



TUGAS AKHIR - SS141501

**PEMODELAN DAN PEMETAAN KASUS DEMAM  
BERDARAH *DENGUE* DI PROVINSI JAWA  
TIMUR TAHUN 2014 DENGAN *GENERALIZED  
POISSON REGRESSION*, REGRESI BINOMIAL  
NEGATIF DAN *FLEXIBLY SHAPED SPATIAL  
SCAN STATISTIC***

**NENDY SEPTI ARNIVA  
NRP 1314 105 007**

**Dosen Pembimbing  
Dr. Purhadi M.Sc**

**PROGRAM STUDI S1  
JURUSAN STATISTIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA 2016**



FINAL PROJECT - SS141501

**MODELING AND MAPPING OF *DENGUE HEMORRHAGIC FEVER* CASE IN EAST JAVA 2014 USING GENERALIZED POISSON REGRESSION, NEGATIVE BINOMIAL REGRESSION AND FLEXIBLY SHAPED SPATIAL SCAN STATISTIC**

**NENDY SEPTI ARNIVA  
NRP 1314 105 007**

**Supervisor  
Dr. Purhadi M.Sc**

**UNDERGRADUATE PROGRAMME  
DEPARTMENT OF STATISTICS  
FACULTY OF MATHEMATICS AND NATURAL SCIENCES  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA 2016**

**LEMBAR PENGESAHAN**

**PEMODELAN DAN PEMETAAN KASUS DEMAM  
BERDARAH *DENGUE* DI PROVINSI JAWA TIMUR  
TAHUN 2014 DENGAN *GENERALIZED POISSON  
REGRESSION*, REGRESI BINOMIAL NEGATIF DAN  
*FLEXIBLY SHAPED SPATIAL SCAN STATISTIC***

**TUGAS AKHIR**

Diajukan Untuk Memenuhi Salah Satu Syarat  
Memperoleh Gelar Sarjana Sains  
pada  
Program Studi S-1 Jurusan Statistika  
Fakultas Matematika dan Ilmu Pengetahuan Alam  
Institut Teknologi Sepuluh Nopember

Oleh :

**NENDY SEPTI ARNIVA  
NRP. 1314 105 007**


Disetujui oleh Pembimbing Tugas Akhir :

**Dr. Purnhadi, M.Sc**  
**NIP. 19620204 198701 1 001**



(.....)

Mengetahui,  
Ketua Jurusan Statistika FMIPA-ITS



**Dr. Suhartono**  
**NIP. 19710929 199512 1 001**

**SURABAYA, JULI 2016**

**PEMODELAN DAN PEMETAAN KASUS DEMAM  
BERDARAH *DENGUE* DI PROVINSI JAWA TIMUR  
TAHUN 2014 DENGAN *GENERALIZED POISSON  
REGRESSION*, REGRESI BINOMIAL NEGATIF  
DAN *FLEXIBLY SHAPED SPATIAL SCAN STATISTIC***

**Nama Mahasiswa** : Nendy Septi Arniva  
**NRP** : 1314 105 007  
**Jurusan** : Statistika FMIPA-ITS  
**Dosen Pembimbing** : Dr. Purhadi, M. Sc

**Abstrak**

*Salah satu penyakit menular yang menjadi masalah utama kesehatan masyarakat di Indonesia adalah penyakit Demam Berdarah Dengue (DBD). DBD ditularkan oleh vektor nyamuk Aedes Aegypti melalui gigitan. Tempat yang paling banyak nyamuk ini berada yaitu di lingkungan yang buruk, lembab. Salah satu Provinsi yang Kasus DBD masih tinggi yaitu Provinsi Jawa Timur. Jumlah kasus DBD di Jawa Timur hingga akhir Januari 2015 sebanyak 1.817 kasus. Dalam penelitian ini dilakukan pemodelan jumlah kasus DBD dengan metode Generalized Poisson Regression (GPR) dan Regresi Binomial Negatif. Metode GPR dan regresi binomial negatif dilakukan untuk menanggapi kasus overdispersi pada kasus jumlah DBD di Jawa Timur tahun 2014. Pendeteksian wilayah yang menjadi kantong DBD menggunakan metode Flexibly Shaped Spatial Scan Statistic. Pemodelan menggunakan GPR dan BNR didapatkan variabel berpengaruh terhadap kasus DBD di Jawa Timur Tahun 2014 adalah persentase rumah sehat dan rasio sarana kesehatan per 1000 penduduk. Hasil pendeteksian kantong DBD berdasarkan model GPR dan BNR didapat sepuluh kantong DBD. Kantong dengan nilai resiko tertinggi terdapat pada Kota Surabaya dengan nilai resiko sebesar 3,32 dengan pemodelan GPR, dan 2,84 dengan pemodelan BNR.*

**Kata kunci** : *DBD, Flexibly Shaped Spatial Scan Statistic, GPR, Regresi Binomial Negatif*

**MODELING AND MAPPING OF *DENGUE*  
HEMORRHAGIC FEVER CASE IN EAST JAVA 2014  
USING GENERALIZED POISSON REGRESSION,  
BINOMIAL NEGATIVE REGRESSION AND FLEXIBLY  
SHAPED SPATIAL SCAN STATISTIC**

**Student Name** : Nendy Septi Arniva  
**NRP** : 1314 105 007  
**Department** : Statistics FMIPA-ITS  
**Supervisor** : Dr. Purhadi, M. Sc

**Abstract**

*One of the infectious diseases that became major public health problem in Indonesia is a disease of Dengue Hemorrhagic Fever (DHF). DHF is transmitted by Aedes Aegypti mosquito vectors through a bite. The most common places that mosquitoes are located in a bad neighborhood, moist. One province is still high which case DBD East Java Province. Number of DHF cases in East Java until the end of January 2015 as many as 1,817 cases. In this research, modeling the number of DHF cases with Generalized Poisson Regression method (GPR) and Binomial Negative Regression. GPR method and binomial negative regression is done in order to tackle cases of overdispersi the number of dengue cases in East Java in 2014. The detection of regions where there is a DHF using Shaped Flexible Spatial Scan Statistic. Modelling using GPR and BNR obtained variables affect the DHF cases in East Java in 2014 is the percentage of healthy house and the ratio of health facilities per 1000 population. The results of the detection of DHF bags by model GPR and BNR gained ten bags of DHF. Bag with the risk value is highest in the city of Surabaya with the risk value of 3.32 with GPR modeling, and 2.84 with modeling BNR.*

**Keywords** : *Binomial Negative Regression, DHF, Flexibly Shaped Spatial Scan Statistic, GPR*

## DAFTAR ISI

	Halaman
<b>HALAMAN JUDUL</b> .....	i
<b>LEMBAR PENGESAHAN</b> .....	v
<b>ABSTRAK</b> .....	vii
<b>ABSTRACT</b> .....	ix
<b>KATA PENGANTAR</b> .....	xi
<b>DAFTAR ISI</b> .....	xiii
<b>DAFTAR GAMBAR</b> .....	xvii
<b>DAFTAR TABEL</b> .....	xix
<b>DAFTAR LAMPIRAN</b> .....	xxi
 <b>BAB I PENDAHULUAN</b>	
1.1 Latar Belakang.....	1
1.2 Rumusan Permasalahan .....	3
1.3 Tujuan Penelitian .....	3
1.4 Manfaat Penelitian .....	4
1.5 Batasan Masalah .....	4
 <b>BAB II TINJAUAN PUSTAKA</b>	
2.1 Statistik Deskriptif .....	5
2.2 Multikolinearitas .....	5
2.3 Regresi Poisson.....	6
2.3.1 Penaksiran Parameter Model Regresi Poisson .....	8
2.3.2 Pengujian Parameter Model Regresi Poisson .....	11
2.3.3 Deteksi Overdispersi Pada Regresi Poisson .....	12
2.4 <i>Generalized Poisson Regression</i> (GPR) .....	13
2.4.1 Penaksiran Parameter Model GPR .....	14
2.4.2 Pengujian Parameter Model GPR.....	16
2.5 Regresi Binomial Negatif .....	17
2.5.1 Penaksiran Parameter Model Regresi Binomial Negatif .....	17
2.5.2 Pengujian Parameter Model Regresi Binomial Negatif .....	21
2.6 Pemilihan Model Terbaik .....	22

2.7 Metode <i>Flexibly Shaped Spatial Scan Statistic</i> .....	23
2.7.1 <i>Likelihood Ratio Test</i> dan Pengujian Hipotesis <i>Monte Carlo</i> .....	25
2.8 Demam Berdarah <i>Dengue</i> (DBD) .....	27
<b>BAB III METODOLOGI PENELITIAN</b>	
3.1 Sumber Data .....	29
3.2 Variabel Penelitian.....	29
3.3 Langkah Analisis .....	32
3.4 Diagram Alir Penelitian .....	34
<b>BAB IV ANALISIS DAN PEMBAHASAN</b>	
4.1 Deskripsi Kasus Demam Berdarah <i>Dengue</i> dan Faktor-Faktor yang diduga Mempengaruhi .....	35
4.1.1 Jumlah Kasus Demam Berdarah <i>Dengue</i> .....	36
4.1.2 Kepadatan Penduduk .....	37
4.1.3 Persentase Rumah Tangga yang Berperilaku Hidup Bersih dan Sehat (PHBS).....	39
4.1.4 Persentase Rumah Sehat.....	40
4.1.5 Rasio Sarana Kesehatan per 1000 Penduduk .....	41
4.1.6 Rasio Tenaga Kesehatan per 1000 Penduduk.....	42
4.2 Pemeriksaan Multikolinearitas.....	44
4.3 Pemodelan Jumlah Kasus DBD tiap Kabupaten/ Kota di Jawa Timur Tahun 2014 Menggunakan Regresi Poisson .....	45
4.4 Pemeriksaan Overdispersi.....	46
4.5 Pemodelan Jumlah Kasus DBD tiap Kabupaten/ Kota di Jawa Timur Tahun 2014 Menggunakan <i>Generalized Poisson Regression</i> (GPR).....	47
4.6 Pemodelan Jumlah Kasus DBD tiap Kabupaten/ Kota di Jawa Timur Tahun 2014 Menggunakan Regresi Binomial Negatif.....	49
4.7 Pemilihan Model Terbaik .....	50
4.8 Pendeteksian Kantong DBD di Jawa Timur Menggunakan <i>Flexibly Shaped Spatial Scan Statistic</i> ...	51

4.8.1 Pendeteksian Kantong dengan Pemodelan Regresi Binomial Negatif.....	51
<b>BAB V KESIMPULAN DAN SARAN</b>	
5.1 Kesimpulan .....	55
5.2 Saran .....	56
<b>DAFTAR PUSTAKA .....</b>	<b>57</b>
<b>LAMPIRAN.....</b>	<b>59</b>



## DAFTAR GAMBAR

	Halaman
<b>Gambar 3.1</b> Provinsi Jawa Timur .....	29
<b>Gambar 3.2</b> Diagram Alir Penelitian .....	34
<b>Gambar 4.1</b> Penyebaran Jumlah Kasus DBD .....	37
<b>Gambar 4.2</b> Penyebaran Kepadatan Penduduk .....	38
<b>Gambar 4.3</b> Penyebaran Persentase Rumah Tangga ber-PHBS .....	39
<b>Gambar 4.4</b> Penyebaran Persentase Rumah Sehat.....	41
<b>Gambar 4.5</b> Penyebaran Sarana Kesehatan per 1000 Penduduk.....	42
<b>Gambar 4.6</b> Penyebaran Tenaga Kesehatan per 1000 Penduduk.....	43
<b>Gambar 4.7</b> Hasil Pendeteksian Kantong DBD dengan Pemodelan Regresi Binomial Negatif .....	53

## DAFTAR TABEL

	Halaman
<b>Tabel 3.1</b> Variabel Penelitian .....	30
<b>Tabel 4.1</b> Statistika Deskriptif Variabel Penelitian .....	36
<b>Tabel 4.2</b> Koefisien Korelasi Antar Variabel Prediktor .....	44
<b>Tabel 4.3</b> Nilai VIF dari Variabel Prediktor.....	44
<b>Tabel 4.4</b> Estimasi Parameter Model Regresi Poisson .....	45
<b>Tabel 4.5</b> Nilai <i>Deviance</i> dan <i>Pearson</i> dari Model Regresi Poisson .....	47
<b>Tabel 4.6</b> Estimasi Parameter Model GPR.....	47
<b>Tabel 4.7</b> Estimasi Parameter Model Regresi Binomial Negatif.....	49
<b>Tabel 4.8</b> Pemilihan Model Terbaik.....	50
<b>Tabel 4.9</b> Hasil Deteksi Kantong DBD dengan Pemodelan Regresi Binomial Negatif.....	52

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Demam Berdarah *Dengue* (DBD) merupakan masalah utama kesehatan masyarakat di Indonesia dan angka kematian DBD selalu meningkat dari tahun ke tahun. DBD sering muncul sebagai kejadian luar biasa (KLB) dengan angka kematian relatif tinggi. DBD ditularkan oleh vektor nyamuk *Aedes Aegypti* melalui gigitan. Penyakit demam berdarah biasanya didahului dengan adanya demam tinggi, nyeri pada kepala dan otot. Tempat yang paling banyak nyamuk ini berada yaitu di lingkungan yang lembab, curah hujan yang tinggi, genang air di dalam maupun di luar rumah. Faktor lain penyebab DBD ini adalah kepadatan penduduk, sanitasi lingkungan yang buruk, perilaku masyarakat yang tidak sehat.

Departemen Kesehatan Republik Indonesia telah mengupayakan berbagai strategi dalam mengatasi kasus DBD yang salah satu diantaranya adalah dengan cara Pemberantasan Sarang Nyamuk Demam Berdarah Dengue (PSN DBD), namun ternyata Kasus DBD di Provinsi Jawa Timur masih tinggi. Menurut kementerian kesehatan sebanyak 1.817 kasus demam berdarah *dengue* (DBD) di Provinsi Jawa Timur hingga akhir Januari 2015. Peningkatan kasus DBD ini hingga 980 kasus dari sebelumnya.

Regresi poisson merupakan analisis regresi nonlinear dari distribusi poisson dan digunakan dalam menganalisis data diskrit (*count*). Salah satu contoh data diskrit (*count*) yaitu jumlah kasus demam berdarah *dengue* di Provinsi Jawa Timur pada penelitian ini. Salah satu permodelan regresi poisson yaitu *Generalized Poisson Regression (GPR)*, model ini dapat mengatasi masalah dengan varian lebih besar dibandingkan dengan rata-rata yang disebut *overdispersion*, karena tidak mengharuskan nilai *mean* yang sama dengan nilai varians. Mengatasi overdispersi pada regresi poisson selain menggunakan *Generalized Poisson*

*Regression* dapat dilakukan dengan menggunakan pendekatan model regresi Binomial Negatif. Penelitian sebelumnya yang telah dilakukan oleh Rahayu (2012) memperoleh hasil dengan penerapan analisis *GPR* dan regresi binomial negatif pada pemodelan pengaruh iklim terhadap angka kejadian demam berdarah *dengue* di Surabaya. Pada penelitian tersebut ditemukan adanya kasus overdispersi pada regresi poisson. Hasil yang diperoleh adalah regresi binomial negatif merupakan model terbaik dalam mengatasi kasus overdispersi pada regresi poisson karena menghasilkan nilai AIC (*Akaike Information Criterion*) terkecil.

Salah satu strategi untuk mengatasi penyebaran penyakit DBD dengan mengetahui daerah yang menjadi kantong-kantong DBD atau *hotspot* dengan cara memetakan daerah tersebut. Salah satu metode yang digunakan untuk mendeteksi *hotspot* yaitu *Flexibly Shaped Spatial Scan Statistic*. Penelitian sebelumnya yang telah dilakukan oleh Hidayanti (2015) memperoleh hasil dengan menggunakan *Geographically Weighted Negative Binomial Regression* (GWNBR) dan *Flexibly Shaped Spatial Scan Statistic* menyebutkan bahwa pemodelan jumlah kasus DBD didapatkan sebanyak 12 kelompok, dari hasil proses pendeteksian kantong DBD menunjukkan bahwa terbentuk tiga kantong DBD. Penelitian lain terkait dengan *Flexibly Shaped Spatial Scan Statistic* dilakukan oleh Sari (2016) memperoleh hasil pendeteksian kantong DBD di provinsi Jawa Timur terbentuk 12 kantong DBD, yang mana Kota Surabaya merupakan daerah yang paling rawan terkena DBD.

Untuk mengetahui wilayah yang paling terkena penyakit Demam Berdarah *Dengue* (DBD), maka dilakukan penelitian mengenai pemetaan wilayah di Jawa Timur berdasarkan penderita Demam Berdarah *Dengue* (DBD). Penanganan kasus overdispersi dilakukan dengan mengaplikasikan metode *Generalized Poisson Regression* (*GPR*) dan Regresi Binomial Negatif untuk mendapatkan model terbaik pada kasus DBD di setiap kabupaten/kota di Provinsi Jawa Timur tahun 2014. Untuk

mengetahui kantong-kantong DBD dengan *Flexibly Shaped Spatial Scan Statistic*, sehingga upaya menangani/ mengatasi wilayah yang paling banyak terjangkit DBD dengan adanya pemetaan wilayah ini, sehingga penanganan lebih efektif dan tepat sasaran dan mengurangi penderita penyakit Demam Berdarah *Dengue* (DBD).

## 1.2 Rumusan Masalah

Demam berdarah *dengue* merupakan salah satu penyakit menular yang mana angka kematian DBD selalu meningkat dan DBD sering muncul sebagai kejadian luar biasa (KLB). Dalam kasus jumlah DBD ini sering terjadi kasus overdispersi maka metode yang digunakan untuk mengatasi overdispersi dan faktor-faktor yang mempengaruhi DBD yaitu dengan membandingkan *Generalized Poisson Regression* dan regresi binomial negatif. Faktor yang diduga mempengaruhi kasus DBD di Provinsi Jawa Timur tahun 2014 yaitu faktor lingkungan, perilaku kesehatan masyarakat serta fasilitas kesehatan. Selanjutnya metode yang digunakan untuk menentukan kantong-kantong DBD di Provinsi Jawa Timur tahun 2014 dengan *Flexibly Shaped Spatial Scan Statistic*

## 1.3 Tujuan Penelitian

Tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut.

1. Membandingkan analisis GPR dan regresi binomial negatif dalam mengatasi overdispersi berdasarkan kriteria nilai AIC pada kasus DBD di kabupaten/kota di Jawa Timur tahun 2014.
2. Mengetahui faktor yang berpengaruh terhadap kasus DBD di Provinsi Jawa Timur tahun 2014.
3. Memetakan *Flexibly Shaped Spatial Scan Statistic* untuk mendeteksi kantong DBD tingkat kabupaten/kota di Jawa Timur.

#### **1.4 Manfaat Penelitian**

Adapun manfaat yang ingin dicapai dari hasil penelitian ini adalah:

1. Memberikan pengetahuan tentang penyakit DBD serta memberikan pemahaman tentang mengatasi kasus overdispersi pada regresi poisson dengan menggunakan analisis GPR dan regresi binomial negatif. Serta pemetaan berdasarkan kantong-kantong DBD yang terbentuk di kabupaten/kota di Jawa Timur dengan metode *Flexibly Shaped Spatial Scan Statistic*.
2. Membantu Dinas Kesehatan dan masyarakat dalam pencegahan meluasnya penyakit DBD di kabupaten/kota di Jawa Timur dari daerah yang paling terjangkit DBD dan rawan terkena DBD.

#### **1.5 Batasan Masalah**

Penelitian ini menggunakan data jumlah kasus DBD di Provinsi Jawa Timur tahun 2014 dengan unit penelitian adalah tiap kabupaten/kota untuk mendapatkan wilayah yang paling banyak terkena DBD dan mendapatkan model dengan menggunakan analisis *GPR* dan regresi binomial negatif dan pembentukan kantong DBD dengan metode *Flexibly Shaped Spatial Scan Statistic*.

