PEMODELAN REGRESI *HURDLE NEGATIVE BINOMIAL* PADA KASUS DIFTERI DI PULAU KALIMANTAN INDONESIA

SYALVA SETIA ANUGRAH
H1011221023

SKRIPSI

PROGRAM STUDI MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS TANJUNGPURA
PONTIANAK
2026

PEMODELAN REGRESI *HURDLE NEGATIVE BINOMIAL* PADA KASUS DIFTERI DI PULAU KALIMANTAN INDONESIA

SYALVA SETIA ANUGRAH
H1011221023

Skripsi
Sebagai salah satu syarat untuk memperoleh gelar
Sarjana Matematika pada Program Studi Matematika

PROGRAM STUDI MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS TANJUNGPURA
PONTIANAK
2026

LEMBAR PENGESAHAN

Judul Tugas Akhir	: Pemodelan Regresi *Hurdle Negative Binomial* Pada Kasus Difteri di Pulau Kalimantan Indonesia
Nama Mahasiswa	: Syalva Setia Anugrah
NIM	: H1011221023
Jurusan/Program Studi	: Matematika/Matematika
Tanggal Lulus	: 18 Mei 2026
SK Pembimbing	: No.3554/UN22.8/TD.06/2025/ 9 Oktober 2025
SK Penguji	: No.2352/DST/UN22.8/DT.00.09/2026/ 20 Mei 2026

Dosen Pembimbing

Pembimbing I	Pembimbing II
Nur'ainul Miftahul Huda, M.Si.	Fransiskus Fran, M.Si.
NIP 199411142020122014	NIP 198804152019031014

Dosen Penguji

Ketua Penguji	Anggota Penguji
Dr. Bayu Prihandono., M.Sc.	Meliana Pasaribu, S.Pd., M.Sc
NIP 197911152005011003	NIP 198710192019032006

Pimpinan Sidang (merangkap anggota penguji)	Sekretaris Sidang (merangkap anggota penguji)
Nur'ainul Miftahul Huda, M.Si	Fransiskus Fran, M.Si.
NIP 199411142020122014	NIP 198804152019031014

Mengesahkan
Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Tanjungpura

Prof. Dr. Gusrizal, M.Si.
NIP 197108022000031001

PERNYATAAN INTEGRITAS AKADEMIK

Saya yang bertanda tangan di bawah ini:

Nama	: Syalva Setia Anugrah
NIM	: H1011221023
Program Studi/ Jurusan	: Matematika/Matematika
Fakultas	: Matematika dan Ilmu Pengetahuan Alam

dengan ini menyatakan bahwa dokumen ilmiah Tugas Akhir yang disajikan ini tidak mengandung unsur pelanggaran integritas akademik sesuai Peraturan Menteri Pendidikan, Kebudayaan, Riset, Dan Teknologi Republik Indonesia Nomor 39 Tahun 2021. Apabila di kemudian hari dokumen ilmiah Tugas Akhir ini mengandung unsur pelanggaran integritas akademik sesuai ketentuan perundangan tersebut, maka saya bersedia menerima sanksi akademik dan/atau sanksi hukum yang berlaku.

Demikian pernyataan ini untuk digunakan sebagaimana mestinya.

Pontianak, 18 Mei 2026



Syalva Setia Anugrah

NIM H1011221023

Pemodelan Regresi *Hurdle Negative Binomial*
Pada Kasus Difteri Di Pulau Kalimantan
Indonesia

Abstrak

Penyebaran kasus difteri di Pulau Kalimantan tahun 2023 menunjukkan ketidakmerataan antarwilayah, yang ditandai dengan banyaknya kabupaten/kota yang mencatatkan nol kasus, sementara wilayah lainnya masih mencatatkan jumlah kasus yang relatif tinggi. Kondisi ini mengindikasikan data bertipe cacah (*count data*) dengan karakteristik *overdispersi* dan proporsi nilai nol yang tinggi (*excess zeros*), sehingga memerlukan model yang mampu memisahkan proses munculnya nilai nol serta jumlah kasus ketika bernilai lebih dari nol. Berdasarkan kondisi tersebut, digunakan model regresi *Hurdle Negative Binomial* yang terdiri atas dua komponen, yaitu *Zero Hurdle* untuk memodelkan peluang terjadinya kasus dan *Truncated negative binomial* untuk memodelkan banyaknya kasus pada nilai lebih dari nol. Penelitian ini bertujuan untuk menganalisis model regresi *Hurdle Negative Binomial* serta mengidentifikasi faktor yang memengaruhi terjadinya kasus difteri di Pulau Kalimantan. Data yang digunakan merupakan data sekunder yang diperoleh dari Badan Pusat Statistik dan Dinas Kesehatan dengan variabel dependen berupa banyaknya kasus difteri (Y) serta variabel independen meliputi Cakupan Imunisasi Dasar Lengkap (X_1), Persentase Air Minum Layak (X_2), Persentase Penduduk Miskin (X_3), Persentase Balita Gizi Kurang (X_4), dan Persentase Indeks Pembangunan Manusia (X_5). Hasil penelitian menunjukkan bahwa Cakupan Imunisasi Dasar Lengkap (X_1), Persentase Penduduk Miskin (X_3), dan Persentase Indeks Pembangunan Manusia (X_5) berpengaruh signifikan terhadap banyaknya kasus difteri di Pulau Kalimantan tahun 2023. Dengan demikian, model regresi *Hurdle Negative Binomial* mampu memberikan pemodelan yang sesuai dengan karakteristik data dan dapat digunakan sebagai dasar kuantitatif dalam perumusan kebijakan pengendalian difteri di Pulau Kalimantan.

