\right)^y \frac{(1+\theta y)^{y-1}}{y!} \exp\left(\frac{-\mu(1+\theta y)}{1+\theta \mu}\right) \quad (7)$$

Dengan mean dan variansnya :

$$E(y) = \mu \text{ dan } Var(y) = \mu(1 + \theta \mu)^2 \quad (8)$$

Keterangan :

$\theta = 0$; Model GPR menjadi regresi Poisson biasa

$\theta > 0$; Model GPR menjelaskan data *count* yang overdispersi

$\theta < 0$; Model GPR menjelaskan data *count* yang underdispersi

Penaksir parameter model GPR dilakukan dengan metode *Maximum Likelihood Estimator* (MLE) sebagai berikut:

$$L(\beta, \theta) = \prod_{i=1}^n f(\beta, \theta) \quad (9)$$

$$L(\beta, \theta) = \prod_{i=1}^n \left\{ \left(\frac{\mu_i}{1+\theta \mu_i}\right)^{y_i} \frac{(1+\theta y_i)^{(y_i-1)}}{y_i!} \exp\left(\frac{-\mu_i(1+\theta \mu_i)}{1+\theta \mu_i}\right) \right\} \quad (10)$$

Jika, diubah dalam bentuk fungsi logaritma natural menjadi :

$$\ln L(\beta, \theta) = \sum_{i=1}^n \{y_i \ln(\mu_i) - y_i \ln(1 + \theta \mu_i) + \Delta\} \quad (11)$$

Dimana

$$\Delta = (y_i - 1) \ln(1 + \theta y_i) - \ln(y_i!) - \mu_i \left(\frac{1+\theta y_i}{1+\theta \mu_i}\right) \quad (12)$$

Untuk mendapatkan taksiran parameter β dan θ maka Persamaan (11) diturunkan terhadap β dan θ menggunakan metode numerik iterasi *Newton-Raphson* sehingga diperoleh penaksir parameter yang konvergen.

Pengujian Parameter *Generalized Poisson Regression* (GPR) dilakukan dengan metode *Maximum Likelihood Ratio Test* (MLRT). Dengan hipotesis sebagai berikut :

$$H_0: \beta_1 = \beta_2 = \beta_3 = \dots = \beta_j$$

$$H_1: \beta_j \neq 0; j = 1, 2, \dots, s$$

Statistik uji:

$$D(\hat{\beta}) = -2 \ln \left(\frac{L(\hat{\omega})}{L(\hat{\Omega})} \right) = 2 \left(\ln(\hat{\Omega}) - \ln L(\hat{\omega}) \right) \quad (13)$$

Tolak H_0 jika $D(\hat{\beta}) > \chi^2_{(s;\alpha)}$ sehingga terdapat paling sedikit $\beta_j \neq 0$ yang menunjukkan pengaruh bahwa X_j berpengaruh secara signifikan terhadap model. Pengujian dilanjutkan menggunakan uji parsial dengan hipotesis berikut :

$$H_0: \beta_1 = 0 \text{ (Pengaruh variabel ke-} j \text{ tidak signifikan)}$$

$$H_1: \beta_j \neq 0 \text{ (Pengaruh variabel ke-} j \text{ signifikan)}$$

Statistik uji yang digunakan mengikuti distribusi z yaitu :

$$t = \frac{\hat{\beta}_j}{se(\hat{\beta}_j)} \quad (14)$$

$se(\hat{\beta}_j)$ adalah nilai *standard error* dari parameter $\hat{\beta}_j$. $se(\hat{\beta}_j)$ diperoleh dari elemen diagonal ke (j+1) dari $[-H^{-1}(\hat{\beta})]$. Tolak H_0 jika $|t_{hit}| > t_{\frac{\alpha}{2},df}$, dimana α adalah taraf signifikansi.

2.7 Pemilihan Model Terbaik

Nilai Kriteria Informasi Akaike (AIC) dihitung untuk memilih model terbaik. Ini dilakukan dengan menghitung nilai *log-likelihood* dari setiap model yang ingin dibandingkan dan kemudian menghitung jumlah parameter prediktor (k). Nilai AIC dihitung menggunakan persamaan berikut.