## **BAB II**

### **TINJAUAN PUSTAKA**

Pada bab ini akan dibahas mengenai konsep dan teori dari metode yang digunakan dalam melakukan analisis untuk menjawab perumusan masalah dari kasus DBD di Provinsi Jawa Timur. Metode yang digunakan diantaranya adalah statistika deskriptif, analisis regresi poisson, *Generalized Poisson Regression* (GPR) dan regresi binomial negatif, serta untuk melihat kantong DBD yang terbentuk menggunakan *Flexibly Shaped Spatial Scan Statistic*. Pembahasan mengenai konsep dan teori yang digunakan dalam analisis disajikan sebagai berikut.

#### **2.1 Statistika Deskriptif**

Statistika deskriptif merupakan bagian statistika yang membahas tentang metode-metode dengan pengumpulan dan penyajian data sehingga menarik dan memberikan informasi. Data dapat didefinisikan menjadi dua, yaitu dideskripsikan menjadi grafik atau tabel dan dideskripsikan secara numerik. Penyajian data dalam bentuk grafik atau tabel memberikan informasi dengan tampilan yang menarik, seperti diagram batang (*bar chart*), diagram lingkaran (*pie chart*), histogram, *scatter plot* dan lainnya. Penyajian data dalam bentuk deskriptif secara numerik memberikan informasi berupa ukuran pemusatan, penyebaran data, serta kecenderungan suatu gugus data seperti rata-rata (*mean*), standar deviasi, nilai maksimum, nilai minimum, serta modus. *Mean* (rata-rata) didefinisikan sebagai jumlah data dibagi dengan banyaknya data (Walpole, 1995). Varians adalah suatu angka yang menunjukkan ukuran variabilitas yang dihitung dengan jalan mengkuadratkan standar deviasi.

#### **2.2 Multikolinieritas**

Multikolinieritas adalah kondisi terdapatnya hubungan linier atau korelasi yang tinggi antara masing-masing variabel independen dalam model regresi. Menurut Draper dan Smith (1992) multikolinieritas biasanya terjadi ketika sebagian besar

variabel yang digunakan saling terkait dalam suatu model regresi. Oleh karena itu masalah multikolinearitas tidak terjadi pada regresi linier sederhana yang hanya melibatkan satu variabel independen. Menurut Hocking (1996) pendeteksian adanya kasus multikolinieritas dapat dilihat melalui.

1. *Variance Inflation Factors* (VIF)  $> 10$

VIF dinyatakan sebagai berikut :

$$VIF = \frac{1}{1 - R_j^2} \quad (2.1)$$

$R_j^2$  merupakan nilai koefisien determinasi antara variabel  $x_j$  dengan variabel prediktor lainnya. Nilai VIF yang lebih besar dari 10 menunjukkan adanya kasus multikolinieritas antar variabel prediktor.

2. Nilai koefisien korelasi Pearson ( $r_{ij}$ ) antar variabel-variabel prediktor. Mengindikasikan adanya multikolinieritas dengan melihat nilai koefisien korelasi Pearson antar variabel prediktor cukup besar (lebih dari 0,95).

Multikolinieritas dapat diatasi dengan mengeluarkan salah satu variabel independen yang berkorelasi tinggi dengan variabel independen yang lain. Pengeluaran variabel ini dapat dilakukan secara manual ataupun otomatis melalui metode *stepwise*.

### 2.3 Regresi Poisson

Regresi poisson merupakan salah satu regresi yang menggambarkan hubungan antara variabel respon (Y) dan variabel prediktor (X) dengan mengasumsikan variabel Y berdistribusi poisson. Menurut Walpole (1995) distribusi poisson menyatakan banyak sukses yang terjadi dalam suatu selang waktu atau daerah tertentu. Menurut Myers (1990) fungsi peluang dari distribusi poisson dapat dinyatakan sebagai berikut.



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

Dengan  $\sim$  merupakan rata-rata variabel random Y yang berdistribusi poisson dimana nilai rata-rata dan varians mempunyai nilai lebih dari nol. Distribusi poisson adalah suatu distribusi yang paling sederhana dalam pemodelan data yang berupa jumlah (*count*). Menurut Agresti (2002) yang menunjukkan bahwa *Generalized Linear Model* (GLM) merupakan pengembangan dari model linear klasik khususnya dalam mengatasi kendala variabel respon yang tidak normal. GLM terdiri dari 3 komponen yaitu komponen random, komponen sistematis, dan *link function*. Komponen random merupakan variabel dependen yaitu variabel  $Y$ . Variabel  $y$  merupakan vektor yang terdiri dari  $[y_1, y_2, \dots, y_n]^T$ . Komponen random dari GLM termasuk ke dalam keluarga eksponensial yang fungsi peluang dapat dinyatakan sebagai berikut.

$$f(y_i, \mu_i) = a(\mu_i) b(y_i) \exp(y_i Q(\mu_i)) \quad (2.3)$$

Komponen sistematis yaitu vektor yang terdiri dari  $[y_1, y_2, \dots, y_n]^T$  menghubungkan dengan variabel X atau sebagai variabel prediktor, sehingga untuk bentuk umum dari  $\eta = X\beta$  dimana  $X$  merupakan suatu matriks dengan elemen yang terdiri

$$\text{variabel prediktor dan dinyatakan dengan } X = \begin{bmatrix} 1 & x_{1,1} & \dots & x_{1,k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n,1} & \dots & x_{n,k} \end{bmatrix},$$

sedangkan  $\beta$  merupakan bentuk vektor dari parameter-parameter model. Masing-masing dari elemen  $\beta$  dapat dinyatakan dengan

$$y_i = \sum_{j=0}^k \beta_j x_{ij} \text{ merupakan kombinasi linier dari variabel prediktor}$$

disebut dengan linear prediktor.

Komponen *link function* yang menghubungkan antara komponen random dan komponen sistematis. Berdasarkan

Agresti (2002) dimisalkan  $\tilde{y}_i = E(Y_i)$  dimana  $i=1,2,\dots,n$ . Model untuk menghubungkan  $\tilde{y}_i$  dengan  $Y_i$  oleh  $g(\tilde{y}_i)=Y_i$ , dimana  $g(\cdot)$  adalah fungsi yang diturunkan. Sehingga  $g(\cdot)$  menghubungkan  $E(Y_i)$  dengan variabel prediktor, dan diformulasikan sebagai berikut.

$$g(\tilde{y}_i) = \sum_{j=0}^k S_j x_{ij} = S_{0i} + S_1 x_{i2} + \dots + S_k x_{ik} \quad (2.4)$$

*Link function* yang digunakan dalam regresi poisson adalah  $\ln$ , sehingga  $\ln(\tilde{y}_i) = Y_i$ . Dengan demikian regresi poisson dapat dinyatakan sebagai berikut.

$$\ln(\tilde{y}_i) = S_0 + \sum_{j=1}^k S_j x_{ij}, i = 1, 2, \dots, n$$

$$\text{dimana } \tilde{y}_i = \tilde{y}_i(\mathbf{x}_i) = \exp\left(S_0 + \sum_{j=1}^k S_j x_{ij}\right) \quad (2.5)$$

### 2.3.1 Penaksiran Parameter Model Regresi Poisson

Penaksiran parameter model regresi poisson dilakukan dengan menggunakan metode *Maximum Likelihood Estimation* (MLE) yaitu dengan cara memaksimumkan fungsi *likelihood* (Myers, 1990). Fungsi *likelihood* dari regresi poisson yaitu.

$$\begin{aligned} \ln L(\cdot) &= \ln\left(\prod_{i=1}^n \frac{\exp(-\tilde{y}_i)^{\tilde{y}_i^{y_i}}}{y_i!}\right) \\ &= \sum_{i=1}^n \ln\left(\frac{\exp(-\tilde{y}_i)^{\tilde{y}_i^{y_i}}}{y_i!}\right) \\ &= -\sum_{i=1}^n e^{\mathbf{x}_i^T} + \sum_{i=1}^n y_i \mathbf{x}_i^T - \sum_{i=1}^n \ln(y_i!) \end{aligned} \quad (2.6)$$

dimana

$$\mathbf{x}_i = [1 \quad x_{i1} \quad x_{i2} \quad \dots \quad x_{ik}]^T$$

$$=[s_0 \ s_1 \ s_2 \ \dots \ s_k]^T$$

Taksiran maksimum *Likelihood* untuk parameter  $\beta_j$  dinyatakan dengan  $\hat{\beta}_j$  yang merupakan penyelesaian dari turunan pertama fungsi logaritma natural dari *likelihood*. Selanjutnya persamaan (2.6) diturunkan terhadap  $\beta_j^T$  menjadi turunan kedua sebagai berikut.

$$\frac{\partial \ln L(\beta_j^T)}{\partial \beta_j^T} = -\sum_{i=1}^n \mathbf{x}_i \exp(\mathbf{x}_i^T \beta_j) + \sum_{i=1}^n y_i \mathbf{x}_i \quad (2.7)$$

Selanjutnya persamaan (2.7) disamakan dengan nol untuk mendapatkan penyelesaiannya. Tetapi penyelesaian dengan cara tersebut sering kali tidak mendapatkan hasil yang eksplisit sehingga alternatif yang bisa digunakan untuk mendapatkan penyelesaian dari MLE adalah dengan metode iterasi numerik Newton-Raphson. Dalam metode iterasi Newton diperlukan turunan kedua dari fungsi *likelihood* untuk pembentukan matriks hessian, turunan kedua tersebut adalah

$$\frac{\partial^2 \ln L(\beta_j^T)}{\partial \beta_j^T \partial \beta_j^T} = -\sum_{i=1}^n \mathbf{x}_i \mathbf{x}_i^T e^{-\mathbf{x}_i^T \beta_j}$$

Menurut Cameron dan Trivedi (1998) algoritma untuk optimalisasi metode Newton-Raphson adalah sebagai berikut.

1. Menentukan nilai taksiran awal parameter  $\hat{\beta}_{(0)}$  yang diperoleh dengan metode *Ordinary Least Square* (OLS) yaitu

$$\hat{\beta}_{(0)} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y} \quad (2.8)$$

dimana

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & \dots & x_{1k} \\ 1 & x_{21} & \dots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \dots & x_{nk} \end{bmatrix} \text{ dan } \mathbf{y} = [Y_1 \ Y_2 \ \dots \ Y_n]^T$$

2. Membentuk vektor gradien  $\mathbf{g}$ ,

$$g(\mathbf{s}) = -\sum_{i=1}^n \mathbf{x}_i \cdot \tilde{y}_i + \sum_{i=1}^n y_i \mathbf{x}_i = \sum_{i=1}^n \mathbf{x}_i (y_i - \tilde{y}_i)$$

$$\mathbf{g}^T(\mathbf{m})_{(k+1) \times 1} = \left[ \frac{\partial \ln L(\mathbf{s})}{\partial s_0} \quad \frac{\partial \ln L(\mathbf{s})}{\partial s_1} \quad \frac{\partial \ln L(\mathbf{s})}{\partial s_2} \quad \dots \quad \frac{\partial \ln L(\mathbf{s})}{\partial s_k} \right]_{\mathbf{s}=\mathbf{s}(m)}$$

dimana  $k$  adalah banyaknya parameter yang ditaksir.

3. Membentuk matriks Hessian  $\mathbf{H}$ .

Matriks Hessian ini disebut juga matriks informasi.

$$H(\mathbf{s}) = -\sum_{i=1}^n \mathbf{x}_i \cdot \mathbf{x}_i^T \cdot \tilde{y}_i = -\sum_{i=1}^n \tilde{y}_i \cdot \mathbf{x}_i \cdot \mathbf{x}_i^T$$

$$\mathbf{H}(\mathbf{m})_{(k+1) \times (k+1)} = \begin{bmatrix} \frac{\partial^2 \ln L(\mathbf{s})}{\partial s_0^2} & \frac{\partial^2 \ln L(\mathbf{s})}{\partial s_0 \partial s_1} & \dots & \frac{\partial^2 \ln L(\mathbf{s})}{\partial s_0 \partial s_k} \\ & \frac{\partial^2 \ln L(\mathbf{s})}{\partial s_1^2} & \dots & \frac{\partial^2 \ln L(\mathbf{s})}{\partial s_1 \partial s_k} \\ & & \ddots & \vdots \\ & & & \frac{\partial^2 \ln L(\mathbf{s})}{\partial s_k^2} \end{bmatrix}_{\mathbf{s}=\mathbf{s}(m)}$$

$$\text{Var}(\hat{\cdot}) = -E[\mathbf{H}^{-1}(\hat{\cdot})]$$

4. Memasukkan nilai  $\hat{\cdot}_{(0)}$  ke dalam elemen-elemen vektor  $\mathbf{g}$  dan matriks  $\mathbf{H}$ , sehingga diperoleh vektor  $\mathbf{g}(\hat{\cdot}_{(0)})$  dan matriks  $\mathbf{H}(\hat{\cdot}_{(0)})$ .

5. Melakukan iterasi pada persamaan

$$\hat{\cdot}_{(m+1)} = \hat{\cdot}_{(m)} - \mathbf{H}^{-1}(\hat{\cdot}_{(m)}) \mathbf{g}(\hat{\cdot}_{(m)})$$

Dimulai dari  $m=0$ , dengan nilai  $\hat{\theta}_{(m)}$  merupakan sekumpulan penaksir parameter yang konvergen pada iterasi ke- $m$ .

6. Penaksir parameter dikatakan konvergen ketika nilai  $\left\| \hat{\theta}_{(m+1)} - \hat{\theta}_{(m)} \right\| \leq \epsilon$  dimana  $\epsilon$  merupakan suatu bilangan yang sangat kecil sekali. Sehingga hampir tidak ada perbedaan antara  $\hat{\theta}_{(m+1)}$  dan  $\hat{\theta}_{(m)}$ . Jika belum didapatkan penaksir parameter yang konvergen, maka dilanjutkan kembali langkah 5 hingga iterasi ke  $m^* = m+1$ .

### 2.3.2 Pengujian Parameter Model Regresi Poisson

Menguji kelayakan model regresi Poisson, terlebih dahulu ditentukan dua buah fungsi likelihood yang berhubungan dengan model regresi yang diperoleh. Fungsi-fungsi likelihood yang dimaksud adalah  $L(\hat{\Omega})$  yaitu nilai likelihood untuk model lengkap dengan melibatkan variabel prediktor dan  $L(\hat{S})$  yaitu nilai likelihood untuk model sederhana tanpa melibatkan variabel prediktor. Salah satu metode yang digunakan untuk menentukan statistik uji dalam pengujian parameter model regresi poisson adalah dengan menggunakan metode *Maximum Likelihood Ratio Test* (MLRT) dengan hipotesis.

$$H_0: \theta_1 = \theta_2 = \dots = \theta_k = 0$$

$$H_1: \text{paling sedikit ada satu } \theta_j \neq 0; j=1,2,\dots,k$$

Statistik uji untuk kelayakan model regresi poisson adalah sebagai berikut.

$$\begin{aligned} D(\hat{\theta}) &= -2 \ln \left[ \frac{L(\hat{S})}{L(\hat{\Omega})} \right] = 2 \left[ \ln(L(\hat{\Omega})) - \ln(L(\hat{S})) \right] \\ &= 2 \sum_{i=1}^n [y_i \ln(y_i) - y_i - y_i \ln(\hat{\theta}_i) + \hat{\theta}_i] \end{aligned} \quad (2.9)$$

$$\text{dimana } \ln L(\hat{\Omega}) = \sum_{i=1}^n \left( -e^{x_i^T \theta} + y_i x_i^T \theta - \ln(y_i!) \right)$$

$$\ln L(\hat{S}) = \sum_{i=1}^n \left( -e^{S_0} + y_i S_0 - \ln(y_i!) \right)$$

Kriteria penolakan: tolak  $H_0$  jika  $D(\hat{\gamma}) > t_{(r;k)}^2$  dengan  $k$  adalah banyaknya parameter model dibawah populasi dikurangi dengan banyaknya parameter dibawah  $H_0$ . Parameter model regresi poisson yang telah dihasilkan dari estimasi parameter belum tentu mempunyai pengaruh yang signifikan terhadap model. Untuk itu perlu dilakukan pengujian terhadap parameter model regresi poisson secara individu. Dengan menggunakan hipotesis sebagai berikut.

$H_0 : \beta_j = 0$  (pengaruh variabel ke-j tidak signifikan)

$H_1 : \beta_j \neq 0$  (pengaruh variabel ke-j signifikan)

Statistik uji yang digunakan adalah:

$$z = \frac{\hat{S}_j}{se(\hat{S}_j)} \quad (2.10)$$

$se(\hat{S}_j)$  adalah nilai *standar error* atau tingkat kesalahan dari parameter  $\beta_j$ , didapat dari akar elemen diagonal ke-(j+1) dari  $\hat{var}(\hat{\gamma})$  dengan  $\hat{var}(\hat{\gamma}) = -E(\mathbf{H}^{-1}(\hat{\gamma}))$ . Kriteria penolakan tolak  $H_0$  jika  $|z_{hitung}| > z_{r/2}$  dimana  $\alpha$  adalah tingkat signifikansi.

### 2.3.3 Deteksi Overdispersi Pada Regresi Poisson

Metode regresi poisson mensyaratkan kondisi dimana nilai varians dan *mean* variabel respon bernilai sama atau kondisi *equidispersion* (Khoshgootaar *et al*, 2004). Sedangkan, regresi poisson dikatakan overdispensi apabila nilai variansnya lebih besar dari nilai meannya. Overdispensi memiliki dampak yang sama dengan pelanggaran asumsi. Dampak ini terlihat pada nilai *standar error* yang menjadi *under estimate*, sehingga kesimpulannya tidak valid. Menurut McCullagh dan Nelder (1983), kondisi overdispensi dapat ditulis  $Var(Y) > E(Y)$ . Misal merupakan parameter dispersi maka jika  $\phi = 0$  artinya tidak terjadi kasus overdispersi, jika  $\phi > 0$  artinya terjadinya kasus overdispersi,

dan jika  $<0$  artinya terjadi kasus underdispersi. Kondisi overdispersi dapat terjadi karena adanya sumber keragaman yang tidak teramat (Cameron dan Trivedi, 1998). Overdispersi dapat terjadi juga karena adanya pengamatan yang *missing value* pada variabel prediktor, adanya pencilaan pada data, variabel prediktor perlu ditransformasi atau kesalahan spesifik *link function* dan perlunya iteraksi dalam model (Hardin dan Hilbe, 2007). *Overdispersion* merupakan nilai dispersi *pearson chi-square* dan *deviance* yang dibagi dengan derajat bebasnya, diperoleh nilai lebih besar dari 1.

#### 2.4 Generalized Poisson Regression (GPR)

Model *GPR* merupakan suatu model yang sesuai digunakan untuk data *count* dimana terjadi *over/under dispersion*. Sehingga selain parameter  $\mu$  juga terdapat  $\nu$  sebagai parameter dispersi.

