

# Pemodelan Regresi Binomial Negatif dan *Geographically Weighted Negative Binomial Regression* untuk Kasus HIV di Provinsi Papua

Fajriyatul Maulidiyah<sup>1</sup>, Tiani Wahyu Utami<sup>2</sup>, Rochdi Wasono<sup>3</sup>

<sup>123</sup>Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Muhammadiyah Semarang

Alamat e-mail : [maulidiyahfajriyatul@gmail.com](mailto:maulidiyahfajriyatul@gmail.com)

## ABSTRAK

Regresi Binomial Negatif merupakan model regresi nonlinier yang dapat digunakan untuk memodelkan data yang mengalami kasus overdispersi. HIV merupakan penyakit menular yang disebabkan oleh infeksi *Human Immunodeficiency Virus* yang menyerang sistem kekebalan tubuh. Salah satu provinsi penyumbang kasus HIV tertinggi di Indonesia adalah Provinsi Papua yang menempati peringkat kelima. Faktor-faktor yang diduga mempengaruhi jumlah kasus HIV antara lain banyaknya layanan PDP, persentase penduduk miskin, banyaknya sarana kesehatan, banyaknya daerah berstatus desa, kepadatan penduduk, dan persentase laki-laki. Penelitian ini bertujuan untuk mengetahui faktor-faktor yang mempengaruhi kasus HIV di Papua dengan pendekatan Regresi Binomial Negatif (RBN) dan *Geographically Weighted Negative Binomial Regression* (GWNBR). Jumlah kasus HIV paling banyak terjadi di Kota Jayapura sebanyak 1.114 kasus dan yang terendah terdapat pada 6 Kabupaten yaitu Kabupaten Asmat, Pegunungan Bintang, Sarmi, Supiori, Mamberamo raya, dan Ndunga dengan 0 kasus. Hasil analisis pemodelan jumlah kasus HIV di Provinsi Papua tahun 2017 dengan pemodelan GWNBR dan fungsi pembobot *adaptive bisquare kernel* diperoleh 9 kelompok berdasarkan variabel-variabel yang signifikan. Hasil penelitian menunjukkan bahwa metode RBN dalam jumlah kasus HIV di Provinsi Papua lebih baik dibandingkan dengan metode GWNBR dengan AIC sebesar 290,320.

**Kata Kunci :** Regresi Binomial Negatif, Overdispersi, HIV, GWNBR

## ABSTRACT

*Negative Binomial Regression is a nonlinear regression model that can be used to model data that has overdispersion cases. HIV is an infectious disease caused by infection with the Human Immunodeficiency Virus which attacks the immune system. One of the provinces contributing the highest HIV cases in Indonesia is Papua Province which is ranked fifth. Factors thought to influence the number of HIV cases include the number of PDP services, the percentage of poor population, the number of health facilities, the number of rural areas, population density, and the percentage of men. This study aims to determine the factors that influence HIV cases in Papua with the approach of Negative Binomial Regression (RBN) and Geographically Weighted Negative Binomial Regression (GWNBR). The highest number of HIV cases occurred in Jayapura City was 1,114 cases and the lowest was in 6 districts namely Asmat, Pegunungan Bintang, Sarmi, Supiori, Mamberamo raya, and Ndunga with 0 cases. The results of the modeling analysis of the number of HIV cases in Papua Province in 2017 by modeling the GWNBR and the adaptive bisquare kernel weighting function obtained 9 groups based on significant variables. The results showed that the RBN method in the number of HIV cases in Papua Province was better than the GWNBR method with AIC of 290.320..*

**Keywords :** Negative Binomial Regression, Overdispersion, HIV, GWNBR

## PENDAHULUAN

Salah satu tujuan analisis data ialah untuk memperkirakan atau memperhitungkan besarnya efek atau pengaruh kuantitatif dari perubahan suatu kejadian terhadap kejadian lainnya. Analisis regresi adalah analisis yang digunakan untuk memprediksikan seberapa jauh perubahan nilai variabel respon, bilai nilai variabel prediktor dimanipulasi/diubah-ubah atau dinaik-turunkan. Manfaat dari hasil analisis regresi adalah untuk membuat keputusan apakah naik atau menurunnya variabel respon dapat dilakukan melalui peningkatan variabel prediktor atau tidak (Sugiyono, 2016). Salah satu syarat yang harus dipenuhi dalam pembentukan model regresi dengan beberapa variabel prediktor adalah tidak ada kasus multikolinearitas. Pendeteksian kasus multikolinearitas yaitu dengan koefisien korelasi dan nilai VIF (*Variance Inflation Factor*). Terdapat kasus multikolinearitas jika nilai VIF lebih besar dari 10.

Distribusi Poisson sering digunakan untuk memodelkan peristiwa yang memiliki peluang kejadian kecil pada interval waktu tertentu (Agresti, 2002). Overdispersi memiliki dampak yang sama dengan pelanggaran asumsi yang akan menyebabkan anak dugaan dari parameter regresinya tetap konsisten namun tidak efisien. Hal ini berdampak pada nilai *standar error* yang menjadi *under estimate*, sehingga kesimpulan yang diperoleh menjadi tidak valid (Cameron & Trivedi, 1990). Salah satu solusi untuk mengatasi adanya kasus overdispersi adalah Regresi Binomial Negatif (RBN) yang merupakan perpaduan antara distribusi Poisson dan Gamma. Metode yang cukup efektif menduga data yang memiliki heterogenitas spasial untuk data *count* yang overdispersi adalah *Geographically Weighted Negative Binomial Regression* (GWNBR). Model GWNBR akan menghasilkan parameter lokal dengan masing-masing lokasi akan memiliki parameter yang berbeda-beda. Salah satu tujuan dalam analisis regresi adalah untuk mendapatkan model terbaik. Model terbaik adalah model yang mampu menjelaskan hubungan antara variabel prediktor dengan variabel respon berdasarkan kriteria tertentu. Kriteria yang sering digunakan dalam pemilihan model terbaik adalah AIC.

