



TUGAS AKHIR - SS091324

**PERBANDINGAN ANALISIS *GENERALIZED POISSON REGRESSION* (GPR) DAN REGRESI BINOMIAL NEGATIF (STUDI KASUS: PEMODELAN JUMLAH PENDERITA KUSTA DI JAWA TIMUR TAHUN 2012)**

Sari Putri  
NRP 1310 100 065

Pembimbing  
Dra. Wiwiek Setya Winahju, M.S

JURUSAN STATISTIKA  
Fakultas Matematika dan Ilmu Pengetahuan Alam  
Institut Teknologi Sepuluh Nopember  
Surabaya 2014



FINAL PROJECT - SS091324

**COMPARISON OF GENERALIZED POISSON REGRESSION  
(GPR) AND NEGATIVE BINOMIAL REGRESSION (CASE  
STUDY : MODELING THE NUMBER OF LEPROSY IN EAST  
JAVA 2012)**

Sari Putri  
NRP 1310 100 065

Supervisor  
Dra. Wiwiek Setya Winahju, M.S

DEPARTMENT OF STATISTICS  
Faculty Of Mathematics And Science  
Sepuluh Nopember Institute Of Technology  
Surabaya 2014

**Perbandingan Analisis *Generalized Poisson Regression* (GPR) dan Regresi Binomial Negatif  
(Studi Kasus: Pemodelan Jumlah Penderita Kusta  
di Jawa Timur Tahun 2012)**

Nama : Sari Putri  
NRP : 1310100065  
Jurusan : Statistika FMIPA – ITS  
Pembimbing : Dra. Wiwiek Setya Winahju, M.S.

**ABSTRAK**

*Kusta merupakan salah satu penyakit dengan jumlah kasus tertinggi di Jawa Timur. Kusta tidak hanya menimbulkan masalah kesehatan bagi penderita tapi juga menyebabkan masalah sosial karena kebanyakan masyarakat masih menganggap kusta sebagai penyakit kutukan dan tidak dapat disembuhkan. Pemodelan jumlah kasus kusta perlu dilakukan untuk mengetahui faktor apa saja yang berpengaruh terhadap jumlah kusta sehingga jumlah penderita kusta bisa diminimalisir. Salah satu metode yang dapat digunakan untuk pemodelan jumlah kasus kusta adalah regresi poisson. Regresi poisson mensyaratkan bahwa nilai mean sama dengan nilai varians (equidispersion). Namun pada kenyataannya kondisi tersebut jarang bisa dipenuhi. *Generalized poisson regression* (GPR) dan regresi binomial negatif adalah metode yang dapat digunakan untuk mengatasi overdispersion pada regresi poisson. Oleh karena itu dalam penelitian ini dilakukan perbandingan antara metode GPR dan regresi binomial negatif untuk mendapatkan model terbaik yang mampu mengatasi kondisi overdispersion pada regresi poisson. Hasil penelitian menunjukkan bahwa model terbaik adalah model regresi binomial negatif. Hal ini ditunjukkan dari nilai  $AIC_C$ ,  $BIC$  dan  $QAIC_C$  pada model regresi binomial negatif yang lebih kecil daripada model GPR. Faktor-faktor yang berpengaruh terhadap jumlah kasus kusta di Jawa Timur adalah, persentase keberadaan puskesmas, persentase penduduk laki-laki, persentase penduduk usia 15-29 tahun, tingkat kepadatan penduduk, dan interaksi anta-*

*ra antara persentase penduduk umur 10 tahun keatas dengan tingkat pendidikan SD/MI kebawah dan persentase keberadaan puskesmas. Selain itu, jumlah kasus kusta di Jawa Timur juga dipengaruhi oleh daerah endemi kusta sehingga untuk penelitian berikutnya disarankan untuk mempertimbangkan faktor lokasi atau efek spasial.*

**Kata Kunci:** *GPR, Jumlah Kasus Kusta, QAIC<sub>C</sub>, Regresi Binomial Negatif, Regresi Poisson*

## **Comparison Of Generalized Poisson Regression (GPR) and Negative Binomial Regression (Case Study: Modeling The Number of Leprosy in East Java 2012)**

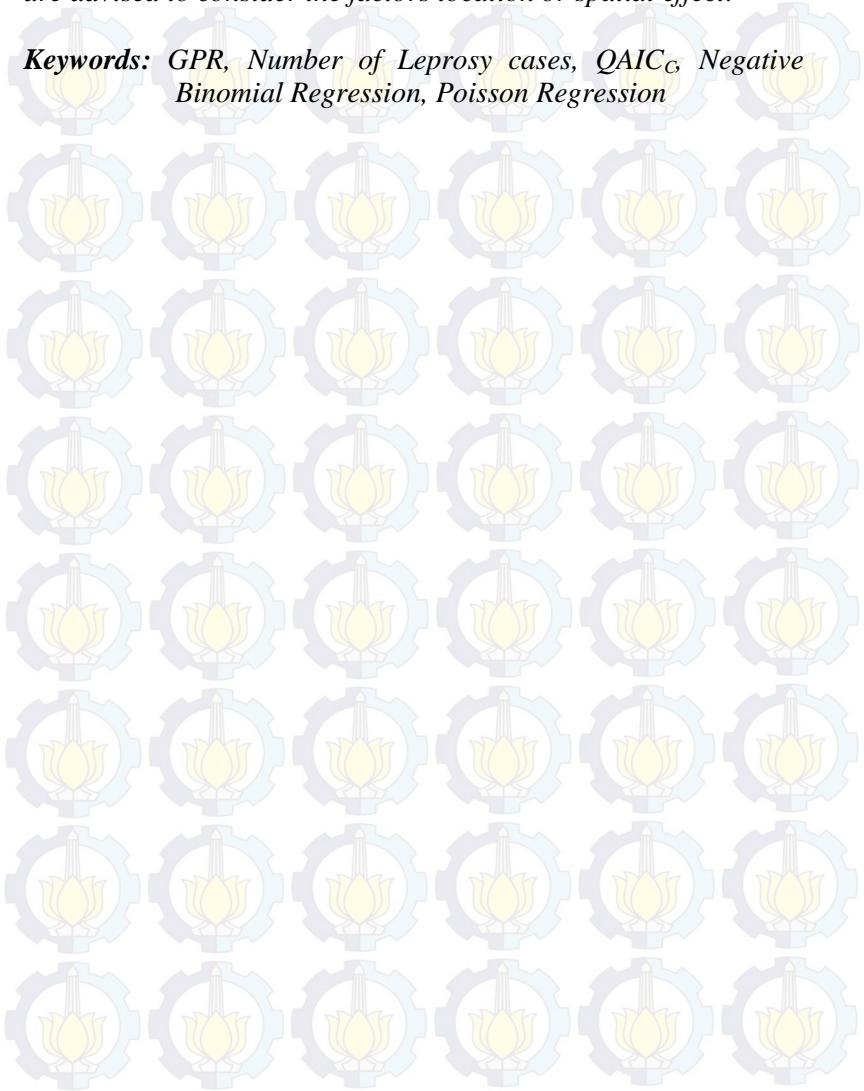
Name : Sari Putri  
NRP : 1310100065  
Department : Statistics FMIPA – ITS  
Supervisor : Dra. Wiwiek Setya Winahju, M.S.

### **ABSTRACT**

*Leprosy is one of the diseases with the highest number in East Java. Leprosy is not only causing any health problem but also cause social problem for the because most people still think of leprosy as a curse and can not be cured. Modelling the number of leprosy case need to be done to know the factors that will affect the number of leprosy so that number can be minimized. One of method that can be used for modeling the number of leprosy is the Poisson regression. Poisson regression requires that the mean value is equal to the variance (equidispersion). But in real cases these conditions rarely can be met. Generalized Poisson regression (GPR) and negative binomial regression is a method that can be used to cope overdispersion in Poisson regression. Therefore, in this study conducted a comparison between GPR and negative binomial regression to get the best model that is able to cope the condition of overdispersion in Poisson regression. The results showed that the best model is the negative binomial regression model. It is shown from the value of  $AIC_C$ ,  $BIC$  and  $QAIC_C$  on negative binomial regression models were smaller than the GPR models. Factors that influence the number of leprosy cases in East Java is the percentage of the existence of health centers, the percentage of the male population, the percentage of people aged 15-29 years, population density, and the interaction between the percentage of population aged 10 years and older with primary education level and percentage existence of health centers . In addition, the number of leprosy case in endemic areas also affect*

*the number of leprosy cases in East Java so for the next research are advised to consider the factors location or spatial effect.*

**Keywords:** *GPR, Number of Leprosy cases, QAIC<sub>G</sub>, Negative Binomial Regression, Poisson Regression*



## LEMBAR PENGESAHAN

### PERBANDINGAN ANALISIS *GENERALIZED POISSON REGRESSION (GPR)* DAN REGRESI BINOMIAL NEGATIF (STUDI KASUS: PEMODELAN JUMLAH PENDERITA KUSTA DI JAWA TIMUR TAHUN 2012)

#### TUGAS AKHIR

Diajukan Untuk Memenuhi Salah Satu Syarat Kelulusan Program Studi S-1 Jurusan Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Institut Teknologi Sepuluh Nopember

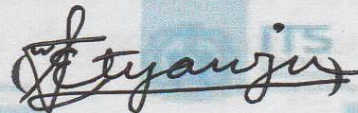
Oleh :

**SARI PUTRI**

NRP 1310 100 065

Disetujui oleh Pembimbing Tugas Akhir

**Dra. Wiwiek Setya Winahju, M.S**  
NIP : 19560424 198303 2 001



Mengetahui  
Ketua Jurusan Statistika FMIPA-ITS



**Dr. Muhammad Mashuri, MT**  
NIP. 19620408 198701 1 001

**SURABAYA, JULI 2014**

## KATA PENGANTAR

Alhamdulillah dan rasa syukur yang sedalam-sedalamnya penulis ucapkan atas nikmat dan kemurahan Allah sehingga penulis dapat menyelesaikan tugas akhir yang berjudul “**Perbandingan Analisis *Generalized Poisson Regression* dan Regresi Binomial Negatif (Studi Kasus: Pemodelan Jumlah Penderita Kusta di Jawa Timur Tahun 2012)** dengan baik.