*GPR* hampir mirip dengan regresi poisson yaitu merupakan suatu model GLM. Akan tetapi pada model *GPR* mengasumsikan bahwa komponen randomnya berdistribusi *Generalized Poisson (GP)*. Misal,  $y=0,1,2,\dots$  merupakan variabel respon. Distribusi *GPR* diberikan Famoye, Wulu, dan Singh (2004) sebagai berikut.

$$f(y; \nu; \mu) = \left( \frac{\nu}{1 + \nu \mu} \right)^y \frac{(1 + \nu \mu)^{y-1}}{y!} \exp\left( \frac{-\nu(1 + \nu \mu)}{1 + \nu \mu} \right), y = 0, 1, 2, \dots \quad (2.11)$$

Mean dan varians model *GPR* adalah sebagai berikut.

$$E(Y) = \mu \quad \text{dan} \quad \text{var}(Y) = \mu(1 + \nu \mu)^2$$

Apabila  $\nu = 0$  maka model *GPR* akan menjadi regresi poisson biasa, apabila  $\nu > 0$  maka model *GPR* merepresentasikan data *count* yang *overdispersion*, dan jika  $\nu < 0$  maka *GPR* merepresentasikan data *count* yang *underdispersion*. Model *GPR* memiliki bentuk yang sama dengan model regresi poisson yaitu sebagai berikut.

$$\ln(\mu_i) = \mathbf{x}_i^T = \beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik}$$

$$\mu_i = \exp(\mathbf{x}_i^T), \quad i=1,2, \dots, n \quad (2.12)$$

dimana

$$\begin{aligned} \mathbf{x}_i &= [1 \quad x_{i1} \quad x_{i2} \quad \dots \quad x_{ik}]^T \\ &= [S_0 \quad S_1 \quad S_2 \quad \dots \quad S_k]^T \end{aligned}$$

#### 2.4.1 Penaksir Parameter Model *Generalized Poisson Regression (GPR)*

Penaksir parameter model *GPR* pada persamaan (2.12) dilakukan dengan metode MLE (*Maximum Likelihood Estimator*). Fungsi likelihood untuk model *GPR* adalah sebagai berikut.

$$\begin{aligned} L(\boldsymbol{\theta}, \mathbf{y}) &= \prod_{i=1}^n f(\boldsymbol{\theta}, y_i) \\ L(\boldsymbol{\theta}, \mathbf{y}) &= \prod_{i=1}^n \left\{ \left( \frac{\tilde{\mu}_i}{1 + \tilde{\mu}_i} \right)^{y_i} \frac{(1 + \tilde{\mu}_i)^{(y_i-1)}}{y_i!} \exp\left( \frac{-\tilde{\mu}_i(1 + \tilde{\mu}_i)}{1 + \tilde{\mu}_i} \right) \right\} \\ L(\boldsymbol{\theta}, \mathbf{y}) &= \prod_{i=1}^n \left( \frac{\tilde{\mu}_i}{1 + \tilde{\mu}_i} \right)^{y_i} \left( \prod_{i=1}^n \frac{(1 + \tilde{\mu}_i)^{(y_i-1)}}{y_i!} \right) \exp\left[ \sum_{i=1}^n \frac{-\tilde{\mu}_i(1 + \tilde{\mu}_i)}{1 + \tilde{\mu}_i} \right] \end{aligned} \quad (2.13)$$

Selanjutnya persamaan (2.13) diubah dalam bentuk fungsi logaritma natural menjadi.

$$\begin{aligned} \ln L(\boldsymbol{\theta}, \mathbf{y}) &= \ln \left\{ \prod_{i=1}^n \left( \frac{\tilde{\mu}_i}{1 + \tilde{\mu}_i} \right)^{y_i} \left( \prod_{i=1}^n \frac{(1 + \tilde{\mu}_i)^{(y_i-1)}}{y_i!} \right) \exp\left[ \sum_{i=1}^n \frac{-\tilde{\mu}_i(1 + \tilde{\mu}_i)}{1 + \tilde{\mu}_i} \right] \right\} \\ \ln L(\boldsymbol{\theta}, \mathbf{y}) &= \sum_{i=1}^n \left\{ y_i \ln(\tilde{\mu}_i) - y_i \ln(1 + \tilde{\mu}_i) + (y_i - 1) \ln(1 + \tilde{\mu}_i) + \Delta \right\} \end{aligned}$$

$$\text{dimana } \Delta = -\ln(y_i!) - \frac{-\tilde{\mu}_i(1 + \tilde{\mu}_i)}{1 + \tilde{\mu}_i}$$

dengan mensubstitusikan nilai  $\tilde{\mu}_i = \exp(\mathbf{x}_i^T)$  maka diperoleh,

$$\ln L(\boldsymbol{\theta}, \mathbf{y}) = \sum_{i=1}^n \left\{ y_i \ln(\exp(\mathbf{x}_i^T)) - y_i \ln(1 + \exp(\mathbf{x}_i^T)) - \ln(y_i!) + \Delta \right\}$$



$$\text{dimana } \Delta = (y_i - 1) \ln(1 + {}_n y_i) - \frac{-\exp(\mathbf{x}_i^T) (1 + {}_n y_i)}{1 + {}_n \exp(\mathbf{x}_i^T)}$$

$$\ln L(\beta, \alpha) = \sum_{i=1}^n \left\{ y_i \mathbf{x}_i^T - y_i \ln(1 + {}_n \exp(\mathbf{x}_i^T)) + (y_i - 1) \ln(1 + {}_n y_i) + \Delta^* \right\}$$

$$\text{dimana } \Delta^* = -\ln(y_i!) - (-\exp(\mathbf{x}_i^T) (1 + {}_n y_i) (1 + {}_n \exp(\mathbf{x}_i^T))^{-1})$$

Kemudian persamaan logaritma natural dari fungsi likelihood diturunkan terhadap  $\beta$  dan disamakan dengan nol untuk mendapatkan parameter  $\hat{\beta}$ . Bentuk turunan persamaan logaritma natural dari likelihood terhadap  $\beta$  yaitu.

$$\frac{\partial \ln L(\beta, \alpha)}{\partial \beta} = \sum_{i=1}^n y_i \mathbf{x}_i - {}_n y_i \mathbf{x}_i \exp(\mathbf{x}_i^T) (1 + {}_n \exp(\mathbf{x}_i^T))^{-1} + \Delta$$

$$\text{dimana } \Delta = -(1 + {}_n y_i) \left\{ \mathbf{x}_i \exp(\mathbf{x}_i^T) (1 + {}_n (\mathbf{x}_i^T))^{-1} + \Delta^* \right\}$$

$$\text{dan } \Delta^* = -{}_n \mathbf{x}_i (\exp(\mathbf{x}_i^T))^2 (1 + {}_n (\mathbf{x}_i^T))^{-2}$$

Jika ingin mendapatkan penaksir parameter  $\hat{\beta}$  maka persamaan tersebut diturunkan terhadap  $\beta$  dan disamakan dengan nol. Hasil turunan terhadap  $\beta$  dinyatakan sebagai berikut.

$$\frac{\partial \ln L(\beta, \alpha)}{\partial \beta} = \sum_{i=1}^n \left[ y_i \exp(\mathbf{x}_i^T) (1 + {}_n \exp(\mathbf{x}_i^T))^{-1} + y_i (y_i - 1) (1 + {}_n y_i)^{-1} + \Delta \right]$$

$$\text{dimana } \Delta = -\exp(\mathbf{x}_i^T) \left\{ y_i (1 + {}_n \exp(\mathbf{x}_i^T))^{-1} + \Delta^* \right\}$$

$$\text{dan } \Delta^* = -(1 + {}_n y_i) \exp(\mathbf{x}_i^T) (1 + {}_n \exp(\mathbf{x}_i^T))^{-2}$$

Penurunan fungsi ln-likelihood terhadap  $\beta$  dan  $\alpha$  seringkali menghasilkan persamaan yang implisit, sehingga digunakan iterasi Newton-Raphson sampai didapatkan penaksir parameter yang konvergen seperti pada subbab 2.3.1.

### 2.4.2 Pengujian Parameter Model GPR

Pengujian parameter *GPR* dilakukan dengan metode *Maximum Likelihood Ratio Test (MLRT)*. Hipotesis yang digunakan adalah sebagai berikut.

$$H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$$

$$H_1: \text{paling sedikit ada satu } \beta_j \neq 0; j=1,2,\dots,k$$

Statistik uji:

$$D(\hat{\beta}) = -2 \ln \Lambda = -2 \ln \left( \frac{L(\check{\beta})}{L(\hat{\beta})} \right) = 2 \left( \ln L(\hat{\beta}) - \ln L(\check{\beta}) \right) \quad (2.14)$$

Dimana

$$\ln L(\hat{\beta}) = \sum_{i=1}^n \left\{ y_i \mathbf{x}_i^T - y_i \ln(1 + \exp(\mathbf{x}_i^T)) + (y_i - 1) \ln(1 + y_i) + \Delta^* \right\}$$

$$\Delta^* = -\ln(y_i!) - (-\exp(\mathbf{x}_i^T))(1 + y_i)(1 + \exp(\mathbf{x}_i^T))^{-1}$$

$$\ln L(\check{\beta}) = \sum_{i=1}^n \left\{ y_i s_0 - y_i \ln(1 + \exp(s_0)) + (y_i - 1) \ln(1 + y_i) + \Delta^{**} \right\}$$

$$\Delta^{**} = -\ln(y_i!) - (-\exp(s_0))(1 + y_i)(1 + \exp(s_0))^{-1}$$

Kriteria penolakan: Tolak  $H_0$  jika nilai  $D(\hat{\beta}) > t_{(k;r)}^2$ . Jika

$H_0$  ditolak berarti paling sedikit ada satu  $\beta_j \neq 0$  yang menunjukkan bahwa  $X_j$  berpengaruh secara signifikan terhadap model. Pengujian dilanjutkan dengan uji secara parsial dengan hipotesis sebagai berikut.

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

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

Statistik uji yang digunakan adalah:

$$z = \frac{\hat{S}_j}{se(\hat{S}_j)} \quad (2.15)$$

$se(\hat{S}_j)$  adalah nilai *standar error* atau tingkat kesalahan dari parameter  $\beta_j$ , didapat dari elemen diagonal ke-(j+1) dari

$\hat{v}ar(\hat{\theta})$  dengan  $\hat{v}ar(\hat{\theta}) = -E(\mathbf{H}^{-1}(\hat{\theta}))$ . Kriteria penolakan tolak  $H_0$  jika  $|z_{hit}| > z_{\alpha/2}$  dimana adalah tingkat signifikansi.

## 2.5 Regresi Binomial Negatif

Penanganan overdispersi pada regresi poisson selain menggunakan *Generalized Poisson Regression* dapat dilakukan dengan menggunakan pendekatan model regresi Binomial Negatif. Masalah overdispersi dapat menyebabkan kesalahan inferensi mengenai estimasi parameter.

Menurut Cameron dan Trivedi (1998), distribusi Binomial Negatif yang mengakomodasi overdispersi memiliki fungsi densitas sebagai berikut.

$$f(y_i; \sim_i, n) = \frac{\Gamma(y_i + n^{-1})}{\Gamma(n^{-1})\Gamma(y_i + 1)} \left( \frac{n^{-1}}{n^{-1} + \sim_i} \right)^{n^{-1}} \left( \frac{\sim_i}{n^{-1} + \sim_i} \right)^{y_i} \quad (2.16)$$

$n \geq 0, y_i = 0, 1, 2, \dots$

dengan  $\Gamma(\cdot)$  adalah fungsi Gamma. Distribusi Binomial Negatif memiliki mean dan varians sebagai berikut.

$$E(y_i | x_i) = \sim_i = \exp(\mathbf{x}_i^T \boldsymbol{\beta}) \text{ dan } \text{var}(y_i | x_i) = \sim_i + n \sim_i^2,$$

jika  $n$  menuju nol maka  $\text{var}(Y_i)$  menuju  $\mu_i$  sehingga Binomial Negatif akan konvergen menuju Poisson. Model regresi Binomial Negatif memiliki bentuk yang sama dengan model regresi Poisson yaitu pada persamaan (2.5) sebagai berikut.

$$\ln(\sim_i) = S_0 + S_1 x_{i1} + S_2 x_{i2} + \dots + S_k x_{ik}$$

dengan,

$$E(y_i | x_i) = \sim_i = \exp(\mathbf{x}_i^T \boldsymbol{\beta})$$

$$\sim_i = \exp(S_0 + S_1 x_{i1} + S_2 x_{i2} + \dots + S_k x_{ik})$$

### 2.5.1 Penaksiran Parameter Model Regresi Binomial Negatif

Penaksiran parameter regresi Binomial Negatif dilakukan dengan metode *Maximum Likelihood Estimation* (MLE) dengan

prosedur Newton Rhapron. Metode ini membutuhkan turunan pertama dan kedua dari fungsi likelihood.  $Y_i$  mempunyai fungsi massa probabilitas distribusi binomial negatif seperti berikut.

$$f(y_i | \sim_i, n) = \frac{\Gamma(y_i + \frac{1}{n})}{\Gamma(\frac{1}{n})\Gamma(y_i + 1)} \left( \frac{1}{1 + n \sim_i} \right)^{\frac{1}{n}} \left( \frac{n \sim_i}{1 + n \sim_i} \right)^{y_i} \quad (2.17)$$

dengan  $y_i = 0, 1, 2, \dots$

Karena fungsinya saling bebas, akan fungsi likelihood adalah:

$$L(, n) = \prod_{i=1}^n \frac{\Gamma(y_i + \frac{1}{n})}{\Gamma(\frac{1}{n})\Gamma(y_i + 1)} \left( \frac{1}{1 + n \sim_i} \right)^{\frac{1}{n}} \left( \frac{n \sim_i}{1 + n \sim_i} \right)^{y_i}$$

dengan  $\frac{\Gamma(y_i + \frac{1}{n})}{\Gamma(\frac{1}{n})} = \prod_{r=1}^{y_i-1} (r + n^{-1})$

$$L(, n) = \prod_{i=1}^n \left( \prod_{r=1}^{y_i-1} (r + n^{-1}) \right) \frac{1}{(y_i!)} + \left( \frac{1}{1 + n \sim_i} \right)^{\frac{1}{n}} \left( \frac{n \sim_i}{1 + n \sim_i} \right)^{y_i}$$

$$L(, n) = \ln \{ l(, n) \}$$

$$= \sum_{i=1}^n \left[ \left( \sum_{r=1}^{y_i-1} \ln(r + n^{-1}) \right) - \ln(y_i!) + y_i \ln(n \sim_i) - (n^{-1} + y_i) \ln(1 + n \sim_i) \right] \quad (2.18)$$

Turunan pertama dari fungsi likelihood terhadap koefisien regresi adalah

$$\frac{\partial L(, n)}{\partial S_0} = \sum_{i=1}^n \left[ y_i - (y_i + n^{-1}) \left( \frac{n \sim_i}{1 + n \sim_i} \right) \right] = \sum_{i=1}^n \left[ \frac{y_i - n \sim_i}{1 + n \sim_i} \right] = 0$$

$$\begin{aligned} \frac{\partial L(, n)}{\partial S_p} &= \sum_{i=1}^n \left[ y_i x_{ip} - (y_i + n^{-1}) \left( \frac{n \sim_i x_{ip}}{1 + n \sim_i} \right) \right] \\ &= \sum_{i=1}^n \left[ \frac{(y_i - n \sim_i) x_{ip}}{1 + n \sim_i} \right] \end{aligned}$$

$$= \sum_{i=1}^n \left[ \frac{\tilde{\gamma}_i}{1 + \tilde{\gamma}_i} \frac{(y_i - \tilde{\gamma}_i)x_{ip}}{\tilde{\gamma}_i} \right] = 0$$

bentuk persamaan matriks dari turunan pertama fungsi likelihood terhadap parameter yaitu:  $\mathbf{q} = \mathbf{X}^T \mathbf{W} \mathbf{z}$ , dengan  $\mathbf{X}$  adalah matriks ( $n \times c$ ) dari variabel prediktor,  $\mathbf{W}$  adalah matriks weight diagonal ke- $i$  dan  $\mathbf{z}$  adalah vektor matriks dengan baris ke- $i$ , dnegan masing-masing elemeny adalah:

$$w_i = \frac{\tilde{\gamma}_i}{1 + \tilde{\gamma}_i} \text{ dan } z_i = \frac{(y_i - \tilde{\gamma}_i)}{\tilde{\gamma}_i}$$

turunan pertama dari fungsi log-likelihood terhadap parameter dispersi adalah

$$f'(\sigma) = \frac{\partial L(\sigma)}{\partial \sigma} = \sum_{i=1}^n \left[ -\frac{1}{\sigma^2} \sum_{r=0}^{y_i-1} \frac{1}{(r + \sigma^{-1})} + \frac{y_i}{\sigma} - \ln(1 + \sigma \tilde{\gamma}_i) - \frac{\left( \frac{-1 + y_i}{\sigma} \right) \tilde{\gamma}_i}{(1 + \sigma \tilde{\gamma}_i)} \right] = 0$$

$$f'(\sigma) = \frac{\partial L(\sigma)}{\partial \sigma} = \sum_{i=1}^n \left[ -\frac{1}{\sigma^2} \sum_{r=0}^{y_i-1} \frac{1}{(r + \sigma^{-1})} + \frac{y_i}{\sigma} - \ln(1 + \sigma \tilde{\gamma}_i) - \left( \frac{y_i - \tilde{\gamma}_i}{\sigma (1 + \sigma \tilde{\gamma}_i)} \right) \right] = 0$$

Turunan parsial kedua fungsi likelihood terhadap parameter koefisien regresi adalah

$$\frac{\partial^2 L(\sigma)}{\partial S_0^2} = - \sum_{i=1}^n \left[ \frac{(1 + \sigma y_i) \tilde{\gamma}_i}{(1 + \sigma \tilde{\gamma}_i)^2} \right]$$

$$\frac{\partial^2 L(\sigma)}{\partial S_0 \partial S_j} = \sum_{i=1}^n \left[ \frac{-\tilde{\gamma}_i x_{ij} (1 + \sigma \tilde{\gamma}_i) - (y_i - \tilde{\gamma}_i) (x_{ij} \tilde{\gamma}_i)}{(1 + \sigma \tilde{\gamma}_i)^2} \right]$$

$$\frac{\partial^2 L(\sigma)}{\partial S_0 \partial S_j} = \sum_{i=1}^n \left[ \frac{(1 + \sigma y_i) x_{ij} \tilde{\gamma}_i}{(1 + \sigma \tilde{\gamma}_i)^2} \right]$$

Misalkan turunan parsial pertama dari  $L(\sigma)$  terhadap  $S_j, j \leq p$  adalah

$$\frac{\partial L(\sigma)}{\partial S_p} = \sum_{i=1}^n \left[ \frac{(y_i - \tilde{\gamma}_i) x_{ij}}{1 + \sigma \tilde{\gamma}_i} \right] = 0$$

Maka turunan parsial kedua terhadap  $S_u, u \leq p$  adalah

$$\frac{\partial^2 L(\beta, \sigma^2)}{\partial \sigma_u \partial \sigma_j} = \sum_{i=1}^n \left[ \frac{-\tilde{y}_i x_{iu} x_{ij} (1 + \sigma^2 \tilde{y}_i) - (y_i - \tilde{y}_i)(x_{iu} x_{ij} \sigma^2 \tilde{y}_i)}{(1 + \sigma^2 \tilde{y}_i)^2} \right] \quad (2.19)$$

$$\frac{\partial^2 L(\beta, \sigma^2)}{\partial \sigma_u \partial \sigma_j} = - \sum_{i=1}^n \left[ \frac{x_{iu} x_{ij} \tilde{y}_i (1 + \sigma^2 y_i)}{(1 + \sigma^2 \tilde{y}_i)^2} \right]$$

Ekspektasi dari turunan kedua log-likelihood adalah

$$E \left( - \frac{\partial^2 L(\beta, \sigma^2)}{\partial \sigma_u \partial \sigma_j} \right) = \sum_{i=1}^n \left[ \frac{x_{iu} x_{ij} \tilde{y}_i}{(1 + \sigma^2 \tilde{y}_i)} \right] \quad (2.20)$$

Jika persamaan (2.20) dinyatakan dalam matriks  $\mathbf{I}$  (matriks informasi) yaitu matriks yang mengandung ekspektasi negatif dari turunan kedua log-likelihood maka.

$$\mathbf{I} = \mathbf{X}^T \mathbf{W} \mathbf{X}$$

Dengan  $\mathbf{X}$  adalah matriks dari variabel prediktor,  $\mathbf{W}$  adalah matriks weight diagonal ke- $i$  dengan elemen:  $w_i = \frac{\tilde{y}_i}{1 + \sigma^2 \tilde{y}_i}$

Turunan kedua fungsi likelihood terhadap parameter dispersi adalah

$$f''(\sigma^2) = \frac{\partial^2 L(\beta, \sigma^2)}{\partial \sigma^2} = \sum_{i=1}^n \left[ -3 \sum_{r=0}^{y_i-1} \frac{(2r + \sigma^2 - 1)}{(r + \sigma^2 - 1)^2} - 2\sigma^2 \ln(1 + \sigma^2 \tilde{y}_i) + \frac{-\sigma^2 \tilde{y}_i}{(1 + \sigma^2 \tilde{y}_i)} - \frac{(y_i - \tilde{y}_i)(1 + 2\sigma^2 \tilde{y}_i)}{(\sigma^2 + \sigma^2 \tilde{y}_i)^2} \right]$$

Langkah-langkah estimasi parameter regresi binomial negatif dilakukan dengan langkah sebagai berikut:

1. Menentukan estimasi awal  $\hat{\sigma}_1^2$ , misal  $\hat{\sigma}_1^2 = 0,1$
2. Menentukan estimasi maksimum *likelihood* dari parameter menggunakan prosedur iterasi Fisher scoring dengan asumsi

$$\hat{\sigma}_{i+1}^2 = \hat{\sigma}_i^2 + (\mathbf{X}^T \mathbf{W}_i \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}_i \mathbf{z}_i$$

Iterasi berakhir jika diperoleh  $\left\| \hat{\sigma}_{i+1}^2 - \hat{\sigma}_i^2 \right\| \leq \epsilon$ .