Memajukan bidang kesehatan merupakan salah satu implementasi pembangunan kesejahteraan umum guna mencapai tujuan nasional. *Millenium Development Goals* (MDGs) merupakan target nasional dalam memperbaiki kondisi kesehatan Indonesia. Salah satu tujuan dalam rangka mencapai MDGs yaitu dengan mengurangi angka penderita HIV dengan cara mencegah meluasnya kasus baru HIV melalui berbagai program kesehatan nasional serta melakukan penanganan profesional terhadap penderita HIV. *Human Immunodeficiency Virus* (HIV) adalah virus yang bila masuk ke dalam tubuh manusia akan menyerang sistem kekebalan tubuh. Tingkat HIV dalam tubuh dan timbulnya berbagai infeksi tertentu merupakan indikator bahwa infeksi HIV telah berkembang menjadi AIDS (*Acquired Immudofficiency Syndrome*). Populasi yang memiliki resiko tertinggi terkena virus HIV adalah LSL (Laki-laki Suka Laki-laki) atau Homoseksual, pengguna narkoba jarum suntik (IDU), dan Heteroseksual.

Jumlah kasus HIV yang dilaporkan dari tahun 2013 sampai tahun 2014 mengalami peningkatan, namu pada tahun 2015 mengalami penurunan kemudian pada tahun 2016 mengalami kenaikan kembali secara signifikan. Jumlah kumulatif infeksi HIV yang dilaporkan sampai dengan bulan desember 2017 sebanyak 48.300 kasus. Adapun provinsi dengan jumlah infeksi HIV tertinggi adalah provinsi Jawa Timur (8.024), diikuti Provinsi DKI Jakarta (6.626), Provinsi Papua (5.819), Provinsi Jawa Tengah (5.425) dan kemudian Provinsi Papua (4.358).

Papua merupakan provinsi dengan jumlah kasus HIV tertinggi kelima di Indonesia. Jumlah kasus HIV di kabupaten/kota di Papua cukup berbeda, hal itu terjadi karena perbedaan topografi dan geografis tiap daerah. Kota Jayapura sendiri merupakan daerah perkotaan terbesar dengan kasus HIV tertinggi di Papua tahun 2017 sebanyak 1.114 kasus. Berbeda dengan daerah kabupaten seperti Kabupaten Asmat, Kabupaten Pegunungan Bintang, Kabupaten Sarmi, Kabupaten Supiori, Kabupaten Mamberamo Raya, dan Kabupaten Ndunga dimana tidak terdapat kasus HIV atau nol (Provinsi Papua dalam Angka, 2017). Perbedaan topografi dan geografis ini menyebabkan

perbedaan jumlah kasus HIV di setiap daerah dan diduga ada keterkaitan hubungan antar kabupaten/kota terdekat di Papua.

Pencegahan merupakan salah satu tindakan pertama yang harus dilakukan oleh semua pihak terutama Pemerintah agar penyebaran HIV tidak semakin meluas. Upaya pencegahan bisa dilakukan melalui pendidikan dan penyuluhan masyarakat terutama ditujukan pada populasi beresiko yang mudah menyebarkan penyakit. Penularan dan penyebaran HIV sangat berhubungan dengan perilaku beresiko, oleh karena itu penganggungan harus memperhatikan faktor-faktor yang berpengaruh terhadap pertumbuhan kasus HIV di suatu daerah.

Penelitian sebelumnya mengenai kasus HIV telah dilakukan oleh Puspitasari dan Purnadi (2013) berjudul *Pemodelan Faktor yang Mempengaruhi jumlah Kasus HIV dan AIDS di Provinsi Jawa Timur Menggunakan Bivariate Poisson Regression*. Berdasarkan penelitian tersebut diketahui bahwa variabel persentase ketersediaan sarana kesehatan yang dibangun, persentase jaminan kesehatan masyarakat miskin, persentase penduduk yang mendonorkan darah, persentase penyuluhan kesehatan, dan persentase tenaga kesehatan masyarakat dibandingkan jumlah penduduk yang ada di setiap kabupaten dan kota berpengaruh terhadap jumlah kasus HIV di Jawa Timur. Untuk penelitian sebelumnya tentang HIV juga dilakukan oleh Anasi (2016) yang berjudul *Analisis Faktor-faktor yang Mempengaruhi jumlah kasus HIV di Jawa Timur tahun 2014 Menggunakan Geographically Weighted Negative Binomial Regression* yang berkesimpulan bahwa model GWNBR memiliki nilai AIC paling kecil jika dibandingkan dengan Regresi Binomial Negatif sehingga GWNBR lebih baik dalam memodelkan jumlah kasus HIV di Jawa Timur tahun 2014. Selain itu, penelitian yang dilakukan oleh Zuhry (2016) yang berjudul *Regresi Binomial Negatif Terboboti Geografis (GWNBR) untuk Mengetahui Faktor-faktor yang Mempengaruhi Jumlah Kasus Malaria di Provinsi Bengkulu*. Dari penelitian tersebut diperoleh 4 pengelompokan dari model regresi binomial negatif terboboti geografis berdasarkan variabel yang berpengaruh signifikan sehingga dapat diketahui faktor apa saja yang mempengaruhi

jumlah kasus malaria di setiap kabupaten/kota di provinsi Bengkulu.