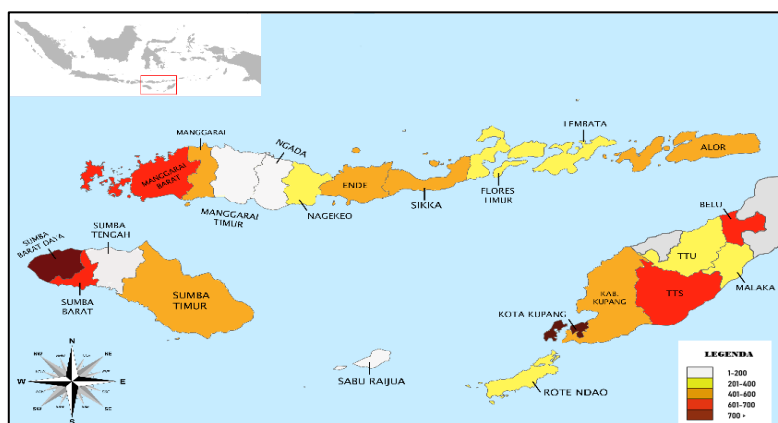
$$AIC = -2\mathcal{L} + 2k \quad (15)$$

Pemilihan model terbaik diperoleh pada nilai AIC terkecil [18].

3. Hasil dan Pembahasan

3.1 Statistika Deskriptif

Berdasarkan Gambar 1 diperoleh bahwa pada tahun 2023, peta penyebaran penyakit TBC dengan jumlah kasus terendah yakni 85 kasus terjadi di Kabupaten Sabu Raijua dan kasus terbanyak yakni 702 kasus di Kabupaten Sumba Barat Daya dan 1.253 kasus terjadi di Kota Kupang.



Gambar 1. Peta Penyebaran TBC di Provinsi NTT tahun 2022

Tabel 1. Statistika Deskriptif

Variabel	N	Min	Max	Mean	Standar Deviasi	Varians
Tuberkulosis(Y)	22	85	1253	433.41	265.290	70378.825
Jumlah Penduduk Miskin(X_1)	22	18.57	119.51	51.8691	27.10003	734.412

Persentase Perokok Akif(X_2)	22	11.16	24.8	17.0341	3.78328	14.313
Jumlah Balita Kurang Gizi(X_3)	22	549	9762	2900.18	2376.732	5648853.489
Jumlah Puskesmas(X_4)	22	6	37	19.82	8.110	65.775
Persentase Akses Sanitasi Layak(X_5)	22	48.22	93.03	74.7755	13.39193	179.344
Persentase Air Bersih(X_6)	22	0.63	59.98	30.7409	17.31589	299.840
Jumlah Kasus HIV/AIDS(X_7)	22	3	113	34.18	27.719	768.346

Pada Tabel 1. ditampilkan nilai minimum, maksimum, rata-rata, standar deviasi dan variansi dari setiap variabel yang digunakan dalam penelitian ini. Rata-rata jumlah kasus *Tuberculosis* di Provinsi Nusa Tenggara Timur selama tahun 2023 sebesar 433.41 atau 433 kasus dengan variansinya sebesar 70378.825.

3.2 Uji Multikolinearitas

Tabel 2. Nilai VIF Variabel Predictor

Variabel	VIF
Jumlah Penduduk Miskin (X_1)	5.351
Persentase Perokok Aktif(X_2)	1.718
Jumlah Balita Kurang gizi(X_3)	4.111
Jumlah Puskesmas(X_4)	2.907
Persentasi Rumah Tangga sanitasi layak(X_5)	1.768
Persentase Air Bersih(X_6)	2.274
Jumlah Kasus HIV/AIDS(X_7)	1.898

Berdasarkan nilai VIF untuk masing-masing variabel tidak lebih dari 10 hal ini menunjukkan bahwa tidak terjadinya multikolinieritas diantara ke tujuh variabel bebas, seperti yang ditunjukkan oleh Tabel 2. Oleh karena itu, analisis dapat dilanjutkan pada tahap berikutnya.

3.3 Uji Overdispersion

Tabel 3. Uji Overdispersi

Tuberkulosis	Value	Df	Value/Df
Deviance	1230.915	14	87.923
Pearson Chi-Square	1287.930	14	91.995

Berdasarkan Tabel 3, terlihat bahwa nilai deviance dari model regresi Poisson adalah 1230,915 dan jika dibagi dengan nilai derajat bebasnya 14 maka didapat 87,923 yang lebih dari 1. Maka dapat disimpulkan bahwa terjadi overdispersi pada data yang diamati. Oleh karena itu untuk mengatasi hal tersebut dapat digunakan model *Generalized Regresi Poisson*.

