ANALISIS FAKTOR PENGARUH JUMLAH KASUS HIV DAN AIDS DI INDONESIA DENGAN METODE *BIVARIATE GENERALIZED POISSON REGRESSION*

ANDI SYAIFUL RAHMAT
H05120106

PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR
2024

ANALISIS FAKTOR PENGARUH JUMLAH KASUS HIV DAN AIDS DI INDONESIA DENGAN METODE *BIVARIATE GENERALIZED POISSON REGRESSION*

ANDI SYAIFUL RAHMAT
H051201060

PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR
2024

ANALISIS FAKTOR PENGARUH JUMLAH KASUS HIV DAN AIDS DI INDONESIA DENGAN METODE *BIVARIATE GENERALIZED POISSON REGRESSION*

ANDI SYAIFUL RAHMAT
H051201060

Skripsi

Diajukan sebagai salah satu syarat untuk mencapai gelar Sarjana Statistika

Program Studi Statistika

pada

PROGRAM STUDI STATISTIKA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
MAKASSAR
2024

SKRIPSI

ANALISIS FAKTOR PENGARUH JUMLAH KASUS HIV DAN AIDS DI INDONESIA DENGAN METODE *BIVARIATE GENERALIZED POISSON REGRESSION*

ANDI SYAIFUL RAHMAT
H051201060

Skripsi,

telah dipertahankan di hadapan Panitia Ujian Sarjana Statistika pada 20 November 2024 dan dinyatakan telah memenuhi syarat kelulusan

pada

Program Studi Statistika
Departemen Statistika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Hasanuddin
Makassar

Mengesahkan:	Mengetahui:
Pembimbing tugas akhir,	Ketua Program Studi,
Andi Kresna Jaya, S.Si., M.Si.	Dr. Anna Islamiyati, S.Si., M.Si.
NIP. 19731228 200003 1 001	NIP. 19770808 200501 2 002

PERNYATAAN KEASLIAN SKRIPSI
DAN PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa, skripsi berjudul "Analisis Faktor Pengaruh Jumlah Kasus HIV dan AIDS di Indonesia dengan Metode *Bivariate Generalized Poisson Regression*" adalah benar karya saya dengan arahan pembimbing skripsi saya (Andi Kresna Jaya, S.Si., M.Si.). Karya ilmiah ini belum diajukan dan tidak sedang diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka skripsi ini. Apabila di kemudian hari terbukti atau dapat dibuktikan bahwa sebagian atau keseluruhan skripsi ini adalah karya orang lain, maka saya bersedia menerima sanksi atas perbuatan tersebut berdasarkan aturan yang berlaku.

Dengan ini saya melimpahkan hak cipta (hak ekonomis) dari karya tulis saya berupa skripsi ini kepada Universitas Hasanuddin.

Makassar, 20 November 2024

Andi Syaiful Rahmat

NIM. H051201060

UCAPAN TERIMA KASIH

Bismillahirrahmanirrahim. *Alhamdulillaahirabbil'aalamiin* puji syukur penulis panjatkan kepada Allah *Subhanahu Wa Ta'ala* atas kebesaran-Nya kemudian semua nikmat yang diberikan sehingga penulis dapat menyelesaikan skripsi dengan judul "**Analisis Faktor Pengaruh Jumlah Kasus HIV dan AIDS di Indonesia dengan Metode *Bivariate Generalized Poisson Regression***". Selawat dan salam semoga senantiasa terucapkan kepada Rasulullah Muhammad *Shallallahu 'Alaihi Wa Sallam,* panutan sekaligus teladan umat dalam kehidupan dunia dan akhirat. Semoga kita termasuk golongan yang mendapat syafaatnya *Ilaa Yaumil Qiyamah*, *Aamiin*.

Penulis menyampaikan ucapan terima kasih, *Syukran Wa Jazakallahu Khayran* kepada Bapak **Andi Kresna Jaya, S.Si., M.Si.**, selaku Pembimbing Utama yang senantiasa meluangkan waktu dalam kesibukannya, membimbing dan mendampingi penulis dalam penyelesaian skripsi ini. Terima kasih kepada Ibu **Prof. Dr. Dr. Georgina Maria Tinungki, M.Si.** dan Bapak **Drs. Raupong, M.Si.**, selaku Tim Penguji atas saran dan kritikan yang berharga serta waktu yang telah diberikan kepada penulis. Terima kasih juga kepada Pimpinan Universitas Hasanuddin beserta Jajarannya, Ketua Departemen Statistika, Jajaran Dosen dan Staf Departemen Statistika atas ilmu yang bermanfaat serta fasilitas yang diberikan selama penulis menempuh studi.

Penghargaan setinggi-tingginya dan ucapan terima kasih serta doa terbaik penulis ucapkan kepada orang tua tercinta, Bapak **Patoari** dan Ibu **Andi Nurjannah, S.Pd., SD.**, yang selalu berjuang dalam mengupayakan yang terbaik untuk penulis dan telah memberikan dukungan, pengorbanan, kasih sayang, serta doa yang senantiasa mengiringi langkah penulis. Terima kasih kepada saudara-saudara penulis beserta keluarga besar lainnya yang selalu memberikan dukungan kepada penulis. Terima kasih kepada **Indah Saputri**, sebagai orang terkasih selama masa perkuliahan dan untuk selamanya *insyaa Allah.* Terima kasih kepada teman-teman Statistika 2020 atas momen, dukungan, hingga ilmu yang diberikan kepada penulis, *Jazakumullahu Khayran*. Terima kasih kepada **YBM BRILiaN RO Makassar** atas bantuan beasiswa dan semua yang telah diberikan selama perkuliahan. Kepada Mentor, Kepala Asrama, dan saudaraku *awardee* **BRIGHT Scholarship Batch 6 Unhas (Fadhil, Faldi, Edward, Taufik, Haikal, Lutfi, Ahlul, Zul, Wahid, Yogi, Dim,** dan **Aksa)**, penulis ucapkan terima kasih, *Jazakumullahu Khayran wa Barakallahu Fiikum*. Terima kasih kepada semua pihak yang tidak dapat disebutkan satu-persatu, semoga Allah membalas kebaikan kalian dengan kebaikan yang lebih baik. Penulis menyadari bahwa masih banyak kekurangan dalam penyusunan tugas akhir ini, untuk itu dengan segala kerendahan hati penulis memohon maaf. Akhir kata, semoga tulisan ini dapat memberikan manfaat untuk berbagai pihak.

Penulis,
Andi Syaiful Rahmat

ABSTRAK

ANDI SYAIFUL RAHMAT. **Analisis Faktor Pengaruh Jumlah Kasus HIV dan AIDS di Indonesia dengan Metode *Bivariate Generalized Poisson Regression*** (dibimbing oleh Andi Kresna Jaya).

Latar belakang. *Generalized Linier Model* (GLM) merupakan perluasan dari model regresi linier klasik yang mampu memodelkan data diskrit yang tidak berdistribusi normal, salah satunya regresi *Poisson*. Dalam regresi *Poisson* sering terjadi pelanggaran asumsi yaitu overdispersi yang dapat diatasi dengan model diperumum seperti regresi *Generalized Poisson* (GPR). Sedangkan untuk sepasang data diskrit *Poisson* yang mengalami overdispersi dapat dimodelkan dengan regresi *Bivariate Generalized Poisson* (BGPR). Salah satu data diskrit yang diasumsikan berdistribusi *Poisson* adalah jumlah kasus *Human Immunodeficiency Virus* (HIV) dan *Acquired Immunodeficiency Syndrome* (AIDS). Penyakit HIV dan AIDS merupakan indikator penting dalam menentukan tingkat kesehatan masyarakat dan termasuk dalam target *Suistanable Development Goals* (SDGs) poin ketiga. Indonesia berada pada peringkat ketiga dengan jumlah kasus HIV/AIDS di Asia Tenggara pada 2022 dan terus mengalami peningkatan selama 10 tahun terakhir sehingga dianalisis menggunakan model BGPR. **Tujuan**. Penelitian ini bertujuan untuk memperoleh estimasi parameter model BGPR dan faktor-faktor yang memengaruhi jumlah kasus HIV dan AIDS di Indonesia menggunakan model tersebut. **Metode**. Penelitian ini meliputi dua tahap yaitu mengestimasi parameter model BGPR menggunakan metode *maximum likelihood estimation* (MLE) dengan algoritma *Broyden-Fletcher-Goldfarb-Shanno* (BFGS) dan *Nelder-Mead* dan studi kasus model BGPR terbaik yang diperoleh dengan AIC pada data jumlah kasus HIV dan AIDS di Indonesia. **Hasil**. Estimasi parameter model BGPR menggunakan metode MLE menghasilkan persamaan yang implisit ketika disamakan dengan nol. Model BGPR terbaik diperoleh dengan menggunakan algoritma BFGS. Pengujian signifikansi menghasilkan empat variabel signifikan. **Kesimpulan**. Faktor-faktor yang berpengaruh signifikan terhadap jumlah kasus HIV adalah persentase penduduk miskin, jumlah penduduk, orang dengan risiko terinfeksi HIV mendapatkan pelayanan deteksi, dan pasangan pengguna kondom. Adapun faktor-faktor yang berpengaruh terhadap jumlah kasus AIDS adalah persentase penduduk miskin, jumlah penduduk, pasangan pengguna kondom, dan jumlah fasilitas kesehatan.

Kata kunci: *Poisson,* Overdispersi, Regresi *Bivariate Generalized Poisson,* Jumlah Kasus HIV dan AIDS.

ABSTRACT

ANDI SYAIFUL RAHMAT. **Analysis of Factors Affecting the Number of HIV and AIDS Cases in Indonesia with Bivariate Generalized Poisson Regression** (supervised by Andi Kresna Jaya).