Kata kunci: Excess Zeros, Overdispersi, Data Cacah

Diphtheria Case Modeling In Kalimantan Island Indonesia Using Hurdle Negative Binomial Regression

Abstract

The distribution of diphtheria cases in Kalimantan in 2023 shows substantial disparities across regions, as indicated by the large number of districts/cities reporting zero cases, while others still record relatively high case counts. This condition suggests that the data are count data characterized by overdispersion and a high proportion of zero values (*excess zeros*), thus requiring a modeling approach capable of separating the process generating zero outcomes from the process governing the number of cases when the values exceed zero. Based on these characteristics, the Hurdle Negative Binomial regression model is employed, consisting of a component for modeling the probability of occurrence (Zero Hurdle) and a component for modeling positive counts (Truncated Negative Binomial). This study aims to analyze the Hurdle Negative Binomial regression model and identify factors influencing the occurrence of diphtheria cases in Kalimantan. The data used are secondary data obtained from the Central Bureau of Statistics and the Health Office, with the dependent variable being the number of diphtheria cases (Y), and the independent variables include of Full Basic Immunization Coverage (X_1), Percentage of Access to Safe Drinking Water (X_2), Poverty Rate (X_3), Percentage of Underweight Toddlers (X_4), and Human Development Index (X_5).The results show that Full Basic Immunization Coverage (X_1), Poverty Rate (X_3), and Human Development Index (X_5) have a significantly effect the number of diphtheria cases in Kalimantan in 2023. Thus, the Hurdle Negative Binomial regression model provides is appropriate for this dataset and can be used as a quantitative basis for formulating diphtheria control policies in Kalimantan.

Keywords: Excess Zeros, Overdispersion, Count Data.

PRAKATA

Puji dan syukur kehadirat Tuhan Yang Maha Esa atas segala Rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul "Pemodelan Regresi *Hurdle Negative Binomial* pada Kasus Difteri di Pulau Kalimantan Indonesia" sebagai salah satu syarat untuk memperoleh gelar Sarjana Matematika pada Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Tanjungpura.

Penulisan skripsi ini tidak terlepas dari bantuan dan dukungan berbagai pihak. Oleh karena itu, penulis ingin menyampaikan terima kasih kepada:

1. Kedua orang tua, Bapak, Mama, dan Kakak, penulis menyampaikan terima kasih atas segala bentuk kasih sayang, dukungan, semangat, serta doa yang tiada henti diberikan selama perkuliahan dan penyelesaian skripsi.
2. Ibu Nur'ainul Miftahul Huda, S.Si., M.Si dan Bapak Fransiskus Fran, S.Si., M.Si selaku dosen pembimbing yang telah memberikan bimbingan dan motivasi dalam proses penyelesaian skripsi.
3. Bapak Dr. Bayu Prihandono., M.Sc dan Ibu Meliana Pasaribu, S.Pd., M.Sc selaku dosen penguji yang telah memberikan motivasi serta arahan yang sangat berharga selama proses penulisan skripsi.
4. Seluruh dosen Jurusan Matematika yang sudah mendidik dan memberikan ilmu yang bermanfaat selama masa perkuliahan.
5. Teman terdekat penulis yang telah memberikan semangat dan mendengarkan keluh kesah penulis selama proses perkuliahan dan penyelesaian skripsi.
6. Semua pihak yang tidak dapat disebutkan satu per satu, yang telah memberikan bantuan dan dukungan secara langsung maupun tidak langsung.