3.4 Model Estimasi Regresi Poisson

Regresi Poisson, model regresi non linier yang sering digunakan untuk menganalisis data count, digunakan untuk memodelkan faktor-faktor yang mempengaruhi jumlah kasus penyakit tuberkulosis di Provinsi NTT tahun 2023. Model regresi poisson untuk jumlah kasus Tuberkulosis pada di Provinsi NTT tahun 2023 umumnya berbentuk seperti berikut:

Model umum regresi Poisson:

$$\mu = \exp (\beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_kx_k) \quad (16)$$

Karena diketahui variable predictor berjumlah 7 variabel maka model regresi poissonnya berbentuk seperti berikut:

$$\mu = \exp (\beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_4x_4 + \beta_5x_5 + \beta_6x_6 + \beta_7x_7) \quad (17)$$

3.5 Uji Serentak

Tabel 4. Uji serentak

$D(\hat{\beta})$	$\chi^2(7; 0.05)$	P-value
1980.914	14.071	0.0001

Berdasarkan Tabel 4. Karena nilai $D(\hat{\beta}) > \chi^2(7; 0,05) = 1980.914 > 14.017$ dan nilai $p - value (0,0001) < \alpha(0,05)$. Keputusan yang dihasilkan yaitu tolak H_0 , sehingga kesimpulannya variabel predictor menunjukkan hubungan yang tepat atau dengan kata lain berpengaruh secara bersama(simultan) terhadap jumlah kasus *Tuberculosis* di Provinsi NTT tahun 2023.

3.6 Uji Parsial

Tabel 5. Uji Parsial

Parameter	Estimasi	Standar Error	Wald	t-hitung	P-value	Keputusan
Jumlah Penduduk Miskin (X_1)	0.011	0.001	124.855	1.180	0.0001	Signifikan
Persentase Perokok Aktif(X_2)	0.002	0.0035	0.365	-0.453	0.546	Tidak Signifikan
Jumlah Balita Kurang gizi(X_3)	-6.824	8.764	0.006	-0.3	0.938	Tidak Signifikan
Jumlah Puskesmas(X_4)	-0.025	0.0024	109.176	-1.331	0.00001	Signifikan
Persentasi Rumah Tangga sanitasi layak(X_5)	0.007	0.0013	28.457	0.432	<0.001	Signifikan
Persentase Air Bersih(X_6)	0.007	0.001	47.985	0.827	<0.001	Signifikan
Jumlah Kasus HIV/ AIDS(X_7)	0.015	0.0005	852.854	3.631	0.0001	Signifikan

Berdasarkan hasil seperti yang dinyatakan pada Tabel 5, dengan tingkat signifikansi 0.05 maka dapat diputuskan bahwa variabel predictor yang berpengaruh secara individu (parsial) terhadap jumlah kasus *tuberculosis* di Provinsi Nusa Tenggara Timur tahun 2023

adalah jumlah penduduk miskin(X_1), jumlah puskesmas(X_4), persentase rumah tangga dengan akses sanitasi yang layak(X_5), persentase air bersih(X_6), dan jumlah HIV dan AIDS(X_7). Dengan demikian diperoleh persamaan yang melibatkan variabel yang signifikan adalah :

$$\mu = \exp(4.639 - 0.011X_1 - 0.025X_4 + 0.007X_5 + 0.007X_6 + 0.015X_7) \quad (18)$$

3.7 Pemodelan Kasus dengan Metode GPR

Untuk pemodelan dengan GPR, beberapa model regresi Poisson akan dipilih berdasarkan nilai *Akaike Information Criteria* (AIC). AIC merupakan alat ukur statistik yang dapat digunakan untuk mengukur kesesuaian model. Model dengan nilai AIC terkecil akan dianggap sebagai model terbaik. Semakin rendah nilai AIC, semakin baik model menjelaskan data prediksi dengan data aktual. Dari 5 variabel prediktor yang berpengaruh signifikan, model GPR yang dapat dibentuk berjumlah 31 model. Dari 31 model GPR yang dibentuk akan diambil model dengan nilai AIC terkecil seperti yang ditunjukkan pada Tabel 6 berikut ini.