Background. The Generalized Linear Model (GLM) is an extension of the classical linear regression model that can model discrete data that do not follow a normal distribution, such as Poisson regression. Poisson regression often encounters assumption violations, particularly overdispersion, which can be addressed using generalized models like Generalized Poisson Regression (GPR). For pairs of discrete Poisson data experiencing overdispersion, Bivariate Generalized Poisson Regression (BGPR) can be used. One type of discrete data assumed to follow a Poisson distribution is the number of Human Immunodeficiency Virus (HIV) and Acquired Immunodeficiency Syndrome (AIDS) cases. HIV and AIDS are crucial indicators of public health levels and are included in the targets of the Sustainable Development Goals (SDGs) point third. Indonesia ranked third in Southeast Asia for the number of HIV/AIDS cases in 2022, with cases increasing consistently over the past decade, warranting analysis using the BGPR model. **Objective**. This study aims to estimate the parameters of the BGPR model and identify the factors influencing the number of HIV and AIDS cases in Indonesia using this model. **Method**. The research comprises two stages: estimating the parameters of the BGPR model using the maximum likelihood estimation (MLE) method with Broyden-Fletcher-Goldfarb-Shanno (BFGS) and Nelder-Mead algorithms, and applying the best BGPR model determined by AIC to the number of HIV and AIDS cases data in Indonesia. **Results**. Parameter estimation using the MLE method produced implicit equations when equated to zero. The best BGPR model was obtained using the BFGS algorithm. The significance tests revealed four significant variables. **Conclusion**. The factors significantly influencing the number of HIV cases include the percentage of the poor population, total population, individuals at risk of HIV infection receiving early detection services, and condom-using partners. The factors influencing the number of AIDS cases include the percentage of the poor population, total population, condom-using partners, and the number of health facilities.

Keywords: Poisson, Overdispersion, Bivariate Generalized Poisson Regression, The Number of HIV and AIDS Cases.

DAFTAR ISTILAH

Istilah	Arti dan Penjelasan
Algoritma	Prosedur yang digunakan untuk menghitung data.
Antiretroviral	Obat yang digunakan untuk mengobati infeksi HIV
Bandwidth	Parameter penghalus untuk mengetahui seberapa jauh pengaruh lokasi pengamatan saling berdekatan.
Case Fatality Rate	Persentase kematian dari total kasus penyakit tertentu.
Diskrit	Jenis data yang hanya dapat mengambil nilai bilangan bulat tertentu yang diperoleh dari hasil perhitungan.
Dispersi	Ukuran variabilitas dari suatu distribusi.
Distribusi	Penyebaran data.
Eksplisit	Kondisi suatu parameter atau variabel tidak dapat dipisahkan secara langsung dengan komponen lain dalam persamaan.
Equidispersi	Kondisi yang memperlihatkan ragam variabel respons dan rataannya bernilai sama pada regresi *Poisson*.
Fungsi kepadatan peluang	Fungsi probabilitas variabel kontinu.
Fungsi *likelihood*	Fungsi yang menggambarkan probabilitas data yang diobservasi berdasarkan parameter model.
Fungsi massa peluang	Fungsi probabilitas variabel diskrit.
Fungsi penghubung	fungsi yang menghubungkan nilai ekspektasi variabel respons dengan kombinasi linier dari variabel prediktor.
Implisit	Kondisi suatu parameter atau variabel dapat dipisahkan secara langsung dengan komponen lain dalam persamaan.
Independen	Saling bebas atau tidak saling bergantung.
Iterasi	Proses pengulangan langkah-langkah dalam sebuah algoritma hingga mencapai hasil konvergen.
Kontinu	Jenis data yang mengambil nilai dalam suatu rentang yang diperoleh dari hasil pengukuran.
Kondom	Alat kontrasepsi untuk mencegah kehamilan dan penyakit menular seksual.
Korelasi	Ukuran statistik yang menunjukkan seberapa kuat dan arah hubungan antara dua variabel.
Linier	Hubungan yang berbentuk garis lurus antara dua variabel.
Mean	Nilai rata-rata.
Multikolinieritas	Kondisi dimana terdapat korelasi antar variabel prediktor.
Multivariat	Analisis statistik yang melibatkan lebih dari dua variabel.
Overdispersi	Kondisi yang memperlihatkan ragam variabel respons yang lebih besar dari rataannya pada regresi *Poisson*
Overestimate	Nilai perkiraan lebih besar daripada nilai sebenarnya.
Parameter	Nilai menggambarkan karakteristik dari suatu populasi.
Parsial	Hanya sebagian dari keseluruhan.
Rasio	Perbandingan antara dua kuantitas.
Regresi	Metode statistika untuk menentukan hubungan antara satu variabel respons dengan satu atau lebih variabel prediktor.
Simpangan baku/ standar deviasi	Ukuran statistik yang menggambarkan seberapa besar penyebaran suatu data dari nilai rata-ratanya.

Istilah	Arti dan Penjelasan
Simultan	Secara keseluruhan.
Standard error	Ukuran kesalahan standar dari suatu estimasi.
Standarisasi	Transformasi data sehingga memiliki rata-rata nol dan variansi satu.
Underdispersi	Kondisi yang memperlihatkan ragam variabel respons yang lebih kecil dari rataannya pada regresi *Poisson*.
Underestimate	Nilai perkiraan lebih kecil daripada nilai sebenarnya.
Variabel	Karakteristik yang dapat diukur atau diobservasi.
Vaksin	Zat yang berfungsi untuk membentuk kekebalan tubuh terhadap suatu penyakit
Vektor	Ruas garis berarah yang memiliki besaran nilai dan arah tertentu.
Variansi	Ukuran penyebaran data menyimpang dari nilai rata-ratanya.

DAFTAR SINGKATAN DAN LAMBANG

Lambang/singkatan	Arti dan Penjelasan
α	“alpha” parameter dispersi, taraf signifikansi
β	“beta” parameter regresi
ε	“epsilon” galat/*error*
θ	“theta” parameter
λ	“lambda” parameter skala
π	“pi” bilangan konstanta
ρ	“rho” koefisien korelasi
τ	“tau” parameter dispersi
ϕ	“phi” nilai dispersi
Ω, ω	“omega” himpunan parameter
U, V, X, Y, Z, v, x, y	Variabel acak
$f(\cdot), h(\cdot)$	Fungsi probabilitas
$g(\cdot)$	Fungsi penghubung
$\log(\cdot)$	logaritma
$\ln(\cdot)$	logaritma natural
$e, \exp(\cdot)$	eksponen/basis dari logaritma natural
$E(\cdot)$	Ekspektasi
$Var(\cdot)$	Variansi
$Cov(\cdot)$	Kovariansi
$\min(\cdot)$	minimum
se	*standard error*
χ^2	Distribusi *chi*-kuadrat
H_0	Hipotesis nol
H_1	Hipotesis alternatif
$p - value$	*probability-value*
k, p	Jumlah variabel prediktor
n	Jumlah pengamatan
r	Bilangan konstanta
t, Z	Nilai statistik uji
BP	*Breusch-Pagan*
$D(\cdot), G^2$	*Deviance*
∂	“do” turunan parsial
!	Faktorial
^{T}	*Transpose* matriks
AIDS	*Acquired Immunodeficiency Syndrome*
BFGS	*Broyden-Fletcher-Goldfarb-Shanno*
BGPR	*Bivariate Generalized Poisson Regression*
BPR	*Bivariate Poisson Regression*
COVID-19	*Coronavirus Disease* 2019
GLM	*Generalized Linier Model*
GPR	*Generalized Poisson Regression*
HIV	*Human Immunodeficiency Virus*
MLE	*Maximum Likelihood Estimation*
MLRT	*Maximum Likelihood Ratio Test*
VIF	*Variance Inflation Factor*
WHO	*World Health Organization*

DAFTAR ISI

Halaman

HALAMAN JUDUL ii
HALAMAN PENGAJUAN ii
HALAMAN PENGESAHAN iii
PERNYATAAN KEASLIAN SKRIPSI DAN PELIMPAHAN HAK CIPTA iv
UCAPAN TERIMA KASIH vi
ABSTRAK vii
ABSTRACT viii
DAFTAR ISTILAH ix
DAFTAR SINGKATAN DAN LAMBANG xi
DAFTAR ISI xii
DAFTAR TABEL xiii
DAFTAR GAMBAR xiv
DAFTAR LAMPIRAN xv
BAB I PENDAHULUAN 1
1.1 Latar Belakang 1
1.2 Batasan Masalah 4
1.3 Tujuan Penelitian 4
1.4 Manfaat Penelitian 4
1.5 Teori 5
BAB II METODOLOGI PENELITIAN 16
2.1 Sumber Data 16
2.2 Identifikasi Variabel 16
2.3 Metode Analisis 17
BAB III HASIL DAN PEMBAHASAN 19
3.1 Estimasi Parameter Model *Bivariate Generalized Poisson Regression* 19
3.2 Penerapan Model *Bivariate Generalized Poisson Regression* pada Jumlah Kasus HIV dan AIDS di Indonesia 27
BAB IV KESIMPULAN DAN SARAN 35
4.1 Kesimpulan 35
4.2 Saran 35
DAFTAR PUSTAKA 36
LAMPIRAN 40

DAFTAR TABEL

Tabel **Halaman**
1. Definisi operasional variabel penelitian.. 16
2. Statistik deskriptif 28
3. Uji overdispersi.. 29
4. Nilai VIF variabel prediktor 30
5. Estimasi dan uji signifikansi parameter model regresi BGP dengan algoritma BFGS... 30
6. Estimasi dan uji signifikansi parameter model regresi BGP dengan algoritma *Nelder-Mead*... 31
7. Nilai AIC untuk model BGPR 31

DAFTAR GAMBAR

Gambar **Halaman**
1. Alur kerja algoritma *Nelder-Mead* 12
2. Peta sebaran jumlah kasus HIV (a) dan AIDS (b) di Indonesia tahun 2022....... 28

DAFTAR LAMPIRAN

Lampiran **Halaman**

1. Data Penelitian 41
2. Peta variabel prediktor 43
3. *Output* pengujian distribusi *Bivariate Generalized Poisson* 45
4. *Output* pengujian korelasi variabel respons 46
5. *Output* pendeteksian overdispersi 47
6. *Output* pengujian multikolinearitas 48
7. *Output* Estimasi dan uji signifikansi parameter model regresi BGP dengan algoritma BFGS 49
8. *Output* Estimasi dan uji signifikansi parameter model regresi BGP dengan algoritma *Nelder-Mead* 50

BAB I
PENDAHULUAN

1.1 Latar Belakang

Analisis regresi merupakan bentuk analisis yang digunakan untuk mengetahui hubungan antara variabel prediktor dengan variabel respons. Secara umum, analisis regresi yang sering digunakan adalah regresi linier klasik dengan variabel respons berupa data kontinu dan mengikuti distribusi normal. Namun, beberapa kasus sering dijumpai variabel respons berupa data diskrit dan tidak berdistribusi normal sehingga regresi linier klasik kurang tepat dalam memodelkan data tersebut. Dalam situasi seperti ini, *Generalized Linear Model* (GLM) dapat digunakan (Sinharay, 2010). GLM merupakan perluasan dari model regresi linier klasik dimana variabel responsnya dapat berdistribusi selain normal tetapi termasuk dalam keluarga eksponensial. GLM memiliki tiga komponen yaitu komponen acak, komponen sistematis, dan fungsi penghubung. Salah satu distribusi dalam keluarga eksponensial adalah distribusi *Poisson* yang merupakan distribusi yang menggambarkan peristiwa yang probabilitas kejadiannya kecil yang bergantung pada interval waktu tertentu atau di suatu daerah tertentu (Agresti, 2007).