Berdasarkan uraian yang telah dijelaskan, maka akan dibuat perbandingan dari dua metode yaitu Regresi Binomial Negatif (RBN) dan *Geographically Weighted Negative Binomial Regression* (GWNBR). Untuk mendapatkan hasil model yang terbaik, maka perbandingan akan menggunakan perhitungan nilai AIC pada kasus HIV di provinsi Papua.

## TINJAUAN PUSTAKA

### 1. Multikolinearitas

Salah satu syarat yang harus dipenuhi dalam pembentukan model regresi dengan beberapa variabel prediktor adalah tidak ada kasus multikolinieritas. Pendeteksian kasus multikolinieritas yaitu dengan koefisien korelasi dan kriteria nilai VIF (Variance Inflation Factor). Terdapat kasus multikolinieritas jika nilai VIF lebih besar dari 10. Nilai VIF dinyatakan sebagai berikut.

$$VIF_k = \frac{1}{1 - R_k^2} \quad (2.1)$$

dimana  $R_k^2 = \frac{SSR}{SST} = \frac{\sum_{i=1}^n (y_i - \bar{y})_k^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$

dengan  $R_k^2$  adalah koefisien determinasi antara satu variabel prediktor dengan variabel prediktor lainnya (Hocking, 1996).

### 2. Regresi Poisson

Model regresi Poisson merupakan model standar untuk data diskrit dan termasuk dalam model regresi nonlinier (Cameron, et.al, 1998). Regresi Poisson digunakan pada data yang berdistribusi poisson. Probabilitas distribusi Poisson diberikan oleh Myers (1990).

$$p(y; \mu) = \frac{e^{-\mu} \mu^y}{y!} \quad (y = 0, 1, 2, \dots) \quad (2.2)$$

dimana  $\mu$  adalah mean distribusi Poisson. Parameter  $\mu$  sangat bergantung pada beberapa unit yang ditetapkan atau periode waktu, jarak, luas, volume, dan lain-lain. Distribusi Poisson digunakan untuk memodelkan peristiwa yang jarang terjadi selama periode tertentu. Probabilitas banyak kejadian  $y$  dalam periode waktu  $t$  yaitu:

$$p(y; \mu) = \frac{e^{-\mu t} (\mu t)^y}{y!} \quad (y = 0, 1, 2, \dots) \quad (2.3)$$

Persamaan tersebut digunakan untuk probabilitas kejadian  $y$ , dan rata-rata jumlah kejadian, berdasarkan asumsi bahwa rata-rata jumlah kejadian per periode waktu adalah konstan. model regresi Poissonnya dapat ditulis sebagai berikut Myers, (1990):

$$y_i = \mu_i + \varepsilon_i (i = 1, 2, \dots, n) \quad (2.4)$$

dimana  $y_i$  adalah jumlah kejadian, dan  $\mu_i$  adalah rata-rata jumlah kejadian dalam periode  $t_i$ .  $\mu_i$  diasumsikan tidak berubah dari data ke data. Persamaan distribusi Poisson dinyatakan dengan persamaan sebagai berikut:

$$p(y_i; \hat{\beta}) = \frac{e^{-[\mu(x_i; \hat{\beta})]} [\mu(x_i; \hat{\beta})]^{y_i}}{y_i!} \quad (2.5)$$

### 3. Overdispersi

Metode regresi Poisson mewajibkan equidispersi, yaitu kondisi dimana nilai mean dan varians dari variabel respon bernilai sama. Namun, adakalanya terjadi fenomena overdispersi dalam data yang dimodelkan dengan distribusi Poisson. Overdispersi berarti varians lebih besar daripada mean. Taksiran dispersi diukur dengan devians atau Pearson's Chi-Square yang dibagi derajat bebas. Data overdispersi jika taksiran dispersi lebih besar dari 1 dan underdispersi jika taksiran dispersi kurang dari 1 (Khoshgoftaar, 2004 dalam Camelia P.S, Nur I.M dan Darsyah M.Y, 2016).

### 4. Aspek Spasial

Pemodelan pada data spasial dapat dikelompokkan berdasarkan tipe data spasial yang digunakan yaitu spasial titik dan spasial area. Masing-masing tipe data spasial tersebut dapat dikelompokkan lagi berdasarkan jenis data yang digunakan yaitu cross sectional dan time series. Pemodelan data spasial selalu melibatkan matriks pembobot spasial. Sedangkan efek spasial pada data dapat berupa error yang saling berkorelasi (dependensi spasial) maupun keragaman (heterogenitas) spasial antar lokasi.

Pengujian heterogenitas spasial dilakukan untuk melihat ciri khas pada setiap lokasi pengamatan yang akan mengakibatkan parameter regresi yang dihasilkan berbeda secara spasial. Pengujian heterogenitas spasial dilakukan dengan statistik uji Breusch-Pagan (BP).

### 5. Matriks Pembobot Spasial

Pembobot memiliki peranan penting pada data spasial, karena nilai suatu pembobot merupakan perwakilan dari lokasi dimana masing-masing data diambil. Informasi mengenai suatu lokasi dapat direpresentasikan oleh sebuah titik koordinat, seperti Garis Lintang dan Garis Bujur.

Berdasarkan informasi spasial tersebut dapat diperhitungkan jarak titik koordinat antar lokasi sehingga diharapkan kekuatan dari dependensi spasial akan menurun dengan adanya jarak tersebut. Lokasi yang berdekatan seharusnya menunjukkan hubungan kemiripan, begitu juga sebaliknya. Lokasi yang berjauhan juga memperlihatkan adanya keragaman spasial. Keragaman spasial antara lokasi yang satu dengan lokasi yang lain ditunjukkan dengan adanya matriks pembobot,  $W(u_i, v_i)$ , yang entri-entrinya merupakan fungsi dari jarak Euclidian antar lokasi. Besarnya pembobotan untuk model GWR di setiap lokasi dapat ditentukan dengan menggunakan fungsi kernel Adaptive Bisquare kernel, Penggunaan metode Adaptive Kernel sangat cocok apabila suatu pengamatan tersebar dengan pola yang acak, tidak beraturan dan berkelompok. Metode Adaptive Kernel memungkinkan untuk mendapatkan nilai bandwidth yang berbeda untuk setiap titik pengamatan. Hal ini dikarenakan metode Adaptive Kernel dapat menyesuaikan dengan kondisi titik pengamatan (Fotheringham, dkk, 2002).