Penulis menyadari bahwa dalam proses pembuatan buku ini tak lepas dari dukungan dan bantuan dari berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih banyak kepada :

1. Ibu Dra. Wiwiek Setya Winahju, M.S, selaku dosen pembimbing yang telah bersedia meluangkan waktu dan dengan sabar memberikan bimbingan, ilmu, saran dan motivasi kepada penulis.
2. Bapak Ir. Dwi Atmono dan Bapak Dr. Heri Kuswanto, S.Si,M.Si selaku dosen penguji atas kritik dan saran yang membangun dalam penyusunan Tugas Akhir ini.
3. Bapak Dr. Muhammad Mashuri, MT selaku Ketua Jurusan dan Ibu Dra.Lucia Aridinanti, MT selaku Ketua Prodi S1 Statistika ITS yang telah memberikan fasilitas demi kelancaran penyelesaian Tugas Akhir ini.
4. Bapak dan ibu yang senantiasa menyebut nama penulis dalam setiap doa-doanya agar penulis dapat meraih kesuksesan, kebahagiaan dan tetap berada dalam lindungan-Nya. Kakak yang senantiasa mencurahkan perhatian dan kasih sayang kepada penulis serta seluruh keluarga besar penulis. Terima kasih, kalian adalah alasan yang membuat penulis bisa sampai ke titik ini.
5. Keluarga besar FORSIS-ITS, terima kasih telah mengembalikan penulis menuju kebahagiaan yang sesungguhnya, *barakallah...*HIMASTA-ITS, kawah candradimuka yang telah memberikan bekal awal sebagai mahasiswa dan memberikan banyak fasilitas kepada penulis untuk mengembangkan diri.
6. Teman-teman BPH, kabinet dan fungsionaris BEM FMIPA-ITS 12/13, terima kasih telah mengajarkan penulis menjadi



*leader, follower* serta mengajarkan *teamwork* yang baik. Sebuah pengalaman yang mengesankan bisa bergabung dengan kalian. *You're rock guys!*





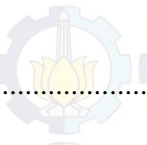










































7. Saudara-saudara  $\Sigma 21$  amazing Statistika 2010, kalian luar biasa serta teman-teman seperjuangan PW110, yang saling mendukung dan memberi semangat dalam proses penyelesaian Tugas Akhir agar bisa bersama-sama menuju Graha.
8. Para penghuni kos *rooftop*, mbak Id, Lala, Hanny, Flashy, dan mbok de Rosna atas kesediaannya berbagi suka dan duka bersama penulis sejak tahun pertama kuliah, dan teman-teman kos funlibin 86 yang telah menyediakan tempat yang nyaman dalam proses penyelesaian tugas akhir ini.

serta semua pihak yang tak bisa penulis sebutkan satu persatu. Selain ucapan terima kasih, penulis juga memohon maaf apabila terdapat kesalahan dalam penulisan tugas akhir ini. Penulis menyadari bahwa penulis masih perlu banyak belajar. Oleh karena itu, penulis mengharapkan kritik dan saran dari pembaca agar kedepannya lebih baik. Akhir kata, semoga buku ini dapat bermanfaat bagi kita semua. Amin.

Surabaya, Juni 2014

Penulis

## DAFTAR GAMBAR

					Halaman
<b>Gambar 3.1</b>	Diagram Alir Analisis Data.....				34
					
					
					
					
					
					
					



# 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> .....	xv
<b>DAFTAR TABEL</b> .....	xvii
<b>DAFTAR LAMPIRAN</b> .....	xix
<b>BAB I PENDAHULUAN</b> .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	4
1.3 Tujuan Penelitian .....	4
1.4 Manfaat Penelitian .....	4
1.5 Batasan Masalah .....	5
<b>BAB II TINJAUAN PUSTAKA</b> .....	7
2.1 Deskripsi Kusta dan Faktor Yang Mempengaruhi .....	7
2.2 Regresi Poisson .....	9
2.2.1 Estimasi Parameter Model Regresi Poisson .....	10
2.2.2 Pengujian Parameter Model Regresi Poisson .....	13
2.2.3 Deteksi Overdispersi Pada Regresi Poisson .....	14
2.3 <i>Generalized Poisson Regression (GPR)</i> .....	15
2.3.1 Estimasi Parameter Model GPR .....	16
2.3.2 Pengujian Parameter Model GPR .....	17
2.4 Regresi Binomial Negatif .....	18
2.4.1 Estimasi Parameter Model Regresi Binomial Negatif .....	19
2.4.2 Pengujian Parameter Model Regresi Binomial Negatif .....	23

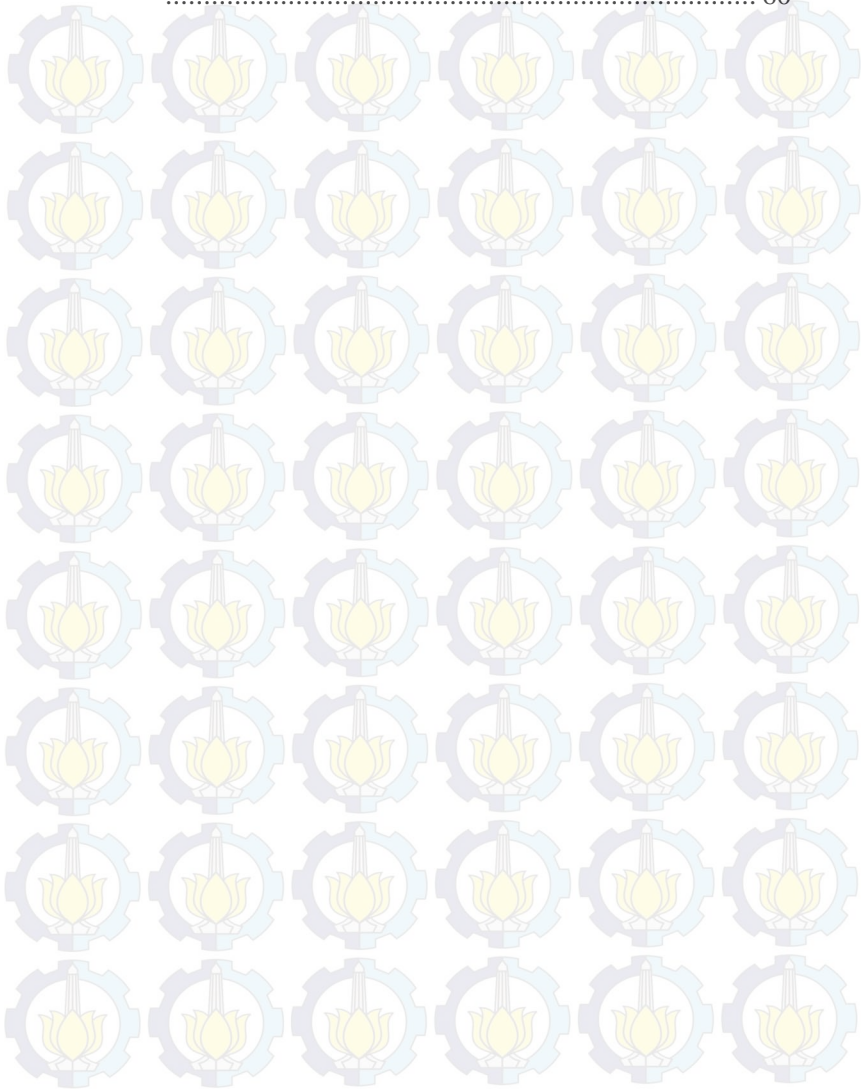
2.5	Kriteria Pemilihan Model Terbaik .....	23
2.6	Multikolinieritas .....	25
2.7	Interpolasi .....	26
<b>BAB III METODOLOGI PENELITIAN .....</b>		<b>29</b>
3.1	Sumber Data .....	29
3.2	Variabel Penelitian .....	29
3.3	Langkah Analisis Data .....	32
<b>BAB IV ANALISIS DAN PEMBAHASAN .....</b>		<b>35</b>
4.1	Karakteristik Jumlah Kasus Kusta di Jawa Timur .....	37
4.2	Pemodelan Jumlah Kasus Kusta di Jawa Timur .....	39
4.2.1	Pemeriksaan Multikolinieritas .....	39
4.2.2	Pemodelan Jumlah Kasus Kusta di Jawa Timur Menggunakan Regresi Poisson .....	40
4.2.3	Pemeriksaan <i>Overdispersion</i> .....	45
4.2.4	Pemodelan Jumlah Kasus Kusta di Jawa Timur Menggunakan GPR .....	46
4.2.5	Pemodelan Jumlah Kasus Kusta di Jawa Timur Menggunakan Regresi Binomial Negatif .....	52
4.3	Pemilihan Model Terbaik .....	59
<b>BAB V PENUTUP .....</b>		<b>63</b>
5.1	Kesimpulan .....	63
5.2	Saran .....	65
<b>DAFTAR PUSTAKA .....</b>		<b>67</b>
<b>LAMPIRAN .....</b>		<b>71</b>

## DAFTAR TABEL

	Halaman
<b>Tabel 2.1</b> Klasifikasi Penyakit Kusta Menurut WHO .....	7
<b>Tabel 3.1</b> Variabel Penelitian .....	29
<b>Tabel 4.1</b> Persentase Penduduk Berdasarkan Golongan Pengeluaran Perkapita Sebulan .....	35
<b>Tabel 4.2</b> Statistika Deskriptif Variabel Penelitian .....	37
<b>Tabel 4.3</b> Nilai VIF dari Sembilan Variabel Prediktor .....	40
<b>Tabel 4.4</b> Estimasi Parameter Model Regresi Poisson .....	41
<b>Tabel 4.5</b> Estimasi Parameter Model Regresi Poisson dengan Interaksi .....	43
<b>Tabel 4.6</b> Estimasi Parameter Model GPR .....	46
<b>Tabel 4.7</b> Kemungkinan Model GPR dengan Interaksi Berdasarkan Nilai QAIC <sub>C</sub> Terkecil .....	48
<b>Tabel 4.8</b> Estimasi Parameter Model GPR dengan Interaksi X <sub>75</sub> .....	49
<b>Tabel 4.9</b> Kemungkinan Model GPR dengan Interaksi Berdasarkan Nilai AIC <sub>C</sub> dan BIC Terkecil .....	51
<b>Tabel 4.10</b> Estimasi Parameter Model Regresi Binomial Negatif .....	53
<b>Tabel 4.11</b> Kemungkinan Model Regresi Binomial Negatif Berdasarkan Nilai QAIC <sub>C</sub> .....	54
<b>Tabel 4.12</b> Estimasi Parameter Model Regresi Binomial Negatif dengan Interaksi X <sub>65</sub> .....	55
<b>Tabel 4.13</b> Kemungkinan Model Regresi Binomial Negatif Berdasarkan Nilai AIC <sub>C</sub> dan BIC .....	58
<b>Tabel 4.14</b> Pemilihan Model Terbaik Berdasarkan Nilai AIC <sub>C</sub> dan BIC .....	60