Pontianak, 18 Mei 2026

Syalva Setia Anugrah

DAFTAR ISI

LEMBAR PENGESAHAN ..... **ii**
PERNYATAAN INTEGRITAS AKADEMIK ..... **iii**
Abstrak ..... **iv**
Abstract ..... **v**
PRAKATA ..... **vi**
DAFTAR ISI ..... **vii**
DAFTAR GAMBAR ..... **ix**
DAFTAR TABEL ..... **x**
DAFTAR LAMPIRAN ..... **xi**
DAFTAR SIMBOL ..... **xii**
BAB 1 PENDAHULUAN **1**
1. 1 Latar Belakang 1
1. 2 Rumusan Masalah 3
1. 3 Tujuan Penelitian 4
1. 4 Batasan Masalah 4
1.5 Tinjauan Pustaka 4
1.6 Metodologi Penelitian 6
BAB II LANDASAN TEORI **9**
2.1 Difteri 9
2.2 Statistika Deskriptif 11
2.3 Teori Probabilitas 13
2.4 Distribusi Probabilitas 14
2.4.1 Variabel Acak 14
2.4.2 Jenis-jenis Distribusi Probabilitas Diskrit 16
2.5 Regresi *Negative Binomial* 19
2.6 *Excess Zeros* 20
2.7 *Overdispersi* 21
2.8 Uji Multikolinearitas 22
BAB III MODEL *REGRESI HURDLE NEGATIVE BINOMIAL* **24**
3.1 Regresi *Hurdle Negative Binomial* 24

3.1.1 Model *Hurdle* 24
3.1.2 Model *Hurdle Negative Binomial* 25
3. 2 Estimasi Parameter 26
3. 3 Uji Signifikansi Parameter 28
3.3.1 Uji Parsial 28
3.3.2 Uji Simultan 29
3. 4 *Akaike's Information Criterion* (AIC) 30
BAB IV HASIL DAN PEMBAHASAN 31
4. 1 Analisis Deskriptif 32
4. 2 Model Regresi *Hurdle Negative Binomial* 41
4. 3 Uji Multikolinearitas 41
4. 4 Uji *Overdispersi* 43
4. 5 Uji *Excess Zeros* 44
4. 6 Pemodelan *Regresi Hurdle Negative Binomial* 44
4.6.1 Estimasi Parameter 45
4.6.2 Uji Signifikansi Parameter 46
4. 7 Model Terbaik 49
4. 8 Hasil dan Interpretasi Model 50
BAB V PENUTUP 54
5. 1 Kesimpulan 54
5. 2 Saran 55
DAFTAR PUSTAKA 56
LAMPIRAN 61

DAFTAR GAMBAR

Gambar 1.1 *Flowchart* model regresi *Hurdle Negative Binomial*..... 8
Gambar 4.1 Persebaran Kasus Difteri pada Kabupaten/Kota di Pulau Kalimantan pada Tahun 2023.....35
Gambar 4.2 Banyaknya Kasus Difteri Antarprovinsi..... 36
Gambar 4.3 Persentase Kabupaten/Kota dengan Kasus dan Tanpa kasus di Pulau Kalimantan..... 36
Gambar 4.4 Persebaran Difteri pada Provinsi di Pulau Kalimantan dengan notasi *i* pada **Tabel 4.1** 40
Gambar 4.5 Simulasi Pengaruh (a). Cakupan Imunisasi Dasar (b) Penduduk Miskin (c) Indeks Pembangunan Manusia..... 51
Gambar 4.6 Simulasi Pengaruh Indeks Pembangunan Manusia 52

DAFTAR TABEL

Tabel 4.1 Data banyaknya kasus Difteri di Pulau Kalimantan Tahun 2023 31
Tabel 4.2 Statistik Deskriptif..... 32
Tabel 4.3 Analisis Deskriptif Per Provinsi..... 37
Tabel 4.4 Uji Multikolinearitas 43
Tabel 4.5 Uji Asumsi *Overdispersi*..... 43
Tabel 4.6 Uji *Excess zeros* 44
Tabel 4.7 Estimasi Parameter model regresi *Hurdle Negative Binomial*..... 45
Tabel 4.8 Estimasi Parameter Variabel Signifikan 46
Tabel 4.9 Estimasi Parameter Signifikan 47
Tabel 4.10 Model Terbaik..... 50

DAFTAR LAMPIRAN

Lampiran 1 Data Banyaknya Kasus Difteri dan variabel indikatornya 61
Lampiran 2 *Syntax* Statistika Deskriptif dan Uji asumsi model..... 64
Lampiran 3 *Syntax* Estimasi Parameter..... 65
Lampiran 4 *Syntax* Simulasi Hasil Model..... 66

DAFTAR SIMBOL

$\bar{X}$: Rata-rata sampel

X_i : Nilai variabel independen untuk observasi ke-i

i : $1,2,\dots,n$

s^2 : Varians sampel

σ^2 : Varians Populasi

n : Jumlah data observasi

N : Jumlah Populasi

$\Gamma(\phi^{-1})$: Fungsi Gamma dari invers parameter dispersi

$\Gamma(y+\phi^{-1})$: Fungsi Gamma dari penjumlahan antara nilai y dan invers parameter dispersi