### 6. Regresi Binomial Negatif

Regresi binomial negatif merupakan salah satu solusi untuk mengatasi adanya kasus overdispersi. Model regresi binomial negatif memiliki fungsi massa peluang sebagai berikut (Greene, 2008).

$$f(y, \mu, \theta) = \frac{\Gamma(y + \theta^{-1})}{\Gamma(\theta^{-1})\Gamma(y + 1)} \left( \frac{1}{1 + \theta\mu} \right)^{\theta^{-1}} \left( \frac{\theta\mu}{1 + \theta\mu} \right)^y, y = 0, 1, 2, \dots$$

Pendugaan parameter model regresi binomial negatif menggunakan metode Maximum Likelihood Estimation (MLE). Kemudian untuk mendapatkan nilai dari parameternya digunakan iterasi Newton Raphson. Fungsi likelihood dari regresi binomial negatif adalah sebagai berikut.

$$L(\beta, \theta) = \prod_{j=1}^n \left\{ \frac{\Gamma(y_j + \theta^{-1})}{\Gamma(\theta^{-1})\Gamma(y_j + 1)} \left( \frac{1}{1 + \theta\mu_j} \right)^{\theta^{-1}} \left( \frac{\theta\mu_j}{1 + \theta\mu_j} \right)^{y_j} \right\} \quad (2.24)$$

### 7. Geographically Weighted Negative Binomial Regression (GWNBR)

Geographically Weighted Negative Binomial Regression (GWNBR) merupakan salah satu metode yang cukup efektif menduga data yang memiliki heterogenitas spasial untuk data count yang over dispersi. Model GWNBR merupakan pengembangan dari model regresi binomial negatif. Model GWNBR menghasilkan parameter lokal dengan masing-masing lokasi akan memiliki parameter yang berbeda-beda. Model GWNBR

dapat dirumuskan sebagai berikut (Ricardo & Carvalho, 2013)

$$y_i \sim NB \left[ \exp(\beta_j(u_i, v_i) x_{ik}), \theta(u_i, v_i) \right], i = 1, 2, \dots, n \quad (2.26)$$

dimana,

$y_i$ : nilai observasi respon ke- $i$

$x_{ik}$ : nilai observasi variabel prediktor ke- $k$  pada pengamatan lokasi ( $u_i, v_i$ )

$\beta_j(u_i, v_i)$ : koefisien regresi variabel prediktor ke- $j$  untuk setiap lokasi ( $u_i, v_i$ )

$\theta(u_i, v_i)$ : parameter disperse untuk setiap lokasi ( $u_i, v_i$ )

Estimasi model GWNBR menggunakan Maximum Likelihood Estimation (MLE). Langkah awal dari metode Likelihood adalah membentuk fungsi Likelihood. Pengujian signifikansi parameter model GWNBR terdiri dari uji serentak dan parsial. Uji signifikansi serentak menggunakan Maximum Likelihood Ratio Test (MLRT).

#### 8. Kriteria Keabakan Model

Akaike Information Criterion (AIC) adalah kriteria kesesuaian model dalam menduga model secara statistik. Kriteria AIC digunakan apabila pemodelan regresi bertujuan untuk mengidentifikasi faktor-faktor yang berpengaruh terhadap model. 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. Nilai devians akan semakin kecil apabila rasio antara fungsi likelihood di bawah  $H_0$  dengan fungsi likelihood di bawah populasi semakin besar. Hal ini mengindikasikan bahwa parameter yang diuji semakin mendekati nilai parameter populasi yang sebenarnya yang berarti dugaan model semakin baik. Oleh karena itu, model terbaik adalah model dengan AIC terkecil dan dengan devians terkecil pula.

### METODE PENELITIAN

#### 1. Sumber Data

Data yang akan digunakan dalam penelitian ini merupakan data sekunder tahun 2017 yang diperoleh dari publikasi Badan Pusat Statistik (BPS) yaitu Provinsi Papua dalam Angka dan data kesehatan dari Dinas Kesehatan yaitu Profil Kesehatan Provinsi Papua. Unit observasi yang digunakan sebanyak 29 Kabupaten/Kota di Provinsi Papua.

#### 2. Variabel Penelitian

Variabel respon penelitian yang digunakan dalam penelitian ini merupakan data jumlah kasus HIV di Provinsi Papua menurut Kabupaten/Kota tahun 2017 sebagai variabel respon dan enam Variabel prediktor yang berhubungan dengan kasus HIV :

$Y$ = Jumlah kasus HIV di Provinsi Papua menurut Kabupaten/Kota tahun 2017

$X_1$ = Banyaknya Layanan PDP di Papua menurut Kabupaten/Kota tahun 2017

$X_2$ = Persentase Penduduk Miskin di Papua menurut Kabupaten/Kota tahun 2017

$X_3$ = Banyaknya Sarana Kesehatan di Papua menurut Kabupaten/Kota tahun 2017

$X_4$ = Banyaknya Daerah Berstatus Desa di Papua menurut Kabupaten/Kota tahun 2017