3. Menggunakan  $\hat{\mu}$  untuk menghasilkan estimasi dari parameter dengan menggunakan prosedur iterasi Newton-Raphson satu variabel, iterasi berakhir jika diperoleh
- $$\left| \hat{\mu}_{i+1} - \hat{\mu}_i \right| \leq \nu$$
- $$\hat{\mu}_{i+1} = \hat{\mu}_i - \frac{f'(\mu_i)}{f''(\mu_i)}$$
4. Jika  $\left| \hat{\mu}_{i+1} - \hat{\mu}_i \right| \leq \nu$  selesai; bila tidak, gunakan parameter  $\mu = \hat{\mu}_{i+1}$  dan kembali ke langkah 2, nilai  $\varepsilon$  merupakan bilangan positif yang sangat kecil.

### 2.5.2 Pengujian Parameter Model Regresi Binomial Negatif

Uji kesesuaian model regresi binomial negatif dengan uji deviansi sebagai berikut.

Hipotesis:

$$H_0: \mu_1 = \mu_2 = \dots = \mu_k = 0$$

$H_1$ : paling sedikit ada satu  $\mu_j > 0, j=1,2,\dots,k$

Statistik Uji:

$$\begin{aligned} D(\hat{s}) &= -2 \ln \Lambda = -2 \ln \left( \frac{L(\check{S})}{L(\hat{\Omega})} \right) = 2(\ln L(\hat{\Omega}) - \ln L(\check{S})) \\ &= 2 \sum_{i=1}^n \left[ y_i \ln \left( \frac{y_i}{\tilde{\mu}_i} \right) - \left( \frac{1}{\mu} + y_i \right) \ln \left( \frac{1 + \mu y_i}{1 + \mu \tilde{\mu}_i} \right) \right] \end{aligned} \quad (2.21)$$

dimana

$$\ln L(\hat{\Omega}) = \sum_{i=1}^n \left[ \left( \sum_{r=1}^{y_i-1} \ln(r + \mu^{-1}) \right) - \ln(y_i!) + y_i \ln(\mu \tilde{\mu}_i) + a \right]$$

dimana

$$a = -(\mu^{-1} + y_i) \ln(1 + \mu \tilde{\mu}_i)$$

$$\ln L(\hat{\mathcal{S}}) = \sum_{i=1}^n \left[ \left( \sum_{r=1}^{y_i-1} \ln(r + n^{-1}) \right) - \ln(y_i!) + y_i \ln(n \exp(\hat{S}_0)) + b \right]$$

dimana

$$b = -(n^{-1} + y_i) \ln(1 + n \exp(\hat{S}_0))$$

Kriteria penolakan: tolak  $H_0$  jika statistik uji  $D(\hat{S}) > t_{r;k}^2$

Uji signifikansi individu variabel prediktornya dengan menggunakan uji z dengan hipotesis sebagai berikut.

$H_0 : \beta_j = 0$  (pengaruh variabel ke-j tidak signifikan)

$H_1 : \beta_j \neq 0$  (pengaruh variabel ke-j signifikan)

Statistik uji yang digunakan adalah:

$$z = \frac{\hat{S}_j}{se(\hat{S}_j)} \quad (2.22)$$

Kriteria penolakan: tolak  $H_0$  jika nilai  $|z_{hitung}| > z_{r/2}$

dimana  $\alpha$  adalah tingkat signifikansi yang digunakan.

## 2.6 Pemilihan Model Terbaik

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* biasanya 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. Menurut Bozdogan (2000) rumus *AIC* didefinisikan sebagai berikut.

$$AIC = -2 \ln L(\hat{\theta}) + 2k \quad (2.23)$$



dimana  $L(\hat{\cdot})$  adalah nilai likelihood yang didapat dari persamaan (2.6) untuk regresi poisson persamaan (2.13) untuk *genelized poisson regression* dan persamaan (2.18) untuk regresi binomial negatif dan  $k$  adalah jumlah parameter.

## 2.7 Metode *Flexibly Shaped Spatial Scan Statistic*

*Spatial Scan Statistic* merupakan salah satu metode statistik yang digunakan untuk mendeteksi *cluster* pada sebuah lokasi yang berupa titik maupun data agregat. Menurut Tango dan Takahashi (2005) metode *Flexibly Shaped Spatial Scan Statistic* mempunyai *power* lebih tinggi daripada metode *Circular Spatial Scan Statistic* saat *cluster* yang dideteksi adalah *non-circular* dan fleksibel terhadap bentuk kantong yang dihasilkan sehingga tidak terbatas pada bentuk lingkaran saja. Untuk mendeteksi *hotspot* dengan metode *Flexibly Shaped Spatial Scan Statistic*, pada awalnya suatu daerah dibagi menjadi  $n$  daerah seperti kabupaten dan sebagainya. Jumlah kasus yang ada di wilayah  $i$  dilambang dengan  $Y_i$  dengan nilai yang diamati  $i = 1, 2, \dots, n$  diasumsikan saling bebas dan mengikuti sebaran poisson dengan fungsi peluang seperti persamaan (2.2). Untuk menentukan letak geografis masing-masing daerah digunakan titik koordinat pusat penduduk administrasi.

Pada awalnya, *Flexibly Scan Statistic* ditempatkan pada kantong yang tidak teratur (*irregularly shaped*) pada setiap wilayah. *Window ke-i* dilambangkan dengan  $W_i$  yang merupakan kumpulan wilayah  $i$  dan wilayah-wilayah yang berbatasan dengan wilayah  $i$  yang kemudian disusun himpunan  $Z$  yang bentuknya tidak teratur dengan panjang  $l$  pada setiap wilayah, terdiri atas  $l$  wilayah. Untuk menghindari pendeteksian kantong yang bentuknya aneh (*unlikely peculiar shape*), wilayah yang berbatasan dibatasi sebagai himpunan bagian dari wilayah  $i$  dan  $(L-1)$  wilayah sekitar yang terdekat dengan wilayah  $i$  dimana  $L$  adalah *pre-specified maximum length* dari *cluster*. Selanjutnya akan berbentuk  $Z$  yang berbeda-beda dan saling *overlapping* (tumpang tindih). Misalkan  $Z_{il(m)}$ ,  $m=1, \dots, m_{il}$  melambangkan  $Z$

ke- $m$  yang merupakan himpunan  $l$  wilayah yang berhubungan dimulai dari wilayah  $i$ , dimana  $m_{il}$  adalah jumlah  $m$  yang memenuhi  $Z_{il(m)} \subseteq Z_{il}$  untuk  $l = 1, 2, \dots, L$ . Kemudian semua yang diperiksa dimasukkan dalam himpunan berikut.

$$Z = \{Z_{il(m)} \mid 1 \leq i \leq n, 1 \leq l \leq L, 1 \leq m \leq m_{il}\} \quad (2.24)$$

Algoritma yang digunakan untuk mendapatkan  $Z$  dengan *pre-specified maximum length*  $L$  (Tango dan Takahashi, 2005) adalah sebagai berikut.

1. Pertama, membuat sebuah matriks  $A = (a_{ih})$  berukuran  $n \times n$  sedemikian hingga

$$a_{ih} = \begin{cases} 1 & , \text{jika wilayah } i \text{ dan } h \text{ bersebelahan atau contiguity} \\ 0 & , \text{untuk lainnya} \end{cases}$$

dan set  $Z_2 = \emptyset$  dan  $i_0 = 0$

2. Misalkan  $i_0 \leftarrow i_0 + 1$  dan  $i_0 (=1, 2, \dots, m)$  menjadi wilayah awal. Kemudian dibentuk  $W_{i_0}$  yang terdiri dari  $(L-1)$  *nearest neighbours* ke wilayah awal  $i_0$  dan  $i_0$  sendiri, sebagai contoh  $W_{i_0} = (i_0, i_1, \dots, i_{L-1})$  dimana  $i_l$  adalah  $l$  wilayah ke- $l$  terdekat terhadap  $i_0$ .

Misalkan wilayah seperti berikut.

1	2	
3	4	5

$W_i = \{1, 2, 3, 4\}$  dimana wilayah awal berada di satu.

3. Pertimbangkan semua himpunan  $Z \subset W_{i_0}$ , dimana termasuk wilayah awal  $i_0$ . Untuk himpunan  $Z$  lainnya, ulangi langkah 4-7. Dimana nilai  $Z$  diperoleh dari anggota himpunan  $W_i$ , sehingga  $Z_i = \{1, 2\} \{1, 3\} [1, 4] \{1, 2, 3\} \{1, 2, 4\}$  dan begitu untuk tiga dan empat anggota himpunan lainnya.
4. Bagi himpunan  $Z$  menjadi dua disjoint :  $Z_0 = \{i_0\}$  dan  $Z_1$  dimana berisi wilayah lain yang terdapat dalam  $Z$ .
5. Buat dua himpunan baru  $Z'_0$  dan  $Z'_1$ .  $Z'_0$  terdiri dari wilayah  $Z_0$  yang berbatasan dengan wilayah  $Z_0$ . Di sisi lain,  $Z'_1$  terdiri

dari wilayah  $Z_1$  yang tidak berbatasan dengan wilayah  $Z_0$ . Kemudian, ganti  $Z_0$  dan  $Z_1$  dengan  $Z'_0$  dan  $Z'_1$ .

Untuk langkah ini dimisalkan  $Z = \{1,2,3\}$  maka  $Z_0 = \{1\}$ ,  $Z_1 = \{2,3\}$ . Di bentuk himpunan baru  $Z'_0 = \{2,3\}$  dan  $Z'_1 = \cdot Z_1$  himpunan kosong karena tidak ada wilayah lain dari anggota dari  $Z$  yang tidak berdekatan dengan wilayah 1.

6. Ulangi langkah 5 secara rekursif sampai  $Z_0$  dan  $Z_1$  menjadi himpunan kosong.
7. Buat kesimpulan sebagai berikut.  
 $Z$  dikatakan berhubungan (*connected*) ketika  $Z_1$  menjadi himpunan kosong terlebih dahulu dan tidak berhubungan ketika  $Z_0$  menjadi himpunan kosong terlebih dahulu, maka  $Z$  dimasukkan dalam set  $Z$ . Jika  $Z$  tidak berhubungan maka  $Z$  dibuang.
8. Ulangi langkah 2-7 sampai pada akhirnya diperoleh himpunan  $Z$  yang terdiri dari  $Z$  berbentuk tertentu dengan *maximum length* adalah  $L$ .

### 2.7.1 Likelihood Ratio Test dan Pengujian Hipotesis Monte Carlo

Untuk setiap wilayah  $i$  dan panjang dari *scanning window*, hipotesis alternatifnya adalah minimal ada satu *window Z* yang mempunyai peluang resiko lebih tinggi (*elevated risk*) daripada di luar *window*. Dengan kata lain, hipotesis yang digunakan adalah sebagai berikut.

$H_0 : E(Y(\mathbf{Z})) = \mu(\mathbf{Z})$  untuk semua  $\mathbf{Z}$

$H_1 : E(Y(\mathbf{Z})) > \mu(\mathbf{Z})$  untuk beberapa  $\mathbf{Z}$

dimana  $Y(\cdot)$  melambangkan jumlah kasus yang random dan  $\mu(\cdot)$  merupakan nilai harapan dari kasus *window* tertentu.

Pada setiap *window*, dapat dihitung *likelihood* untuk mengetahui jumlah kasus di dalam dan di luar *window*. Persamaan *likelihood* untuk setiap *window* mengikuti dengan asumsi Poisson, uji statistik yang disusun dengan *likelihood ratio test* (Kulldorff, 1997) adalah sebagai berikut.

$$J = \sup_{z \in Z} \left( \frac{y(Z)}{-Z} \right)^{n(Z)} \left( \frac{y(Z^c)}{-Z^c} \right)^{n(Z^c)} I \left( \frac{y(Z)}{-Z} > \frac{y(Z^c)}{-Z^c} \right) \quad (2.25)$$

dimana  $Z^c$  melambangkan semua wilayah di luar *window*  $Z$ , dan  $y(\cdot)$  melambangkan jumlah kasus dalam *window* yang ditentukan dan  $I(\cdot)$  merupakan fungsi indikator. Ketika memeriksa kantong dengan *high rates*, maka  $I(\cdot)$  bernilai 1 jika  $y(Z^c)$  dibagi  $(Z^c)$  memiliki peluang lebih besar .

Uji statistik yang dilakukan menggunakan pengujian hipotesis Monte Carlo. *P-value* diperoleh dengan membandingkan *rank* dari *likelihood* yang maksimal dari data yang sebenarnya dengan *likelihood* yang maksimal dari himpunan data acak. Jika *rank* dilambangkan dengan  $R$  maka  $p\text{-value} = R/(1+\#\text{simulasi})$  dimana  $\#\text{simulasi}$  menunjukkan banyaknya replikasi yang digunakan.

Prosedur untuk mendapatkan *p-value* dengan pendekatan *Monte Carlo* adalah sebagai berikut.

1. Hitung penjumlahan nilai *log likelihood ratio* tertinggi  $t_0$  untuk data riil.
2. Membangun data acak yang ukurannya sama dengan data riil yang dibangun di bawah kondisi  $H_0$ .
3. Melakukan proses pembentukan *scanning window*  $Z$  dari data acak yang dibangun berdasarkan kondisi  $H_0$ .
4. Mencari nilai *log likelihood ratio* dari setiap *scanning window*, dan dicatat apakah jumlah kasus yang diamati lebih besar atau lebih kecil dari yang diestimasi, kemudian menjumlahkan nilai *log likelihood ratio* yang jumlah kasusnya lebih besar dari jumlah yang diestimasi, untuk setiap *scanning window*. Langkah selanjutnya, mendapatkan penjumlahan nilai *log likelihood ratio* yang tertinggi dari simulasi pertama pembangunan data acak tersebut.
5. Mengulang langkah 2, 3, dan 4 sebanyak  $m$  kali pengulangan/simulasi, sehingga memperoleh  $m$  penjumlahan nilai *log likelihood ratio* tertinggi dari data acak dan data riil.

6. Hitung *p-value*,  $p = \frac{\text{banyaknya } (T(y) \geq t_0)}{m+1}$

$t_0$  menyatakan penjumlahan nilai *log likelihood ratio* tertinggi yang dimiliki suatu *scanning window Z* dari data riil.  $T(y)$  adalah penjumlahan nilai *log likelihood ratio* dari data acak yang dibangun di bawah kondisi  $H_0$ .  $m$  adalah banyaknya simulasi untuk membangun data di bawah kondisi  $H_0$ .

Kantong-kantong yang terbentuk (*window Z*) diurutkan berdasarkan nilai *likelihood*-nya dan *window Z* yang memiliki nilai *likelihood* lebih besar daripada yang lain dan nyata pada tingkat signifikansi tertentu akan membentuk *hotspot*.

## 2.8 Demam Berdarah *Dengue* (DBD)

Demam Berdarah *Dengue* (DBD) adalah penyakit menular yang disebabkan oleh infeksi virus DEN-1, DEN-2, DEN-3, atau DEN-4 yang ditularkan melalui gigitan nyamuk *Aedes Aegypti* atau *aedes albopictus* yang sebelumnya telah terinfeksi oleh virus *dengue* dari penderita DBD lainnya.

GINANJAR (2007) Masa inkubasi penyakit DBD, yaitu periode sejak virus *dengue* menginfeksi manusia hingga menimbulkan gejala klinis, antara 3-14 hari, rata-rata antara 4-7 hari. Penyakit DBD tidak ditularkan langsung dari orang ke orang. Penderita menjadi infeksiif bagi nyamuk pada saat viremia yaitu beberapa saat menjelang timbulnya demam hingga saat masa demam berakhir. Nyamuk *aedes aegypti* merupakan penyebar penyakit DBD yang paling efektif dan utama karena tinggal di sekitar permukiman penduduk. Adapun nyamuk *aedes albopictus*, banyak terdapat di daerah perkebunan dan semak-semak.

Penyakit DBD pertama kali di Surabaya pada tahun 1968. Pada tahun 1994 DBD telah menyebar ke seluruh provinsi (pada waktu itu berjumlah 27 provinsi) di Indonesia. Penyakit DBD mesti mendapatkan perhatian serius dari semua pihak, mengingat jumlah kasusnya yang cenderung meningkat setiap tahun. Menurut data Departemen Kesehatan Republik Indonesia, pada

awal 2007 jumlah penderita DBD telah mencapai 16.803 orang dan 267 orang yang diantaranya meninggal dunia. Jumlah orang yang meninggal tersebut jauh lebih banyak dibandingkan kasus kematian manusia karena flu burung.

Penularan penyakit DBD ini juga dipengaruhi oleh interaksi tiga faktor, yaitu sebagai berikut.

- a. Faktor penjamu (Target penyakit, inang), dalam hal ini adalah manusia yang rentan tertular penyakit DBD. Adapun indikator yang termasuk faktor penjamu ini adalah kepadatan penduduk.
- b. Faktor penyebar (vektor) dan penyebab penyakit (agen), dalam hal ini adalah virus DEN tipe 1-4 sebagai agen penyebab penyakit, sedangkan nyamuk *Aedes Aegypti* dan *Aedes albopictus* yang berperan sebagai vektor penyebar penyakit DBD. Adapun indikator yang termasuk faktor penyebab ini adalah jentik nyamuk.
- c. Faktor lingkungan yaitu lingkungan yang memudahkan terjadinya kontak penularan penyakit DBD. Adapun indikator yang termasuk faktor lingkungan ini adalah rumah tangga berperilaku hidup bersih dan sehat.

Berbagai upaya untuk memutus mata rantai penularan penyakit DBD dapat ditempuh dengan cara memodifikasi faktor-faktor yang terlibat di dalamnya. Perbaikan kualitas kebersihan (sanitasi) lingkungan, serta pencegahan penyakit dan pengobatan segera bagi penderita penyakit DBD. Namun, yang paling penting diperhatikan adalah peningkatan pemahaman, kesadaran, sikap dan perubahan perilaku masyarakat terhadap penyakit DBD ini, sakan sangat mendukung percepatan upaya memutuskan mata rantai penularan penyakit DBD.

## **BAB III**

### **METODOLOGI PENELITIAN**

#### **3.1 Sumber Data**

Data yang digunakan dalam penelitian ini adalah data sekunder yaitu data variabel respon (jumlah penderita penyakit Demam Berdarah *Dengue* (DBD) tiap kabupaten/kota di Provinsi Jawa Timur tahun 2014) diperoleh dari profil kesehatan Dinas Kesehatan Provinsi Jawa Timur. Sedangkan untuk variabel prediktor diperoleh dari profil kesehatan Dinas Kesehatan Provinsi Jawa Timur dan Jawa Timur Dalam Angka Badan Pusat Statistik.



**Gambar 3.1** Provinsi Jawa Timur

#### **3.2 Variabel Penelitian**

Variabel yang digunakan dalam penelitian terdiri dari satu variabel respon (Y) yaitu jumlah kasus DBD dan tujuh variabel prediktor (X). Variabel yang digunakan dalam penelitian ini dapat dituliskan dalam Tabel 3.1 sebagai berikut.

**Tabel 3.1** Variabel Penelitian

Variabel	Keterangan	Satuan	Kelompok
Y	Banyak penderita DBD	Jiwa	
X <sub>1</sub>	Kepadatan penduduk	Jiwa/Km <sup>2</sup>	Target penyakit
X <sub>2</sub>	Persentase rumah tangga yang berperilaku hidup bersih dan sehat	Persen	Lingkungan
X <sub>3</sub>	Persentase rumah sehat	Persen	Lingkungan
X <sub>4</sub>	Rasio sarana kesehatan	Per 1000 Penduduk	Lingkungan
X <sub>5</sub>	Rasio tenaga kesehatan	Per 1000 Penduduk	Lingkungan

Berikut merupakan definisi operasional variabel yang digunakan dalam penelitian.