Pendekatan dengan model regresi *Poisson* merupakan model regresi non-linier yang menggambarkan adanya hubungan antara variabel respons diskrit dan berdistribusi *Poisson* dengan variabel prediktor (Agresti, 2007). Sedangkan jika sepasang variabel respon berdistribusi *Poisson* yang saling berkorelasi dapat dimodelkan dengan regresi *Bivariate Poisson*. Menurut Karlis dan Ntzoufras (2005), Regresi *Poisson Bivariate* direkomendasikan untuk memodelkan data berpasangan yang menunjukkan nilai korelasi tinggi. Hal ini mengindikasikan adanya hubungan yang kuat antara kedua variabel dependen, sehingga distribusi yang tepat untuk digunakan adalah distribusi *Bivariate Poisson.*

Dalam regresi *Poisson*, salah satu asumsi yang harus dipenuhi adalah equidispersi, yaitu nilai rataan dan variansi dari variabel respons harus bernilai sama. Namun dalam analisis sering dijumpai data diskrit memperlihatkan variansi yang lebih besar dari rataannya (overdispersi) atau variansi lebih kecil dari rataannya (underdispersi). Kasus overdispersi lebih sering terjadi karena adanya keragaman dalam variabel respons serta adanya korelasi positif antar variabel respons yang terjadi pada analisis multivariat (Hilbe, 2011). Ketika asumsi equidispersi dilanggar, model regresi *Poisson* standar tidak dapat memberikan hasil yang akurat. Oleh karena itu, diperlukan metode alternatif untuk menangani data dengan pelanggaran asumsi ini. Salah satu metode yang umum digunakan adalah *Generalized Poisson Regression* (GPR). Famoye, dkk (2004) mendemonstrasikan aplikasi GPR pada data kecelakaan mobil. Hasil penelitian mereka menunjukkan bahwa *Generalized Poisson Regression* menghasilkan model yang lebih tepat dalam menggambarkan data dibandingkan dengan model regresi *Poisson* standar.

Bivariate Generalized Poisson Regression (BGPR) merupakan pengembangan dari regresi *Bivariate Poisson* yang dirancang untuk menangani data hitung *bivariate* yang mengalami kasus overdispersi atau underdispersi. Menurut Zamani, dkk (2016), BGPR memiliki beberapa keunggulan yakni kemampuan menangani berbagai jenis korelasi (korelasi positif, negatif maupun tidak berkorelasi), dan fleksibilitas terhadap overdispersi dan underdispersi. Wardani (2016) telah melakukan pendugaan parameter dan pengujian hipotesis BGPR terhadap faktor-faktor yang berpengaruh terhadap kematian bayi dan ibu di Provinsi Jawa Timur. Dari hasil penelitian diperoleh estimasi parameter model BGPR, uji hipotesis serentak terhadap model, serta faktor-faktor yang berpengaruh signifikan terhadap kematian bayi dan ibu yaitu persentase persalinan oleh tenaga kesehatan, persentase ibu hamil mendapatkan tablet Fe3, persentase wanita kawin dengan umur perkawinan di bawah usia 17 tahun, dan persentase ibu hamil melaksanakan program K4.

Penaksiran parameter BGPR dilakukan dengan menggunakan metode *Maximum Likelihood Estimation* (MLE). MLE dapat digunakan dengan memaksimumkan fungsi *likelihood* dari distribusi. Namun beberapa fungsi *likelihood* tidak dapat diselesaikan secara analitik menggunakan metode MLE karena diperoleh bentuk implisit sehingga harus diselesaikan secara iteratif. Dalam hal ini, metode algoritma *derivatif* seperti *Newton-Raphson,* dan *Broyden-Fletcher-Goldfarb-Shanno* (BFGS)*,* serta algoritma *non-derivatif* seperti *Nelder-Mead* sering digunakan untuk mendapatkan penaksir parameter yang optimal. Setelah parameter model diestimasikan, langkah selanjutnya adalah membangun model BGPR. Model yang telah dibangun kemudian diuji dengan menggunakan metode *Maximum Likelihood Ratio Test* (MLRT) untuk meninjau validitas model dan signifikansi parameter.

Penelitian ini akan dilakukan dengan menggunakan metode BGPR pada data jumlah kasus HIV dan AIDS di Indonesia tahun 2022. Metode ini dipilih karena mempertimbangkan dua variabel respons bertipe hitung yang berkorelasi serta mengalami overdispersi. Famoye, dkk (2004) menyatakan bahwa regresi *Poisson* tidak tepat digunakan untuk data yang mengalami overdispersi atau underdispersi. Dalam situasi ini, diperlukan model regresi alternatif untuk mendapatkan hasil yang lebih akurat. Beberapa model yang dapat digunakan untuk memodelkan data tersebut adalah GPR*, Negative Binomial Regression, Zero-Inflated Poisson Regression,* dan sebagainya. Secara khusus, model *Zero-Inflated Poisson Regression* dan *Zero-Inflated Negative Binomial Regression* cocok digunakan ketika variabel respons memiliki banyak nilai nol berlebih (*excess zero*). Sebelum memulai analisis data, pengujian dilakukan terlebih dahulu untuk mengetahui kesamaan antara rata-rata dan ragam data. Apabila diperoleh rata-rata dan ragam yang sama (equidispersi) maka digunakan model regresi *Poisson*, akan tetapi jika diperoleh ragam lebih besar dari rata-rata (overdispersi) atau ragam lebih kecil dari rata-rata (underdispersi), maka perlu dilakukan analisis dengan model regresi yang lebih sesuai.

Human Immunodeficiency Virus (HIV) dan *Acquired Immunodeficiency Syndrome* (AIDS) merupakan salah satu penyakit yang menjadi perhatian di Indonesia saat ini. Hal ini tidak luput dari masuknya penyakit ini sebagai indikator penting dalam menentukan tingkat kesehatan masyarakat. Keberhasilan pembangunan khususnya dalam bidang kesehatan juga dilihat dari angka jumlah penyakit menular secara langsung, salah satunya ialah HIV/AIDS. Salah satu agenda yang harus dipenuhi dalam *Suistanable Development Goals* (SDGs) poin ke 3 ialah kehidupan sehat dan sejahtera di mana target yang ingin dicapai ialah nol infeksi baru HIV, nol diskriminasi, dan nol kematian AIDS pada tahun 2030. Adanya target tersebut menunjukkan betapa pentingnya untuk menjadi perhatian pemerintah terhadap berbagai upaya penurunan angka kasus HIV/AIDS. Indonesia berada pada peringkat 3 dengan jumlah orang yang hidup dengan HIV di Asia Tenggara, dengan perkiraan sebanyak 52.955 orang pada 2022. Ironisnya, angka ini terus mengalami peningkatan selama 10 tahun terakhir dan hanya mengalami penurunan pada tahun 2020 hingga 2021 pada saat diberlakukan pembatasan aktivitas masyarakat akibat Covid-19.

Dalam upaya penurunan HIV dan AIDS untuk mempercepat capaian SDGs, pemerintah telah meluncurkan Program Nasional Pencegahan dan Pengendalian HIV, AIDS, dan Penyakit Infeksi Menular Seksual (PIMS). Turunan dari program tersebut tertuang pada Rencana Aksi Nasional (RAN) HIV/AIDS dan PIMS untuk periode 2020-2023 yaitu: 90% orang dengan HIV (ODHIV) mengetahui status, 70% ODHIV memperoleh ART, dan 75% ODHIV diperiksa beban virusnya. Tidak hanya itu, pemerintah juga telah membentuk Komisi Penanggulangan AIDS mulai tingkat nasional hingga kabupaten/kota yang bertugas mulai dari mengkoordinasikan perumusan penyusunan kebijakan, strategi, dan langkah-langkah yang sesuai, hingga memimpin, mengelola, mengendalikan, memantau, dan mengevaluasi pelaksanaan penanggulangan HIV dan AIDS di setiap tingkatan. (P2PM Kemenkes, 2023).