$X_5$ = Kepadatan Penduduk di Papua menurut Kabupaten/Kota tahun 2017

$X_6$ = Persentase Laki-laki di Papua menurut Kabupaten/Kota tahun 2017

#### 3. Analisis Data

Berikut merupakan langkah-langkah untuk analisis faktor-faktor yang mempengaruhi jumlah kasus HIV di Provinsi Papua :

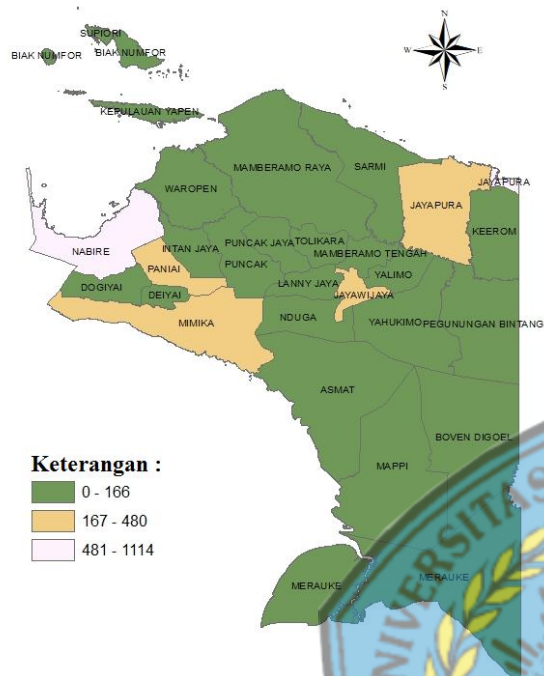
1. Mendeskripsikan gambaran umum dari data jumlah kasus HIV di Provinsi Papua serta variabel-variabel yang diduga mempengaruhi.
2. Mendeteksi adanya multikolineritas.
3. Melakukan uji distribusi Poisson.
4. Melakukan pengujian overdispersi.
5. Melakukan pemodelan untuk Regresi Binomial Negatif.
6. Melakukan pengujian aspek data spasial. Uji Moran's I untuk dependensi spasial data dan Breusch Pagan untuk heterogenitas spasial data.
7. Mendapatkan model *Geographically Weighted Negative Binomial Regression* (GWNBR)
8. Membandingkan model Regresi Binomial Negative dengan GWNBR dengan menggunakan nilai AIC.
9. Menarik kesimpulan.

### HASIL DAN PEMBAHASAN

1. Persebaran Jumlah Kasus HIV per Kabupaten dan Kota di Papua

Jumlah kasus HIV dan faktor yang mempengaruhinya dikategorikan menjadi tiga kelompok yaitu tinggi, sedang, dan rendah. Hal ini bertujuan untuk mempermudah menginterpretasi

hasil dari pemetaan. Berikut ini hasil dari pemetaan variabel respon yang digunakan dalam penelitian dengan menggunakan Kabupaten dan kota yang ada di Kabupaten/Kota Papua.



#### Persebaran HIV di Papua Berdasarkan Kabupaten/Kota tahun 2017

Wilayah yang masuk dalam kategori tinggi dalam jumlah kasus HIV adalah Kabupaten Nabire dan Kota Jayapura dengan 481-1114 kasus.

#### 2. Multikolinearitas

Uji multikolinearitas merupakan penyujian untuk melihat apakah terdapat hubungan antara masing-masing variabel bebas dengan variabel bebas lainnya. Apabila jika terdapat multikolinearitas, maka harus ada variabel yang harus direduksi sampai tidak terdapat korelasi antar variabel bebasnya. Asumsi Multikolinearitas ini harus terpenuhi agar bisa dilakukan pengujian menggunakan pendekatan Binomial Negatif Seperti yang sudah dijelaskan pada subbab 2.2.1 terdapat tiga kriteria yang bisa digunakan untuk melihat adanya multikolinearitas, salah satunya adalah nilai VIF. Berikut adalah table hasil nilai VIF dari masing-masing variabel bebas.

Tabel 4.1 Hasil VIF

| Variabel | Nilai VIF |
|----------|-----------|
| $X_1$    | 1,790     |
| $X_2$    | 1,655     |
| $X_3$    | 2,407     |
| $X_4$    | 2,255     |
| $X_5$    | 1,204     |

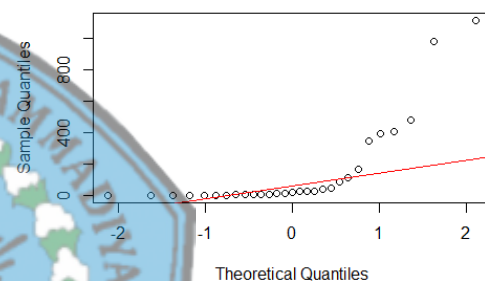
|       |       |
|-------|-------|
| $X_6$ | 1,425 |
|-------|-------|

Dari tabel diatas, dapat dilihat bahwa tidak terjadi multikolinearitas atau tidak terdapat hubungan antar variabel bebas, karena semua variabel bebas menunjukkan nilai VIF yang kurang dari 10 sehingga variabel di atas dapat digunakan untuk membentuk regresi poisson, maka seluruh variabel prediktor dapat digunakan untuk pembentukan model regresi poisson dan regresi binomial negatif.

#### 3. Uji Distribusi Poisson untuk Variabel Respon

Uji distribusi poisson untuk variabel predictor pada data HIV di provinsi Papua dapat dibuktikan dengan gambar 4.3 dibawah :

Normal Q-Q Plot



Plot Kuantil-kuantil Normal

Gambar diatas merupakan plot kuantil-kuantil normal untuk variabel jumlah kasus HIV di Provinsi Papua. Dari Plot di atas menunjukkan bahwa sebaran data tidak mengikuti garis lurus, ini menunjukkan penyimpangan data dari sebaran normal. selanjutnya adalah melakukan pengujian Overdispersi.