**Tabel 4.15** Pemilihan Model Terbaik Berdasarkan Nilai QAIC<sub>c</sub>

60



# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Kusta merupakan penyakit menular menahun yang disebabkan oleh kuman kusta (*mycobacterium leprae*) yang menyerang kulit, saraf tepi dan jaringan tubuh lainnya. Bila tidak terdiagnosis dan diobati secara dini, akan menimbulkan kecacatan menetap (Dinas Kesehatan, 2014). Penyakit kusta menjadi hal penting yang harus diperhatikan oleh Indonesia karena jumlahnya masih cukup tinggi. Menurut data WHO (2013), pada tahun 2012 Indonesia merupakan negara dengan jumlah penderita kusta baru tertinggi ketiga di dunia setelah India dan Brazil dengan kontribusi sebesar 18.994 orang (8,7% di dunia). Angka tersebut otomatis menempatkan Indonesia menjadi negara dengan jumlah penderita kusta tertinggi di tingkat negara-negara Asia. Jawa Timur merupakan provinsi yang memberikan kontribusi terbesar terhadap jumlah penderita kusta di Indonesia. Penyebaran penderita kusta merata di seluruh kabupaten/kota di Jawa Timur terutama di pantai utara pulau Jawa dan Madura. Pada tahun 2012 penderita kusta baru di Jawa Timur mencapai 4.807 orang (Dinas Kesehatan, 2014). Artinya, hampir sepertiga jumlah penderita kusta di Indonesia berada di Jawa Timur sehingga jika jumlah penderita kusta di Jawa Timur menurun maka jumlah penderita kusta nasional juga akan turun drastis.

Penyakit kusta tidak hanya menimbulkan dampak kesehatan dan psikologis bagi para penderitanya tetapi juga dampak secara sosial ekonomi. Pandangan masyarakat yang masih menganggap kusta sebagai penyakit menular yang tidak bisa disembuhkan mengakibatkan penderita kusta mengalami diskriminasi dalam kehidupan sosial. Dari hasil penelitian tentang terjadinya diskriminasi pada penderita kusta dan Orang Yang Pernah Mengalami Kusta (OYPMK) didapatkan hasil bahwa penderita kusta dan OYPMK akan ditolak dan dikeluarkan dari pekerjaan, sekolah, tidak bisa menikah bahkan ditolak di pelayanan kesehatan sehingga



mereka akan merasa takut, depresi, tidak percaya diri, malu, merasa diri tidak berharga, dan khawatir akan dikucilkan (Dinas Kesehatan, 2014). Penolakan dari pekerjaan juga mengakibatkan kondisi ekonomi penderita kusta dan OYPMK yang secara umum berasal dari kalangan ekonomi menengah bawah menjadi semakin parah (Dinas Kesehatan, 2014).

Analisis mengenai faktor-faktor yang mempengaruhi penyebaran kusta telah dilakukan di bidang kedokteran, epidemiologi, dan kesehatan masyarakat. Susanto (2006) dalam penelitiannya yang berjudul faktor-faktor yang berhubungan dengan tingkat kecacatan penderita kusta menyatakan bahwa ada hubungan antara umur, pendidikan, tipe kusta, reaksi, pengetahuan, pengobatan, diagnosis, dan perawatan diri dengan tingkat kecacatan penderita kusta. Penelitian tersebut melakukan analisis terhadap faktor internal dan eksternal yang diduga berpengaruh terhadap tingkat kecacatan penderita kusta di wilayah kecamatan Tirto kabupaten Pekalongan. Penelitian lain juga pernah dilakukan Norlatifah *et al* (2010) yang meneliti terkait hubungan kondisi fisik rumah, sarana air bersih dan karakteristik masyarakat dengan kejadian kusta di kabupaten Tapin Kalimantan Selatan. Kesimpulan yang didapat dari penelitian tersebut adalah kondisi fisik rumah, interaksi masyarakat dan tingkat pendidikan adalah faktor dominan yang mempengaruhi penyebaran kusta.

Jumlah penderita kusta merupakan salah satu contoh data diskrit (*count*). Metode yang dapat digunakan untuk memodelkan data diskrit (*count*) adalah regresi poisson. Pada regresi poisson terdapat asumsi bahwa rata-rata (*mean*) dan varians variabel respon harus sama (kondisi *equidispersion*). Namun pada kenyataannya kondisi tersebut jarang terpenuhi. Penelitian mengenai pemodelan data *count* telah dilakukan oleh Winkelmann dan Zimmermann (1994). Pada penelitian tersebut dilakukan pemodelan data demografi seperti usia dan jenis kelamin terhadap jumlah kelahiran, jumlah pernikahan dan perceraian dengan menggunakan regresi poisson. Kesimpulan dari penelitian tersebut adalah model yang dihasilkan mengalami overdispersi. Keadaan tersebut berpe-

ngaruh terhadap nilai penduga bagi kesalahan baku yang lebih kecil (*underestimate*) sehingga dapat mengakibatkan kesalahan penarikan kesimpulan bagi parameternya. Metode yang bisa digunakan untuk mengatasi masalah overdispersi adalah *Generalized Poisson Regression* (GPR) dan model regresi binomial negatif. Kedua metode tersebut tidak mensyaratkan nilai varian dan *mean* sama. Metode GPR pernah digunakan dalam penelitian Famoye *et al* (2004). Pemodelan dilakukan pada data jumlah kecelakaan terhadap faktor-faktor yang mempengaruhinya yaitu faktor demografi, kebiasaan mengemudi dan penggunaan obat-obatan pada pengemudi usia lanjut. Kesimpulan dari penelitian tersebut adalah model GPR mampu mengatasi overdispersi pada model regresi poisson dan mampu memberikan hasil yang lebih baik. Hal ini dapat diketahui dari nilai *standard error* pada pemodelan dengan GPR tidak mengalami *underestimate*. Selanjutnya, Chamidah (2008) menerapkan regresi binomial negatif untuk mengatasi overdispersi regresi poisson pada kasus kematian ibu melahirkan di Jawa Timur. Penelitian tersebut juga menyimpulkan bahwa pemodelan regresi binomial negatif dapat mengatasi overdispersi model regresi poisson yang ditunjukkan dengan penurunan nilai deviansi.

Berdasarkan uraian diatas maka dalam tugas akhir ini akan dilakukan penelitian jumlah kasus kusta di Jawa Timur pada tahun 2012 dengan pendekatan regresi poisson, *Generalized Poisson Regression* dan regresi binomial negatif serta membandingkan hasil antara *Generalized Poisson Regression* dan regresi binomial negatif untuk menentukan model terbaik yang mampu mengatasi overdispersi. Penelitian ini diharapkan dapat dijadikan referensi tambahan bagi pengembangan keilmuan. Selain itu, diharapkan dapat memberikan tambahan informasi kepada pihak terkait tentang faktor-faktor yang secara signifikan mempengaruhi terjadinya kasus kusta sehingga dapat diambil kebijakan yang tepat sebagai upaya meminimalisir jumlah kejadian kusta di Jawa Timur.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang, permasalahan yang diselesaikan pada penelitian ini adalah sebagai berikut.

1. Bagaimana karakteristik jumlah kasus kusta di Jawa Timur pada tahun 2012?
2. Bagaimana pemodelan jumlah kasus kusta di Jawa Timur tahun 2012 dan faktor-faktor yang diduga mempengaruhi dengan metode regresi poisson, *Generalized Poisson Regression* (GPR) dan regresi binomial negatif ?
3. Bagaimana perbandingan antara model *Generalized Poisson Regression* (GPR) dan regresi binomial negatif dalam memodelkan faktor-faktor yang mempengaruhi jumlah kasus kusta di Jawa Timur pada tahun 2012?

## 1.3 Tujuan Penelitian

Tujuan penelitian yang dicapai pada penelitian ini adalah

1. Mendeskripsikan jumlah kasus kusta di Jawa Timur pada tahun 2012.
2. Mendapatkan model regresi poisson, *Generalized Poisson Regression* (GPR) dan regresi binomial negatif terbaik untuk jumlah kasus kusta di Jawa Timur pada tahun 2012 dan faktor-faktor yang diduga mempengaruhi.
3. Membandingkan analisis *Generalized Poisson Regression* (GPR) dan regresi binomial negatif untuk jumlah kasus kusta di Jawa Timur pada tahun 2012.

## 1.4 Manfaat Penelitian

Berdasarkan tujuan penelitian maka manfaat dari penelitian ini adalah:

1. Bagi peneliti dapat menambah wawasan tentang penyakit kusta dan faktor-faktor yang mempengaruhinya. Selain itu, peneliti dapat mengetahui metode terbaik yang dapat digunakan untuk mengatasi overdispersi pada pemodelan data *count* dengan regresi poisson serta aplikasinya secara langsung.

2. Bagi dinas kesehatan, penelitian ini dapat memberikan tambahan informasi mengenai penyakit kusta sehingga dapat dijadikan sebagai acuan dan bahan pertimbangan dalam pengambilan kebijakan untuk meminimalisir jumlah penderita kusta.

### **1.5 Batasan Masalah**

Pada penelitian ini pemodelan dilakukan pada jumlah kasus kusta pada tahun 2012 dengan unit penelitian 29 kabupaten dan 8 kota di Jawa Timur. Kriteria pemilihan model terbaik yang digunakan adalah AIC<sub>c</sub> (*Corrected Akaike's Information Criterion*), BIC (*Bayesian Information Criterion*) dan QAIC<sub>c</sub> (*Quasi-Corrected Akai-ke's Information Criterion*).