Beberapa penelitian yang membahas tentang kasus ini yaitu Tantri (2017) yang memodelkan kasus HIV dan Tuberkolosis (TB) di Jawa Timur dengan regresi *Poisson* bivariat. Hasil penelitian didapatkan bahwa faktor yang berpengaruh adalah persentase penduduk tercatat KB menggunakan kondom. Sedangkan Nuchaila (2020) melakukan penelitian dengan model regresi binomial negatif pada kasus HIV/AIDS di Jawa Timur menggunakan estimator lokal linier. Faktor yang mempengaruhi adalah banyak penduduk yang mendonorkan darah dan banyaknya pengguna narkotika, psikotropika, dan zat aditif (NAPZA). Pada penelitian oleh Thahar & Sirait (2023) terkait variabel-variabel yang memengaruhi jumlah kasus baru HIV/AIDS di Jawa Timur dengan metode *Geographically Weighted Negative Binomial Regression* (GWNBR). Penelitian ini menggunakan tujuh variabel yakni persentase pengguna alat kontrasepsi kondom, kepadatan penduduk, persentase penduduk miskin, rata-rata lama sekolah, persentase desa yang pernah terdapat tindak kejahatan pengedaran dan penyalahgunaan narkoba, rasio jumlah fasilitas kesehatan per 10.000 penduduk, dan rasio jumlah tenaga kesehatan per 10.000 penduduk. Dari hasil pengolahan data disimpulkan bahwa penyebaran kasus cenderung mengalami pengelompokan serta terdapat efek spasial pada data penelitian yang menyebabkan tiap wilayah memiliki model GWNBR dengan parameter berbeda untuk setiap variabel.

Kasus HIV dan AIDS merupakan dua kejadian yang saling berhubungan. HIV merupakan sebuah virus yang dapat menurunkan kekebalan tubuh manusia karena telah menginfeksi sel-sel darah putih. Sementara itu, gejala-gejala penyakit yang timbul dari HIV itu sendiri disebut sebagai AIDS (Kementerian Kesehatan, 2023). Penyakit AIDS tergolong mematikan karena memiliki *case fatality rate* dalam jangka waktu 5 tahun. Hal tersebut mengartikan bahwa dalam jangka waktu rata-rata 5 tahun setelah menderita AIDS, pasien tersebut dapat meninggal. Tingginya angka jumlah penderita HIV/AIDS saat ini di Indonesia menjadi sebuah hal yang memilukan namun sayangnya, sampai saat ini belum ditemukan metode pengobatan ataupun vaksin yang dapat mencegah penyakit ini. Oleh karena itu, penelitian terkait HIV dan AIDS merupakan salah satu hal yang penting untuk dilakukan guna mencari tahu berbagai faktor terkait penyebaran penyakit ini.

Pada penelitian ini akan mengkaji dua variabel respons yakni jumlah kasus HIV dan AIDS di Indonesia pada tahun 2022. Analisis dilakukan dengan menggunakan pendekatan BGPR dengan dua algoritma dalam mengestimasi parameter yaitu algoritma BFGS dan *Nelder-Mead.* Metode ini digunakan karena data diduga terjadi kasus equidispersi dan data tidak mengandung banyak nol, serta variabel yang digunakan merupakan peristiwa yang diduga mengikuti distribusi *Poisson.* Hasil kajian ini diharapkan dapat menentukan faktor-faktor yang berpengaruh secara signifikan terhadap jumlah kasus HIV dan AIDS di Indonesia.

1.2 Batasan Masalah

Batasan masalah dalam penelitian ini adalah sebagai berikut.

1. Ruang lingkup penelitian adalah semua provinsi di Indonesia pada tahun 2022.
2. Metode estimasi parameter model regresi *Bivariate Generalized Poisson* menggunakan metode *maximum likelihood estimation* dengan iterasi *Broyden-Fletcher-Goldfarb-Shanno* dan iterasi *Nelder-Mead*.
3. Pemilihan model terbaik menggunakan *Akaike Information Criterion* (AIC).

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah, maka tujuan penelitian adalah sebagai berikut.

1. Memperoleh estimasi parameter model regresi *Bivariate Generalized Poisson* pada jumlah kasus HIV dan AIDS di Indonesia.
2. Memperoleh faktor-faktor yang memengaruhi jumlah kasus HIV dan AIDS di Indonesia menggunakan model regresi *Bivariate Generalized Poisson*.

1.4 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini adalah dapat memperluas pemahaman teoritis dan praktis terkait model regresi *Bivariate Generalized Poisson* terutama dalam estimasi parameter dan penerapannya untuk analisis data kesehatan. Selain itu, penelitian ini dapat menjadi bahan informasi bagi masyarakat maupun instansi pemerintah untuk mengevaluasi upaya penurunan angka kasus HIV dan AIDS serta menerapkan strategi yang lebih efektif untuk penanganan penyakit ini.

1.5 Teori

1.5.1 *Generalized Linear Model*

Generalized Linear Model (GLM) adalah perluasan dari model regresi linier klasik dimana variabel responnya dapat berdistribusi selain normal tetapi termasuk dalam keluarga eksponensial (McCullagh & Nelder, 1989). Menurut Agresti (2007), GLM memiliki tiga komponen yaitu:

1. Komponen Acak

Komponen acak dari GLM mengidentifikasi variabel respon Y beserta distribusi peluangnya. Misalkan suatu pengamatan pada Y dengan $(Y_1, Y_2, Y_3, \dots, Y_n)$, maka GLM menyatakan $(Y_1, Y_2, Y_3, \dots, Y_n)$ saling bebas.

2. Komponen Sistematis

Komponen sistematis dari GLM menentukan variabel penjelas yang dimasukkan secara linier sebagai variabel prediktor dalam persamaan model, dengan kata lain menentukan variabel x_j pada persamaan berikut.

$$\beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p \qquad (1)$$

Kombinasi linier dari variabel penjelas ini disebut dengan *linear predictor*.

3. Fungsi Penghubung (*link function*)

Fungsi penghubung dari GLM menentukan sebuah fungsi $g(\cdot)$ yang menghubungkan λ dengan prediktor linier sebagai berikut.

$$g(\lambda) = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p \qquad (2)$$

Fungsi penghubung $g(\cdot)$ menggabungkan komponen acak dan komponen sistematis. Fungsi penghubung yang paling sederhana adalah $g(\lambda) = \lambda$ yang memodelkan rata-rata secara langsung yang disebut *identity link*. Fungsi penghubung lainnya memungkinkan λ berhubungan secara non-linier dengan variabel prediktor seperti $g(\lambda) = \ln(\lambda)$ yang memodelkan logaritma natural dari rata-rata yang disebut *log link*.

1.5.2 Regresi *Poisson*

Regresi *Poisson* merupakan model regresi non-linier dengan variabel responsnya berupa data hitung (diskrit) dan berdistribusi *Poisson*. Distribusi *Poisson* adalah suatu distribusi untuk peristiwa yang probabilitas kejadiannya kecil, dimana kejadian tersebut bergantung pada interval waktu tertentu atau di suatu daerah tertentu. Jika variabel acak diskrit (Y) berdistribusi *Poisson* dengan parameter $\lambda > 0$, maka fungsi probabilitasnya dinyatakan sebagai berikut.

$$f(y; \lambda) = \frac{e^{-\lambda}\lambda^y}{y!}, y = 0,1,2, \dots \qquad (3)$$

Distribusi *Poisson* memiliki nilai mean dan variansi yang sama, yaitu $E(Y) = Var(Y) = \lambda$ (Arkandi & Winahju, 2015).

Regresi *Poisson* menggunakan fungsi penghubung $\ln(\cdot)$ dari GLM agar modelnya dapat digunakan dalam data pengamatan tanpa mengharuskan variabel responnya berdistribusi normal. Oleh karena itu, model regresi *Poisson* dapat dituliskan sebagai berikut (Cahyandari, 2014).

$$g(\lambda_i) = \ln(\lambda_i)$$
$$\ln(\lambda_i) = \boldsymbol{x}_i^T \boldsymbol{\beta} \tag{4}$$
$$\lambda_i = \exp(\boldsymbol{x}_i^T \boldsymbol{\beta})\,;\, i = 1,2,3,\dots,n$$

dengan

λ_i menunjukkan rata-rata variabel respon berdistribusi *Poisson* pengamatan ke-i.

$\boldsymbol{x}_i^T = [1 \quad x_{i1} \quad x_{i2} \quad \cdots \quad x_{ip}]$ menunjukkan vektor dari variabel prediktor yang berukuran $1 \times (p+1)$ pada pengamatan ke-i.

$\boldsymbol{\beta} = [\beta_0 \quad \beta_1 \quad \beta_2 \quad \cdots \quad \beta_p]^T$ menunjukkan vektor dari parameter regresi yang berukuran $(p+1) \times 1$.

1.5.3 Regresi *Bivariate Poisson*

Regresi *Bivariate Poisson* (BP) adalah sebuah metode yang digunakan untuk memodelkan sepasang data hitung (diskrit) yang berdistribusi *Poisson* dan saling berkorelasi (Karlis & Ntzoufras, 2005). Karena terdapat dua variabel respons yang saling berkorelasi, maka distribusinya akan mengikuti distribusi BP.

Distribusi BP terjadi saat variabel acak Z_0, Z_1, dan Z_2 masing-masing berdistribusi *Poisson* dengan parameter λ_0, λ_1, dan λ_2. Terdapat variabel acak Y_1 dan Y_2 yang terbentuk dari variabel Z_0, Z_1, dan Z_2 yang saling bebas (independen) seperti berikut.

$$Y_1 = Z_1 + Z_0$$
$$Y_2 = Z_2 + Z_0$$

Maka variabel acak Y_1 dan Y_2 secara bersama-sama berdistribusi *Bivariate Poisson* dengan fungsi massa peluang bersamanya sebagai berikut (Kawamura, 1973).

$$f(y_1, y_2) = \exp -(\lambda_1 + \lambda_2 + \lambda_0) \sum_{k=0}^{\min(y_1,y_2)} \frac{\lambda_1^{y_1-k} \lambda_2^{y_2-k} \lambda_0^k}{(y_1-k)!\,(y_2-k)!\,k!}; y_1, y_2 \geq 0 \tag{5}$$

dengan k adalah bilangan yang membuat nilai non-negatif bagi $y_1 - k; y_2 - k; k$. Nilai *mean* dan variansi dari kedua variabel acak tersebut adalah.