#### 4. Overdispersi

Salah satu asumsi yang harus dipenuhi dalam regresi binomial negatif adalah kondisi ekuidispersi. Namun dapat diketahui bahwa nilai overdispersi sebesar 0,002. Sehingga dapat disimpulkan bahwa pada untuk data jumlah kasus HIV di provinsi Papua mengalami kasus overdispersi.

#### 5. Pengujian Aspek Data Spasial

Pengujian aspek data spasial dilakukan dengan pengujian, yaitu pengujian heterogenitas spasial. Dari uji Breusch-Pagan diperoleh nilai p-value sebesar 0.002. Jika dibandingkan dengan  $\alpha$  sebesar 5% maka dapat disimpulkan bahwa  $H_0$  ditolak atau variansi antar lokasi berbeda. Sehingga terdapat perbedaan karakteristik antara

satu titik pengamatan dengan titik pengamatan lainnya.

#### 6. Regresi Binomial Negatif

Metode regresi binomial negatif juga dapat digunakan untuk menangani kasus overdispersi yang terjadi pada regresi poisson. Semua variabel prediktor dalam penelitian akan digunakan untuk memodelkan jumlah kasus HIV di Provinsi Papua menggunakan regresi binomial negatif. Hasil estimasi ditunjukkan dalam tabel 4.3.

| Parameter       | Estimasi | SE    | Z hitung |
|-----------------|----------|-------|----------|
| $\hat{\beta}_0$ | 2,010    | 5,167 | 0,389    |
| $\hat{\beta}_1$ | 0,329    | 0,157 | 2,098    |
| $\hat{\beta}_2$ | 0,015    | 0,035 | 0,438    |
| $\hat{\beta}_3$ | 0,142    | 0,055 | 2,557    |
| $\hat{\beta}_4$ | -0,002   | 0,003 | -0,506   |
| $\hat{\beta}_5$ | 0,008    | 0,005 | 1,673    |
| $\hat{\beta}_6$ | -0,010   | 0,043 | -0,227   |
| Devians         | 34,477   | -     | -        |
| AIC             | 292,320  | -     | -        |

Pengujian parameter model regresi binomial negatif secara serentak menggunakan nilai devians ditunjukkan yang oleh tabel di atas sebesar 34,477. Dengan menggunakan  $\alpha$  sebesar 5% sehingga nilai  $\chi^2_{(0,05;6)}$  sebesar 12,592. Hal ini menunjukkan bahwa nilai devians lebih besar dari nilai  $\chi^2_{(0,05;6)}$ . Sehingga dapat disimpulkan bahwa  $H_0$  ditolak yang berarti minimal terdapat satu variabel prediktor yang berpengaruh terhadap variabel respon.

Pengujian parameter secara parsial menggunakan statistik uji  $|Z_{hitung}|$  yang dibandingkan dengan  $Z_{0,05/2} = 1,965$ . Tabel di atas menunjukkan bahwa nilai  $|Z_{hitung}|$  dari variabel  $X_1$  dan  $X_3 > 1,965$ . Maka dapat diambil keputusan bahwa dari keenam variabel predictor hanya variabel  $X_1$  dan  $X_3$  yang gagal tolak  $H_0$  atau berpengaruh secara signifikan terhadap variabel  $Y$  dengan menggunakan model regresi binomial negatif. Sehingga persamaan model regresi binomial negatif yang diperoleh sebagai berikut:

$$\mu = \exp(2,010 + 0,329 X_1 + 0,015 X_2 + 0,142 X_3 - 0,002 X_4 + 0,008 X_5 - 0,010 X_6)$$

#### 7. Pemodelan Geographically Weighted Negative Binomial Regression (GWNBR)

Pemodelan GWNBR merupakan bentuk lokal dari regresi binomial negatif karena memperhatikan letak geografis sehingga perlu adanya pembobot lokasi. Sebelum menghitung matriks pembobot yang harus dilakukan yaitu mencari jarak Euclidean ( $d_{ij}$ ) antar provinsi

terlebih dahulu. Untuk nilai bandwidth optimum (h) pemodelan GWNBR didapatkan dengan melihat hasil *Cross Validation* (CV) terkecil. Dari hasil pada lampiran didapatkan bandwidth nilai CV terkecil yaitu 0,343. Setelah mendapatkan nilai  $d_{ij}$  pada wilayah yang ditaksir parameternya, selanjutnya menyusun matriks pembobot spasial dengan menggunakan bandwidth optimum.

matriks pembobot *Adaptive Bisquare Kernel* adalah sebagai berikut:

$$W(i)(u_i, v_i) = \text{diag}[1; 4,379; 1,667; 1,760; \dots \dots \dots; W_{29}(u_i, v_i)0; \dots \dots \dots; W_{27}(u_i, v_i)]$$

Dalam penentuan matriks pembobot akan dipilih salah satu pembobot yang terbaik dengan melihat nilai AIC terkecil.

Pembentukan matriks pembobot digunakan untuk menaksir parameter berdasarkan lokasi ( $u_2, v_2$ ). Jadi untuk penaksiran parameter dari lokasi ( $u_2, v_2$ ) sampai ( $u_{29}, v_{29}$ ) menggunakan matriks yang berbeda.

#### 8. Estimasi Parameter GWNBR

##### a. Pengujian Kesamaan Model

Berdasarkan hasil perhitungan didapatkan nilai  $f_{hitung}$  sebesar 4,345. Dengan menggunakan  $\alpha$  sebesar 5% maka diperoleh nilai  $f_{(0,05;29;22)} = 2,58$  yang dapat diartikan bahwa nilai  $f_{hitung} > f_{(0,05;29;22)}$ . Sehingga  $H_0$  ditolak atau terdapat perbedaan yang signifikan antara model regresi binomial negatif dengan model GWNBR.