Tabel 6. Model GPR untuk Data Jumlah Kasus TC di Provinsi NTT Tahun 2023

Variabel	Model	AIC
X_7	$\mu = \exp(5.530 + 0.014X_7)$	1704.258
X_1X_7	$\mu = \exp(5.263 + 0.005X_1 + 0.014X_7)$	1537.838
$X_1X_4X_7$	$\mu = \exp(5.379 + 0.007X_1 - 0.010X_4 + 0.013X_7)$	1498.922
$X_1X_4X_6X_7$	$\mu = \exp(5.194 + 0.009X_1 - 0.021X_4 + 0.007X_6 + 0.015X_7)$	1446.976
$X_1X_4X_5X_6X_7$	$\mu = \exp(4.704 + 0.011X_1 - 0.025X_4 + 0.006X_5 + 0.007X_6 + 0.015X_7)$	1413.285

3.8 Pemilihan Model Terbaik

Berdasarkan Tabel 6 dapat dijelaskan bahwa model terbaik yang dipilih berdasarkan kriteria nilai AIC terkecil adalah model yang melibatkan variabel X_1 , X_4 , X_5 , X_6 , X_7 sebagai variabel predictor dalam model yang mana menghasilkan nilai AIC yang paling kecil dari 31 model GPR yang terbentuk. Dengan demikian model *Generalized Poisson Regression* untuk kasus TBC di Provinsi Nusa Tenggara Timur tahun 2023 sebagai berikut :

$$\mu = \exp(4.704 + 0.011X_1 - 0.025X_4 + 0.006X_5 + 0.007X_6 + 0.015X_7) \quad (19)$$

3.9 Interpretasi Model

Berdasarkan model yang diperoleh maka dapat dikatakan bahwa variabel yang berpengaruh secara signifikan terhadap model penyebaran jumlah kasus TBC di Provinsi NTT tahun 2023 adalah jumlah penduduk miskin (X_1), jumlah puskesmas (X_4), persentase rumah tangga dengan akses sanitasi yang layak (X_5), persentase air bersih (X_6), dan jumlah HIV dan AIDS (X_7). Dari model yang diperoleh dapat dikatakan bahwa setiap penambahan seribu jiwa penduduk Miskin (X_1) akan menyebabkan kenaikan rata-rata jumlah kasus *tuberculosis* sebesar $\exp(0.011)$ atau 1.01106 dari rata-rata jumlah kasus *tuberculosis* semula. Semakin banyak jumlah penduduk miskin maka akan meningkatkan

jumlah *tuberculosis* di setiap wilayah kabupaten/kota sebanyak 1 kasus. Hal ini menunjukkan bahwa faktor ekonomi berperan penting dalam peningkatan kasus TBC

Jumlah puskesmas (X_4) memberi kontribusi yang signifikan terhadap jumlah *tuberculosis* di Provinsi Nusa Tenggara Timur, dimana setiap penambahan 1 puskesmas akan menyebabkan penurunan rata-rata jumlah *tuberculosis* sebesar $\exp(-0.025)$ atau 0.97531 dari rata-rata jumlah kasus *tuberculosis* semula. Semakin banyak jumlah puskesmas akan mengurangi jumlah kasus *tuberculosis* di setiap wilayah kabupaten/kota sebanyak 1 kasus. Hal ini menunjukkan bahwa peningkatan jumlah puskesmas berkontribusi terhadap penurunan kasus *tuberculosis*.

Persentase rumah tangga dengan akses sanitasi layak (X_5) memberi kontribusi yang signifikan terhadap jumlah *tuberculosis* di Provinsi Nusa Tenggara Timur, dimana setiap penambahan 1 rumah tangga dengan akses sanitasi layak akan menyebabkan kenaikan rata-rata jumlah *tuberculosis* sebesar $\exp(0.006)$ atau 1.00602 dari rata-rata jumlah kasus *tuberculosis* semula. Semakin banyak rumah tangga dengan akses sanitasi layak akan meningkatkan jumlah kasus *tuberculosis* di setiap wilayah kabupaten/kota sebanyak 1 kasus. Secara teori, sanitasi yang baik seharusnya mengurangi penyebaran *tuberculosis*. Namun, hubungan positif ini bisa terjadi karena daerah dengan sanitasi lebih baik juga memiliki pencatatan kesehatan yang lebih akurat, sehingga lebih banyak kasus terdeteksi.