$E(Y_1) = E(Z_1 + Z_0) = E(Z_1) + E(Z_0) = \lambda_1 + \lambda_0$

$E(Y_2) = E(Z_2 + Z_0) = E(Z_2) + E(Z_0) = \lambda_2 + \lambda_0$

$E(Y_1Y_2) = E(Z_1Z_0) + E(Z_1Z_2) + E(Z_0Z_2) + E(Z_0^2) = (\lambda_1 + \lambda_0)(\lambda_2 + \lambda_0) + \lambda_0$

$Var(Y_1) = Var(Z_1 + Z_0) = Var(Z_1) + Var(Z_0) = \lambda_1 + \lambda_0$

$Var(Y_2) = Var(Z_2 + Z_0) = Var(Z_2) + Var(Z_0) = \lambda_2 + \lambda_0$

$Cov(Y_1Y_2) = E(Y_1Y_2) - E(Y_1)E(Y_2) = \lambda_0$

Menurut Karlis dan Ntzoufras (2005), terdapat 3 model regresi BP dengan nilai λ_0 yang berbeda, diantaranya:

a. Model dengan nilai λ_0 adalah suatu konstanta (tidak ada kovariat di dalam λ_0).

b. Model dengan nilai λ_0 merupakan fungsi dari variabel bebas (kovariat) sehingga membentuk suatu persamaan berikut.

$$\lambda_0 = \exp\left(\beta_{00} + \beta_{10}x_1 + \beta_{20}x_2 + \cdots + \beta_{p0}x_p\right) = \exp\,(\boldsymbol{x\beta_0})$$

c. Model dengan nilai λ_0 adalah nol, dengan tidak ada kovariat dari kedua variabel atau kedua variabel respons saling bebas, biasanya disebut model *Double Poisson*.

Parameter λ_j dihubungkan dengan variabel prediktor menggunakan fungsi penghubung $\ln(\cdot)$. Sehingga model regresi BP adalah sebagai berikut.

$$\begin{aligned} \ln(\lambda_{ji}) &= \boldsymbol{x}_i^T\boldsymbol{\beta}_j \\ \lambda_{ji} &= \exp(\boldsymbol{x}_i^T\boldsymbol{\beta}_j)\,;\, i = 1,2,\ldots,n;\ j = 1,2 \end{aligned} \tag{6}$$

dengan

λ_{ji} menunjukkan rata-rata variabel respons ke-j pada pengamatan ke-i.

$\boldsymbol{x}_i^T = [1 \quad x_{i1} \quad x_{i2} \quad \cdots \quad x_{ip}]$ menunjukkan vektor variabel prediktor yang berukuran $1 \times (p+1)$ pada pengamatan ke-i.

$\boldsymbol{\beta}_j = \left[\beta_{0j} \quad \beta_{1j} \quad \beta_{2j} \quad \cdots \quad \beta_{pj}\right]^T$ menunjukkan vektor parameter regresi yang berukuran $(p+1) \times 1$ pada variabel respons ke-j.

1.5.4 Overdispersi

Regresi *Poisson* mensyaratkan adanya asumsi equidispersi, di mana *mean* dan variansi variabel respons memiliki nilai yang sama. Namun, di dalam praktiknya asumsi ini sering kali dilanggar. Terdapat dua kondisi yang dapat terjadi, yaitu overdispersi dan underdispersi. Overdispersi terjadi ketika variansi lebih besar dari *mean*, sedangkan underdispersi terjadi ketika variansi lebih kecil dari *mean* (Cameron & Trivedi, 1998).

Overdispersi pada regresi *Poisson* dapat berakibat fatal. Estimasi simpangan baku parameter menjadi lebih kecil dari nilai sebenarnya (*underestimate*), dan uji signifikansi variabel prediktor menjadi lebih besar dari nilai sebenarnya (*overestimate*). Hal ini menyebabkan kesimpulan yang dihasilkan menjadi tidak valid (Ismail & Jemain, 2007).

Menurut Cahyandari (2014), penyebab lain overdispersi seperti banyaknya pengamatan yang bernilai nol (*excess zero*), adanya sumber keragaman antar individu yang tidak teramati (*unobserved heterogeneity*), terdapat pengamatan yang hilang (*data missing*), adanya pencilan (*outlier*) pada data, atau kesalahan spesifikasi fungsi penghubung. Jika terjadi overdispersi, nilai simpangan baku dari estimasi parameter jauh lebih kecil dari nilai sebenarnya (*underestimate*) dan uji signifikansi dari parameter jauh lebih besar dari nilai sebenarnya (*overestimate*) sehingga model regresi *Poisson* tidak efisien dengan nilai *standard error* yang tinggi (Ismail & Jemain, 2007).

Beberapa faktor yang dapat menyebabkan terjadinya overdispersi, di antaranya kelebihan nilai nol (*excess zero*), heterogenitas antar individu yang tidak teramati (*unobserved heterogeneity*), data hilang (*data missing*), pencilan (*outlier*) pada data, hingga kesalahan spesifikasi fungsi penghubung. Oleh karena itu, penting untuk memeriksa keberadaan overdispersi sebelum menggunakan regresi *Poisson.* Jika overdispersi ditemukan, perlu dilakukan langkah-langkah korektif untuk mengatasinya, seperti menggunakan model regresi alternatif (misalnya, regresi binomial negatif) atau memasukkan variabel kontrol tambahan ke dalam model (Cahyandari, 2014).

Overdispersi pada variabel respon Y_1 dan Y_2 dapat dideteksi menggunakan uji *Deviance* dengan hipotesis dan statistik uji sebagai berikut (McCullagh & Nelder, 1989).

H_0: $\phi = 0$ (variabel respon Y_1 dan Y_2 tidak mengalami overdispersi).

H_1: $\phi > 0$ (variabel respon Y_1 dan Y_2 mengalami overdispersi).

$$\hat{\phi} = \frac{D}{(n-k)} \tag{7}$$

dengan $D(y;\mu) = 2\sum_{i=1}^{n}\left(y_i \ln\left(\frac{y_i}{\lambda_i}\right) - (y_i - \lambda_i)\right)$. Kriteria pengambilan keputusan dalam uji *Deviance* adalah menolak H_0 jika $\hat{\phi} > 0$.

1.5.5 Regresi *Generalized Poisson*

Model regresi *Generalized Poisson* (GP) merupakan suatu model yang digunakan apabila pada data hitung terjadi pelanggaran asumsi equidispersi pada sebaran *Poisson* atau terjadi overdispersi atau underdispersi. Model regresi GP hampir sama dengan model regresi *Poisson* yaitu merupakan suatu model GLM, akan tetapi pada model GP mengasumsikan bahwa komponen randomnya berdistribusi GP.

Pada model regresi GP selain terdapat parameter λ juga terdapat paramater α sebagai parameter dispersi, dimana apabila nilai $\alpha = 0$ maka sama dengan model regresi *Poisson*. Kondisi overdispersi pada data ditunjukkan dengan nilai $\alpha > 0$, sedangkan kondisi underdispersi pada data ditunjukkan dengan nilai $\alpha < 0$. Adapun fungsi kepekatan peluang sebaran regresi GP terdapat pada persamaan berikut ini (Famoye, 1993).

$$f(Y = y) = \frac{\lambda(\lambda + y\alpha)^{y-1}}{y!}\exp(-\lambda - y\alpha) \tag{8}$$

dengan rata-rata dan ragam sebagai berikut:

$$E(y) = \lambda(1-\alpha)^{-1}$$
$$var(y) = \lambda(1-\alpha)^{-3}$$

sehingga model regresi GP memiliki bentuk yaitu:

$$\begin{gathered} \ln(\lambda_i) = \boldsymbol{x}_i^T\boldsymbol{\beta} = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p \\ \lambda_i = \exp\,(\boldsymbol{x}_i^T\boldsymbol{\beta}); i = 1,2,\dots,n \end{gathered} \tag{9}$$

dimana:

$\boldsymbol{x}_i^T = [1 \quad x_{i1} \quad x_{i2} \quad \cdots \quad x_{ip}]$ menunjukkan vektor dari variabel prediktor yang berukuran $1 \times (p+1)$ pada pengamatan ke-i.

$\boldsymbol{\beta} = [\beta_0 \quad \beta_1 \quad \beta_2 \quad \cdots \quad \beta_p]^T$ menunjukkan vektor dari parameter regresi yang berukuran $(p+1) \times 1$.

1.5.6 Regresi *Bivariate Generalized Poisson*

Menurut Karlis dan Ntzoufras (2005), regresi *Bivariate Generalized Poisson* (BGP) merupakan suatu metode yang digunakan untuk memodelkan sepasang data hitung yang memiliki korelasi dengan beberapa variabel prediktor. Metode ini adalah pengembangan regresi BP pada data yang mengalami kasus overdispersi atau underdispersi.

Menurut Vernic (1997) misalkan Z_1, Z_2, dan Z_0 adalah variabel acak saling bebas yang masing-masing berdistribusi GP dan $Z_i \sim GPD(\lambda_i, \alpha_i)$ $i = 0,1,2$. $Y_1 = Z_1 + Z_0$ dan $Y_2 = Z_2 + Z_0$, maka fungsi peluang bersama dari Y_1 dan Y_2 adalah,

$$f(Y_1 = y_1, Y_2 = y_2) = \sum_{k}^{\min(y_1,y_2)} f_1(y_1 - k) f_2(y_2 - k) f_3(k) \tag{10}$$

Consul dan Shoukri (1985) dalam Vernic (1997) menjelaskan bahwa $X_i \sim GPD(\lambda, \alpha)$ maka fungsi kepadatan peluang dari distribusi BGP adalah sebagai berikut.

$$\begin{aligned} f(y_1, y_2) = {} & \lambda_0 \lambda_1 \lambda_2 \exp(-(\lambda_0 + \lambda_1 + \lambda_2) - y_1\alpha_1 - y_2\alpha_2) \\ & \sum_{k}^{\min(y_1,y_2)} \left(\frac{(\lambda_1 + (y_1 - k)\alpha_1)^{y_1-k-1}}{(y_1 - k)!} \frac{(\lambda_2 + (y_2 - k)\alpha_2)^{y_2-k-1}}{(y_2 - k)!} \right. \\ & \left. \frac{(\lambda_0 + k\alpha_0)^{k-1}}{k!} \exp\big(k(\alpha_1 + \alpha_2 - \alpha_0)\big) \right) \end{aligned} \tag{11}$$

Jika $(Y_1, Y_2) \sim BGP(\lambda_0, \beta_1, \beta_2, \alpha_1, \alpha_2)$ maka model dari regresi BGP dengan menggunakan fungsi penghubung $\ln(\cdot)$ serta λ_0 sebagai suatu konstanta adalah sebagai berikut,

$$\begin{aligned} \ln(\lambda_{ji} + \lambda_0) &= \boldsymbol{x}_i^T \boldsymbol{\beta}_j = \beta_{0j} + \beta_{1j}x_{i1} + \beta_{2j}x_{i2} + \cdots + \beta_{pj}x_{ip}\,;\, i = 1,2,\dots,n;\; j = 1,2 \\ \lambda_{ji} + \lambda_0 &= \exp(\boldsymbol{x}_i^T \boldsymbol{\beta}_j) = \exp(\beta_{0j} + \beta_{1j}x_{i1} + \beta_{2j}x_{i2} + \cdots + \beta_{pj}x_{ip}) \end{aligned} \tag{12}$$

dengan

λ_{ji} menunjukkan rata-rata variabel respons ke-j pada pengamatan ke-i.