##### b. Pengujian Serentak

Nilai devians yang dihasilkan sebesar 9307,564 dengan menggunakan  $\alpha$  sebesar 5% sehingga nilai  $\chi^2_{(0,05;6)}$  sebesar 12,292. Hal ini menunjukkan bahwa nilai devians lebih besar dari  $\chi^2_{(0,05;6)}$ . Maka dapat disimpulkan bahwa  $H_0$  ditolak yang artinya minimal ada satu variabel prediktor yang berpengaruh signifikan terhadap model regresi GWNBR

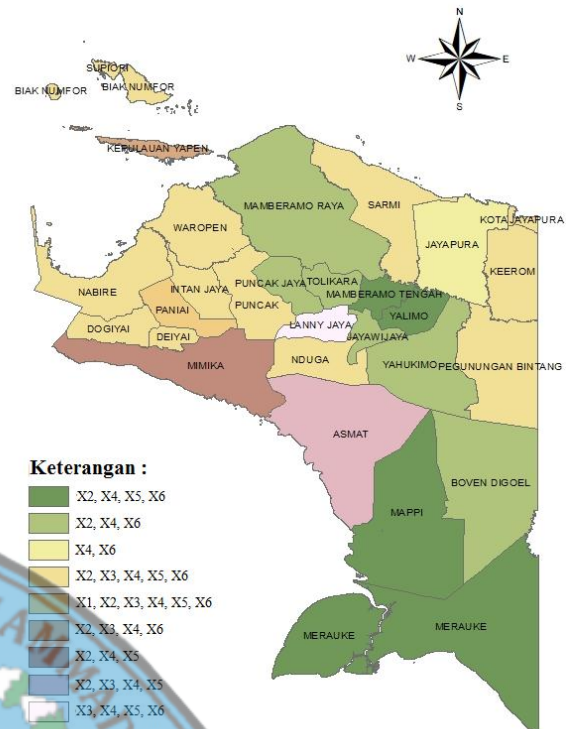
##### c. Pengujian Parsial

Hasil pengujian secara parsial menghasilkan parameter yang berbeda untuk setiap kabupaten/kota. Nilai  $|Z_{hitung}|$  yang dibandingkan dengan  $Z_{0,05/2} = 1,965$ .  $H_0$  ditolak jika nilai  $|Z_{hitung}| > 1,965$ . Variabel yang signifikan di setiap kabupaten/kota di Papua disajikan dalam table berikut :

| Variabel Signifikan            | Kabupaten/Kota              | Jumlah |
|--------------------------------|-----------------------------|--------|
| $X_2, X_4, X_5, X_6$           | Kabupaten Merauke,          | 4      |
|                                | Kabupaten Mappi,            |        |
|                                | Kabupaten Mamberamo Tengah, |        |
|                                | Kabupaten Yalimo            |        |
| $X_2, X_4, X_6$                | Kabupaten Jayawijaya,       | 6      |
|                                | Kabupaten Puncak Jaya,      |        |
|                                | Kabupaten Boven Digoel,     |        |
|                                | Kabupaten Yahukimo,         |        |
|                                | Kabupaten Tolikara,         |        |
|                                | Kabupaten Mamberamo Raya    |        |
| $X_4, X_6$                     | Kabupaten Jayapura          | 1      |
| $X_2, X_3, X_4, X_5, X_6$      | Kabupaten Nabire,           | 13     |
|                                | Kabupaten Biak Numfor,      |        |
|                                | Pegunungan Bintang,         |        |
|                                | Kabupaten Sarmi,            |        |
|                                | Kabupaten Keerom,           |        |
|                                | Kabupaten Waropen,          |        |
|                                | Kabupaten Supiori,          |        |
|                                | Kabupaten Nduga,            |        |
|                                | Kabupaten Puncak,           |        |
|                                | Kabupaten Dogiyai,          |        |
|                                | Kota Jayapura,              |        |
|                                | Kabupaten Deiyai,           |        |
|                                | Kabupaten Intan Jaya        |        |
| $X_1, X_2, X_3, X_4, X_5, X_6$ | Kabupaten Paniai            | 1      |
| $X_2, X_3, X_4, X_6$           | Kabupaten Kepulauan Yapen   | 1      |
| $X_2, X_4, X_5$                | Kabupaten Mimika            | 1      |
| $X_2, X_3, X_4, X_5$           | Kabupaten Asmat             | 1      |
| $X_3, X_4, X_5, X_6$           | Kabupaten Lanny Jaya        | 1      |

Berdasarkan hasil pengujian parsial untuk pemodelan GWNBR diperoleh 9 pengelompokan seperti pada tabel 4.5. Hasil pengelompokan menunjukkan bahwa terdapat satu variabel yang berpengaruh signifikan untuk semua kabupaten/kota di Papua yaitu variabel  $X_4$  (banyaknya daerah berstatus desa). Sedangkan untuk variabel  $X_1$  (banyaknya layanan PDP) hanya berpengaruh secara signifikan terhadap jumlah kasus HIV untuk semua kabupaten/kota di

Papua untuk kelompok 5 berdasarkan variabel yang signifikan.



## 9. Pemilihan Model Terbaik

Pemilihan model terbaik didasarkan pada nilai AIC. Model yang baik akan memiliki nilai AIC yang kecil. Nilai AIC untuk setiap model regresi ditampilkan dalam table berikut.

| Metode                   | AIC     |
|--------------------------|---------|
| Regresi Binomial Negatif | 290,320 |
| GWNBR                    | 346,499 |

Dari tabel 4.7 dapat diketahui bahwa nilai AIC yang dihasilkan oleh metode Regresi Binomial Negatif (NBR) dan GWNBR berbeda, dimana metode regresi binomial negatif memiliki nilai AIC lebih kecil. Sehingga metode yang sesuai untuk memodelkan faktor-faktor yang mempengaruhi jumlah kasus HIV di setiap kabupaten/kota di provinsi Papua adalah metode Regresi Binomial Negatif (RBN).