*(Halaman ini sengaja dikosongkan)*

## BAB II

### TINJAUAN PUSTAKA

#### 2.1 Deskripsi Kusta dan Faktor yang Mempengaruhi

Istilah kusta berasal dari bahasa sansekerta, yakni kushtha yang berarti kumpulan gejala-gejala kulit secara umum. Kusta adalah penyakit menular menahun yang disebabkan oleh kuman kusta (*mycobacterium leprae*) yang menyerang saraf tepi, kulit, dan jaringan tubuh lainnya yang dalam jangka panjang dapat mengakibatkan sebagian anggota tubuh penderita tidak berfungsi sebagaimana mestinya. Kusta terdiri dari dua tipe yaitu kusta tipe *pausibaciler* (PB) dan kusta tipe *multibaciler* (MB). Tipe kusta yang sangat mudah menular dan banyak diderita oleh penduduk Indonesia adalah kusta tipe *multibaciler*. Klasifikasi kusta menurut WHO ditampilkan pada Tabel 2.1 berikut.

**Tabel 2.1** Klasifikasi Penyakit Kusta Menurut WHO

Kelainan kulit dan hasil pemeriksaan	Tipe <i>pausibaciler</i>	Tipe <i>multibaciler</i>
a. Ukuran	a. Kecil dan besar	a. Kecil
b. Distribusi	b. Unilateral atau bilateral simetris	b. Bilateral simetris
c. Konsistensi	c. Kering dan kasar	c. Halus, berkilat
d. Batas	d. Tegas	d. Kurang tegas
e. Kehilangan rasa pada bercak	e. Selalu ada dan jelas	e. Biasanya tidak jelas, jika ada, terjadi pada yang sudah lanjut
f. Kehilangan kemampuan berkeringat, rambut rontok pada bercak	f. Selalu ada dan jelas	f. Biasanya tidak jelas, jika ada, terjadi pada yang sudah lanjut.

Tanda-tanda penyakit kusta bermacam-macam, tergantung dari tingkat atau tipe dari penyakit kusta. Secara umum tanda-tanda seseorang menderita kusta antara lain, kulit mengalami bercak putih, merah, ada bagian tubuh yang tidak berkeringat, rasa kesemutan pada anggota badan atau bagian raut muka dan mati rasa karena kerusakan syaraf tepi. Penyebab kusta adalah kuman *Mycobacterium leprae*. Microbacterium ini adalah kuman aerob, tidak membentuk spora, berbentuk batang, dikelilingi oleh membran sel lilin yang merupakan ciri dari spesies *Mycobacterium*, berukuran panjang 1–8 micro, lebar 0,2–0,5 micro biasanya berkelompok dan ada yang tersebar satu-satu, hidup dalam sel dan bersifat tahan asam (BTA), tidak mudah diwarnai namun jika diwarnai akan tahan terhadap dekolorisasi oleh asam atau alkohol. Cara penularan kusta belum diketahui secara pasti dengan jelas namun beberapa ahli berpendapat bahwa penyakit kusta ditularkan melalui saluran pernafasan dan kulit.

Kelompok yang beresiko tinggi terkena kusta adalah kelompok yang tinggal di daerah endemi dengan kondisi buruk seperti tempat tidur yang tidak memadai, air yang tidak bersih, asupan gizi yang buruk dan adanya penyertaan penyakit lain seperti HIV yang dapat menekan sistem imun. Daerah endemi merupakan suatu wilayah tertentu dimana suatu penyakit berasal, menyebar dan sering atau terus-menerus ada dalam wilayah tersebut. Menurut Kerr-pontes (2006) insiden kusta juga dipengaruhi oleh faktor sosial ekonomi, faktor lingkungan, faktor demografi dan faktor perilaku. Sementara itu, Bernadus (2010) menyatakan bahwa faktor-faktor yang menyebabkan timbulnya kusta adalah sebagai berikut.

a. Faktor kuman kusta

Hasil penelitian membuktikan bahwa kuman kusta yang masih utuh (solid) bentuknya, lebih besar kemungkinan menyebabkan penularan daripada kuman yang tidak utuh lagi. Kuman kusta dapat hidup di luar tubuh manusia antara 1 sampai 9 hari tergantung suhu atau cuaca dan diketahui hanya kuman kusta yang utuh saja dapat menimbulkan penularan.

b. Faktor imunitas

Sebagian manusia kebal terhadap penyakit kusta (95%). Berdasarkan hasil penelitian menunjukkan bahwa dari 100 orang yang terpapar, 95 orang tidak menjadi sakit, 3 orang sembuh tanpa obat dan 2 orang menjadi sakit. Hal ini belum memper-timbangkan pengaruh pengobatan.

c. Keadaan lingkungan

Keadaan rumah yang berjejal biasanya berkaitan dengan kemiskinan yang merupakan faktor penyebab tingginya jumlah kusta. Lingkungan yang tidak bersih dengan sanitasi buruk juga merupakan sarana penyebaran kuman *mycobacterium leprae*.

d. Faktor usia

Penyakit kusta jarang sekali ditemukan pada bayi. Angka kejadian penyakit kusta meningkat sesuai umur dengan puncak kejadian pada umur 10-20 tahun dan kemudian menurun.

e. Faktor jenis kelamin

Insiden maupun prevalensi pada laki-laki lebih banyak daripada wanita kecuali di Afrika dimana wanita lebih banyak menderita kusta daripada laki-laki.

## 2.2 Regresi Poisson

Regresi poisson merupakan salah satu komponen *link function* pada *generalized linier model*. Regresi poisson adalah salah satu regresi yang menggambarkan hubungan antara variabel respon (Y) dan variabel prediktor dengan mengasumsikan variabel Y berdistribusi poisson. Fungsi peluang dari distribusi poisson dinyatakan dalam persamaan (2.1) sebagai berikut (Myers, 1990).

$$f(y; \mu) = \frac{e^{-\mu} \mu^y}{y!}, \quad y = 0, 1, 2, \dots \quad (2.1)$$

dengan  $\mu$  adalah rata-rata variabel random Y yang berdistribusi poisson dengan nilai rata-rata dan varians lebih dari nol. Distribusi poisson merupakan salah satu distribusi yang paling sederhana



dalam pemodelan data *count*. Menurut Cameron dan Trivedi (1998), beberapa karakteristik percobaan yang berdistribusi poisson adalah sebagai berikut.

1. Kejadian yang terjadi pada jumlah anggota populasi yang besar dengan probabilitas yang kecil (*rare event*)
2. Kejadian yang termasuk dalam *counting process* atau termasuk dalam lingkungan proses stokastik
3. Bergantung pada interval waktu tertentu
4. Perulangan dari kejadian yang mengikuti sebaran distribusi binomial.

*Generalized linier model* (GLM) sendiri merupakan perluasan dari model regresi umum untuk respon berdistribusi dalam keluarga eksponensial dan model yang dimiliki merupakan fungsi dari nilai harapannya. *Generalized linier model* (GLM) terdiri dari tiga komponen yaitu komponen random, komponen sistematis dan *link function*. *Link function* yang digunakan dalam model regresi poisson adalah  $\ln$  yaitu  $\ln(\mu)$ , sehingga model regresi poisson untuk  $y_i$  dengan  $i=1,2,\dots,n$  dimana  $n$  menyatakan banyaknya pengamatan dengan  $k$  variabel prediktor dapat dituliskan pada persamaan (2.2) berikut.

$$y_i \sim \text{poisson}(\mu_i)$$

$$E(y_i|x_i) = \mu_i = \exp(x_i^T \beta)$$

$$\mu_i = \exp(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki})$$

maka,

$$\ln(\mu_i) = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} \quad (2.2)$$

### 2.2.1 Estimasi Parameter Model Regresi Poisson

Metode estimasi parameter model regresi poisson yang digunakan adalah *maximum likelihood estimation* yaitu dengan cara memaksimalkan fungsi likelihood (Myers, 1990). Bentuk logaritma natural dari fungsi likelihood regresi poisson adalah sebagai berikut.

$$\ln L(\beta) = \ln \left( \prod_{i=1}^n \frac{\exp(-\mu_i) \mu_i^{y_i}}{y_i!} \right)$$

$$\begin{aligned}
&= \sum_{i=1}^n \ln \left( \frac{\exp(-\mu_i) \mu_i^{y_i}}{y_i!} \right) \\
&= \sum_{i=1}^n (\ln(e^{-\mu_i}) + \ln(\mu_i^{y_i}) - \ln(y_i!)) \\
&= \sum_{i=1}^n (-\mu_i + y_i \ln(\mu_i) - \ln(y_i!)) \\
\ln L(\boldsymbol{\beta}) &= \sum_{i=1}^n \left( -e^{(x_i^T \boldsymbol{\beta})} + y_i \ln(e^{(x_i^T \boldsymbol{\beta})}) - \ln(y_i!) \right) \quad (2.3) \\
\ln L(\boldsymbol{\beta}) &= - \sum_{i=1}^n e^{(x_i^T \boldsymbol{\beta})} + \sum_{i=1}^n y_i x_i^T \boldsymbol{\beta} - \sum_{i=1}^n \ln(y_i!)
\end{aligned}$$

Estimasi parameter model regresi poisson disimbolkan dengan  $\hat{\boldsymbol{\beta}}$ . Nilai  $\hat{\boldsymbol{\beta}}$  diperoleh dari turunan pertama persamaan (2.3) terhadap  $\boldsymbol{\beta}^T$  sebagai berikut.

$$\frac{\partial \ln L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}^T} = - \sum_{i=1}^n \mathbf{x}_i \exp(\mathbf{x}_i^T \boldsymbol{\beta}) + \sum_{i=1}^n y_i \mathbf{x}_i \quad (2.4)$$

Kemudian persamaan (2.4) disamadengankan nol untuk mendapatkan penyelesaiannya. Tetapi penyelesaian tersebut seringkali tidak menghasilkan solusi karena permasalahannya berbentuk nonlinier sehingga digunakan metode iterasi *Newton Raphson* dengan algoritma sebagai berikut (Cameron & Trivedi, 1998).