- a. Banyak penderita Demam Berdarah *Dengue* (Y).  
Banyak penderita DBD di setiap kabupaten/kota di Provinsi Jawa Timur tahun 2014.
- b. Kepadatan penduduk (X<sub>1</sub>)  
Hasil bagi jumlah penduduk terhadap luas wilayah di setiap kabupaten/kota Provinsi Jawa Timur dalam satuan (jiwa/km<sup>2</sup>).
- c. Persentase rumah tangga yang berperilaku hidup bersih dan sehat (X<sub>2</sub>)  
Hasil bagi antar jumlah rumah tangga yang berperilaku melaksanakan 10 indikator PHBS pada kurun waktu tertentu dengan jumlah tangga yang diperiksa pada kurun waktu yang sama di kabupaten/kota Provinsi Jawa Timur dikali dengan 100%.  
Perilaku hidup bersih dan sehat (PHBS) adalah sekumpulan perilaku yang dipraktikkan atas dasar kesadaran sebagai hasil pembelajaran, yang menjadikan seseorang, keluarga, kelompok dan masyarakat mampu menolong dirinya sendiri( mandiri) di bidang kesehatan dan berperan



aktif dalam mewujudkan kesehatan masyarakat. PHBS merupakan suatu tindakan pencegahan agar masyarakat terhindar dari penyakit dan gangguan kesehatan (Depkes, 2013). Adapun 10 indikator PHBS yaitu pertolongan persalinan oleh tenaga kesehatan, bayi diberi ASI eksklusif, balita ditimbang setiap bulan, menggunakan air bersih, mencuci tangan dengan air bersih dan sabun, menggunakan jamban sehat, memberantas jentik dirumah sekali seminggu, makan sayur dan buah setiap hari, melakukan aktifitas fisik setiap hari dan tidak merokok di dalam rumah.

d. Persentase rumah sehat ( $X_3$ )

Merupakan hasil bagi antara jumlah rumah kategori sehat dengan jumlah rumah yang dipantau dikalikan 100%. Rumah sehat merupakan salah satu sarana untuk mencapai derajat kesehatan yang optimum. Rumah sehat adalah bangunan rumah tinggal yang memenuhi syarat kesehatan yaitu rumah yang memiliki jamban sehat, sarana air bersih, tempat pembuangan sampah, sarana pembuangan air limbah, ventilasi rumah yang baik, kepadatan hunian rumah yang sesuai dan lantai rumah yang tidak terbuat dari tanah.

e. Rasio sarana kesehatan Per 1000 Penduduk ( $X_4$ )

Merupakan hasil bagi dari jumlah penduduk dengan jumlah rumah sakit dan puskesmas di kabupaten/kota di Jawa Timur per 1000 penduduk. Dengan adanya sarana kesehatan diharapkan dapat menjangkau masyarakat terutama masyarakat di pedesaan agar mendapatkan pelayanan kesehatan dengan mudah dan bermutu.

f. Rasio tenaga kesehatan Per 1000 Penduduk ( $X_5$ )

Merupakan hasil bagi dari jumlah penduduk dengan jumlah dokter umum dan perawat di rumah sakit dan puskesmas di kabupaten/kota di Jawa Timur per 1000 penduduk. Tenaga kesehatan berperan aktif dalam penanganan dan penyembuhan penyakit penderita DBD.

### 3.3 Langkah Analisis

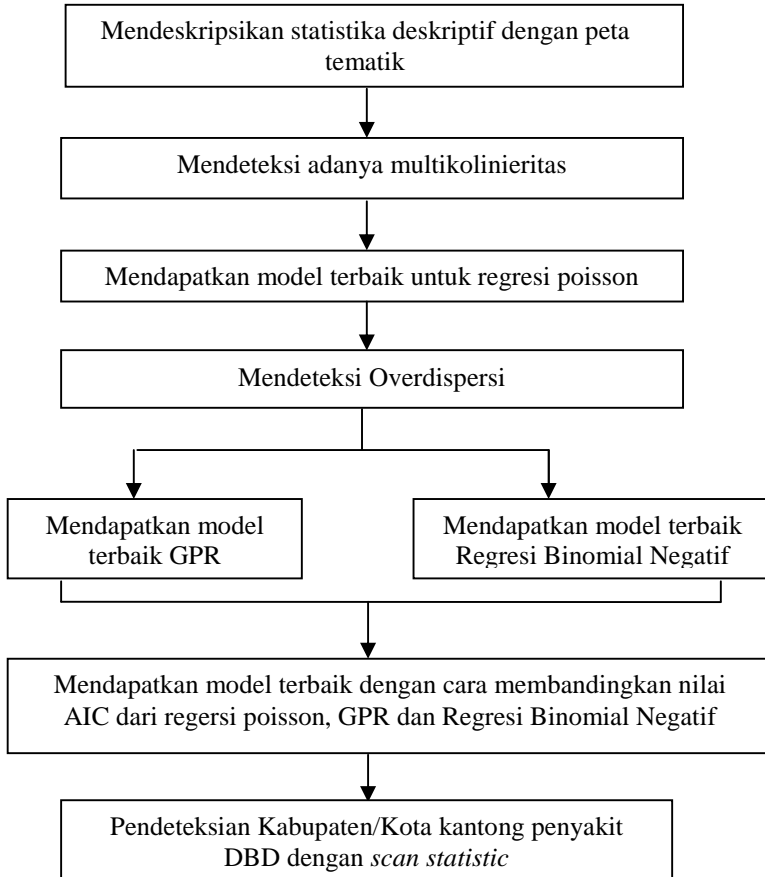
Langkah analisis data yang dilakukan pada penelitian ini dapat dituliskan sebagai berikut.

- A. Langkah-langkah mendapatkan model terbaik dan faktor-faktor yang berpengaruh terhadap jumlah penderita DBD di kabupaten/kota di Jawa Timur adalah sebagai berikut.
  1. Mendeskripsikan jumlah kasus DBD dan faktor-faktor yang mempengaruhinya dengan peta tematik.
  2. Menganalisis korelasi antar variabel-variabel prediktor untuk mendeteksi adanya kasus multikolinieritas dengan melihat nilai koefisien korelasi *pearson* dan VIF pada persamaan 2.1.
  3. Mendapatkan model untuk regresi poisson pada jumlah penderita DBD yang terjadi di kabupaten/kota di Jawa Timur pada persamaan 2.5.
  4. Mendeteksi adanya overdispersi pada data dengan melihat nilai *Pearson Chi-squares* dan *Deviance* yang dibagi dengan derajat bebasnya.
  5. Mendapatkan model untuk *GPR* pada pemodelan jumlah kasus DBD yang terjadi di kabupaten/kota di Jawa Timur pada persamaan 2.12 dengan langkah-langkah sebagai berikut.
    - a. Menaksir parameter model *GPR*
    - b. Menguji signifikansi parameter model *GPR* secara serentak dan parsial pada persamaan 2.15.
  6. Mendapatkan model untuk regresi binomial negatif pada pemodelan jumlah kasus DBD yang terjadi di kabupaten/kota di Jawa Timur dengan langkah-langkah sebagai berikut.
    - a. Menaksir parameter model regresi binomial negatif
    - b. Menguji signifikansi parameter model regresi binomial negatif secara serentak dan parsial pada persamaan 2.22.

7. Memilih model terbaik dari regresi poisson, *GPR* dan regresi binomial negatif dengan melihat nilai AIC terkecil pada persamaan 2.23.
- B. Langkah-langkah untuk mendeteksi kantong-kantong kasus DBD di kabupaten/kota di Jawa Timur dengan pendekatan *Flexibly Shaped Spatial Scan Statistic* dengan menggunakan *software FlexScan 3.1* ([https://sites.google.com/site/flexscansoftware/download\\_e](https://sites.google.com/site/flexscansoftware/download_e)) sebagai berikut :
1. Mengambil data banyaknya kasus DBD per kabupaten/kota sebagai kasus dan jumlah penduduk per kabupaten/kota sebagai populasi.
  2. Membuat prediksi jumlah kasus DBD per kabupaten/kota ( $\hat{y}_i$ ) berdasarkan model terpilih pada langkah (A7).
  3. Mengidentifikasi kandidat *cluster/hotspot* dengan algoritma yang ada di metode *Flexibly Shaped Spatial Scan Statistic* pada sub bab 2.7. Penentuan *pre-set maximum number* mengacu pada *pre-set maximum number* yang digunakan pada metode *Flexibly Shaped Spatial Scan Statistic*.
  4. Menghitung jumlah populasi dan jumlah kasus untuk setiap *cluster*.
  5. Menghitung *log likelihood function* dari setiap *cluster* kemudian menguji signifikansi *cluster* dengan simulasi *Monte Carlo* untuk mendapatkan *p-value*.
  6. Menghitung resiko relatif untuk setiap *cluster*.
  7. Membuat peta kantong kasus DBD berdasarkan *cluster* yang signifikan pada sebesar 0,10.
  8. Menginterpretasikan hasil dengan membuat deskripsi kantong kasus DBD sehingga diperoleh prioritas lokasi pengendalian penyakit dan faktor-faktor yang signifikan mempengaruhi kasus di wilayah kantong-kantong.

### 3.4 Diagram Alir Penelitian

Berdasarkan langkah-langkah analisis data dapat disajikan dalam bentuk diagram alir sehingga lebih mudah dipahami seperti pada Gambar 3.2 berikut ini.



**Gambar 3.2** Diagram Alir Penelitian

## **BAB IV**

### **ANALISIS DAN PEMBAHASAN**

Pada bab analisis dan pembahasan ini akan di paparkan mengenai hasil pengolahan data penyakit Demam Berdarah *Dengue* di Provinsi Jawa Timur tahun 2014. Hal-hal yang dibahas meliputi deskripsi banyaknya kasus DBD yang terjadi di Provinsi Jawa Timur berdasarkan Kabupaten/Kota serta faktor-faktor yang mempengaruhinya, deteksi kasus multikolinieritas, melakukan permodelan regresi Poisson, pendeteksian overdispersi pada model regresi poisson. Apabila terjadi kasus overdispersi, maka dilanjutkan dengan metode *Generalized Poisson Regression* (GPR) dan regresi binomial negatif. Selanjutnya, dimungkinkan adanya aspek spasial maka di lanjutkan dengan pendekatan *Flexibly Shaped Spasial Scan Statistic* untuk mendeteksi kantong-kantong DBD di Jawa Timur.

#### **4.1 Deskripsi Kasus Demam Berdarah *Dengue* dan Faktor-Faktor yang diduga Mempengaruhi**

Seperti yang telah dijelaskan pada sub bab 3.2 bahwa penelitian ini menggunakan data jumlah kasus DBD di Provinsi Jawa Timur. Berdasarkan data Dinas Kesehatan Provinsi Jawa Timur, kasus DBD terjadi di 38 Kabupaten/Kota di Jawa Timur tahun 2014. Data yang digunakan dideskripsikan berdasarkan nilai rata-rata, varians, standar deviasi, maksimum dan minimum yang ditampilkan pada Tabel 4.1.

Pada Tabel 4.1 terlihat rata-rata variabel kepadatan penduduk ( $X_1$ ) di Provinsi Jawa Timur sebesar 1893,16 jiwa/km<sup>2</sup>, dengan varians yang cukup besar 5169638,17 yang artinya keragaman kepadatan penduduk di kabupaten/kota di Provinsi Jawa Timur sangat beragam. Rata-rata rumah tangga berperilaku hidup bersih dan sehat ( $X_2$ ) sebesar 46,29 % kurang dari 50% rumah tangga yang dipantau telah berperilaku hidup bersih dan sehat. Persentase rumah tangga berperilaku hidup bersih dan sehat tertinggi sebesar 68,7% yaitu di Kabupaten Gresik. Rata-rata

persentase rumah tangga ( $X_3$ )sehat di Provinsi Jawa Timur sebesar 54,16% berarti lebih dari 50% rumah tangga yang dipantau berstatus rumah sehat. Persentase rumah sehat yang tertinggi terdapat di Kabupaten Banyuwangi sebesar 132,98%. Rata-rata variabel rasio saran kesehatan ( $X_4$ ) sebesar 31,15 per 1000 penduduk. Selanjutnya rata-rata variabel rasio tenaga kesehatan ( $X_5$ ) di Provinsi Jawa Timur sebesar 1,87 per 1000 penduduk. Berikut statistika deskriptif disajikan dalam Tabel 4.1

**Tabel 4.1** Statistika Deskriptif Variabel Penelitian

Variabel	Rata-Rata	Varians	Maksimum	Minimum
Jumlah Kasus DBD	244,39	43405,49	901,00	9,00
Kepadatan penduduk	1893,16	5169638,17	8683,16	274,70
Persentase rumah tangga ber-PHBS	46,29	209,51	68,7	20,1
Persentase rumah sehat	54,16	821,08	132,98	4,20
Rasio sarana kesehatan	31,15	75,34	52,10	12,47
Rasio tenaga kesehatan	1,87	6,40	15,93	0,21

Penyebaran jumlah kasus DBD dan faktor-faktor yang diduga mempengaruhinya ditampilkan juga dalam bentuk peta tematik dengan menggunakan *Arcview GIS* 3.3. Penyebaran data dibagi menjadi tiga kelompok kategori yaitu kategori rendah, sedang dan tinggi dengan menggunakan *confidence interval*.

#### 4.1.1 Jumlah Kasus Demam Berdarah *Dengue*

Penyakit DBD merupakan salah satu penyakit menular. Penyakit DBD sering muncul sebagai kejadian luar biasa (KLB) dengan angka kematian relatif tinggi. Adapun penyebaran jumlah kasus DBD di Provinsi Jawa Timur sebagai berikut.



**Gambar 4.1** Penyebaran Jumlah Kasus DBD di Jawa Timur

Gambar 4.1 diatas menunjukkan penyebaran jumlah kasus DBD di setiap kabupaten/kota di Jawa Timur. Jumlah DBD dengan kategori tinggi di tunjukkan dengan warna oren yaitu terdapat di Kota Surabaya, Kota Probolinggo, Kabupaten Sumenep, Bondowoso, Jember, Banyuwangi, Malang dan Kabupaten Ponorogo dengan jumlah kasus berada di interval 299,82 hingga 901 kasus. Selanjutnya, untuk kategori jumlah kasus DBD sedang terdapat di sembilan kabupaten/kota yaitu, Kabupaten Sampang, Bangkalan, Gresik, Probolinggo, Situbondo, Jombang, Tulungagung, Trenggalek dan Kabupaten Pacitan dengan jumlah kasus DBD berada di interval 188,97 hingga 299,82 kasus. Jumlah kasus DBD dengan kategori rendah terdapat di 21 kabupaten/kota yang belum disebutkan.

#### **4.1.2 Kepadatan Penduduk**

Kepadatan penduduk adalah perbandingan antara jumlah penduduk dengan luas wilayahnya. Kepadatan penduduk merupakan salah satu faktor yang diduga mempengaruhi

banyaknya kasus DBD. Penyakit DBD merupakan penyakit menular, sehingga semakin padat penduduk suatu wilayah maka penularan penyakit DBD semakin mudah mengakibatkan jumlah kasus DBD di wilayah tersebut semakin tinggi. Penyebaran kepadatan penduduk di Provinsi Jawa Timur sebagai berikut.



**Gambar 4.2** Penyebaran Kepadatan Penduduk di Jawa Timur

Gambar 4.2 diatas menunjukkan penyebaran kepadatan penduduk di setiap kabupaten/kota di Jawa Timur. Kepadatan penduduk dengan kategori tinggi terdapat di sembilan kabupaten/kota yaitu Kota Surabaya, Mojokerto, Madiun, Kediri, Blitar, Malang, Probolinggo, Pasuruan dan Kabupaten Sidoarjo. Kabupaten/kota yang termasuk dalam kategori tinggi ini merupakan kota dengan luas wilayah kecil dan jumlah penduduk di wilayah tersebut sangat banyak, sehingga kepadatan penduduk di wilayah ini tinggi. Selanjutnya, untuk kepadatan penduduk dengan kategori sedang terdapat di tujuh kabupaten/kota yaitu Kabupaten Mojokerto. Selanjutnya kepadatan penduduk dengan kategori rendah terdapat di 28 kabupaten/kota yang belum





Madiun, Magetan, Ponorogo, Pacitan dan Kabupaten Jember dengan persentase rumah tangga ber-PHBS dalam rentang 50,14 hingga 68,7 persen. Selanjutnya, untuk rumah tangga ber-PHBS dengan kategori sedang terdapat Kabupaten Tuban, Pasuruan, Blitar dan Kabupaten Lumajang. Selanjutnya untuk wilayah yang termasuk dalam kategori rendah terdapat 17 kabupaten/kota yang berada di interval 20,1 hingga 42,44 persen. Wilayah ini jadi perhatian khusus bagi pemerintah, karena masih banyak rumah tangga dengan perilaku hidup yang kurang bersih dan kurang sehat. Salah satu cara yang dapat dilakukan pemerintah adalah menghimbau masyarakat dan memberikan penyuluhan kepada masyarakat untuk memperhatikan lingkungan sekitar. Jika lingkungan disekitar rumah bersih dan sehat maka sarang wabah penyakit DBD akan berkurang dan mengurangi jumlah kasus penyakit DBD.

#### **4.1.4 Persentase Rumah Sehat**

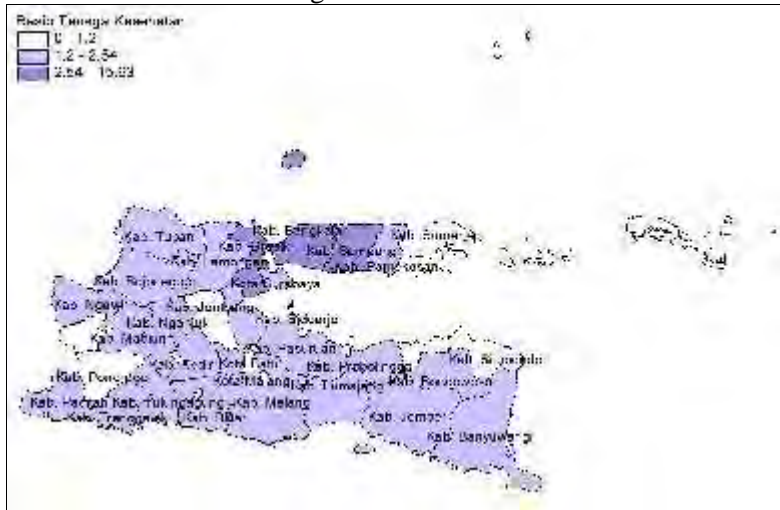
Rumah Sehat adalah bangunan rumah tinggal yang memenuhi syarat kesehatan yaitu memiliki jamban sehat, tempat pembuangan sampah, sarana air bersih, sarana pembuangan air limbah, ventilasi baik, kepadatan hunian rumah sesuai dan lantai rumah tidak dari tanah. Adapun penyebaran kondisi rumah sehat di Provinsi Jawa Timur terlihat pada Gambar 4.4.

Pada Gambar 4.4 diatas menunjukkan penyebaran persentase rumah sehat di setiap kabupaten/kota di Jawa Timur. Wilayah yang termasuk dalam kategori tinggi terhadap persentase rumah sehat yaitu Kota Surabaya, Mojokerto, Pasuruan, Probolinggo, Malang, Kediri, Madiun, Kabupaten Gresik, Sidoarjo, Lamongan, Tuban, Bojonegoro, Magetan, Madiun, Jombang, Ponorogo, Pacitan, Tulungagung, Jember dan Kabupaten Banyuwangi. Selanjutnya, wilayah yang termasuk dalam kategori rendah terdapat di 18 Kabupaten/Kota yang belum disebutkan. Wilayah ini perlu perhatian khusus terhadap lingkungan rumah. Untuk meningkatkan cakupan rumah sehat di Jawa Timur, bisa dengan cara pemberdayaan masyarakat dengan





Tenaga Kesehatan dalam penelitian ini terdiri dari dokter umum, dan tenaga keperawatan dari rumah sakit dan puskesmas di Kabupaten/Kota. Adapun penyebaran tenaga kesehatan di Provinsi Jawa Timur sebagai berikut.