ϕ : Parameter dispersi

$L(\phi,\beta,\gamma)$: Fungsi *likelihood* pada parameter model

$\ln L(\phi,\beta,\gamma)$: Fungsi *log-likelihood* pada parameter model

$\hat{\phi}$: Estimasi Parameter Dispersi

Y_i : Banyaknya kasus difteri pada observasi ke-i

k : Jumlah parameter yang diestimasi dalam model

X_{ij} : Nilai variabel ke-j untuk observasi ke-i

$\hat{X}_{ij}$: Nilai prediksi untuk observasi ke-i pada variabel ke-j

β_0 : Intercept pada komponen *Truncated Negative Binomial*

β_j : koefisien regresi variabel ke-j, dengan $j = 1,2,3,4,5$

μ_i : Rata-rata kasus difteri pada observasi ke-i

γ_0 : Intercept pada komponen *Zero Hurdle*

γ_j : koefisien regresi variabel ke-j, dengan $j = 1,2,3,4,5$

π_i : Peluang banyak kasus difteri ke-i

$\exp(X_i^T\beta)$: Fungsi Eksponensial dari linear predictor pada model regresi komponen *Truncated*

BAB 1
PENDAHULUAN

1. 1 Latar Belakang

Kesehatan merupakan salah satu aspek fundamental dalam meningkatkan kualitas hidup masyarakat sekaligus mendukung keberlangsungan pembangunan suatu bangsa (Pradono & Sulistyowati, 2014). Meskipun berbagai upaya peningkatan kesehatan telah dilakukan di Indonesia, penyakit menular masih menjadi salah satu tantangan utama kesehatan masyarakat (Simbolon dkk., 2025). Penyakit menular memiliki kemampuan penyebaran yang cepat serta berpotensi menimbulkan kejadian luar biasa dan kematian apabila tidak dikendalikan dengan baik (Masriadi & Km, 2017). Salah satu penyakit menular yang hingga kini masih menjadi perhatian penting adalah difteri (Simbolon dkk., 2025).

Difteri merupakan infeksi akut yang disebabkan oleh bakteri *Corynebacterium diphtheriae*, yang menyerang saluran pernapasan bagian atas seperti tonsil, nasofaring, dan laring (Kemenkes RI, 2023; Saputra, 2018). Penyakit ini menular melalui droplet saat penderita batuk atau bersin, sehingga mudah menyebar terutama pada kelompok dengan daya tahan tubuh rendah (Pattipeilohy dkk., 2025). Apabila tidak ditangani secara tepat, difteri dapat menimbulkan komplikasi serius hingga kematian, khususnya pada anak-anak (Kemenkes RI, 2023; Mustikasari dkk., 2025). Meskipun program imunisasi telah terbukti menurunkan angka kejadian difteri secara signifikan sejak diterapkan secara luas, penyakit ini masih ditemukan di sejumlah wilayah, terutama daerah berkembang dengan cakupan imunisasi yang belum merata serta keterbatasan sistem kesehatan (Mahdang dkk., 2022; Pattipeilohy dkk., 2025; WHO, 2009).

Hingga saat ini, difteri masih menjadi ancaman kesehatan masyarakat di Indonesia (Achjar dkk., 2024). Selama periode 2017-2021 tercatat sebanyak 2.969 kasus difteri dengan puncak kejadian pada tahun 2017 sebanyak 954 kasus dan 44 kematian (Kemenkes RI, 2023). Pada tahun-tahun berikutnya, jumlah kasus dan kematian akibat difteri kembali menunjukkan peningkatan (Nugrahaeni, 2024). Kondisi ini mengindikasikan bahwa risiko penularan difteri masih cukup tinggi,

sehingga diperlukan penguatan upaya pengendalian melalui cakupan imunisasi serta optimalisasi surveilans kesehatan masyarakat (Mahdang dkk., 2022).

Kondisi alam dan karakteristik geografis Indonesia turut memengaruhi penyebaran kasus difteri (Pracoyo, 2020; Salsabila dkk., 2024). Indonesia merupakan negara kepulauan yang terletak di wilayah beriklim tropis dengan suhu dan tingkat kelembapan yang relatif tinggi, sehingga berpotensi mendukung penyebaran penyakit menular, terutama di wilayah dengan kepadatan penduduk tinggi dan kondisi lingkungan yang kurang mendukung kesehatan (Achjar dkk., 2024). Selain itu, perbedaan kondisi geografis antarwilayah, seperti keterpencilan daerah, perbedaan topografi, serta keterbatasan akses transportasi, berdampak pada ketidakmerataan akses dan kualitas layanan kesehatan (Muta'ali, 2024). Hal ini menyebabkan pola penyebaran kasus difteri yang tidak merata antarwilayah di Indonesia. Pulau Kalimantan merupakan salah satu wilayah yang mencerminkan dampak kondisi geografis tersebut. Dengan wilayah daratan yang luas dan akses antarwilayah yang belum merata, distribusi kasus difteri di Kalimantan menunjukkan pola yang beragam. Pada beberapa kabupaten/kota, kasus difteri dilaporkan sangat rendah bahkan tidak ditemukan, sementara di wilayah lain masih mencatat jumlah kasus yang relatif lebih tinggi (Dinkes Kalbar, 2023; Dinkes Kalsel, 2023; Dinkes Kaltara, 2023; Dinkes Kalteng, 2023; Dinkes Kaltim, 2023). Ketidakseimbangan ini dipengaruhi oleh kombinasi faktor lingkungan, sosial, serta keterjangkauan layanan kesehatan, termasuk pelaksanaan imunisasi dan ketepatan pelaporan kasus (Firmansyah & Danardono, 2025). Akibatnya, data kejadian difteri di Pulau Kalimantan didominasi oleh nilai nol dan menunjukkan variasi jumlah kasus yang cukup tinggi antarwilayah (Kemenkes RI, 2023).