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

$$\hat{\boldsymbol{\beta}}_{(0)} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y} \quad (2.5)$$

dengan

$$X = \begin{bmatrix} 1 & x_{11} & \dots & x_{k1} \\ 1 & x_{12} & \dots & x_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & \dots & x_{kn} \end{bmatrix} \text{ dan } Y = [Y_1 \ Y_2 \ \dots \ Y_n]^T$$

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

$$g(\beta) = - \sum_{i=1}^n x_i \mu_i + \sum_{i=1}^n x_i y_i = \sum_{i=1}^n x_i (y_i - \mu_i)$$

$$g^T(\beta^{(m)})_{(k+1) \times 1} = \left[ \frac{\partial \ln L(\beta)}{\partial \beta_0} \quad \frac{\partial \ln L(\beta)}{\partial \beta_1} \quad \dots \quad \frac{\partial \ln L(\beta)}{\partial \beta_k} \right]_{\beta = \beta_m}$$

3. Membentuk matrik hessian H

$$H(\beta) = - \sum_{i=1}^n x_i x_i^T \mu_i = - \sum_{i=1}^n \mu_i x_i x_i^T$$

$$H(\beta^{(m)})_{(k+1) \times (k+1)} = \Lambda$$

$$\Lambda = \begin{bmatrix} \frac{\partial^2 \ln L(\beta)}{\partial \beta_0^2} & \frac{\partial^2 \ln L(\beta)}{\partial \beta_0 \partial \beta_1} & \dots & \frac{\partial^2 \ln L(\beta)}{\partial \beta_0 \partial \beta_k} \\ \frac{\partial^2 \ln L(\beta)}{\partial \beta_0 \partial \beta_1} & \frac{\partial^2 \ln L(\beta)}{\partial \beta_1^2} & \dots & \frac{\partial^2 \ln L(\beta)}{\partial \beta_1 \partial \beta_k} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \ln L(\beta)}{\partial \beta_0 \partial \beta_k} & \frac{\partial^2 \ln L(\beta)}{\partial \beta_1 \partial \beta_k} & \dots & \frac{\partial^2 \ln L(\beta)}{\partial \beta_k^2} \end{bmatrix}_{\beta = \beta_m}$$

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

4. Memasukan nilai  $\hat{\beta}_{(0)}$  kedalam elemen-elemen vektor G dan matrik H, sehingga diperoleh  $g(\hat{\beta}_{(0)})$  dan  $H(\hat{\beta}_{(0)})$
5. Melakukan iterasi pada persamaan

$$\beta_{(m+1)} = \beta_{(m)} - H^{-1}(\beta_{(m)})g(\beta_{(m)})$$

Dimulai dari  $m=0$ , dengan  $\beta_{(m)}$  adalah kumpulan penaksir parameter yang konvergen pada iterasi ke- $m$

6. Penaksir parameter dikatakan konvergen jika nilai  $\|\beta_{(m+1)} - \beta_{(m)}\| \leq \varepsilon$ , dimana  $\varepsilon$  merupakan suatu bilangan yang nilainya sangat kecil sekali sehingga hampir tidak ada perbedaan antara  $\beta_{(m+1)}$  dan  $\beta_{(m)}$ . Jika belum didapatkan penaksir parameter yang konvergen, maka dilakukan pengulangan langkah 5 dengan  $m^* = m + 1$ .

### 2.2.2 Pengujian Parameter Model Regresi Poisson

Pengujian parameter model regresi poisson dilakukan secara serentak dan secara parsial. Pengujian secara serentak dilakukan untuk mengetahui pengaruh variabel secara keseluruhan terhadap model dengan tingkat signifikansi tertentu sedangkan pengujian parameter secara parsial dilakukan untuk mengetahui efek atau pengaruh setiap variabel terhadap model. Salah satu metode yang digunakan untuk menentukan statistik uji dalam pengujian parameter model regresi poisson adalah *Maximum Likelihood Ratio Test* (MLRT) sehingga sebelum menentukan statistik uji, terlebih dahulu ditentukan dua buah fungsi likelihood yang berhubungan dengan model regresi yang diperoleh. Fungsi likelihood tersebut adalah  $L(\hat{\Omega})$  yaitu nilai maksimum likelihood untuk model lengkap dengan melibatkan variabel prediktor dan  $L(\hat{\omega})$  yaitu nilai maksimum likelihood untuk model sederhana tanpa melibatkan variabel prediktor. Hipotesis yang digunakan adalah sebagai berikut.

$$H_0: \beta_1 = \beta_2 = \beta_3 = \dots = \beta_k = 0$$

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

Statistik uji yang digunakan adalah sebagai berikut.

$$D(\hat{\beta}) = -2 \ln \left[ \frac{L(\hat{\omega})}{L(\hat{\Omega})} \right] = 2 \left[ \ln (L(\hat{\Omega})) - \ln(L(\hat{\omega})) \right] \quad (2.6)$$

Keputusan yang diambil akan tolak  $H_0$  jika  $D(\hat{\beta}) > \chi_{v,\alpha}^2$  dengan  $v$  adalah banyaknya parameter model dibawah populasi dikurangi dengan banyaknya parameter model dibawah  $H_0$ . Tolak  $H_0$  berarti ada salah satu variabel yang berpengaruh signifikan terhadap model sehingga dilanjutkan dengan 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)

Statistik uji yang digunakan adalah

$$z = \frac{\hat{\beta}_j}{se(\hat{\beta}_j)} \quad (2.7)$$

dengan  $se(\hat{\beta}_j)$  adalah tingkat kesalahan  $\beta_j$ . Keputusan yang diambil adalah tolak  $H_0$  jika  $|z_{hit}| > z_{\frac{\alpha}{2}}$  dengan  $\alpha$  adalah tingkat signifikansi.

### 2.2.3 Deteksi Overdispersi pada Regresi Poisson

Metode regresi poisson mensyaratkan kondisi dimana nilai *mean* dan varians variabel respon bernilai sama atau kondisi *equi-dispersion* (Khoshgootaar *et al*, 2004). Namun, dalam penerapannya kondisi tersebut sulit dipenuhi oleh data. Dengan kata lain, sering terjadi kondisi *overdispersion* dalam data yang dimodelkan. Kondisi *Overdispersion* dapat ditulis  $Var(Y) > E(Y)$ . Sebaliknya, data yang memiliki varians variabel respon lebih kecil dari *mean* variabel respon disebut dengan kondisi *underdispersion*. Menurut Cameron dan Trivedi (1998) kondisi *overdispersion* dapat terjadi karena adanya sumber keragaman yang tidak teramat. *Overdispersion* dapat pula terjadi karena adanya pengamatan *missing* pada variabel prediktor, adanya pencilan pada data, perlunya interaksi dalam model, variabel prediktor perlu ditransformasi atau kesalahan spesifik *link function*. Kondisi *overdispersion* dapat menyebabkan nilai penaksir simpangan baku koefisien regresi  $\hat{\beta}_j$  terlalu kecil, sehingga nilai statistik uji untuk pengujian hipotesis pengaruh prediktor menjadi lebih besar dan menjadikan predi-

ktor terlalu mudah dianggap berpengaruh (Cameron & Trivedi, 1998).

Deteksi overdispersion dapat dilakukan dengan menyatakan distribusi poisson termasuk dalam keluarga distribusi katz dengan parameter  $\lambda$  dan rasio antara varians terhadap *mean* adalah 1 sehingga  $\sigma^2 = \mu + c\mu$  (Ghahfarokhi, Iravani dan Sepehri, 2008). Jadi hipotesis untuk menguji adanya *overdispersion* pada regresi poisson adalah

$$H_0: c = 0$$

$$H_1: c > 0$$

Statistik uji yang digunakan adalah sebagai berikut

$$L = \sum_{i=1}^n \frac{(y_i - \hat{\mu}_i)^2 - y_i}{2\hat{\mu}_i} \quad (2.8)$$

Dean (1992) menyatakan bahwa  $\sqrt{2}L \sim N(0,1)$ , jadi keputusan akan menolak  $H_0$  jika  $\sqrt{2}L > Z_{\alpha}$ . Artinya ada indikasi terjadi *overdispersion* pada model regresi poisson

### 2.3 Generalized Poisson Regression (GPR)

*Generalized poisson regression* merupakan pengembangan dari regresi poisson yang digunakan untuk mengatasi kondisi *over/underdispersion* sehingga model *Generalized Poisson Regression* (GPR) hampir sama dengan regresi poisson tetapi model GPR mengasumsikan bahwa komponen randomnya berdistribusi *generalized poisson*. Pada model GPR selain terdapat parameter  $\mu$  juga terdapat parameter  $\theta$  sebagai parameter dispersi. Menurut Famoye *et al* (2004) fungsi distribusi *Generalized poisson* adalah

$$f(y; \mu; \theta) = \left( \frac{\mu}{1 + \theta\mu} \right)^y \frac{(1 + \theta y)^{y-1}}{y!} \exp\left( \frac{-\mu(1 + \theta y)}{1 + \theta\mu} \right), \quad (2.9)$$

$$y = 0, 1, 2, \dots$$

*Mean* model GPR adalah  $E(y) = \mu$  dan variansnya adalah  $var(y) = \mu(1 + \theta\mu)^2$ .

Jika  $\theta = 0$  maka model GPR akan menjadi model regresi poisson biasa, jika  $\theta > 0$  maka model GPR merepresentasikan data *count* yang mengandung *overdispersion* dan jika  $\theta < 0$  maka model GPR merepresentasikan data *count* yang mengandung *underdispersion*. Model GPR ditunjukkan pada persamaan (2.10).

$$\begin{aligned} \ln(\mu_i) &= \mathbf{x}_i^T \boldsymbol{\beta} = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} \\ \mu_i &= \exp(\mathbf{x}_i^T \boldsymbol{\beta}), i = 1, 2, \dots, n \end{aligned} \quad (2.10)$$

### 2.3.1 Estimasi Parameter Model GPR

Estimasi parameter model *Generalized Poisson Regression* dilakukan dengan metode *maximum likelihood estimation*. Fungsi likelihood untuk model GPR adalah sebagai berikut.

$$\begin{aligned} L(\boldsymbol{\beta}, \theta) &= \prod_{i=1}^n \left\{ \left( \frac{\mu_i}{1 + \theta \mu_i} \right)^{y_i} \frac{(1 + \theta y_i)^{y_i - 1}}{y_i!} \exp \left( \frac{-\mu_i (1 + \theta y_i)}{1 + \theta \mu_i} \right) \right\} \\ &= \prod_{i=1}^n \left( \frac{\mu_i}{1 + \theta \mu_i} \right)^{y_i} \left( \prod_{i=1}^n \frac{(1 + \theta y_i)^{y_i - 1}}{y_i!} \right) \exp \left[ \sum_{i=1}^n \frac{-\mu_i (1 + \theta y_i)}{1 + \theta \mu_i} \right] \end{aligned} \quad (2.11)$$

Kemudian persamaan (2.11) diubah dalam bentuk fungsi logaritma natural sebagai berikut.