**Gambar 4.6** Penyebaran Tenaga Kesehatan per 1000 Penduduk di Jawa Timur

Gambar 4.6 diatas menunjukkan penyebaran tenaga kesehatan di setiap kabupaten/kota di Jawa Timur. Wilayah yang memiliki tenaga kesehatan dengan kategori tinggi terdapat di tiga kabupaten/kota yaitu Kabupaten Gresik, Kabupaten Bangkalan dan Kabupaten Sampang. Selanjutnya, untuk sarana kesehatan dengan kategori rendah terdapat di Kota Pasuruan, Kota Probolinggo, Kota Malang, Kota Batu, Kota Kediri, Kota Blitar, Kota Mojokerto, Kota Madiun, Kabupaten Magetan, Ponorogo, Situbondo, Sidoarjo, Jombang dan Kabupaten Sumenep. Untuk yang termasuk dalam wilayah ini sangat dibutuhkan tenaga kesehatan yang lebih banyak, sehingga bisa lebih cepat menangani pasien. Selanjutnya, wilayah yang termasuk dalam

kategori sedang terdapat di 21 Kabupaten/Kota yang belum disebutkan.

#### 4.2 Pemeriksaan Multikolinearitas

Sebelum melakukan analisis dengan tiga metode, yaitu regresi poisson, *Generalized Poisson Regression* dan regresi binomial negatif maka perlu dilakukan pengujian multikolinieritas terhadap data yang digunakan untuk melihat antar variabel prediktor apakah ada tidaknya korelasi yang tinggi. Ada beberapa cara untuk mendeteksi kasus multikolinieritas yaitu dengan matriks koefisien korelasi dan nilai VIF. Berikut ini merupakan matrik koefisien korelasi antar variabel prediktor

**Tabel 4.2** Koefisien Korelasi Antar Variabel Prediktor

	$X_1$	$X_2$	$X_3$	$X_4$
$X_2$	0,25701			
$X_3$	0,37686	0,51574		
$X_4$	-0,32628	-0,19944	-0,32806	
$X_5$	-0,23274	0,02750	-0,36282	0,23569

Adanya kasus multikolinieritas dilihat dari koefisien korelasi pearson jika nilainya lebih besar dari 0,95. Berdasarkan Tabel 4.2 semua koefisien korelasi antar variabel prediktor kurang dari 0,95 maka tidak adanya kasus multikolinieritas pada data yang digunakan. Kriteria lain yang dapat mendeteksi kasus multikolinieritas adalah nilai VIF. Terjadinya multikolinieritas jika nilai VIF lebih besar dari 10. Nilai VIF masing-masing variabel prediktor ditampilkan pada Tabel 4.3.

**Tabel 4.3** Nilai VIF dari Variabel Prediktor

Variabel	VIF
$X_1$	1,25267
$X_2$	1,49194
$X_3$	1,80543
$X_4$	1,20212
$X_5$	1,28413

Berdasarkan Tabel 4.3 dapat diketahui bahwa nilai VIF masing-masing variabel prediktor bernilai kurang dari 10, maka

dapat disimpulkan tidak ada kasus multikolinieritas. Sehingga dapat dilanjutkan ke pemodelan regresi poisson, GPR dan regresi binomial negatif.

#### 4.3 Pemodelan Jumlah Kasus DBD tiap Kabupaten/Kota di Jawa Timur Tahun 2014 Menggunakan Regresi Poisson

Setelah dilakukan pemeriksaan kasus multikolinieritas diperoleh hasil bahwa tidak terdapat korelasi yang tinggi antar variabel prediktor, maka dapat dilanjutkan pada pemodelan regresi poisson. Berikut adalah hasil estimasi parameter model regresi poisson.

**Tabel 4.4** Estimasi Parameter Model Regresi Poisson

Parameter	Estimasi	Standart Error	Z <sub>hitung</sub>	P Value
0	3,7552	0,05786	62,90	< 0,0001
1	0,000014	4,824 x 10 <sup>-6</sup>	2,98	0,0051
2	-0,00458	0,000866	-5,29	<0,0001
3	0,01177	0,000426	27,61	<0,0001
4	0,03537	0,001204	29,38	<0,0001
5	0,06248	0,004246	14,71	<0,0001
<b>Devians : 3750,9852</b>		DF: 32		
<b>AIC : 4032,6</b>				

<sup>\*)</sup> signifikan dengan taraf nyata 10%

Selanjutnya dilakukan pengujian secara serentak yang bertujuan untuk mengetahui apakah secara serentak variabel prediktor memberikan pengaruh terhadap variabel respon. Berikut adalah hipotesis pengujian parameter regresi poisson secara serentak.

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$$

$$H_1 : \text{paling sedikit ada satu } \beta_j \neq 0; j=1,2,\dots,5$$

Berdasarkan Tabel 4.3 menunjukkan bahwa dari hasil analisis diperoleh nilai devians sebesar 3750,9852, hal ini berarti nilai devians lebih besar dari  $t_{(0,1;5)}^2$  sebesar 9,236. Sehingga diperoleh keputusan tolak  $H_0$  yang berarti bahwa paling sedikit

ada satu variabel prediktor yang berpengaruh terhadap model. Oleh karena itu, perlu dilanjutkan pada pengujian secara parsial dengan hipotesis sebagai berikut.

$H_0 : \beta_j = 0$  (pengaruh variabel ke-j tidak signifikan)

$H_1 : \beta_j \neq 0$  (pengaruh variabel ke-j signifikan)

Pengujian ini menggunakan statistik uji  $|Z_{hitung}|$  yang kemudian dibandingkan dengan  $Z_{tabel}$ , dengan menggunakan taraf signifikansi 10% didapat  $Z_{(0,1/2)}$  sebesar 1,64. Berdasarkan Tabel 4.3 menunjukkan bahwa  $|Z_{hitung}|$  dari masing-masing variabel prediktor lebih besar dari 1,64 sehingga diputuskan tolak  $H_0$ . Jadi semua variabel prediktor secara individu memberikan pengaruh signifikan terhadap jumlah kasus DBD di Jawa Timur. Berikut model regresi poisson yang dihasilkan.

$$\hat{\lambda} = \exp(3,7552 + 0,000014X_1 - 0,00458X_2 + 0,01177X_3 + 0,03537X_4 + 0,06248X_5)$$

Dari model regresi poisson dapat diketahui bahwa setiap kenaikan satu persen rumah sehat ( $X_3$ ), maka akan meningkatkan rata-rata jumlah kasus DBD di Jawa Timur tahun 2014 sebesar  $\exp(0,01177) = 1,01$  kasus dengan syarat semua variabel lain konstan, selain itu interpretasi yang sama untuk kepadatan penduduk ( $X_1$ ), rasio sarana kesehatan ( $X_4$ ) dan rasio tenaga kesehatan ( $X_5$ ). Jika persentase rumah tangga berperilaku PHBS ( $X_2$ ) meningkat satu persen, maka rata-rata jumlah kasus DBD di Jawa Timur tahun 2014 akan menurun sebesar  $\exp(-0,00458) = 0.995$  kasus dengan syarat semua variabel lain konstan.

#### 4.4 Pemeriksaan Overdispersi

Dalam regresi poisson terdapat asumsi yang harus dipenuhi yaitu nilai rata-rata dan varians harus sama disebut *equidispersion*. Untuk memenuhi kondisi *equidispersion* jarang terpenuhi sehingga muncul adanya kasus overdispersi. Untuk mendeteksi, dapat dilihat dari nilai *deviance/db* atau *pearson/db*. Jika nilai *deviance/db* atau *pearson/db* lebih besar dari 1, maka dapat dikatakan terjadinya overdispersi. Dapat diketahui bahwa nilai *deviance/db* atau *pearson/db* lebih besar dari 1, sehingga



dapat disimpulkan pada kasus DBD di Jawa Timur tahun 2014 terjadi overdispersi. Adanya overdispersi tersebut menyebabkan model yang terbentuk akan menghasilkan estimasi parameter yang bias. Oleh karena itu, metode yang sesuai digunakan untuk mengatasi overdispersi adalah *Generalized Poisson Regression* (GPR) dan Regresi Binomial Negatif. Untuk lebih rincinya ditampilkan pada Tabel 4.5 sebagai berikut.

**Tabel 4.5** Nilai *Deviance* dan *Pearson* dari Model Regresi Poisson

Kriteria	Nilai	db	Nilai/db
<i>Deviance</i>	3750,9852	32	117,2183
<i>Pearson Chi-Square</i>	4211,6855	32	131,6152

#### 4.5 Pemodelan Jumlah Kasus DBD tiap Kabupaten/Kota di Jawa Timur Tahun 2014 Menggunakan *Generalized Poisson Regression* (GPR)

Model GPR merupakan suatu model yang sesuai digunakan untuk data *count* dan bertujuan untuk mengatasi kasus overdispersi pada regresi poisson. Sebelum membuat model GPR, dilakukan penaksiran parameter dan pengujian parameter secara serentak dan parsial. Berikut ditampilkan hasil estimasi parameter model GPR.

**Tabel 4.6** Estimasi Parameter Model GPR

Parameter	Estimasi	Standart Error	Z <sub>hitung</sub>	P Value
0	3,8077	0,6966	5,47	<0,0001
1	-0,00002	0,000055	-0,28	0,7811
2	-0,00768	0,01034	-0,74	0,4623
3	0,01312	0,007201	1,82 <sup>*)</sup>	0,0763
4	0,03751	0,01814	2,07 <sup>*)</sup>	0,0455
5	0,06860	0,06687	1,03	0,3114

**Devians : 480,2**

**AIC : 494,2**

<sup>\*)</sup> signifikan dengan taraf nyata 10%

Berikut adalah pengujian secara serentak pada model GPR yang bertujuan untuk mengetahui apakah secara serentak variabel

prediktor memberikan pengaruh terhadap variabel respon. Berikut adalah hipotesis pengujian parameter GPR secara serentak.

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$$

$$H_1 : \text{paling sedikit ada satu } \beta_j \neq 0; j=1,2,\dots,5$$

Berdasarkan Tabel 4.6 menunjukkan bahwa dari hasil analisis diperoleh nilai devians sebesar 480,2, hal ini berarti nilai devians lebih besar dari  $t_{(0,1;5)}^2$  sebesar 9,236. Sehingga diperoleh

keputusan tolak  $H_0$  yang berarti bahwa paling sedikit ada satu variabel prediktor yang berpengaruh terhadap model. Oleh karena itu, perlu dilanjutkan pada pengujian secara parsial dengan hipotesis sebagai berikut.

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

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

Pengujian ini menggunakan statistik uji  $|Z_{hitung}|$  yang kemudian dibandingkan dengan  $Z_{tabel}$ , dengan menggunakan taraf signifikansi 10% didapat  $Z_{(0,1/2)}$  sebesar 1,64. Berdasarkan Tabel 4.5 menunjukkan bahwa variabel yang berpengaruh adalah persentase rumah sehat ( $X_3$ ) dan rasio sarana kesehatan ( $X_4$ ) sehingga diperoleh model sebagai berikut.

$$\hat{\lambda} = \exp(3,8077 - 0,00002X_1 - 0,00768X_2 + 0,01312X_3 + 0,03751X_4 + 0,06860X_5)$$

Dari model GPR dapat diketahui bahwa setiap kenaikan satu persen rumah sehat ( $X_3$ ), maka akan meningkatkan rata-rata jumlah kasus DBD di Jawa Timur tahun 2014 sebesar  $\exp(0,01312) = 1,01$  kasus dengan syarat semua variabel lain konstan, selain itu interpretasi yang sama untuk, rasio sarana kesehatan ( $X_4$ ) dan rasio tenaga kesehatan ( $X_5$ ). Jika persentase rumah tangga berperilaku PHBS ( $X_2$ ) meningkat satu persen, maka rata-rata jumlah kasus DBD di Jawa Timur tahun 2014 akan menurun sebesar  $\exp(-0,00768) = 0,992$  kasus dengan syarat semua variabel lain konstan, selain itu interpretasi yang sama untuk kepadatan penduduk ( $X_1$ ).

#### 4.6 Pemodelan Jumlah Kasus DBD tiap Kabupaten/Kota di Jawa Timur Tahun 2014 Menggunakan Regresi Binomial Negatif

Selain menggunakan GPR untuk mengatasi overdispersi pada model regresi poisson, dapat juga menggunakan model regresi binomial negatif. Sebelum membuat model regresi binomial negatif, dilakukan penaksiran parameter dan pengujian parameter secara serentak dan parsial. Berikut ditampilkan hasil estimasi parameter model regresi binomial negatif.

**Tabel 4.7** Estimasi Parameter Model Regresi Binomial Negatif

Parameter	Estimasi	Standart Error	Z <sub>hitung</sub>	P Value
0	3,834	0,6079	6,307	2,84 x10 <sup>-10</sup>
1	2,797 x10 <sup>-6</sup>	5,199 x10 <sup>-5</sup>	0,054	0,95709
2	-0,006894	0,008919	-0,773	0,43956
3	0,01225	0,004953	2,473 <sup>*)</sup>	0,01339
4	0,03601	0,01334	2,700 <sup>*)</sup>	0,00693
5	0,06405	0,04729	1,354	0,17558

**Devians : 40,643**  
**AIC : 489,21**

<sup>\*)</sup> signifikan dengan taraf nyata 10%

Pengujian secara serentak pada model regresi binomial negatif yang bertujuan untuk mengetahui apakah secara serentak variabel prediktor memberikan pengaruh terhadap variabel respon. Berikut adalah hipotesis pengujian parameter regresi binomial negatif secara serentak.

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$$

$$H_1 : \text{paling sedikit ada satu } \beta_j \neq 0; j=1,2,\dots,5$$

Berdasarkan Tabel 4.7 menunjukkan bahwa dari hasil analisis diperoleh nilai devians sebesar 40,643, hal ini berarti nilai devians lebih besar dari  $t_{(0,1;5)}^2$  sebesar 9,236. Sehingga diperoleh keputusan tolak  $H_0$  yang berarti bahwa paling sedikit ada satu variabel prediktor yang berpengaruh terhadap model. Oleh karena itu, perlu dilanjutkan pada pengujian secara parsial dengan hipotesis sebagai berikut.

$H_0 : \beta_j = 0$  (pengaruh variabel ke- $j$  tidak signifikan)

$H_1 : \beta_j \neq 0$  (pengaruh variabel ke- $j$  signifikan)

Pengujian ini menggunakan statistik uji  $|Z_{hitung}|$  yang kemudian dibandingkan dengan  $Z_{tabel}$ , dengan menggunakan taraf signifikansi 10% didapat  $Z_{(0,1/2)}$  sebesar 1,64. Berdasarkan Tabel 4.6 menunjukkan bahwa variabel yang berpengaruh adalah persentase rumah sehat ( $X_3$ ) dan rasio sarana kesehatan ( $X_4$ ) sehingga diperoleh model sebagai berikut.

$$\hat{\lambda} = \exp(3,834 + 2,797 \times 10^{-6} X_1 - 0,006894 X_2 + 0,01225 X_3 + 0,03601 X_4 + 0,06405 X_5)$$

Dari model regresi binomial negatif dapat diketahui bahwa setiap kenaikan satu persen rumah sehat ( $X_3$ ), maka akan meningkatkan rata-rata jumlah kasus DBD di Jawa Timur tahun 2014 sebesar  $\exp(0,01225) = 1,01$  kasus dengan syarat semua variabel lain konstan, selain itu interpretasi yang sama untuk kepadatan penduduk ( $X_1$ ), rasio sarana kesehatan ( $X_4$ ) dan rasio tenaga kesehatan ( $X_5$ ). Jika persentase rumah tangga berperilaku PHBS ( $X_2$ ) meningkat satu persen, maka rata-rata jumlah kasus DBD di Jawa Timur tahun 2014 akan menurun sebesar  $\exp(-0,006894) = 0,993$  kasus dengan syarat semua variabel lain konstan.

#### 4.7 Pemilihan Model Terbaik

Perbandingan model regresi poisson, GPR dan regresi binomial negatif dilakukan untuk mengetahui model yang lebih baik digunakan dalam pemodelan jumlah kasus DBD di Jawa Timur tahun 2014. Kriteria pemilihan model terbaik yang digunakan dengan melihat nilai AIC. Model terbaik adalah model dengan memiliki nilai AIC terkecil.

**Tabel 4.8** Pemilihan Model Terbaik

Model	Variabel Signifikan	Nilai AIC
Regresi Poisson	$X_1 X_2 X_3 X_4 X_5$	4032,6
GPR	$X_3 X_4$	494,2
Regresi Binomial Negatif	$X_3 X_4$	489,21

Berdasarkan Tabel 4.8 nilai AIC terkecil diperoleh pada model regresi binomial negatif. Maka dapat disimpulkan model terbaik untuk jumlah kasus DBD per Kabupaten/Kota di Jawa Timur tahun 2014 adalah model regresi binomial negatif. Hal ini menunjukkan bahwa regresi binomial negatif lebih sesuai dalam memodelkan kasus overdispersi pada regresi poisson

#### **4.8 Pendeteksian Kantong DBD di Jawa Timur Menggunakan *Flexibly Shaped Spatial Scan Statistic***

*Flexibly Spatial Scan Statistic* merupakan salah satu metode statistik yang digunakan untuk mendeteksi *hospot* /kantong DBD dimana dalam penelitian ini yang menjadi *hospot* adalah kabupaten/kota. Dengan adanya informasi tentang *hospot* atau kantong DBD maka dapat diketahui wilayah yang jadi perhatian khusus dan prioritas utama dalam penanggulangan dan pencegahan penyakit DBD.

##### **4.8.1 Pendeteksian Kantong dengan Pemodelan Regresi Binomial Negatif**

Pendeteksian kantong DBD di Provinsi Jawa Timur dengan pemodelan regresi binomial negatif menggunakan data prediksi. Hasil pendeteksian kantong dijelaskan pada Tabel 4.9 dan Gambar 4.7 sebagai berikut.

Jumlah kabupaten/kota untuk setiap kantong DBD di batasi 15 kabupaten/kota yang berbatasan dan jarak terdekat termasuk kabupaten awal. Pengujian tingkat signifikansi dilakukan dengan teknik simulasi monte carlo dimana pengulangan sebanyak 99, 999 dan 9999 menghasilkan kesimpulan yang sama baik dari sisi jumlah kantong DBD yang dideteksi maupun resiko relatifnya. Perbedaan hasil hanya pada keakuratan *p-value*. Hasil yang ditampilkan pada Tabel 4.9 adalah hasil dengan pengulangan sebanyak 9999.

Deteksi kantong DBD di Kabupaten/Kota di Jawa Timur menghasilkan sepuluh kantong seperti Tabel 4.9 yang dibentuk dalam sebuah peta pada Gambar 4.7. Peta yang dibuat

menggambarkan kabupaten/kota yang paling beresiko relatif tinggi yang masuk pada lima kantong paling beresiko. Berikut hasil pendeteksian kantong DBD dengan pemodelan regresi binomial negatif.

**Tabel 4.9** Hasil Deteksi Kantong DBD dengan Pemodelan Regresi Binomial Negatif

Kantong DBD	Jumlah Kabupaten/Kota	Jumlah Kasus	Harapan	Resiko Relatif	P-Value
1	1	816	287,022	2,84	0,0001
2	1	834	337,457	2,47	0,0001
3	1	389	169,413	2,30	0,0001
4	2	1412	686,567	2,06	0,0001
5	1	255	143,937	1,77	0,0001
6	1	319	232,871	1,37	0,0001
7	1	213	167,184	1,27	0,0231
8	1	216	172,768	1,25	0,0452
9	1	176	145,52	1,21	0,3015
10	1	161	142,933	1,13	0,9548

Keterangan :

- Kantong DBD : Kumpulan satu atau beberapa kabupaten/kota yang membentuk *hotspot* atau kantong DBD.
- Jumlah Kabupaten/Kota : Jumlah kabupaten/kota dalam satu kantong
- Jumlah Kasus : Jumlah total kasus DBD dalam satu kantong
- Harapan : Jumlah kasus DBD yang diharapkan dalam satu kantong
- Resiko Relatif : Nilai resiko relatif di dalam kantong DBD dibanding nilai resiko relatif di luar kantong, dimana didapatkan dari jumlah kasus dalam satu kantong dibagi dengan harapan.