Kondisi tersebut mengindikasikan bahwa data jumlah kasus difteri di Pulau Kalimantan merupakan data bertipe *count* yang cenderung mengalami overdispersi serta mengandung *excess zeros*. Namun, sebagian besar penelitian sebelumnya masih menggunakan metode analisis statistik yang umum, seperti regresi *Poisson*, yang mengasumsikan mean dan variansi sama (Ginting, 2018). Asumsi ini sering kali tidak terpenuhi pada data kasus penyakit sehingga dapat menghasilkan estimasi yang kurang tepat. Penelitian Mustafidah & Sofro (2024) menjelaskan bahwa pada

kasus kematian akibat filariasis di Jawa Timur menunjukkan bahwa data penyakit di dominasi nilai nol dan variasi kasus yang tinggi memerlukan pendekatan pemodelan yang mampu memisahkan proses pembentukan nilai nol dan nilai positif. Karakteristik tersebut sejalan dengan kondisi data difteri di Pulau Kalimantan, sehingga diperlukan pendekatan pemodelan yang sesuai. Salah satu pendekatan yang dirancang khusus untuk menangani data dengan karakteristik *excess zeros* dan overdispersi adalah model *Hurdle Negative Binomial* (HNB) (Rahayu dkk., 2018). Model ini memisahkan proses pembentukan data menjadi dua tahap, yaitu komponen *Zero Hurdle* memodelkan peluang suatu wilayah memiliki kasus, dan komponen *Truncated Negative Binomial* untuk jumlah kasus pada wilayah yang mengalami kejadian difteri (Aida, 2021; Rahayu dkk., 2018). Pendekatan ini memungkinkan analisis yang lebih mendalam terhadap faktor-faktor yang memengaruhi kejadian difteri baik dari sisi peluang terjadinya kasus maupun dari sisi banyaknya jumlah kasus (Hilbe, 2011). Namun, hingga saat ini penerapan model *Hurdle Negative Binomial* untuk menganalisis kasus difteri, khususnya pada wilayah dengan karakteristik geografis seperti Pulau Kalimantan masih sangat terbatas dalam literatur ilmiah.

Berdasarkan karakteristik tersebut, model *Hurdle Negative Binomial* dinilai relevan untuk menganalisis faktor-faktor yang memengaruhi banyaknya kasus difteri di Pulau Kalimantan. Oleh karena itu, penelitian ini menerapkan model *Hurdle Negative Binomial* pada data kasus difteri di Pulau Kalimantan tahun 2023. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi baru mengenai pengendalian difteri, terutama dalam menangani ketidakmerataan penyebaran kasus dan tantangan dalam surveilans kesehatan di wilayah Kalimantan.

1. 2 Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan, maka permasalahan yang dapat dibahas yaitu bagaimana model regresi *Hurdle Negative Binomial* pada data kasus difteri di Pulau Kalimantan dan indikator apa saja yang mempengaruhi terjadinya kasus difteri di Pulau Kalimantan.

1. 3 Tujuan Penelitian

Berdasarkan rumusan masalah tersebut, maka tujuan dari penelitian ini adalah memodelkan regresi *Hurdle Negative Binomial* pada data kasus difteri di Pulau Kalimantan dan menganalisis variabel indikator apa saja yang mempengaruhi terjadinya kasus difteri di Pulau Kalimantan.

1. 4 Batasan Masalah

Batasan masalah sangat diperlukan untuk memastikan bahwa kesimpulan yang akan diperoleh dapat dipertanggungjawabkan. Dengan demikian, batasan masalah dalam penelitian ini adalah sebagai berikut:

a. Data yang digunakan merupakan data sekunder yang bersumber dari publikasi resmi Dinas kesehatan dan Badan Pusat Statistik.
b. Penelitian ini menggunakan data banyaknya kasus difteri di Pulau Kalimantan pada Tahun 2023
c. Variabel independen dalam penelitian difokuskan pada Cakupan Imunisasi dasar lengkap, air minum layak, penduduk miskin, balita gizi kurang, dan indeks pembangunan manusia, sehingga variabel lain tidak digunakan.
d. Model yang digunakan pada penelitian ini adalah regresi *Hurdle Negative Binomial*.