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

$$\text{dengan } \Delta = -\ln(y_i!) - \frac{\mu_i(1 + \theta y_i)}{1 + \theta \mu_i}$$

dengan mensubstitusikan nilai  $\mu_i = \exp(\mathbf{x}_i^T \boldsymbol{\beta})$  maka diperoleh

$$\begin{aligned} \ln L(\boldsymbol{\beta}, \theta) &= \sum_{i=1}^n \{ y_i \ln(\exp(\mathbf{x}_i^T \boldsymbol{\beta})) - y_i \ln(1 + \theta \exp(\mathbf{x}_i^T \boldsymbol{\beta})) \\ &\quad - \ln(y_i!) + \Delta^* \} \end{aligned} \quad (2.12)$$

dengan

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

Selanjutnya, estimasi parameter  $\hat{\boldsymbol{\beta}}$  diperoleh dengan menurunkan persamaan logaritma natural dari fungsi likelihood terhadap  $\boldsymbol{\beta}^T$  sedangkan estimasi parameter  $\theta$  diperoleh dengan menurunkan persamaan logaritma natural dari fungsi likelihood terhadap  $\theta$  dan disamadengankan nol. Penurunan fungsi ln-likelihood dilakukan dengan menggunakan metode iterasi *Newton-Raphson* sampai didapatkan estimasi parameter yang konvergen.

### 2.3.2 Pengujian Parameter Model GPR

Pengujian parameter model *Generalized Poisson Regression* sama dengan pengujian parameter yang digunakan dalam model regresi poisson yaitu dengan metode *Maximum Likelihood Ratio Test* (MLRT). Hipotesis yang digunakan adalah

$$H_0: \beta_1 = \beta_2 = \beta_3 = \dots = \beta_k = 0$$

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

Statistik uji yang digunakan adalah  $D(\hat{\boldsymbol{\beta}})$  pada persamaan (2.6). Keputusan yang diambil akan tolak  $H_0$  jika  $D(\hat{\boldsymbol{\beta}}) > \chi_{v;\alpha}^2$  dengan  $v$  adalah banyaknya parameter model dibawah populasi dikurangi dengan banyaknya parameter model dibawah  $H_0$ . Tolak  $H_0$  berarti ada salah satu variabel yang berpengaruh signifikan terhadap model sehingga dilanjutkan dengan pengujian secara parsial dengan hipotesis sebagai berikut.

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

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

dan

$$H_0: \theta = 0 \text{ (tidak terdapat kasus } \textit{overdispersion})$$

$$H_1: \theta \neq 0 \text{ (terdapat kasus } \textit{overdispersion})$$

Statistik uji yang digunakan seperti pada persamaan (2.7) dengan daerah penolakan menolak  $H_0$  jika  $|z_{hit}| > z_{\frac{\alpha}{2}}$  dengan  $\alpha$  adalah tingkat signifikansi.



## 2.4 Regresi Binomial Negatif

Regresi binomial negatif merupakan salah satu metode yang dapat digunakan untuk mengatasi adanya *over/underdispersion* pada model regresi poisson. Regresi binomial negatif adalah regresi terapan dari GLM sehingga distribusi binomial negatif memiliki tiga komponen GLM sebagai berikut.

i. Komponen random

Variabel respon  $Y_i$  pada regresi binomial negatif diasumsikan berdistribusi binomial negatif yang dihasilkan dari distribusi *mixture* distribusi poisson dan gamma. Fungsi kepadatan peluang binomial negatif hasil distribusi *mixture* poisson-gamma adalah sebagai berikut.

$$P(Y|\alpha, \beta) = \frac{\Gamma(y + \alpha)}{y! \Gamma(\alpha)} \left( \frac{1}{1 + \beta} \right)^\alpha \left( 1 - \frac{1}{1 + \beta} \right)^y, \quad (2.13)$$

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

Nilai *mean* dan varian dari distribusi *mixture* poisson-gamma tersebut adalah  $E[Y] = \alpha\beta$  dan  $V[Y] = \alpha\beta + \alpha\beta^2$

ii. Komponen sistematis

Kontribusi variabel prediktor dalam model regresi binomial negatif dinyatakan dalam bentuk kombinasi linier antara parameter ( $\eta$ ) dengan parameter regresi yang akan diestimasi.

$$\eta_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}$$

iii. Fungsi link

Dalam model binomial negatif, nilai ekspektasi variabel respon  $Y$  adalah bilangan bulat non-negatif. Untuk mentransformasikan nilai  $\eta_i$  ke rentang yang sesuai dengan rentang pada respon  $y$  maka diperlukan suatu fungsi link. Regresi binomial negatif umumnya menggunakan fungsi link logaritma atau log link yaitu :

$$\ln \mu_i = \mathbf{x}_i^T \boldsymbol{\beta}$$

### 2.4.1 Estimasi Parameter Model Regresi Binomial Negatif

Untuk membentuk model regresi pada distribusi binomial negatif maka nilai parameter dari distribusi *mixture* dinyatakan dalam bentuk  $\mu = \alpha\beta$  dan  $\theta = \frac{1}{\alpha}$  sehingga  $E[Y] = \mu$  dan  $V[Y] = \mu + \theta\mu^2$ . Fungsi massa peluang binomial negatif menjadi sebagai berikut.

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

Jika  $\theta$  menuju nol maka  $V[Y] = E[Y] = \mu$  sehingga binomial negatif akan mendekati distribusi poisson.

Metode yang digunakan untuk estimasi parameter model regresi binomial negatif adalah metode *maximum likelihood estimation*. Untuk variabel respon  $Y_i$ , fungsi massa probabilitas distribusi binomial negatifnya adalah

$$f(y_i; \mu_i, \theta) = \frac{\Gamma\left(y_i + \frac{1}{\theta}\right)}{\Gamma\left(\frac{1}{\theta}\right) \Gamma(y_i + 1)} \left(\frac{1}{1 + \theta\mu_i}\right)^{\frac{1}{\theta}} \left(\frac{\theta\mu_i}{1 + \theta\mu_i}\right)^{y_i}, y_i = 0, 1, 2, \dots \quad (2.14)$$

karena fungsinya saling independen maka fungsi likelihoodnya adalah

$$L(\boldsymbol{\beta}, \theta) = \prod_{i=1}^n \frac{\Gamma\left(\theta + \frac{1}{\theta}\right)}{\Gamma\left(\frac{1}{\theta}\right) \Gamma(y_i + 1)} \left(\frac{1}{1 + \theta\mu_i}\right)^{\frac{1}{\theta}} \left(\frac{\theta\mu_i}{1 + \theta\mu_i}\right)^{y_i} \quad (2.15)$$

dengan  $\frac{\Gamma\left(y + \frac{1}{\theta}\right)}{\Gamma\left(\frac{1}{\theta}\right)} = \prod_{r=1}^{y-1} (r + \theta^{-1})$

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

$$L(\boldsymbol{\beta}, \theta) = \ln\{L(\boldsymbol{\beta}, \theta)\}$$

$$L(\boldsymbol{\beta}, \theta) = \sum_{i=1}^n \left[ \left( \sum_{r=0}^{y_i-1} \ln(r + \theta^{-1}) \right) - \ln(y_i!) + y_i \ln(\theta \mu_i) - (\theta^{-1} + y_i) \ln(1 + \theta \mu_i) \right]$$

Selanjutnya, estimasi parameter untuk model regresi binomial negatif diperoleh dengan menurunkan fungsi ln likelihood terhadap parameter yang akan diestimasi kemudian disamadengankan nol. Turunan pertama terhadap koefisien  $\beta$  ditunjukkan sebagai berikut.

$$\frac{\partial L(\boldsymbol{\beta}, \theta)}{\partial \beta_0} = \sum_{i=1}^n \left[ y_i - (\theta^{-1} + y_i) \left( \frac{\theta \mu_i}{1 + \theta \mu_i} \right) \right] = \sum_{i=1}^n \left[ \frac{y_i - \mu_i}{1 + \theta \mu_i} \right]$$

$$= 0$$

⋮

$$\frac{\partial L(\boldsymbol{\beta}, \theta)}{\partial \beta_k} = \sum_{i=1}^n \left[ y_i x_{ik} - (\theta^{-1} + y_i) \left( \frac{\theta \mu_i x_{ik}}{1 + \theta \mu_i} \right) \right]$$

$$\frac{\partial L(\boldsymbol{\beta}, \theta)}{\partial \beta_k} = \sum_{i=1}^n \left[ \frac{\mu_i}{1 + \theta \mu_i} \frac{(y_i - \mu_i) x_{ik}}{\mu_i} \right] = 0$$

Jika ditulis dalam bentuk matrik maka turunan pertama fungsi likelihood terhadap  $\beta$  adalah  $\mathbf{q} = \mathbf{X}^T \mathbf{Wz}$ , dengan  $\mathbf{X}$  adalah matriks ( $n \times (k+1)$ ) dari variabel prediktor yang ditampilkan pada halaman 12.  $\mathbf{W}$  adalah matrik *weigh* dengan elemen diagonal  $w_i$  dan  $\mathbf{z}$  adalah vektor baris.  $w_i$  dan elemen  $\mathbf{z}$  adalah sebagai berikut.