Daerah yang berwarna merah merupakan daerah kantong DBD yang paling rawan terkena DBD, daerah itu Kota Surabaya.



Nilai resiko relatif untuk kantong DBD ini sebesar 2,84. Artinya Kota Surabaya memiliki resiko ditemukannya kasus DBD 2,84 kali lebih beresiko dibanding dengan daerah lain yang berada diluar kantong 1.

Daerah yang menjadi kantong bernilai resiko relatif tertinggi nomor dua yaitu Kabupaten Malang. Nilai resiko relatifnya sebesar 2,47. Artinya Kabupaten Malang memiliki resiko ditemukannya kasus DBD 2,47 kali lebih tinggi dibanding dengan daerah lain yang berada diluar kantong 2. Selanjutnya daerah yang termasuk pada kantong ketiga yaitu Kabupaten Ponorogo. Nilai resiko relatif terkena kasus DBD sebesar 2,30, artinya Kabupaten Ponorogo memiliki resiko 2,30 kali lebih tinggi dibanding dengan daerah lain yang berada diluar kantong 3. Daerah yang termasuk dalam kantong keempat yaitu Kabupaten Jember dan kabupaten Bondowoso.

Nilai resiko relatif sebesar 2,06, artinya Kabupaten Jember dan Kabupaten Bondowoso memiliki resiko terkena kasus DBD 2,06 kali lebih tinggi dibanding dengan daerah lain yang berada diluar kantong 4. Daerah untuk kantong kelima yaitu Kabupaten Trenggalek. Nilai resiko relatif sebesar 1,77, artinya Kabupaten Trenggalek memiliki resiko terkena kasus DBD 1,77 kali lebih besar dibanding dengan daerah lain yang berada diluar kantong ini.



## **BAB V**

### **KESIMPULAN DAN SARAN**

#### **5.1 Kesimpulan**

Berdasarkan hasil analisis dan pembahasan yang telah dilakukan, diperoleh kesimpulan terkait hasil penelitian sebagai berikut.

1. Menentukan model terbaik untuk kasus DBD di Jawa Timur tahun 2014 dengan cara melihat nilai AIC terkecil. Nilai AIC untuk model regresi poisson sebesar 4032,6, nilai AIC untuk model GPR sebesar 494,2 dan nilai AIC untuk model regresi binomial negatif sebesar 489,21. Dari melihat nilai AIC didapatkan model terbaik yaitu model regresi binomial negatif.
2. Hasil pemodelan GPR dan BNR sama-sama diperoleh faktor yang signifikan yaitu persentase rumah sehat ( $X_3$ ) dan rasio sarana kesehatan ( $X_4$ ). Artinya, persentase rumah sehat ( $X_3$ ) dan rasio sarana kesehatan ( $X_4$ ) memberikan pengaruh pada kasus DBD di Jawa Timur Tahun 2014.
3. Hasil pendeteksian kantong DBD di Jawa Timur tahun 2014 dengan pemodelan GPR dan BNR menggunakan *Flexibly Shaped Spatial Scan Statistic* memberikan informasi bahwa sama-sama didapatkan 10 kantong DBD dengan masing-masing terdiri dari 1 dan 2 kabupaten/kota. Adapun kantong-kantong yang terbentuk dijelaskan sebagai berikut.
  - a. Kantong yang terbentuk dengan pemodelan GPR berdasarkan nilai resiko relatif tertinggi yaitu Kantong DBD 1 terdiri dari Kota Surabaya dengan nilai resiko relatif sebesar 3,32 kali lebih tinggi dibanding dengan kabupaten/kota yang berada diluar kantong 1. Kantong DBD 2 terdiri dari Kabupaten Malang yang memiliki nilai resiko relatif terkena DBD sebesar 2,40 kali lebih tinggi dibandingkan daerah lain yang berada diluar kantong 2. Kantong DBD 3 yaitu Kabupaten Ponorogo dimana nilai resiko relatif sebesar 2,30 . Kantong DBD 4

yaitu Kabupaten Jember dan Kabupaten Bondowoso dimana nilai resiko relatif sebesar 1,97. Kemudian yang termasuk dalam kantong kelima adalah Kabupaten Trenggalek dengan nilai resiko relatif sebesar 1,78

- b. Kantong yang terbentuk dengan pemodelan regresi binomial negatif berdasarkan nilai resiko relatif tertinggi yaitu kantong DBD 1 yaitu Kota Surabaya nilai resiko relatif untuk kantong DBD ini sebesar 2,84 kali lebih beresiko dibanding dengan daerah lain yang berada diluar kantong 1. Kantong DBD 2 yaitu Kabupaten Malang nilai resiko relatifnya sebesar 2,47 kali lebih tinggi dibanding dengan daerah lain yang berada diluar kantong 2. Kantong DBD 3 yaitu Kabupaten Ponorogo dengan nilai resiko relatif terkena kasus DBD sebesar 2,30. Kantong DBD 4 yaitu Kabupaten Jember dan kabupaten Bondowoso dengan ilai resiko relatif sebesar 2,06. Kantong DBD 5 yaitu Kabupaten Trenggalek nilai resiko relatif sebesar 1,77 kali lebih besar dibanding dengan daerah lain yang berada diluar kantong.

## 5.2 Saran

Dari kesimpulan yang diperoleh, maka terdapat beberapa hal yang dapat disarankan anatar lain sebagai berikut.

1. Untuk penelitian selanjutnya pada penelitian ini masih belum mencakup banyak variabel sehingga berdasarkan hasil analisis masih kurang mampu memberikan informasi yang jelas terhadap faktor-faktor yang mempengaruhi. Oleh karena itu dapat menambahkan beberapa variabel untuk memperoleh hasil yang lebih akurat dan informatif, seperti penambahan variabel persentase rumah tangga bebas jentik.
2. Bagi Dinas Kesehatan Jawa Timur perlu memprioritaskan dan memberikan arahan kepada masyarakat yang berada pada kantong-kantong DBD yang memiliki nilai resiko lebih tinggi.

## LAMPIRAN

**Lampiran 1.** Data Jumlah Kasus DBD dan Faktor-Faktor yang Mempengaruhi di Jawa Timur Tahun 2014

No	Kabupaten / Kota	Y	X1	X2	X3	X4	X5
1	Kab. Pacitan	213	396.00	60.3	70.34	21.13	1.22
2	Kab. Ponorogo	389	631.16	63.5	68.14	23.40	0.92
3	Kab. Trenggalek	255	544.46	27.7	20.58	27.47	1.33
4	Kab. Tulungagung	229	883.15	38.1	64.84	26.74	2.53
5	Kab. Blitar	126	718.02	46.0	8.65	40.74	2.10
6	Kab. Kediri	161	1110.30	55.8	10.53	35.79	1.46
∴	∴	∴	∴	∴	∴	∴	∴
33	Kota Probolinggo	319	4001.92	59.2	78.26	28.35	0.55
34	Kota Pasuruan	123	5285.10	40.8	77.87	21.48	1.03
35	Kota Mojokerto	9	7577.10	53.4	85.29	12.47	0.33
36	Kota Madiun	176	5247.46	62.1	79.75	15.85	0.21
37	Kota Surabaya	816	8683.16	67.1	83.46	30.80	2.09
38	Kota Batu	62	997.58	28.1	45.59	28.37	0.46

**Lampiran 1.** Data Jumlah Kasus DBD dan Faktor-Faktor yang Mempengaruhi di Jawa Timur Tahun 2014 (Lanjutan)

**Keterangan:**

Y : Jumlah kasus DBD menurut Kabupaten/Kota di Provinsi Jawa Timur tahun 2014

X<sub>1</sub> : Kepadatan Penduduk (jiwa/Km<sup>2</sup>)

X<sub>2</sub> : Persentase rumah tangga berperilaku hidup bersih dan sehat

X<sub>3</sub> : Persentase rumah sehat

X<sub>4</sub> : Rasio sarana kesehatan per 1000 penduduk

X<sub>5</sub> : Rasio tenaga medis per 1000 penduduk

## Lampiran 2. Pemeriksaan Multikolinieritas

### Lampiran 2a. Matriks Korelasi 5 Variabel Prediktor

```
MTB > Name m1 "CORR1"
MTB > Correlation 'X1'-'X5' 'CORR1'.
MTB > print m1
```

#### Data Display

Matrix CORR1

1,00000	0,25701	0,37686	-0,32628	-0,23274
0,25701	1,00000	0,51574	-0,19944	0,02750
0,37686	0,51574	1,00000	-0,32806	-0,36282
-0,32628	-0,19944	-0,32806	1,00000	0,23569
-0,23274	0,02750	-0,36282	0,23569	1,00000

### Lampiran 2b. Nilai VIF 5 Variabel Prediktor

```
MTB > name m2 'VIF'
MTB > Invert 'CORR1' 'VIF'.
MTB > print m2
```

#### Data Display

Matrix VIF

1,25267	-0,13616	-0,26624	0,26195	0,13696
-0,13616	1,49194	-0,83753	0,07109	-0,39335
-0,26624	-0,83753	1,80543	0,20452	0,56792
0,26195	0,07109	0,20452	1,20212	-0,15012
0,13696	-0,39335	0,56792	-0,15012	1,28413

**Lampiran 3.** Macro SAS Untuk Regresi Poisson

```

Data poisson;
input y x1 x2 x3 x4 x5;
datalines;
213 396.00 60.3 70.34 21.13 1.22
389 631.16 63.5 68.14 23.40 0.92
255 544.46 27.7 20.58 27.47 1.33
229 883.15 38.1 64.84 26.74 2.53
126 718.02 46.0 8.65 40.74 2.10
161 1110.30 55.8 10.53 35.79 1.46
834 780.38 28.3 33.19 46.80 1.41
  :      :      :      :      :      :
  :      :      :      :      :      :
142 4385.99 52.6 77.08 19.86 0.31
86  4203.35 40.4 37.31 19.56 0.28
160 7683.68 41.4 76.56 36.78 0.45
319 4001.92 59.2 78.26 28.35 0.55
123 5285.10 40.8 77.87 21.48 1.03
9   7577.10 53.4 85.29 12.47 0.33
176 5247.46 62.1 79.75 15.85 0.21
816 8683.16 67.1 83.46 30.80 2.09
62  997.58  28.1 45.59 28.37 0.46
;
run;
title 'poisson model';
proc nlmixed data=poisson start hess tech=newwrap;
parms a0=0 a1=0 a2=0 a3=0 a4=0 a5=0;
eta=a0+a1*x1+a2*x2+a3*x3+a4*x4+a5*x5;
lambda=exp(eta);
model y ~ poisson(lambda);
predict _ll out=LL_1;
run;

```

**Lampiran 4.** Output SAS Regresi Poisson

Fit Statistics					
	-2 Log Likelihood				4018.6
	AIC (smaller is better)				4032.6
	AICC (smaller is better)				4036.3
	BIC (smaller is better)				4044.0
Parameter Estimates					
Parameter	Estimate	Standard Error	DF	t Value	Pr >  t
a0	3.7552	0.05786	38	64.90	<.0001
a1	0.000014	4.824E-6	38	2.98	0.0051
a2	-0.00458	0.000866	38	-5.29	<.0001
a3	0.01177	0.000426	38	27.61	<.0001
a4	0.03537	0.001204	38	29.38	<.0001
a5	0.06248	0.004246	38	14.71	<.0001
_11	1.0000	0	38	Infy	<.0001

**Lampiran 5.** Macro SAS Mendapatkan Nilai *Deviance* dan *Pearson Chi-Square* Regresi Poisson

```

data cari_devians;
input y x1 x2 x3 x4 x5;
datalines;
213 396.00 60.3 70.34 21.13 1.22
389 631.16 63.5 68.14 23.40 0.92
255 544.46 27.7 20.58 27.47 1.33
229 883.15 38.1 64.84 26.74 2.53
126 718.02 46.0 8.65 40.74 2.10
161 1110.30 55.8 10.53 35.79 1.46
834 780.38 28.3 33.19 46.80 1.41
: : : : : :
: : : : : :
142 4385.99 52.6 77.08 19.86 0.31
86 4203.35 40.4 37.31 19.56 0.28
160 7683.68 41.4 76.56 36.78 0.45
319 4001.92 59.2 78.26 28.35 0.55
123 5285.10 40.8 77.87 21.48 1.03
9 7577.10 53.4 85.29 12.47 0.33
176 5247.46 62.1 79.75 15.85 0.21
816 8683.16 67.1 83.46 30.80 2.09
62 997.58 28.1 45.59 28.37 0.46
;
run;
proc genmod data=cari_devians;
model y = x1 x2 x3 x4 X5/dist = poisson
link = log
type1
type3 wald
scale=deviance;
run;

```



**Lampiran 6.** Output SAS Nilai *Deviance* dan *Pearson Chi-Square* Regresi Poisson

poisson model			
The GENMOD Procedure			
Model Information			
Data Set	WORK.CARI_DEVIANS		
Distribution	Poisson		
Link Function	Log		
Dependent Variable	y		
Observations Used	38		
Criteria For Assessing Goodness Of Fit			
Criterion	DF	Value	Value/DF
Deviance	32	3750.9852	117.2183
Scaled Deviance	32	32.0000	1.0000
Pearson Chi-Square	32	4211.6855	131.6152
Scaled Pearson X2	32	35.9303	1.1228
Log Likelihood		362.5563	

## Lampiran 7. Macro SAS Untuk *Generalized Poisson Regression* (GPR)

```

data GPR;
input y x1 x2 x3 x4 X5;
datalines;
213      396.00      60.3      70.34      21.13      1.22
389      631.16      63.5      68.14      23.40      0.92
255      544.46      27.7      20.58      27.47      1.33
229      883.15      38.1      64.84      26.74      2.53
  ⋮        ⋮          ⋮        ⋮          ⋮        ⋮
9         7577.10      53.4      85.29      12.47      0.33
176      5247.46      62.1      79.75      15.85      0.21
816      8683.16      67.1      83.46      30.80      2.09
62        997.58       28.1      45.59      28.37      0.46
;
run;
/*generalized poisson regression model*/
title 'Model GPR';
proc nlmixed data=GPR tech=newwrap maxiter=500 maxfunc=5000;
parms a0=0 a1=0 a2=0 a3=0 a4=0 a5=0 teta=0;
lambda = exp(a0+a1*x1+a2*x2+a3*x3+a4*x4+a5*x5);
ll=y*log(lambda/(1+teta *lambda))+(y-1)*log(1+teta *y)-
lambda*(1+teta*y)/(1+teta*lambda)-lgamma(y+1);
model y~general(ll);
run;

```

**Lampiran 8.** Output SAS *Generalized Poisson Regression*  
(GPR)

Model GPR					
The NL MIXED Procedure					
Fit Statistics					
	-2 Log Likelihood				480.2
	AIC (smaller is better)				494.2
	AICC (smaller is better)				498.0
	BIC (smaller is better)				505.7
Parameter Estimates					
Parameter	Estimate	Standard Error	DF	t Value	Pr >  t
a0	3.8077	0.6966	38	5.47	<.0001
a1	-0.00002	0.000055	38	-0.28	0.7811
a2	-0.00768	0.01034	38	-0.74	0.4623
a3	0.01312	0.007201	38	1.82	0.0763
a4	0.03751	0.01814	38	2.07	0.0455
a5	0.06860	0.06687	38	1.03	0.3114
teta	0.04456	0.006182	38	7.21	<.0001

**Lampiran 9.** Sintax R dan Output Regresi Binomial Negatif

```

> library(MASS)
> binom=glm.nb(y~x1+x2+x3+x4+x5,data=TA)
> summary(binom)

Call:
glm.nb(formula = y ~ x1 + x2 + x3 + x4 + x5, data = TA, init.theta
= 2.447403482,
link = log)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.9320 -0.8486 -0.2700  0.2732  1.9666

Coefficients:
            Estimate      Std. Error  z value    Pr(>|z|)
(Intercept) 3.834e+00    6.079e-01   6.307    2.84e-10 ***
x1          2.797e-06    5.199e-05   0.054    0.95709
x2         -6.894e-03    8.919e-03  -0.773    0.43956
x3          1.225e-02    4.953e-03   2.473    0.01339 *
x4          3.601e-02    1.334e-02   2.700    0.00693 **
x5          6.405e-02    4.729e-02   1.354    0.17558
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Negative Binomial(2.4474) family taken
to be 1)

Null deviance: 54.133  on 37  degrees of freedom
Residual deviance: 40.643  on 32  degrees of freedom
AIC: 489.21

Number of Fisher Scoring iterations: 1

Theta: 2.447
Std. Err.: 0.538

2 x log-likelihood: -475.206

```

**Lampiran 10.** *Output Flexscan* Hasil Deteksi Kantong DBD dengan Model Regresi Binomial Negatif dan Jumlah Replikasi 99

----- FlexScan ver3.1.2 -- purely spatial -----	
<b>MOST LIKELY CLUSTER</b>	
1.Census areas included .:	KotaSurabaya
Maximum distance.....:	0 km (areas: KotaSurabaya to KotaSurabaya)
Number of cases .....	816
(Expected number of cases:	287.022)
Overall relative risk .:	2.84299
Statistic value .....	339.483
Monte Carlo rank .....	1/100
P-value .....	0.01
-----	
<b>SECONDARY CLUSTERS</b>	
2.Census areas included .:	Kab.Jember, Kab.Bondowoso
Maximum distance.....:	63.67 km (areas: Kab.Jember to Kab.Bondowoso)
Number of cases .....	1412
(Expected number of cases:	686.567)
Overall relative risk .:	2.05661
Statistic value .....	324.194
Monte Carlo rank .....	1/100
P-value .....	0.01
3.Census areas included .:	Kab.Malang
Maximum distance.....:	0 km (areas: Kab.Malang to Kab.Malang)
Number of cases .....	834
(Expected number of cases:	337.457)
Overall relative risk .:	2.47143
Statistic value .....	272.093
Monte Carlo rank .....	1/100
P-value .....	0.01

**Lampiran 10.** *Output Flexscan* Hasil Deteksi Kantong DBD dengan Model Regresi Binomial Negatif dan Jumlah Replikasi 99 (Lanjutan)

4.Census areas included .: Kab.Ponorogo  
 Maximum distance.....: 0 km (areas: Kab.Ponorogo to Kab.Ponorogo)  
 Number of cases .....: 389  
 (Expected number of cases: 169.413)  
 Overall relative risk .: 2.29616  
 Statistic value .....: 106.431  
 Monte Carlo rank .....: 1/100  
 P-value .....: 0.01

5.Census areas included .: Kab.Trenggalek  
 Maximum distance.....: 0 km (areas: Kab.Trenggalek to Kab.Trenggalek)  
 Number of cases .....: 255  
 (Expected number of cases: 143.937)  
 Overall relative risk .: 1.77161  
 Statistic value .....: 35.4461  
 Monte Carlo rank .....: 1/100  
 P-value .....: 0.01

6.Census areas included .: KotaProbolinggo  
 Maximum distance.....: 0 km (areas: KotaProbolinggo to KotaProbolinggo)  
 Number of cases .....: 319  
 (Expected number of cases: 232.871)  
 Overall relative risk .: 1.36986  
 Statistic value .....: 14.6732  
 Monte Carlo rank .....: 1/100  
 P-value .....: 0.01

7.Census areas included .: Kab.Pacitan  
 Maximum distance.....: 0 km (areas: Kab.Pacitan to Kab.Pacitan)  
 Number of cases .....: 213  
 (Expected number of cases: 167.184)  
 Overall relative risk .: 1.27404  
 Statistic value .....: 5.88722  
 Monte Carlo rank .....: 3/100  
 P-value .....: 0.03

**Lampiran 10.** *Output Flexscan* Hasil Deteksi Kantong DBD dengan Model Regresi Binomial Negatif dan Jumlah Replikasi 99 (Lanjutan)