Persentase air bersih (mata air terlindungi) (X_6) memberi kontribusi yang signifikan terhadap jumlah *tuberculosis* di Provinsi Nusa Tenggara Timur, dimana setiap penambahan 1% air bersih (mata air terlindungi) akan menyebabkan kenaikan rata-rata jumlah *tuberculosis* sebesar $\exp(0.007)$ atau 1,00702 dari rata-rata jumlah kasus *tuberculosis* semula. Semakin banyak persentase air bersih (mata air terlindungi) akan meningkatkan jumlah kasus *tuberculosis* di setiap wilayah kabupaten/kota sebanyak 1 kasus. Secara statistik, hubungan positif ini bisa disebabkan oleh kondisi di mana daerah dengan akses air bersih yang lebih baik juga memiliki populasi yang lebih padat, sehingga potensi penularan penyakit lebih tinggi.

Jumlah HIV/AIDS (X_7) memberi kontribusi yang signifikan terhadap jumlah *tuberculosis* di Provinsi Nusa Tenggara Timur, dimana setiap penambahan 1 puskesmas akan menyebabkan kenaikan rata-rata jumlah *tuberculosis* sebesar $\exp(0.015)$ atau 1,01511 dari rata-rata jumlah kasus *tuberculosis* semula. Semakin banyak jumlah kasus HIV/AIDS akan meningkatkan jumlah *tuberculosis* di setiap wilayah kabupaten/kota sebanyak 1 kasus. Hal ini menunjukkan bahwa *tuberculosis* lebih rentan menyerang individu dengan HIV/AIDS karena daya tahan tubuh mereka yang lemah.

4 Kesimpulan

Hasil dari analisis dan pembahasan yang telah dijelaskan dapat diambil kesimpulan sebagai berikut:

1. Model *Generalized Poisson Regression (GPR)* pada kasus *tuberculosis* di Provinsi Nusa Tenggara Timur pada Tahun 2023 yaitu sebagai berikut:
$$\mu = \exp(4.704 + 0.011X_1 - 0.025X_4 + 0.006X_5 + 0.007X_6 + 0.015X_7)$$
2. Faktor-faktor yang berpengaruh signifikan terhadap jumlah kasus *tuberculosis* di Provinsi Nusa Tenggara Timur pada Tahun 2023 berdasarkan model *Generalized Poisson Regression (GPR)* adalah jumlah penduduk miskin (X_1), jumlah puskesmas

(X_4), persentase rumah tangga dengan akses sanitasi yang layak (X_5), persentase air bersih (X_6), dan jumlah HIV dan AIDS(X_7)