## KESIMPULAN

Kesimpulan yang diperoleh berdasarkan analisis yang dilakukan sebagai berikut :

1. Kota Jayapura memiliki jumlah kasus HIV paling tinggi yaitu sebanyak 1.114 kasus. Sedangkan Kabupaten Asmat, Kabupaten Pegunungan Bintang, Kabupaten Sarmi, Kabupaten Supiori, Kabupaten Mamberamo

raya, dan Kabupaten Ndunga memiliki jumlah kasus HIV terendah yaitu sebanyak 0 atau tidak terdapat kasus HIV pada keenam kabupaten tersebut pada tahun 2017.

2. Hasil pemodelan metode Regresi Binomial Negatif (RBN) menghasilkan 2 variabel yang signifikan terhadap jumlah kasus HIV yaitu variabel  $X_1$  (banyaknya lahan PDP) dan  $X_3$  (banyaknya sarana kesehatan). Sedangkan dengan metode GWNBR terbentuk sembilan kelompok kabupaten/kota menurut variabel yang berpengaruh signifikan terhadap jumlah kasus HIV di Provinsi Papua. Hasil pemodelan GWNBR menunjukkan bahwa terdapat satu variabel yang berpengaruh signifikan untuk seluruh kabupaten/kota di Provinsi Papua yaitu variabel  $X_4$  (banyaknya daerah berstatus desa).
3. Kriteria AIC menunjukkan bahwa metode Regresi Binomial Negatif (RBN) lebih sesuai untuk memodelkan faktor-faktor yang mempengaruhi jumlah kasus HIV di setiap kabupaten/kota di provinsi Papua jika dibandingkan dengan metode GWNBR karena metode RBN memiliki nilai AIC lebih kecil.

#### DAFTAR PUSTAKA

- Agresti, A. (2002). *Categorical Data Analysis Second Edition*. New York: John Wiley & Sons.
- Badan Informasi Geospasial. (2017). Diambil kembali dari Badan Informasi Geospasial: <http://www.big.go.id/peta-tematik/>
- Badan Pusat Statistik Provinsi Papua. (2017). *Provinsi Provinsi Papua dalam Angka 2017*. <https://papua.bps.go.id/publication/download.html>
- Bappenas, 2014. *Prakarsa Strategis MDGs Acceleration Framework Penanggulangan HIV dan AIDS di Indonesia Tahun 2014*. Jakarta : Direktorat Kesehatan dan Gizi Masyarakat.
- Cameron, A. C., & Trivedi, P. K. (1990). Regression-Based Test For Overdispersion In The Poisson Model. *Journal of Econometrics*, vol. 46 No. 1, 347-346.
- Crisana, C. W. (2014). *Analisis Perbandingan Metode Klasifikasi Autocorrelation Based Regioclassification (ACRC) dan Non-ACRD untuk Data Spasial*. Bogor: FMIPA IPB.
- Departemen Kesehatan RI. (2018). *Hari AIDS Sedunia, Momen Stop Penularan HIV*. Diambil kembali
- <http://www.depkes.go.id/article/view/1812030000>
- Dinas Kesehatan Provinsi Papua. *Profil Kesehatan Provinsi Papua Tahun 2012*. Semarang: Dinas Kesehatan Provinsi Papua
- Draper, N.R., & Smith, H. (1996). *Applied Regression Analysis (2nd ed.)* New York: John Wiley & Sons.
- Kementrian Kesehatan Republik Indonesia Direktorat Jenderal Pencegahan dan Penanganan Penyakit. (2018). Jakarta : CV. Mitra Karya Putra
- Kurniawan, Deny. (2008). Tabel Distribusi Dilengkapi Metode Untuk Membaca Tabel Distribusi. Forum Statistika. [http://lkeb.umm.ac.id/files/file/tabel\\_distribusi.pdf](http://lkeb.umm.ac.id/files/file/tabel_distribusi.pdf).
- Puspitasari, Lucy Dian & Purhadi. (2013). *Pemodelan Faktor-faktor yang Mempengaruhi Jumlah Kasus HIV & AIDS di Provinsi Jawa Timur Tahun 2013 Menggunakan Bivariate Poisson Regression*. Diambil kembali dari [http://ejurnal.its.ac.id/index.php/sains\\_seni/article/view/9976](http://ejurnal.its.ac.id/index.php/sains_seni/article/view/9976).
- Ricardo, A., & Carvalho, T. (2013). *Geographically Weighted Negative Binomial Regression-Incorporating Overdispersion*. Business Media New York: Springer Science.
- Sugiyono. (2016). *Statistik untuk Penelitian*. Bandung: Alfabeta.
- Walpole R. (1995). *Pengantar Metode Statistika. Dalam Edisi Ketiga, Alih Bahasa* : Bambang Sumantri, Jakarta: PT Gramedia Pusaka Utama.
- Yunitasari, Isca. (2017). *Penerapan Model Geographically Weighted Poisson Regression pada Jumlah Kematian Ibu di Provinsi Jawa Tengah*. Diambil kembali dari <https://eprints.uns.ac.id/38234/>
- Zuhry, Ummu Kholishoh. (2016). *Regresi Binomial Negatif Terboboti Geografis untuk Mengetahui Faktor-faktor yang Mempengaruhi Jumlah Kasus Malaria di Provinsi Bengkulu*.