$\boldsymbol{x}_i^T = [1 \quad x_{i1} \quad x_{i2} \quad \cdots \quad x_{ip}]$ menunjukkan vektor variabel prediktor yang berukuran $1 \times (p + 1)$ pada pengamatan ke-i.

$\boldsymbol{\beta}_j = [\beta_{0j} \quad \beta_{1j} \quad \beta_{2j} \quad \cdots \quad \beta_{pj}]^T$ menunjukkan vektor parameter regresi yang berukuran $(p + 1) \times 1$ pada variabel respons ke-j.

Untuk mengetahui apakah variabel Y_1 dan Y_2 mengikuti distribusi BGP, maka dilakukan uji distribusi BGP menggunakan *Crockett's test* dengan hipotesis sebagai berikut.

H_0 : Variabel respons Y_1 dan Y_2 mengikuti distribusi BGP

H_1 : Variabel respons Y_1 dan Y_2 tidak mengikuti distribusi BGP

Statistik uji:

$$T = \boldsymbol{Z}^{\boldsymbol{T}} \widehat{\boldsymbol{V}}^{-1} \boldsymbol{Z} \tag{13}$$

dimana,

$\boldsymbol{Z}^{\boldsymbol{T}} = [Z_{Y_1} \quad Z_{Y_2}]\,; Z_h = var[Y_h] - \bar{Y}_h,\; h = 1,2$ dan

$$\widehat{\boldsymbol{V}} = \frac{2}{n}\begin{bmatrix} \hat{\lambda}_1 & \hat{\lambda}_{12} \\ \hat{\lambda}_{21} & \hat{\lambda}_2 \end{bmatrix} ; n = 1,2 ; \hat{\lambda}_h = v\hat{a}r(Y_h) ; \hat{\lambda}_{gh} = c\hat{o}v\left(Y_g, Y_h\right), g, h = 1,2 ; g \neq h$$

Tolak H_0 apabila $T > \chi^2_{(n,\alpha)}$ (Triyanto, 2017).

Penaksiran parameter model regresi BGP dilakukan menggunakan metode *Maximum Likelihood Estimation* (MLE) dengan memaksimumkan fungsi *likelihood*. Kemudian akan dilakukan pengujian parameter, dan penghitungan ukuran kebaikan model yang akan dibahas lebih lanjut pada bagian analisis dan pembahasan.

1.5.7 *Maximum Likelihood Estimation*

Metode *Maximum Likelihood Estimation* (MLE) adalah salah satu metode estimasi titik yang berperan dalam statistik. Konsep dasar dari estimasi parameter metode MLE yaitu dengan memaksimalkan fungsi *likelihood* untuk menentukan nilai penduga (Wibisono, 2009). Metode MLE dikenal sebagai metode sampel besar yang merupakan aplikasi yang lebih luas diterapkan dalam model regresi nonlinier dalam suatu parameter (Gujarati & Porter, 2010).

Diketahui n data pasangan $\{(x_1.y_1),(x_2.y_2),\dots,(x_n.y_n)\}$ yang saling bebas (independen) dan diasumsikan bahwa untuk $X = x$ fungsi kepadatan peluang Y adalah

$$Y|x = f(y|\boldsymbol{\theta}) \tag{14}$$

dengan $\boldsymbol{\theta}$ adalah parameter yang merupakan fungsi x, yaitu $\boldsymbol{\theta} = s(x)$ di mana $s(x)$ adalah fungsi penghalus, maka fungsi *likelihood* dari $Y|x$ adalah

$$L(\boldsymbol{\theta}|x,y) = \prod_{i=1}^{n} f\left(y_i|s(x_i)\right) \tag{15}$$

Untuk $x = \{x_1, x_2, \dots, x_n\}$ dan $y = \{y_1, y_2, \dots, y_n\}$

Estimator maksimum *likelihood* $\widehat{\boldsymbol{\theta}}$ adalah nilai $\boldsymbol{\theta}$ yang memaksimumkan fungsi *likelihood* $L(\boldsymbol{\theta}|x,y)$. Namun, lebih mudah bekerja dengan logaritma natural dari fungsi *likelihood,* yaitu $l(\boldsymbol{\theta}|x,y) = \ln L(\boldsymbol{\theta}|x,y)$. Karena fungsi logaritma natural adalah fungsi yang monoton naik, maka nilai yang memaksimumkan fungsi $l(\boldsymbol{\theta}|x,y)$ sama dengan memaksimum $L(\boldsymbol{\theta}|x,y)$ sehingga dapat dituliskan sebagai berikut.

$$\begin{aligned} \ell(\boldsymbol{\theta}|x,y) &= \ln L(\boldsymbol{\theta}|x,y) \\ \ell(\boldsymbol{\theta}|x,y) &= \ln\left\{\prod_{i=1}^{n} f(y_i|\boldsymbol{\theta})\right\} = \sum_{i=1}^{n} \ln f(y_i|\boldsymbol{\theta}) \end{aligned} \tag{16}$$

Untuk memperoleh nilai $\boldsymbol{\theta}$ yang memaksimumkan fungsi $\ell(\boldsymbol{\theta}|x,y)$, maka $\ell(\boldsymbol{\theta}|x,y)$ diturunkan terhadap $\boldsymbol{\theta}$ dan kemudian menyamakannya dengan nol seperti pada persamaan berikut.

$$\ell'(\boldsymbol{\theta}|x,y) = \frac{\partial \ell(\boldsymbol{\theta}|x,y)}{\partial \boldsymbol{\theta}} = \mathbf{0} \tag{17}$$

(Hogg, McKean, & Craig, 2013)

1.5.8 Algoritma *Broyden-Fletcher-Goldfarb-Shanno*

Dalam optimasi numerik, algoritma *Broyden-Fletcher-Goldfarb-Shanno* (BFGS) adalah salah satu algoritma untuk menyelesaikan optimasi *nonlinear* tanpa kendala. Algoritma ini dikembangkan oleh *Broyden-Fletcher-Goldfarb-Shanno* dari algoritma Quasi-Newton. Algoritma ini membutuhkan gradien dari fungsi *likelihood* yang dimaksimalkan pada setiap iterasi. Matriks *hessian* didekati melalui iterasi evaluasi gradien. Gradien diperoleh dengan melakukan turunan terhadap fungsi *ln-likelihood* terhadap tiap koefisien parameter yang digunakan. (Yuan, G dan Lu, X, 2011).

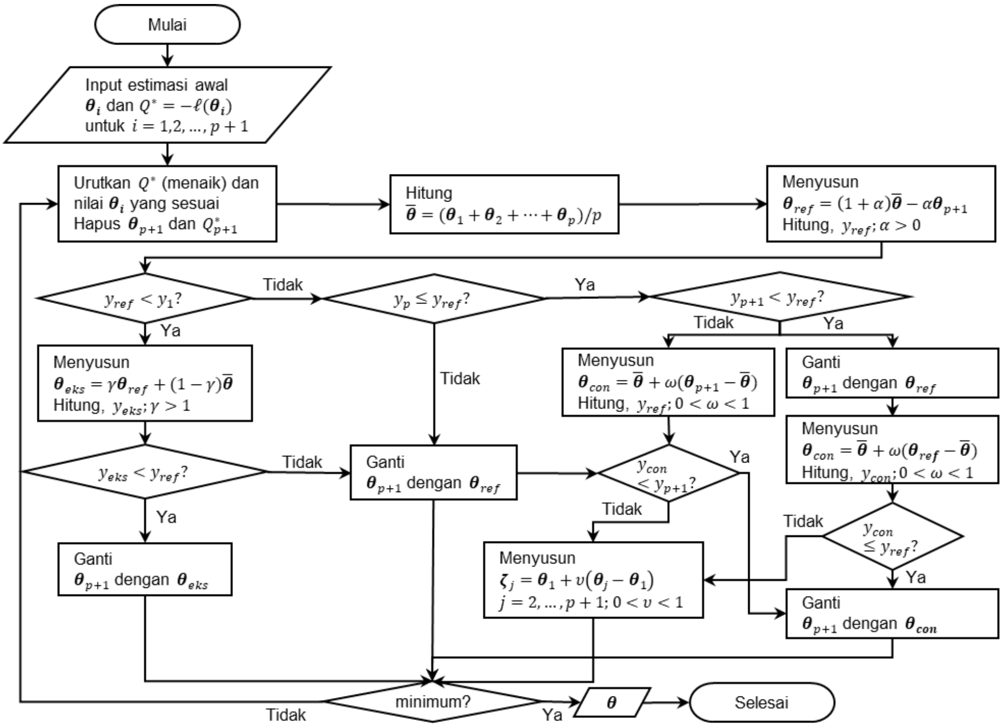
Algoritma BFGS dimulai dengan nilai awal $\boldsymbol{\theta}_0$ dan matriks *hessian* $\boldsymbol{H_0}$, algoritma ini akan mengulangi langkah-langkah berikut ini hingga $\boldsymbol{\theta}_r$ mencapai konvergensi.