1.5 Tinjauan Pustaka

Untuk mendukung penelitian ini, maka akan disajikan hasil penelitian terdahulu yang relevan sebagai acuan. Penelitian-penelitian terdahulu memberikan gambaran umum mengenai topik yang diteliti serta menjadi dasar dalam memahami variabel-variabel yang berhubungan. Selain itu, kajian hasil penelitian sebelumnya juga membantu dalam mengidentifikasi kesenjangan penelitian yang masih ada, sehingga penelitian ini dapat memberikan kontribusi baru dalam bidang yang sama. Dengan demikian, pembahasan mengenai penelitian terdahulu menjadi bagian penting untuk memperkuat analisis dan memperjelas arah penelitian yang dilakukan.

Penelitian terdahulu mengenai pemodelan *Hurdle Negative Binomial* dilakukan oleh Mustafidah & Sofro (2024). Penelitian tersebut membahas tentang pemodelan jumlah kematian akibat filariasis di Jawa Timur. Hasil dari penelitian menunjukkan bahwa variabel independen berpengaruh secara simultan. Kemudian, uji signifikansi parsial mengidentifikasi beberapa faktor yang berpengaruh, yaitu persentase penduduk miskin, persentase sanitasi layak, persentase air bersih, jumlah sarana kesehatan, persentase indeks pembangunan manusia, serta persentase keluhan kesehatan. Dengan demikian, penelitian ini menguatkan bahwa faktor sosial-ekonomi dan lingkungan tersebut merupakan faktor yang menyebabkan angka kematian akibat filariasis di Jawa Timur.

Penelitian tentang kasus Difteri ini pernah dilakukan oleh Tasya Azzahra (2023) dengan tujuan untuk menganalisis faktor-faktor yang mempengaruhi jumlah kasus difteri di Jawa Barat pada tahun 2020 menggunakan model regresi *Hurdle Negative Binomial*. Hasil dari penelitian ini, mengidentifikasi rendahnya Cakupan Imunisasi dasar lengkap dan tingginya jumlah balita gizi kurang variabel yang berpengaruh signifikan. Selain itu, penelitian tersebut juga menyimpulkan bahwa model regresi *Hurdle Negative Binomial* sangat sesuai untuk data penyakit seperti difteri karena mampu menangani overdispersi dan kelebihan nilai nol (*excess zeros*) dalam data.

Purnama (2021) melakukan penelitian untuk memodelkan konsumsi rokok di Provinsi Sulawesi Tengah. Penelitian ini membandingkan beberapa pendekatan pemodelan, yaitu regresi *Poisson*, *Zero Inflated Poisson* (ZIP), dan *Hurdle Negative Binomial* (HNB). Berdasarkan kriteria uji, regresi *Hurdle Negative Binomial* merupakan model terbaik karena lebih sesuai dengan karakteristik data konsumsi rokok yang mengalami overdispersi. Hasil penelitian ini, mengindikasi beberapa faktor yang berpengaruh secara signifikan, yaitu tingkat pendapatan rumah tangga, harga rokok, dan variabel lainnya seperti usia serta pendidikan.

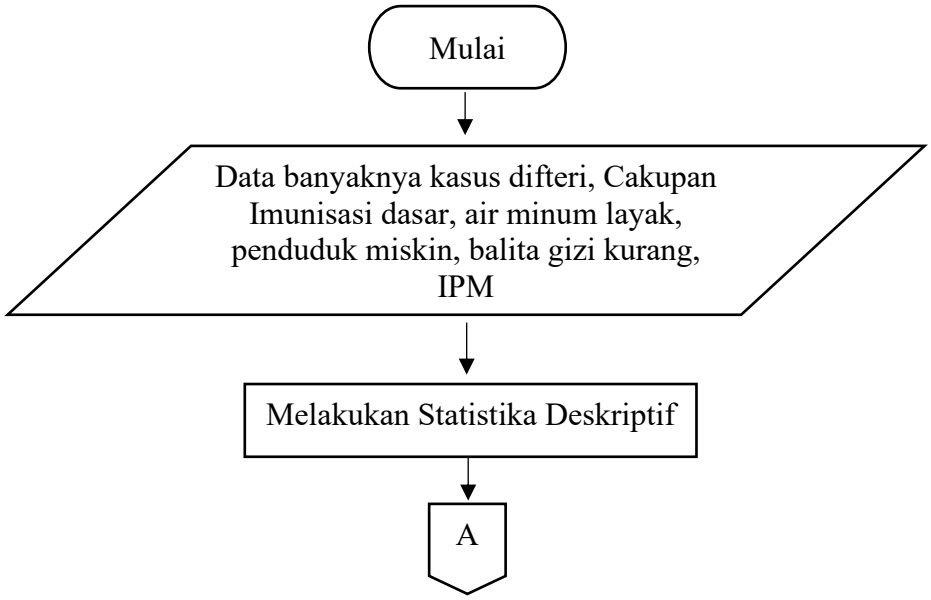
Berdasarkan pada penelitian-penelitian terdahulu, dapat disimpulkan bahwa model *Hurdle Negative Binomial* telah terbukti efektif untuk memodelkan data yang mengandung overdispersi dan *excess zeros*. Hasil dari penelitian ini menunjukkan bahwa faktor sosial-ekonomi, lingkungan, dan cakupan kesehatan merupakan faktor penting. Namun, penelitian terdahulu masih terbatas pada satu