8.Census areas included .: Kab.Probolinggo  
Maximum distance.....: 0 km (areas: Kab.Probolinggo to Kab.Probolinggo)  
Number of cases .....: 216  
(Expected number of cases: 172.768)  
Overall relative risk .: 1.25023  
Statistic value .....: 5.10967  
Monte Carlo rank .....: 5/100  
P-value .....: 0.05

9.Census areas included .: KotaMadiun  
Maximum distance.....: 0 km (areas: KotaMadiun to KotaMadiun)  
Number of cases .....: 176  
(Expected number of cases: 145.52)  
Overall relative risk .: 1.20945  
Statistic value .....: 3.04079  
Monte Carlo rank .....: 28/100  
P-value .....: 0.28

10.Census areas included .: Kab.Kediri  
Maximum distance.....: 0 km (areas: Kab.Kediri to Kab.Kediri)  
Number of cases .....: 161  
(Expected number of cases: 142.933)  
Overall relative risk .: 1.12641  
Statistic value .....: 1.11449  
Monte Carlo rank .....: 96/100  
P-value .....: 0.96

**Lampiran 11.** *Output Flexscan* Hasil Deteksi Kantong DBD dengan Model Regresi Binomial Negatif dan Jumlah Replikasi 999

-----  
 FleXScan ver3.1.2 -- purely spatial  
 -----

MOST LIKELY CLUSTER

1.Census areas included .: KotaSurabaya  
 Maximum distance.....: 0 km (areas: KotaSurabaya to KotaSurabaya)  
 Number of cases .....: 816  
 (Expected number of cases: 287.022)  
 Overall relative risk .: 2.84299  
 Statistic value .....: 339.483  
 Monte Carlo rank .....: 1/1000  
 P-value .....: 0.001

-----  
 SECONDARY CLUSTERS

2.Census areas included .: Kab.Jember, Kab.Bondowoso  
 Maximum distance.....: 63.67 km (areas: Kab.Jember to Kab.Bondowoso)  
 Number of cases .....: 1412  
 (Expected number of cases: 686.567)  
 Overall relative risk .: 2.05661  
 Statistic value .....: 324.194  
 Monte Carlo rank .....: 1/1000  
 P-value .....: 0.001

3.Census areas included .: Kab.Malang  
 Maximum distance.....: 0 km (areas: Kab.Malang to Kab.Malang)  
 Number of cases .....: 834  
 (Expected number of cases: 337.457)  
 Overall relative risk .: 2.47143  
 Statistic value .....: 272.093  
 Monte Carlo rank .....: 1/1000  
 P-value .....: 0.001



**Lampiran 11. Output Flexscan Hasil Deteksi Kantong DBD dengan Model Regresi Binomial Negatif dan Jumlah Replikasi 999 (lanjutan)**

4.Census areas included .: Kab.Ponorogo  
 Maximum distance.....: 0 km (areas: Kab.Ponorogo to Kab.Ponorogo)  
 Number of cases .....: 389  
 (Expected number of cases: 169.413)  
 Overall relative risk .: 2.29616  
 Statistic value .....: 106.431  
 Monte Carlo rank .....: 1/1000  
 P-value .....: 0.001

5.Census areas included .: Kab.Trenggalek  
 Maximum distance.....: 0 km (areas: Kab.Trenggalek to Kab.Trenggalek)  
 Number of cases .....: 255  
 (Expected number of cases: 143.937)  
 Overall relative risk .: 1.77161  
 Statistic value .....: 35.4461  
 Monte Carlo rank .....: 1/1000  
 P-value .....: 0.001

6.Census areas included .: KotaProbolinggo  
 Maximum distance.....: 0 km (areas: KotaProbolinggo to KotaProbolinggo)  
 Number of cases .....: 319  
 (Expected number of cases: 232.871)  
 Overall relative risk .: 1.36986  
 Statistic value .....: 14.6732  
 Monte Carlo rank .....: 1/1000  
 P-value .....: 0.001

7.Census areas included .: Kab.Pacitan  
 Maximum distance.....: 0 km (areas: Kab.Pacitan to Kab.Pacitan)  
 Number of cases .....: 213  
 (Expected number of cases: 167.184)  
 Overall relative risk .: 1.27404  
 Statistic value .....: 5.88722  
 Monte Carlo rank .....: 29/1000  
 P-value .....: 0.029

**Lampiran 11.** *Output Flexscan* Hasil Deteksi Kantong DBD dengan Model Regresi Binomial Negatif dan Jumlah Replikasi 999 (lanjutan)

8.Census areas included .: Kab.Probolinggo  
Maximum distance.....: 0 km (areas: Kab.Probolinggo to Kab.Probolinggo)  
Number of cases .....: 216  
(Expected number of cases: 172.768)  
Overall relative risk .: 1.25023  
Statistic value .....: 5.10967  
Monte Carlo rank .....: 54/1000  
P-value .....: 0.054

9.Census areas included .: KotaMadiun  
Maximum distance.....: 0 km (areas: KotaMadiun to KotaMadiun)  
Number of cases .....: 176  
(Expected number of cases: 145.52)  
Overall relative risk .: 1.20945  
Statistic value .....: 3.04079  
Monte Carlo rank .....: 326/1000  
P-value .....: 0.326

10.Census areas included .: Kab.Kediri  
Maximum distance.....: 0 km (areas: Kab.Kediri to Kab.Kediri)  
Number of cases .....: 161  
(Expected number of cases: 142.933)  
Overall relative risk .: 1.12641  
Statistic value .....: 1.11449  
Monte Carlo rank .....: 953/1000  
P-value .....: 0.953

**Lampiran 12.** *Output Flexscan* Hasil Deteksi Kantong DBD dengan Model Regresi Binomial Negatif dan Jumlah Replikasi 9999

----- FlexScan ver3.1.2 -- purely spatial -----	
<b>MOST LIKELY CLUSTER</b>	
1.Census areas included .:	KotaSurabaya
Maximum distance.....:	0 km (areas: KotaSurabaya to KotaSurabaya)
Number of cases .....	816
(Expected number of cases:	287.022)
Overall relative risk .:	2.84299
Statistic value .....	339.483
Monte Carlo rank .....	1/10000
P-value .....	0.0001
-----	
<b>SECONDARY CLUSTERS</b>	
2.Census areas included .:	Kab.Jember, Kab.Bondowoso
Maximum distance.....:	63.67 km (areas: Kab.Jember to Kab.Bondowoso)
Number of cases .....	1412
(Expected number of cases:	686.567)
Overall relative risk .:	2.05661
Statistic value .....	324.194
Monte Carlo rank .....	1/10000
P-value .....	0.0001
3.Census areas included .:	Kab.Malang
Maximum distance.....:	0 km (areas: Kab.Malang to Kab.Malang)
Number of cases .....	834
(Expected number of cases:	337.457)
Overall relative risk .:	2.47143
Statistic value .....	272.093
Monte Carlo rank .....	1/10000
P-value .....	0.0001

**Lampiran 12.** *Output Flexscan* Hasil Deteksi Kantong DBD dengan Model Regresi Binomial Negatif dan Jumlah Replikasi 9999 (Lanjutan)

4.Census areas included .: Kab.Ponorogo  
 Maximum distance.....: 0 km (areas: Kab.Ponorogo to Kab.Ponorogo)  
 Number of cases .....: 389  
 (Expected number of cases: 169.413)  
 Overall relative risk .: 2.29616  
 Statistic value .....: 106.431  
 Monte Carlo rank .....: 1/10000  
 P-value .....: 0.0001

5.Census areas included .: Kab.Trenggalek  
 Maximum distance.....: 0 km (areas: Kab.Trenggalek to Kab.Trenggalek)  
 Number of cases .....: 255  
 (Expected number of cases: 143.937)  
 Overall relative risk .: 1.77161  
 Statistic value .....: 35.4461  
 Monte Carlo rank .....: 1/10000  
 P-value .....: 0.0001

6.Census areas included .: KotaProbolinggo  
 Maximum distance.....: 0 km (areas: KotaProbolinggo to KotaProbolinggo)  
 Number of cases .....: 319  
 (Expected number of cases: 232.871)  
 Overall relative risk .: 1.36986  
 Statistic value .....: 14.6732  
 Monte Carlo rank .....: 1/10000  
 P-value .....: 0.0001

7.Census areas included .: Kab.Pacitan  
 Maximum distance.....: 0 km (areas: Kab.Pacitan to Kab.Pacitan)  
 Number of cases .....: 213  
 (Expected number of cases: 167.184)  
 Overall relative risk .: 1.27404  
 Statistic value .....: 5.88722  
 Monte Carlo rank .....: 231/10000  
 P-value .....: 0.0231

**Lampiran 12.** *Output Flexscan* Hasil Deteksi Kantong DBD dengan Model Regresi Binomial Negatif dan Jumlah Replikasi 9999 (Lanjutan)

8.Census areas included .: Kab.Probolinggo  
Maximum distance.....: 0 km (areas: Kab.Probolinggo to Kab.Probolinggo)  
Number of cases .....: 216  
(Expected number of cases: 172.768)  
Overall relative risk .: 1.25023  
Statistic value .....: 5.10967  
Monte Carlo rank .....: 452/10000  
P-value .....: 0.0452

9.Census areas included .: KotaMadiun  
Maximum distance.....: 0 km (areas: KotaMadiun to KotaMadiun)  
Number of cases .....: 176  
(Expected number of cases: 145.52)  
Overall relative risk .: 1.20945  
Statistic value .....: 3.04079  
Monte Carlo rank .....: 3015/10000  
P-value .....: 0.3015

10.Census areas included .: Kab.Kediri  
Maximum distance.....: 0 km (areas: Kab.Kediri to Kab.Kediri)  
Number of cases .....: 161  
(Expected number of cases: 142.933)  
Overall relative risk .: 1.12641  
Statistic value .....: 1.11449  
Monte Carlo rank .....: 9548/10000  
P-value .....: 0.9548

*(Halaman ini sengaja dikosongkan)*

## DAFTAR LAMPIRAN

	Halaman
<b>Lampiran 1</b>	Data Jumlah Kasus DBD dan Faktor-Faktor yang Mempengaruhi di Jawa Timur Tahun 2014 .....59
<b>Lampiran 2</b>	Pemeriksaan Multikolinieritas ..... 61
<b>Lampiran 3</b>	Macro SAS Untuk Regresi Poisson ..... 62
<b>Lampiran 4</b>	Output SAS Regresi Poisson..... 63
<b>Lampiran 5</b>	Macro SAS Mendapatkan Nilai <i>Deviance</i> dan <i>Pearson Chi-Square</i> Regresi Poisson ..... 64
<b>Lampiran 6</b>	Output SAS Nilai <i>Deviance</i> dan <i>Pearson Chi-Square</i> Regresi Poisson..... 65
<b>Lampiran 7</b>	Macro SAS Untuk <i>Generalized Poisson Regression</i> (GPR) ..... 66
<b>Lampiran 8</b>	Output SAS <i>Generalized Poisson Regression</i> (GPR) ..... 67
<b>Lampiran 9</b>	Syntax R dan Output Regresi Binomial Negatif... 68
<b>Lampiran 10</b>	Output <i>Flexscan</i> Hasil Deteksi kantong DBD dengan Model Regresi Binomial Negatif dan Jumlah Replikasi 99 ..... 69
<b>Lampiran 11</b>	Output <i>Flexscan</i> Hasil Deteksi kantong DBD dengan Model Regresi Binomial Negatif dan Jumlah Replikasi 999 ..... 72
<b>Lampiran 12</b>	Output <i>Flexscan</i> Hasil Deteksi kantong DBD dengan Model Regresi Binomial Negatif dan Jumlah Replikasi 9999 ..... 75

## **BAB V**

### **KESIMPULAN DAN SARAN**

#### **5.1 Kesimpulan**

Berdasarkan hasil analisis dan pembahasan yang telah dilakukan, diperoleh kesimpulan terkait hasil penelitian sebagai berikut.

1. Menentukan model terbaik untuk kasus DBD di Jawa Timur tahun 2014 dengan cara melihat nilai AIC terkecil. Nilai AIC untuk model regresi poisson sebesar 4032,6, nilai AIC untuk model GPR sebesar 494,2 dan nilai AIC untuk model regresi binomial negatif sebesar 489,21. Dari melihat nilai AIC didapatkan model terbaik yaitu model regresi binomial negatif.
2. Hasil pemodelan GPR dan BNR sama-sama diperoleh faktor yang signifikan yaitu persentase rumah sehat ( $X_3$ ) dan rasio sarana kesehatan ( $X_4$ ). Artinya, persentase rumah sehat ( $X_3$ ) dan rasio sarana kesehatan ( $X_4$ ) memberikan pengaruh pada kasus DBD di Jawa Timur Tahun 2014.
3. Hasil pendeteksian kantong DBD di Jawa Timur tahun 2014 dengan pemodelan GPR dan BNR menggunakan *Flexibly Shaped Spatial Scan Statistic* memberikan informasi bahwa sama-sama didapatkan 10 kantong DBD dengan masing-masing terdiri dari 1 dan 2 kabupaten/kota. Adapun kantong-kantong yang terbentuk dijelaskan sebagai berikut.
  - a. Kantong yang terbentuk dengan pemodelan GPR berdasarkan nilai resiko relatif tertinggi yaitu Kantong DBD 1 terdiri dari Kota Surabaya dengan nilai resiko relatif sebesar 3,32 kali lebih tinggi dibanding dengan kabupaten/kota yang berada diluar kantong 1. Kantong DBD 2 terdiri dari Kabupaten Malang yang memiliki nilai resiko relatif terkena DBD sebesar 2,40 kali lebih tinggi dibandingkan daerah lain yang berada diluar kantong 2. Kantong DBD 3 yaitu Kabupaten Ponorogo dimana nilai resiko relatif sebesar 2,30 . Kantong DBD 4



yaitu Kabupaten Jember dan Kabupaten Bondowoso dimana nilai resiko relatif sebesar 1,97. Kemudian yang termasuk dalam kantong kelima adalah Kabupaten Trenggalek dengan nilai resiko relatif sebesar 1,78

- b. Kantong yang terbentuk dengan pemodelan regresi binomial negatif berdasarkan nilai resiko relatif tertinggi yaitu kantong DBD 1 yaitu Kota Surabaya nilai resiko relatif untuk kantong DBD ini sebesar 2,84 kali lebih beresiko dibanding dengan daerah lain yang berada diluar kantong 1. Kantong DBD 2 yaitu Kabupaten Malang nilai resiko relatifnya sebesar 2,47 kali lebih tinggi dibanding dengan daerah lain yang berada diluar kantong 2. Kantong DBD 3 yaitu Kabupaten Ponorogo dengan nilai resiko relatif terkena kasus DBD sebesar 2,30. Kantong DBD 4 yaitu Kabupaten Jember dan kabupaten Bondowoso dengan ilai resiko relatif sebesar 2,06. Kantong DBD 5 yaitu Kabupaten Trenggalek nilai resiko relatif sebesar 1,77 kali lebih besar dibanding dengan daerah lain yang berada diluar kantong.

## 5.2 **Saran**

Dari kesimpulan yang diperoleh, maka terdapat beberapa hal yang dapat disarankan anatar lain sebagai berikut.

1. Untuk penelitian selanjutnya pada penelitian ini masih belum mencakup banyak variabel sehingga berdasarkan hasil analisis masih kurang mampu memberikan informasi yang jelas terhadap faktor-faktor yang mempengaruhi. Oleh karena itu dapat menambahkan beberapa variabel untuk memperoleh hasil yang lebih akurat dan informatif, seperti penambahan variabel persentase rumah tangga bebas jentik.
2. Bagi Dinas Kesehatan Jawa Timur perlu memprioritaskan dan memberikan arahan kepada masyarakat yang berada pada kantong-kantong DBD yang memiliki nilai resiko lebih tinggi.

## DAFTAR PUSTAKA

- Agresti, A. (2002). *Categorical Data Analysis Second Edition*. John Wiley & Sons, New York.
- Bozdogan, H. (2000). *Akaike's Information Criterion and Recent Developments in Information Complexity*, *Mathematical Psychology*, 44, 62-91.
- Cameron, A.C. dan Trivedi, P.K. (1998). *Regression Analysis of Count Data*. Cambridge: Cambridge University Press.
- [Depkes] Departemen Kesehatan. 2015. Kemenkes Terima Laporan Peningkatan Kasus DBD di Jawa Timur. (Diakses tanggal 8 November 2015).
- Draper, N. R dan Smith, H. (1992). *Analisis Regresi Terapan*. Jakarta: PT Gramedia Pustaka Utama.
- Famoye, F., Wulu, J.T. da Singh, K.P. (2004). *On The Generalized Poisson Regression Model with an Application to Accident Data*. *Journal of Data Science* 2 (2004) 287-295.
- Ginanjari, Genis. (2007). *Apa yang Dokter Anda Tidak Katakan Tentang Demam Berdarah*. B first, Yogyakarta.
- Hardin JW dan Hilbe JM. (2007). *Generalized Linier Models and Extensions*. Texas: A Strata Perss Publication.
- Hidayanti, Urifah. (2015). *Pemodelan dan Pemetaan Jumlah Kasus Demam Berdarah Dengue (DBD) di Kota Surabaya dengan Geographically Weighted Negative Binomial Regression (GWNBR) dan Flexibly Shaped Spatial Scan Statistic*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Hocking, R. (1996). *Methods and Application of Linear Models*. New York: John Wiley & Sons.
- Khoshgoftaar, T.M., Gao, K, Szabo, R.M. (2004). Comparing software fault predictions of pure and zero-inflated poisson regression models. *International Journal of System Science* 36 (11), 705-715.
- Kuldorff, M. (1997). *A Spatial Scan Statistic. Communication In Statistics Theory And Method*, 26(6),1481-1496.

- McCullagh P dan Nelder JA. (1983). *Generalized Linier Models*. London: Chapman and Hall.
- Myers, R. H. (1990). *Classical and Modern Regression with Applications, second edition*. Boston: PWS-KENT Publishing Company.
- Rahayu, Dian. (2012). *Pemodelan Pengaruh Iklim Terhadap Angka Kejadian Demam Berdarah Dengue di Surabaya*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Sari, Fefy Dita. (2016). *Identifikasi Wilayah Kantong Penyakit Demam Berdarah Dengue (DBD) dengan Flexibly Shaped Spasial Scan Statistic Melalui Pemodelan Geographically Weighted Negative Binomial Regression (GWNBR) (Studi Kasus Jumlah Kasus DBD di Jawa Timur)*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Tango, T. dan Takahashi, K. (2005) *A Flexibly Shaped Spatial Scan Statistic For Detecting Clusters*. International Journal of Health Geographics, Volume: 4 :11.
- Walpole, R. E. (1995). *Pengantar Statistika Edisi Ketiga*. (B. Sumantri, Trans.) Jakarta: PT. Gramedia Pustaka Utama.
- Zucchini, Walter. (2000). An introduction to model selection. *Journal of Mathematical Psychology* 44: 41-61.



## **BIODATA PENULIS**

Nendy Septi Arniva lahir di Padang pada tanggal 1 September 1992, merupakan anak kedua dari tiga bersaudara dari Syafruddin dan Murni. Penulis telah menempuh pendidikan formal di TK Pertiwi I Kantor Gubernur Padang (1996-1998), SD Pertiwi 2 Padang (1998-2004), SMP Negeri 34 Padang (2004-2007), SMA Negeri 8 Padang (2007-2010), Diploma

III Statistika FMIPA-Universitas Negeri Padang (2010-2014). Pada tahun 2014 penulis melanjutkan studi jenjang Sarjana di Jurusan Statistika FMIPA-ITS. Selama kuliah D-3 Penulis menjadi asisten dosen pada mata kuliah Aplikasi Komputer dan aktif dalam organisasi kemahasiswaan sebagai anggota departement pengabdian masyarakat. Segala saran dan kritik yang membangun serta yang ingin berdiskusi lebih lanjut dengan Penulis mengenai Tugas Akhir ini dapat dikirimkan melalui email [nendyseptiarniva@gmail.com](mailto:nendyseptiarniva@gmail.com)