a. Tentukan nilai awal dari parameter yang akan ditaksir ($\boldsymbol{\theta}_0$) dan $\boldsymbol{H}_0$, $\boldsymbol{H}_0 = \boldsymbol{I}$.
b. Membentuk vektor gradien ($\boldsymbol{g}$) dengan menggunakan turunan pertama dari fungsi *ln-likelihood* terhadap setiap parameter dalam $\boldsymbol{\theta}$.
$$\boldsymbol{g}_r = \left(\frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right)^T \tag{18}$$
$\boldsymbol{H}_r$ adalah matriks *hessian* pada iterasi ke-r.
c. Melakukan iterasi dengan $r = 0,1,2,\ldots$ pada persamaan
$$\boldsymbol{\theta}_{r+1} = \boldsymbol{\theta}_r + \lambda_r (\boldsymbol{H}_r)^{-1} \boldsymbol{g}_r$$
d. Jika nilai estimasi yang diperoleh tidak memenuhi $\|\boldsymbol{\theta}_{r+1} - \boldsymbol{\theta}_r\| \leq 10^{-3}$, maka iterasi dilanjutkan dengan menghitung matriks pembaruan BFGS
$$\boldsymbol{H}_{r+1} = \left(\boldsymbol{I} - \frac{\boldsymbol{s_r}(\boldsymbol{y_r})^T}{(\boldsymbol{y}^r)^T \boldsymbol{s_r}}\right) \boldsymbol{H}_r \left(\boldsymbol{I} - \frac{(\boldsymbol{s_r})^T \boldsymbol{y_r}}{(\boldsymbol{y}^r)^T \boldsymbol{s_r}}\right) + \frac{\boldsymbol{s_r}(\boldsymbol{s_r})^T}{(\boldsymbol{y}^r)^T \boldsymbol{s_r}} \tag{19}$$
dimana
$\boldsymbol{s}_r = \boldsymbol{\theta}_{k+1} - \boldsymbol{\theta}_k$,
$\boldsymbol{y}_r = \boldsymbol{g}_{k+1} - \boldsymbol{g}_k$, dan
$\boldsymbol{H}_0 =$ matriks identitas $n \times n$

1.5.9 Algoritma Nelder-Mead

Algoritma *Nelder-Mead* adalah algoritma tanpa turunan. Algoritma ini mencari koefisien dan estimasi indeks dispersi menggunakan proses refleksi, ekspansi, kontraksi, dan pengecilan. Optimasi *Nelder-Mead* mudah diimplementasikan dan sangat cepat konvergen pada sembarang nilai awal yang digunakan. Algoritma *Nelder-Mead* dapat digunakan untuk menentukan nilai minimum dari suatu fungsi multivariat tanpa menentukan nilai diferensialnya terlebih dahulu.

Misalkan $\boldsymbol{\Theta}$ adalah ruang parameter berdimensi p dan $\ell(\boldsymbol{\theta})$ adalah fungsi *ln-likelihood* atau $Q^* = -\ell(\boldsymbol{\theta})$ adalah fungsi objektif yang harus diminimalkan, di mana $\boldsymbol{\theta} \in \boldsymbol{\Theta}$. Setiap iterasi algoritma dimulai dari sebuah *simplex* di ruang $\boldsymbol{\Theta}$, yaitu struktur yang dibentuk oleh $p + 1$ simpul, yaitu $\widehat{\boldsymbol{\theta}}_{(1)} \ldots \widehat{\boldsymbol{\theta}}_{(p+1)}$, di mana $p = mp + 1$. Jika solusi optimal ditemukan, algoritma menggantikan simpul terburuk dalam *simplex* dengan titik baru yang memiliki nilai negatif *ln-likelihood* yang lebih kecil menggunakan salah satu dari operasi *reflection*, *expantion*, *constraction,* dan *shrink*. Jika operasi-operasi tersebut gagal menemukan simpul terburuk pada simplex, maka operasi pengecilan dilakukan. Berikut ini alur kerja algoritma *Nelder-Mead.*



Gambar 1. Alur kerja algoritma *Nelder-Mead*

1.5.10 Pengujian Signifikansi Model *Bivariate Generalized Poisson Regression*

Pengujian signifikansi model regresi BGP dilakukan untuk mengevaluasi apakah secara keseluruhan, variabel prediktor dalam model memiliki pengaruh signifikan terhadap variabel dependen. Fungsi *likelihood* yang berhubungan dengan model regresi BGP adalah $L(\hat{\Omega})$ dan $L(\hat{\omega})$. $L(\hat{\Omega})$ adalah nilai *maximum likelihood* untuk model dengan melibatkan variabel prediktor dan $L(\hat{\omega})$ adalah nilai *maximum likelihood* untuk model sederhana tanpa melibatkan variabel prediktor. Metode yang digunakan untuk pengujian simultan pada parameter model regresi BGP adalah *Maximum Likelihood Rasio Test* (MLRT). Hipotesis yang digunakan adalah sebagai berikut.

$H_0 : \beta_{j1} = \beta_{j2} = \cdots = \beta_{jp} = 0$

H_1 : paling sedikit ada satu $\beta_{jl} \neq 0$; dengan $j = 1,2$ dan $p = 1,2,\dots,k$

Dengan statistik uji:

$$G^2 = -2\ln\left(\frac{L^*(\hat{\omega})}{L^*(\hat{\Omega})}\right) = 2(\ln L^*(\hat{\omega}) - \ln L^*(\hat{\Omega})) \sim \chi^2_{(a,v)} \tag{20}$$

Daerah penolakan H_0 adalah $G^2 > \chi^2_{(\alpha,v)}$ maka terdapat variabel prediktor yang berpengaruh terhadap variabel respons dengan α adalah taraf signifikansi (Agresti, 2002).

$$G^2 = -2\ln\left(\frac{L(\widehat{\omega})}{L(\widehat{\Omega})}\right) \tag{21}$$

dengan $L(\widehat{\Omega})$ dibentuk dari himpunan parameter di bawah populasi yaitu $\mathbf{\Omega} = \{\boldsymbol{\beta}_{jl}; \tau; j = 1,2, p = 1,2, \dots, k;\}$ dan $L(\widehat{\omega})$ dibentuk dari himpunan parameter di bawah H_0 yaitu $\omega = \{\beta_{j.0}; \tau; j = 1,2\}$.

Pengujian signifikansi model regresi BGP secara parsial dilakukan untuk mengetahui faktor-faktor apa saja yang berpengaruh signifikan terhadap variabel dependen. Pengujian secara parsial dilakukan menggunakan uji *Wald*. Hipotesis yang digunakan adalah sebagai berikut.

$H_0 : \beta_{j1} = \beta_{j2} = \cdots = \beta_{jp} = 0$

H_1 : paling sedikit ada satu $\beta_{jl} \neq 0$; dengan $j = 1,2$ dan $p = 1,2, \dots, k$

Dengan statistik uji:

$$Z_\beta = \frac{\hat{\beta}_{jp}}{\sqrt{Var(\hat{\beta}_{jp})}} \tag{22}$$

dimana $\hat{\beta}_{jp}$ diperoleh dari diagonal utama matriks *varian kovarian* dan $\sqrt{Var(\hat{\beta}_{jp})}$ adalah varians dari estimasi koefisien tersebut.

Kriteria pengujian hipotesis yaitu tolak H_0 jika $|Z| > Z_{\frac{\alpha}{2}}$ yang berarti parameter β_{jp} signifikan secara statistik dan memiliki pengaruh terhadap variabel dependen.

1.5.11 Koefisien Korelasi

Koefisien korelasi merupakan suatu indikator atau suatu nilai dalam hubungan linear antara dua variabel (Draper & Smith, 1992). Koefisien korelasi didefinisikan seperti pada persamaan berikut.

$$\rho_{Y_1,Y_2} = \frac{Cov(Y_1, Y_2)}{\sqrt{Var(Y_1)Var(Y_2}} \tag{23}$$

Koefisien korelasi dapat menunjukkan dua hubungan, yaitu positif dan negatif. Nilai positif dan negatif ini dikarenakan nilai korelasi berkisar antara -1 hingga 1 atau dapat ditulis $-1 \leq \rho_{Y_1,Y_2} \leq 1$. Apabila nilai korelasi mendekati 1, baik itu positif maupun negatif hal tersebut berarti kedua variabel memiliki hubungan yang erat. Nilai korelasi 0 menunjukkan bahwa kedua variabel tidak memiliki hubungan erat secara linier. Nilai korelasi yang positif menunjukkan adanya hubungan berbanding lurus pada dua variabel tersebut, sedangkan nilai korelasi yang negatif menunjukkan hubungan yang berbanding terbalik. Pengujian korelasi untuk variabel respons dilakukan dengan hipotesis sebagai berikut:

H_0 : $\rho = 0$ (Tidak terdapat hubungan antara Y_1dan Y_2)
H_1 : $\rho \neq 0$ (Terdapat hubungan antara Y_1dan Y_2)
Statistik uji korelasi *Pearson* yang digunakan pada pengujian ini adalah sebagai berikut (Walpole et al., 2011).

$$t = \frac{r_{y_1,y_2}\sqrt{n-2}}{\sqrt{1-\left(r_{y_1,y_2}\right)^2}} \tag{24}$$

dimana:

$$r_{y_1,y_2} = \frac{\sum_{i=1}^{n}(Y_{1i}-\bar{Y}_1)\,(Y_{2i}-\bar{Y}_2)}{\sqrt{\sum_{i=1}^{n}(Y_{1i}-\bar{Y}_1)^2\,(Y_{2i}-\bar{Y}_2)^2}}$$

Daerah penolakan H_0 adalah $|t_{hit}| > t_{\left(\frac{\alpha}{2};(n-2)\right)}$

1.5.12 Multikolinearitas

Multikolinieritas merupakan kondisi dimana terdapat korelasi antar variabel prediktor dalam hal ini antar variabel prediktor tidak saling bebas. Multikolinieritas dapat menyebabkan nilai variansi estimasi parameter *overestimate*, kesalahan tanda koefisien parameter, dan masalah pada selang kepercayaan estimasi parameter sehingga perlu dideteksi. Multikolinieritas dapat dideteksi menggunakan *Variance Inflation Factor* (VIF) dengan hipotesis dan statistik uji sebagai berikut (Ryan, 2009).
H_0: $VIF_p < 10$ (tidak terdapat gejala multikolinieritas pada variabel prediktor).
H_1: $VIF_p \geq 10$ (terdapat gejala multikolinieritas pada variabel prediktor).

$$VIF_p = \frac{1}{1-R_p^2}; p = 1,2,\dots,k \tag{25}$$

dengan $R_p^2 = \frac{\sum_{i=1}^{n}\left(\hat{x}_{ip}-\bar{x}_p\right)^2}{\sum_{i=1}^{n}\left(x_{ip}-\bar{x}_p\right)^2}$. Kriteria pengambilan keputusan dalam uji multikolinieritas adalah menolak H_0 jika $VIF_p \geq 10$.