$$w_i = \frac{\mu_i}{1 + \theta \mu_i} \text{ dan } \mathbf{z} = \frac{(y_i - \mu_i)}{\mu_i}$$

turunan pertama fungsi likelihood terhadap  $\theta$  adalah

$$\begin{aligned} f'(\theta) &= \frac{\partial L(\boldsymbol{\beta}, \theta)}{\partial \theta} \\ &= \sum_{i=1}^n \left[ -\theta^{-2} \sum_{r=0}^{y_i-1} \frac{1}{r + \theta^{-1}} + \frac{y_i}{\theta} + \theta^{-2} \ln(1 + \theta \mu_i) - \frac{(\theta^{-1} + y_i) \mu_i}{1 + \theta \mu_i} \right] = 0 \\ &= \sum_{i=1}^n \left[ -\theta^{-2} \sum_{r=0}^{y_i-1} \frac{1}{r + \theta^{-1}} + \frac{y_i}{\theta} + \theta^{-2} \ln(1 + \theta \mu_i) + \left( \frac{y_i - \mu_i}{\theta(1 + \theta \mu_i)} \right) \right] = 0 \end{aligned}$$

Turunan parsial kedua untuk parameter  $\boldsymbol{\beta}$  adalah sebagai berikut.

$$\begin{aligned} \frac{\partial l(\boldsymbol{\beta}, \theta)}{\partial \beta_0^2} &= - \sum_{i=1}^n \left[ \frac{(1 + \theta y_i) \mu_i}{(1 + \theta \mu_i)^2} \right] \\ \frac{\partial l(\boldsymbol{\beta}, \theta)}{\partial \beta_0 \partial \beta_j} &= \sum_{i=1}^n \left[ \frac{-\mu_i x_{ij} (1 + \theta \mu_i) - (y_i - \mu_i) (x_{ij} \theta \mu_i)}{(1 + \theta \mu_i)^2} \right] \\ \frac{\partial l(\boldsymbol{\beta}, \theta)}{\partial \beta_0 \partial \beta_j} &= - \sum_{i=1}^n \left[ \frac{(1 + \theta y_i) x_{ij} \mu_i}{(1 + \theta \mu_i)^2} \right] \end{aligned}$$

Misalkan turunan parsial pertama dari  $L(\boldsymbol{\beta}, \theta)$  terhadap  $\beta_j$  dengan  $j \leq k$  adalah

$$\frac{\partial l(\boldsymbol{\beta}, \theta)}{\partial \beta_k} = \sum_{i=1}^n \left[ \frac{(y_i - \mu_i) x_{ij}}{1 + \theta \mu_i} \right] = 0$$

Maka turunan parsial kedua terhadap  $\beta_u$  dengan  $u \leq k$  adalah sebagai berikut.

$$\begin{aligned} \frac{\partial^2 l(\boldsymbol{\beta}, \theta)}{\partial \beta_u \partial \beta_j} &= \sum_{i=1}^n \left[ \frac{-\mu_i x_{iu} x_{ij} (1 + \theta \mu_i) - (y_i - \mu_i) (x_{iu} x_{ij} \theta \mu_i)}{(1 + \theta \mu_i)^2} \right] \\ \frac{\partial^2 l(\boldsymbol{\beta}, \theta)}{\partial \beta_u \partial \beta_j} &= - \sum_{i=1}^n \left[ \frac{(1 + \theta y_i) x_{iu} x_{ij} \mu_i}{(1 + \theta \mu_i)^2} \right] \end{aligned}$$

Ekspektasi untuk turunan kedua fungsi ln-likelihood adalah

$$E \left( \frac{\partial^2 l(\boldsymbol{\beta}, \theta)}{\partial \beta_u \partial \beta_j} \right) = \sum_{i=1}^n \left[ \frac{x_{iu} x_{ij} \mu_i}{(1 + \theta \mu_i)} \right]$$

Jika ditulis dalam bentuk matriks  $\mathbf{I}$  (matriks informasi) maka persamaan tersebut menjadi  $\mathbf{I} = \mathbf{X}^T \mathbf{W} \mathbf{X}$  dengan  $\mathbf{X}$  adalah matriks dari variabel prediktor,  $\mathbf{W}$  adalah matriks *weight* dengan elemen diagonal ke- $i$  adalah

$$w_i = \frac{\mu_i}{1 + \theta \mu_i}$$

Turunan kedua fungsi likelihood terhadap parameter  $\theta$  ditunjukkan pada persamaan berikut.

$$f''(\theta) = \sum_{i=1}^n \left[ \theta^{-3} \left( \sum_{r=0}^{y_i-1} \frac{(2r + \theta^{-1})}{(r + \theta^{-1})^2} - 2\theta^{-3} \ln(1 + \theta \mu_i) + \Lambda^* \right) \right]$$

dengan

$$\Lambda^* = \frac{\theta^{-2} \mu (y_i - \mu_i)(1 + 2\theta \mu_i)}{1 + \theta \mu_i (\theta + \theta^2 \mu_i)^2}$$

Estimasi parameter model regresi binomial negatif dengan MLE tidak bisa menghasilkan solusi secara langsung sehingga perlu dilakukan iterasi dengan langkah-langkah sebagai berikut.

- i. Menentukan nilai taksiran awal untuk  $\theta$ , misal  $\hat{\theta}_1 = 0$
- ii. Menentukan estimasi parameter  $\boldsymbol{\beta}$  dengan menggunakan informasi Fisher *Scoring* dengan asumsi  $\theta = \hat{\theta}_1$  dan persamaan estimasi parameter berikut.

$$\hat{\boldsymbol{\beta}}_{m+1} = \hat{\boldsymbol{\beta}}_m + (\mathbf{X}^T \mathbf{W} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}_i \mathbf{z}_i$$

Iterasi dilakukan sampai mendapat  $\|\hat{\boldsymbol{\beta}}_{(m+1)} - \hat{\boldsymbol{\beta}}_{(m)}\| \leq \varepsilon$ .

- iii. Melakukan iterasi *Newton-Raphson* satu dimensi untuk mendapatkan estimasi parameter  $\theta$  dengan menggunakan  $\hat{\boldsymbol{\beta}}$ .

$$\hat{\theta}_{m+1} = m - \frac{f'(\theta_m)}{f''(\theta_m)}$$

Iterasi dilakukan sampai mendapat  $|\hat{\theta}_{m+1} - \hat{\theta}_m| < \varepsilon$ .

- iv. Jika tidak mendapatkan nilai  $|\hat{\theta}_{m+1} - \hat{\theta}_m| < \varepsilon$  maka diulangi lagi langkah ii dengan  $\theta = \hat{\theta}_{m+1}$ .

#### 2.4.2 Pengujian Parameter Model Regresi Binomial Negatif

Pengujian parameter model secara serentak dilakukan menggunakan statistik uji  $D(\hat{\beta})$  pada persamaan (2.6) dengan hipotesis sebagai berikut.

$$H_0: \beta_1 = \beta_2 = \beta_3 = \dots = \beta_k = 0$$

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

Keputusan yang diambil akan tolak  $H_0$  jika  $D(\hat{\beta}) > \chi_{v;\alpha}^2$  dengan  $v$  adalah banyaknya parameter model dibawah populasi dikurangi dengan banyaknya parameter model dibawah  $H_0$ . Tolak  $H_0$  berarti ada salah satu variabel yang berpengaruh signifikan terhadap model sehingga dilanjutkan dengan pengujian secara parsial. Berikut adalah hipotesis yang digunakan untuk pengujian parameter secara parsial.

$$H_0: \beta_j = 0$$

$$H_1: \beta_j \neq 0$$

dan

$$H_0: \theta = 0$$

$$H_1: \theta > 0$$

Statistik uji yang digunakan seperti pada persamaan (2.7). Keputusan yang diambil adalah tolak  $H_0$  jika  $|z_{hit}| > z_{\frac{\alpha}{2}}$  dengan  $\alpha$  adalah tingkat signifikansi.

#### 2.5 Kriteria Pemilihan Model Terbaik

Menurut Zucchini (2000) ada beberapa metode yang dapat digunakan untuk memilih model terbaik dalam analisis regresi antara lain *Akaike Information Criterion* (AIC), *bootstrap*, *cross validation*, dan *Bayes Information Criterion* (BIC). Kriteria yang paling banyak digunakan untuk pemilihan model terbaik adalah *Akaike Information Criterion* (AIC) dan (*Bayesian Information*

*Criterion*) BIC. Zucchini (2000) mendefinisikan AIC dan BIC sebagai berikut.

$$AIC = -2 \log(L) + 2p \quad (2.16)$$

$$BIC = -2 \log(L) + p \log(n) \quad (2.17)$$

dengan  $\log(L)$  adalah nilai likelihood,  $p$  adalah jumlah parameter dalam model dan  $n$  adalah jumlah unit observasi. Model terbaik adalah model yang memiliki nilai AIC terkecil. Namun, Burnham dan Anderson (2002) menyatakan bahwa untuk ukuran sampel kecil kriteria AIC kurang baik untuk digunakan karena akan memberikan hasil yang bias sehingga perlu adanya faktor pengkoreksi. Kriteria pemilihan model dengan memberikan faktor koreksi pada AIC disebut dengan *Corrected Akaike Information Criterion* ( $AIC_C$ ).  $AIC_C$  dinyatakan dalam persamaan berikut.

$$AIC_C = AIC + \frac{2p(p+1)}{n-p-1} \quad (2.18)$$

dengan  $n$  adalah jumlah sampel dalam penelitian.  $AIC_C$  digunakan jika rasio antara jumlah sampel dan jumlah parameter dalam model ( $n/p$ ) kecil ( $< 40$ ).

Pada model regresi untuk data *count* seperti *generalized poisson regression* dan regresi binomial negatif terdapat parameter dispersi untuk mengatasi kondisi overdispersi sehingga jumlah parameter dalam model bertambah. Hal ini akan berpengaruh terhadap nilai likelihood model sehingga perlu dilakukan modifikasi pemilihan kriteria model terbaik. Modifikasi AIC untuk data *count* disebut dengan QAIC yang didefinisikan dalam persamaan 2.19.

$$QAIC = -2 \left( \frac{\log(L)}{c} \right) + 2p \quad (2.19)$$

sehingga kriteria pemilihan model untuk data *count* dengan jumlah sampel kecil dikembangkan menjadi QAIC<sub>C</sub>. Persamaan 2.20 adalah persamaan untuk QAIC<sub>C</sub>.

$$QAIC_C = QAIC + \frac{2p(p+1)}{n-p-1} \quad (2.20)$$

dengan  $c$  adalah taksiran untuk parameter dispersi. Dispersi pada regresi poisson ditaksir dengan nilai *pearson chi-square* dibagi dengan derajat bebas. Model yang terbaik adalah model yang memiliki nilai QAIC<sub>C</sub> terkecil.