Daftar Pustaka

- [1] Getu, H. F. Wolde, Y. Animut, and A. A. Kibret, "Incidence and Predictors of Tuberculosis among Patients enrolled in Anti-Retroviral Therapy after Universal Test and Treat Program, Addis Ababa, Ethiopia. A Retrospective follow -up Study," *PLoS ONE*, vol. 17, no. 8, p. e0272358, 2022, doi: 10.1371/journal.pone.0272358.
- [2] R. Syakur, J. Usman, and H. Asyng, "Faktor yang Berhubungan dengan Kejadian Tuberkulosis (TBC) di Balai Besar Kesehatan Paru Masyarakat Kota Makassar," *JKKM*, vol. 1, no. 1, pp. 17–24, 2019, doi: 10.36090/jkkm.v1i1.284.
- [3] H. Helmy, M. T. Kamaluddin, I. Iskandar, S. Suheryanto, I. Irfannuddin, and N. Novrikasari, "Spatial Modelling of Pulmonary TB Distribution in Indonesia using on Environmental and Socio-economic Variables," in *Proceedings of the 3rd Sriwijaya International Conference on Environmental Issues, SRICOENV 2022, October 5th, 2022, Palembang, South Sumatera, Indonesia*, Palembang, Indonesia: EAI, 2023. doi: 10.4108/eai.5-10-2022.2328785.
- [4] WHO. 2022. Global Tuberculosis Report 2022. Geneva: World Health Organization, Licence: CC BY-NC-SA 3.0 IGO, 2023.
- [5] Sulistya, A., Kurniawan, B. and Pramesona, B.A., Factors Related to the Incidence of Pulmonary TB in Rural Communities in South Lampung, Indonesia: A Case-Control Study. *International Journal of Current Science Research and Review*. Vol.6(12) pp 7457-7463, 2023.
- [6] Kementerian Kesehatan RI., Profil Kesehatan Indonesia. Jakarta: Kementerian Kesehatan RI. 2023.
- [7] Dinkes NTT, Profil Kesehatan Provinsi NTT, Dinas Kesehatan Provinsi NTT. 2022.
- [8] Hayati, M., Sadik, K. and Kurnia, A, Conway-Maxwell Poisson Distribution: Approach for Over-and-Under-Dispersed Count Data Modelling. *IOP Conference Series: Earth and Environmental Science*. 187(1) : 012039, 2018. <https://dx.doi.org/10.1088/1755-1315/187/1/012039>
- [9] M. F. Lais, A. Atti, R. M. Pangaribuan, and R. D. Guntur, "Model Generalized Poisson Regression (GPR) pada Kasus Stunting di Provinsi Nusa Tenggara Timur," *JD*, vol. 5, no. 2, pp. 68–75, Aug. 2023, doi: 10.35508/jd.v5i2.11562.
- [10] Guntur, R. D., & Da Rato, M. R., Generalized Poisson Regression Modeling on the Number of Infant Deaths in East Nusa Tenggara Province in 2022. *J Statistika: Jurnal Ilmiah Teori dan Aplikasi Statistika*, 17(2), 779–788, 2024, <https://jurnal.unipasby.ac.id/index.php/jstatistika/article/view/9318>
- [11] Guntur, R. D., & Dappa, J. S. B., Pemodelan Generalized Poisson Regression pada Kasus AIDS di Provinsi Nusa Tenggara Timur Tahun 2023. *VARIANSI: Journal of Statistics and Its Application on Teaching and Research*, 6(2), 95–104, 2024. <https://jurnalvariansi.unm.ac.id/index.php/variansi/article/view/168>

- [12] Guntur, R. D., & Njudang, C. C. I. A., Pemodelan Generalized Poisson Regression (GPR) untuk Mengatasi Pelanggaran Equidisersi pada Regresi Poisson Kasus Pneumonia di Provinsi NTT. *MATHunesa: Jurnal Ilmiah Matematika*, 13(1), 2025. <https://ejournal.unesa.ac.id/index.php/mathunesa/article/view/62516>
- [13] Esra, R., Nohe, D. A. dan Fathurahman, M. Pemilihan Model Terbaik pada Generalized Poisson Regression menggunakan Akaike Information Criterion. *Statistika*. 23(1) : 73-87, 2023. <https://doi.org/10.29313/statistika.v23i1.1925>
- [14] Y. Tiara, M. N. Aidi, E. Erfiani, and R. Rachmawati, "Overdispersion Handling in Poisson Regression Model by Applying Negative Binomial Regression," *Barekeng: Journal of Mathematics and Its Applications.*, vol. 17, no. 1, pp. 0417–0426, 2023, doi: 10.30598/barekengvol17iss1pp0417-0426.
- [15] M. P. Putri and P. Purhadi, "Analisis Faktor-faktor yang Berpengaruh terhadap Jumlah Kematian Ibu dan Jumlah Kematian Bayi di Provinsi Jawa Tengah dengan Bivariate Generalized Poisson Regression," *JSSITS*, vol. 6, no. 1, pp. 101–107, 2017, doi: 10.12962/j23373520.v6i1.22467.
- [16] M. Majore, D. T. Salaki, and J. D. Prang, "Penerapan Regresi Binomial Negatif dalam Mengatasi Overdispersi Regresi Poisson pada Kasus Jumlah Kematian Ibu," *dC*, vol. 9, no. 2, p. 133, 2021, doi: 10.35799/dc.9.2.2020.29150.
- [17] W. Wang and F. Famoye, "Modeling Household Fertility Decisions with Generalized Poisson regression," *Journal of Population Economics*, vol. 10, no. 3, pp. 273–283, 1997, doi: 10.1007/s001480050043.
- [18] K. Kristy, J. Jajang, and N. Nurhayati, "Analisis Regresi Count Data untuk Pemodelan Jumlah Kasus Penyakit Tuberkulosis di Kabupaten Banyumas," *JMP*, vol. 13, no. 2, p. 57, 2021, doi: 10.20884/1.jmp.2021.13.2.4919.