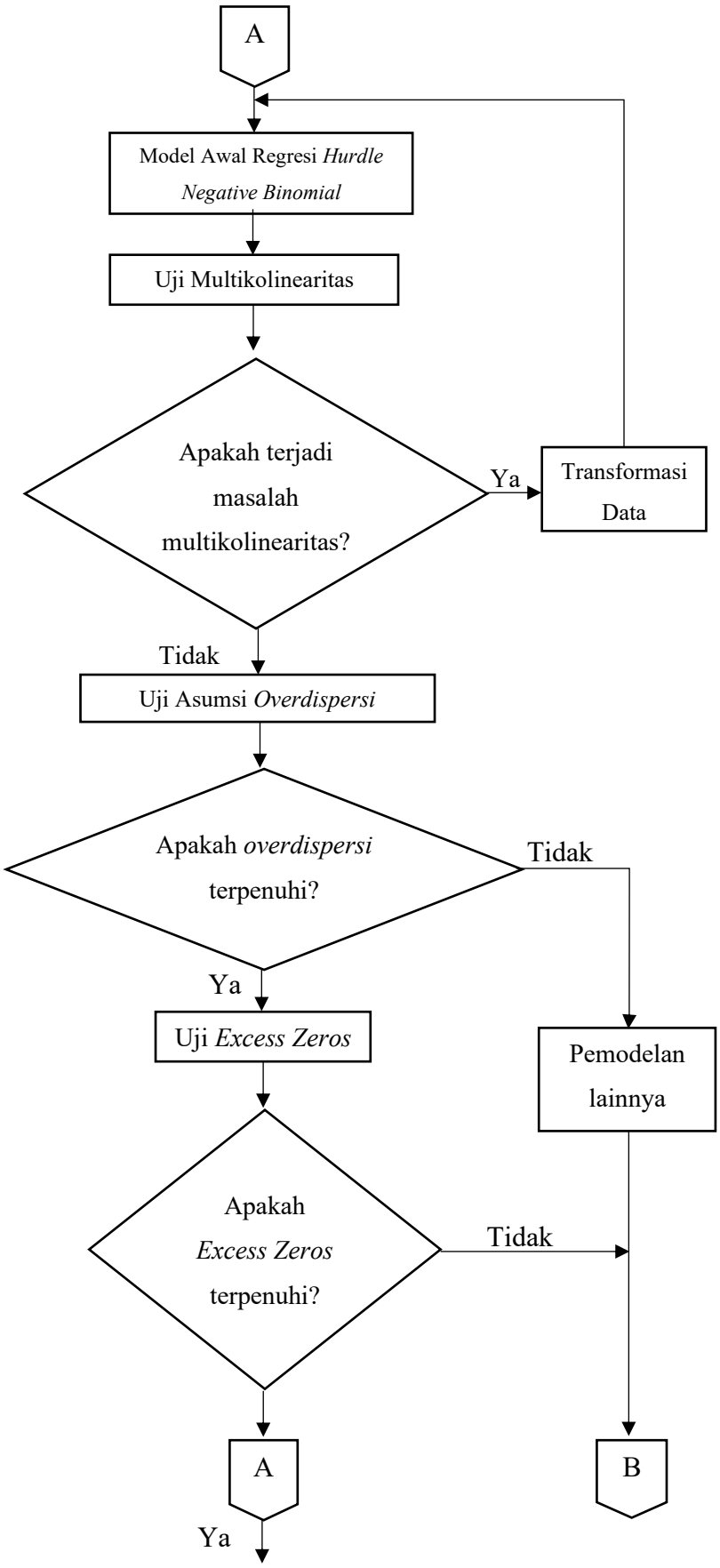
wilayah tertentu, serta belum menguji pengaruh variabel independen dalam konteks dinamika kejadian difteri di suatu provinsi. Oleh karena itu, penelitian ini akan menganalisis faktor-faktor yang mempengaruhi banyaknya kasus difteri di Pulau Kalimantan Indonesia pada Tahun 2023 dengan menggunakan pendekatan regresi *Hurdle Negative Binomial* dan memperluas cakupan variabel independen guna memberikan pemahaman serta lebih tepat sasaran.