Solusi untuk mengatasi adanya kasus multikolinearitas yaitu dengan cara mengeluarkan variabel prediktor yang tidak signifikan dari dalam model satu per satu dimulai dari variabel prediktor yang memiliki nilai VIF terbesar dan meregresikan kembali variabel-variabel prediktor yang signifikan.

1.5.13 Pemilihan Model Terbaik

Terdapat beberapa metode dalam menentukan model terbaik, salah satunya adalah *Akaike Information Criterion* (AIC). AIC merupakan kriteria kesesuaian model dalam mengestimasi model secara statistik. Kriteria AIC dapat digunakan apabila pembentukan model regresi bertujuan untuk mendapatkan faktor-faktor yang berpengaruh terhadap model bukan untuk melakukan suatu prediksi.

Besarnya nilai AIC sejalan dengan nilai devians dari model. Semakin kecil nilai devians maka akan semakin kecil pula tingkat kesalahan yang dihasilkan model sehingga model yang diperoleh menjadi semakin tepat. Oleh karena itu, model terbaik adalah model yang mempunyai nilai AIC terkecil.

Tujuan dari penelitian ini adalah pemilihan model terbaik. Pemilihan model terbaik dari BGPR dilakukan menggunakan nilai AIC. Metode ini didasarkan pada metode *Maximum Likelihood Estimation* (MLE). Perhitungan nilai AIC menggunakan rumus sebagai berikut (Akaike, 1978).

$$AIC = -2\ln(maximum\ likelihood) + 2(number\ of\ parameters) \tag{26}$$

1.5.14 *Human Immunodeficiency Virus* dan *Acquired Immune Deficiency Syndrome*

Human Immunodeficiency Virus (HIV) adalah virus yang menginfeksi sel darah putih yang menyebabkan turunnya kekebalan tubuh manusia. Infeksi tersebut menyebabkan penderita mengalami penurunan kekebalan sehingga sangat mudah untuk terinfeksi berbagai macam penyakit lain. Sedangkan *Acquired Immune Deficiency Syndrome* (AIDS) adalah sekumpulan gejala yang timbul karena turunnya kekebalan tubuh yang disebabkan infeksi oleh HIV.

Orang yang terinfeksi HIV memerlukan pengobatan Antiretroviral (ARV) untuk menekan jumlah virus HIV di dalam tubuh. Virus yang tertekan (tersupresi) tidak berpotensi menular kepada lain, dan orang dengan HIV akan memiliki kualitas hidup yang lebih baik. Penemuan kasus pada stadium awal dan segera mendapatkan pengobatan ARV, membuat seseorang tidak jatuh pada HIV stadium lanjut (AIDS). Pada HIV stadium lanjut (AIDS) terjadi potensi masuknya infeksi-infeksi lainnya yang dikenal dengan infeksi oportunistik (Kemenkes RI, 2023).

Menurut Nyoman (1990), faktor-faktor yang diperkirakan meningkatkan angka kejadian HIV/AIDS antara lain lingkungan sosial ekonomi khususnya kemiskinan, latar belakang kebudayaan dan keadaan demografi masyarakat (banyaknya pelabuhan dan persinggahan orang asing). Kelompok masyarakat dengan potensi tinggi terkena HIV adalah penerima transfusi darah dari pendonor darah jika alat tidak steril, bayi dari ibu dengan status AIDS, pecandu narkotika, serta orang dengan banyak pasangan seks. Sebagian besar sebaran kasus HIV dan AIDS terdapat pada kelompok umur produktif 20-59 tahun. Hal ini disebabkan oleh rentang usia produktif lebih rentan terhadap perilaku berisiko seperti perilaku seks yang tidak aman atau penggunaan NAPZA suntik. Sementara itu masih ditemukan kasus HIV dan AIDS pada kelompok usia 1-4 tahun yang menunjukkan penularan HIV dari ibu ke anak (Kemenkes RI, 2023).

BAB II
METODOLOGI PENELITIAN

2.1 Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari laporan dan publikasi Kementerian Kesehatan Republik Indonesia (Kemenkes RI) dan Badan Pusat Statistik (BPS) Indonesia pada tahun 2022 dan 2023. Unit pengamatan data berupa provinsi di seluruh Indonesia pada tahun 2022 sebanyak 34 provinsi.

2.2 Identifikasi Variabel

Variabel yang digunakan dalam penelitian ini terdiri dari variabel respons dan variabel prediktor yang diuraikan pada Tabel 1 sebagai berikut.

Tabel 1. Definisi operasional variabel penelitian

Jenis Variabel	Nama Variabel	Definisi Operasional	Satuan
Respons	HIV (Y_1)	Jumlah kasus HIV tahun 2022.	Kasus
Respons	AIDS (Y_2)	Jumlah kasus AIDS tahun 2022.	Kasus
Prediktor	Persentase penduduk miskin (X_1)	Rasio antara jumlah penduduk yang memiliki rata-rata pengeluaran per kapita per bulan di bawah garis kemiskinan dengan jumlah penduduk dikali 100%.	Persen (%)
Prediktor	Persentase penduduk (X_2)	Rasio antara jumlah penduduk setiap wilayah dengan jumlah penduduk dikali 100%.	Persen (%)
Prediktor	Persentase orang dengan risiko terinfeksi HIV mendapatkan pelayanan deteksi dini HIV sesuai standar (X_3)	Rasio antara jumlah orang yang masuk kategori berisiko dan telah menjalani tes HIV sesuai standar dengan total populasi yang berisiko dikali 100%.	Persen (%)
Prediktor	Persentase kasus baru AIDS pada pengguna NAPZA suntikan (IDU) (X_4)	Rasio antara jumlah kasus baru AIDS pada pengguna NAPZA suntikan (IDU) dengan total kasus baru AIDS dikali 100%.	Persen (%)
Prediktor	Persentase pasangan usia produktif yang menggunakan kondom (X_5)	Rasio antara jumlah pasangan usia produktif yang melaporkan penggunaan kondom dengan total pasangan usia produktif dalam populasi tertentu dikali 100%.	Persen (%)

Lanjutan Tabel 1

Jenis Variabel	Nama Variabel	Definisi Operasional	Satuan
Prediktor	Persentase jumlah fasilitas kesehatan tempat pelayanan dan pengobatan HIV AIDS (X_6)	Rasio antara jumlah fasilitas kesehatan tempat pelayanan dan pengobatan HIV AIDS (puskesmas, RS tipe D, RS Pemerintah, RS Swasta) di setiap wilayah dengan jumlah fasilitas kesehatan tiap wilayah dikali 100%	Persen (%)

2.3 Metode Analisis

2.3.1 Estimasi Parameter Model *Bivariate Generalized Poisson Regression*

1. Membentuk fungsi *likelihood* dari model regresi BGP.
2. Melakukan transformasi bentuk persamaan $\lambda_{ij} + \lambda_0 = e^{\boldsymbol{x}_i^T \boldsymbol{\beta}_j}$ terhadap fungsi *likelihood.*
3. Membentuk fungsi *ln likelihood.*
4. Menurunkan fungsi *ln likelihood* terhadap parameter $\lambda_0, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \alpha_1, \alpha_2, \alpha_0$ kemudian disamakan dengan nol.
5. Hasil penurunan pertama terhadap masing-masing parameternya menghasilkan bentuk yang implisit maka untuk menyelesaikannya menggunakan algoritma BGFS dan Nelder-Mead.
6. Menentukan estimasi parameter dengan menggunakan algoritma BGFS dan Nelder-Mead.

2.3.2 Penerapan Model *Bivariate Generalized Poisson Regression*

1. Menguji distribusi *Bivariate Generalized Poisson* pada variabel respon Y_1 dan Y_2 menggunakan *Crockett's test* pada Persamaan (13).
2. Menguji korelasi antara variabel respon Y_1 dan Y_2 menggunakan uji korelasi *Pearson* pada Persamaan (24).
3. Mendeteksi overdispersi pada variabel respon Y_1 dan Y_2 menggunakan nilai *Deviance* pada Persamaan (7).
4. Menguji asumsi non-multikolinieritas pada variabel prediktor menggunakan nilai *Variance Inflation Factor* (VIF) pada Persamaan (25).
5. Memodelkan data dengan regresi BGPR.
6. Mengestimasi parameter model BGPR menggunakan metode MLE dengan algoritma BFGS dan Nelder-Mead.
7. Melakukan pemilihan model terbaik berdasarkan nilai *Akaike Information Criterion* (AIC) pada Persamaan (26)

8. Melakukan uji signifikansi secara simultan terhadap parameter model BGPR menggunakan MLRT pada Persamaan (20).
9. Melakukan uji signifikansi secara parsial terhadap parameter $\boldsymbol{\beta}_j(u_i, v_i)$ menggunakan uji *Wald* pada Persamaan (22).
10. Menginterpretasikan model yang diperoleh dan mengelompokkan variabel yang signifikan.