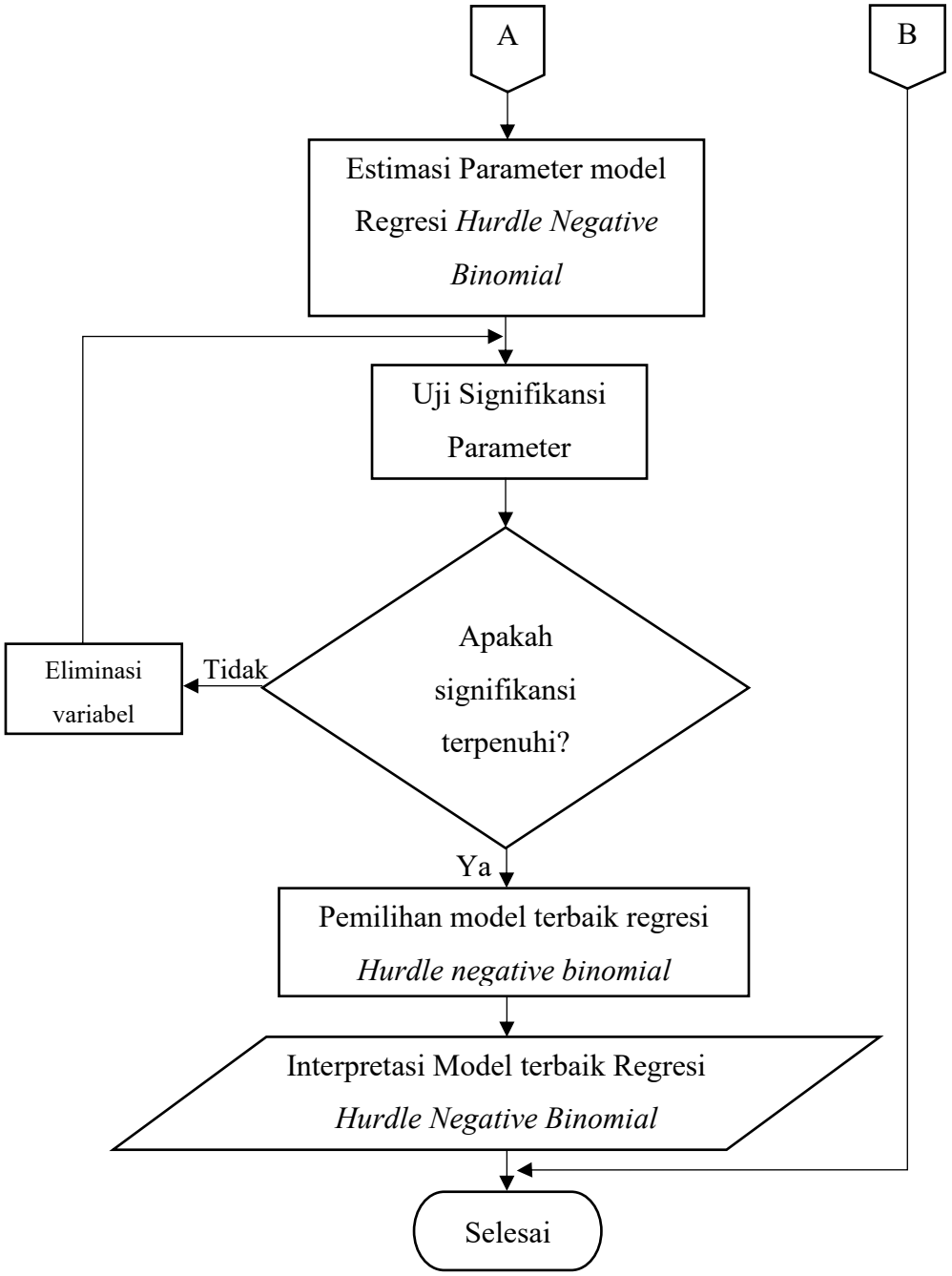
1.6 Metodologi Penelitian

Penelitian ini tentang pemodelan analisis regresi *Hurdle Negative Binomial*, metode penelitian yang digunakan merupakan metode yang melakukan kajian teoritis tentang regresi *Hurdle Negative Binomial* dan menerapkan data kasus Difteri di Pulau Kalimantan menggunakan regresi *Hurdle Negative Binomial*. Penelitian ini menggunakan data sekunder yaitu data banyaknya kasus difteri di Pulau Kalimantan Indonesia.

Variabel dependen yang digunakan merupakan variabel banyaknya kasus difteri Pulau Kalimantan Indonesia dan variabel independen yang digunakan adalah persentase cakupan imunisasi dasar, persentase air minum layak, persentase penduduk miskin, persentase balita gizi kurang, dan persentase Indeks Pembangunan Manusia (IPM). Terdapat beberapa tahap yang dilakukan yaitu:




A
Model Awal Regresi *Hurdle Negative Binomial*
Uji Multikolinearitas
Apakah terjadi masalah multikolinearitas?
Ya
Transformasi Data
Tidak
Uji Asumsi *Overdispersi*
Apakah *overdispersi* terpenuhi?
Tidak
Ya
Uji *Excess Zeros*
Pemodelan lainnya
Apakah *Excess Zeros* terpenuhi?
Tidak
A
Ya
B



Gambar 1.1 *Flowchart* model regresi *Hurdle Negative Binomial*