## 2.6 Multikolinieritas

Salah satu asumsi yang harus dipenuhi dalam model regresi adalah tidak ada multikolinieritas. Multikolinieritas merupakan kondisi antar variabel prediktor yang digunakan dalam pembentukan model regresi memiliki korelasi yang kuat. Salah satu akibat adanya kondisi multikolinieritas adalah nilai *standar error* yang besar (*overestimated*). Menurut Hocking (1996) ada dua cara untuk mendeteksi adanya multikolinieritas.

### 1. *Variance Inflation Factor* (VIF)

Suatu model regresi menunjukkan adanya kasus multikolinieritas antar variabel prediktor jika nilai VIF lebih dari 10. Nilai VIF dinyatakan dengan rumus 2.21.

$$VIF = \frac{1}{1 - R_j^2} \quad (2.21)$$

dengan  $R_j^2$  adalah nilai koefisien determinasi antara variabel  $x_j$  dengan variabel  $x$  lainnya.

### 2. Nilai koefisien korelasi pearson ( $r_{ij}$ ) antar variabel prediktor.

indikasi adanya kasus multikolinieritas dapat diketahui dari nilai koefisien korelasi *pearson* antar variabel prediktor yang cukup tinggi (lebih dari 0,95).



Kasus multikolinieritas dapat diatasi dengan mengeluarkan variabel prediktor yang memiliki korelasi tinggi kemudian meregresikan kembali antara variabel respon dengan variabel prediktor yang tersisa.

## 2.8 Interpolasi

Interpolasi adalah metode untuk menghasilkan titik-titik data dalam jangkauan dari suatu data-data yang diketahui. Bentuk interpolasi yang paling sering digunakan adalah interpolasi polinomial newton dan interpolasi polinomial lagrange (Chapra, 2008).

### 1. Interpolasi polinomial *newton*

Persamaan interpolasi polinomial newton secara umum adalah sebagai berikut.

$$f_{n-1}(x) = b_1 + b_2(x - x_1) + \dots + b_n(x - x_1) \dots (x - x_{n-1}) \quad (2.22)$$

Untuk interpolasi polinomial orde ke  $(n-1)$  dibutuhkan  $n$  titik data yaitu  $[x_1, f(x_1)], [x_2, f(x_2)], \dots, [x_n, f(x_n)]$ . Titik data tersebut digunakan untuk memperoleh koefisien  $b_1, \dots, b_n$ .

$$b_1 = f(x_1)$$

$$b_1 = f[x_2, x_1]$$

$$\vdots$$

$$b_n = f[x_n, x_{n-1}, \dots, x_2, x_1]$$

Fungsi dalam kurung siku adalah fungsi beda bagi. Fungsi beda bagi pertama secara umum didefinisikan sebagai berikut.

$$f[x_i, x_j] = \frac{f(x_i) - f(x_j)}{x_i - x_j}$$

Sedangkan untuk beda bagi ke- $n$  adalah

$$f[x_n, x_{n-1}, \dots, x_2, x_1] = B$$

dengan

$$B = \frac{f(x_n, x_{n-1}, \dots, x_2) - f(x_n, x_{n-2}, \dots, x_1)}{x_n - x_1}$$

Jika masing-masing koefisien  $b_1, \dots, b_n$  disubstitusikan kedalam persamaan (2.22) maka persamaan umum interpolasi polinomial newton menjadi

$$f_{n-1}(x) = f(x_1) + (x - x_1)f[x_2, x_1] + \dots + (x - x_1) \dots (x - x_{n-1})f[x_n, x_{n-1}, \dots, x_2, x_1] \quad (2.23)$$

## 2. Interpolasi polinomial lagrange

Secara umum interpolasi lagrange dinyatakan dengan persamaan (2.24) (Chapra, 2008).

$$f_{n-1}(x) = \sum_{i=1}^n L_i(x)f(x_i) \quad (2.24)$$

dengan

$$L_i(x) = \prod_{\substack{j=1 \\ j \neq i}}^n \frac{x - x_j}{x_i - x_j}$$

Jumlah titik data dinotasikan dengan  $n$ . Persamaan (2.24) menunjukkan bahwa setiap suku  $L_i(x)$  bernilai 1 saat  $x = x_i$  dan bernilai nol pada semua titik lainnya sehingga hasil penjumlahan semua suku akan menghasilkan kurva unik yang menghubungkan setiap titik data yang digunakan. Interpolasi polinomial lagrange orde  $n-1$  adalah interpolasi yang melibatkan  $n$  titik data.



## BAB III

### METODOLOGI PENELITIAN

#### 3.1 Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder. Data jumlah penderita kusta pada tahun 2012 dan faktor-faktor yang diduga mempengaruhi (kecuali persentase penduduk miskin) diperoleh dari Profil Kesehatan Provinsi Jawa Timur tahun 2012 sedangkan data persentase penduduk menurut kabupaten/kota dan golongan pengeluaran perkapita sebulan diperoleh dari publikasi Provinsi Jawa Timur dalam Angka 2013. Unit pengamatan dalam penelitian ini adalah 29 kabupaten dan 8 kota di Jawa Timur.

#### 3.2 Variabel Penelitian

Variabel yang digunakan dalam penelitian ini terdiri dari variabel dependen (Y) yaitu jumlah penderita kusta pada tahun 2012 dan variabel independen yaitu faktor-faktor yang mempengaruhi jumlah penderita kusta (X). Variabel yang digunakan ditunjukkan pada Tabel 3.1.

**Tabel 3.1** Variabel Penelitian

Kode	Variabel
Y	Jumlah penderita kusta
X <sub>1</sub>	Persentase rumah tangga yang memiliki rumah sehat
X <sub>2</sub>	Persentase rumah tangga berperilaku hidup bersih dan sehat
X <sub>3</sub>	Persentase keberadaan puskesmas
X <sub>4</sub>	Persentase penduduk laki-laki
X <sub>5</sub>	Persentase penduduk usia 15-29 tahun
X <sub>6</sub>	Persentase penduduk miskin
X <sub>7</sub>	Tingkat kepadatan penduduk
X <sub>8</sub>	Persentase penduduk umur 10 tahun keatas dengan tingkat pendidikan SD/MI kebawah

Penjelasan untuk masing-masing variabel dan dasar penentuan variabel prediktor tersebut adalah sebagai berikut.

1. Jumlah penderita kusta di masing-masing kabupaten/kota di Jawa Timur tahun 2012 (Y)  
Jumlah penderita kusta yang dimaksud adalah penduduk yang menderita kusta baik itu kasus baru atau penderita lama. Penelitian dilakukan pada 29 kabupaten dan 8 kota di Jawa Timur dengan tidak menyertakan kabupaten Mojokerto dalam analisis karena Dinas Kesehatan Jawa Timur tidak menerima laporan data yang lengkap dari Dinas kesehatan kabupaten dan Badan Pusat Statistik kabupaten tersebut.
2. Persentase rumah sehat ( $X_1$ )  
Rumah sehat yang dimaksud adalah rumah yang memenuhi kriteria minimal: akses air minum, akses jamban sehat, lantai, ventilasi dan pencahayaan yang dihitung kumulatif dari tahun sebelumnya (Kemenkes RI, 2011). Pemilihan letak rumah yang jauh dari sumber penyakit dan pemeliharaan rumah merupakan elemen penting dari manajemen untuk mengontrol penyebaran penyakit. Rumah sehat merupakan faktor kunci bagi kesehatan (Hansen & Chaignat, 2013).
3. Persentase rumah tangga berperilaku hidup bersih dan sehat (PHBS)( $X_2$ )  
Rumah tangga ber-PHBS adalah rumah tangga yang seluruh anggotanya berperilaku hidup bersih dan sehat, yang meliputi sepuluh indikator. Rumah tangga ber-PHBS adalah rumah tangga yang memenuhi minimal 7 indikator (Kemenkes RI, 2011). Kerr-pontes (2006) dalam penelitiannya menyatakan bahwa perilaku yang meliputi tidak mengkonsumsi makanan bergizi, mandi di tempat terbuka dan jarang mengganti sprei tempat tidur selama 10 tahun terakhir berpengaruh signifikan terhadap kejadian kusta.
4. Persentase puskesmas ( $X_3$ )  
Puskesmas merupakan unit pelaksana teknis dinas kesehatan kabupaten/kota yang bertanggungjawab menyelenggarakan pembangunan kesehatan di wilayah kerja tertentu se-

hingga kesehatan di suatu wilayah dapat diketahui berdasarkan peran serta puskesmas di wilayah tersebut. Menurut Naik *et al* (1999) angka kusta tertinggi masih berada pada daerah-daerah dengan akses kesehatan terbatas.

5. Persentase penduduk laki-laki ( $X_4$ )

Kejadian penyakit kusta pada laki-laki lebih banyak terjadi daripada wanita, kecuali di Afrika, wanita lebih banyak terkena penyakit kusta dari pada laki-laki. Dengan kata lain, laki-laki 2 kali lebih beresiko tertular kusta daripada perempuan Bakker (2006).

6. Persentase penduduk usia 15-29 tahun ( $X_5$ )

Penyakit kusta jarang sekali ditemukan pada bayi dan penduduk usia diatas 70 tahun. Beberapa penelitian menyebutkan bahwa kusta lebih banyak terjadi pada usia produktif. Menurut Bakker (2006) Angka kejadian kusta banyak terjadi pada kelompok umur 15-29 tahun.

7. Persentase penduduk miskin ( $X_6$ )

Sekitar 90% penderita kusta menyerang kelompok sosial ekonomi lemah atau miskin. Sosial ekonomi rendah akan menyebabkan buruknya lingkungan, masalah kurang gizi dan rendahnya kemampuan untuk mendapatkan pelayanan kesehatan yang layak (WHO, 2013). BPS (2012) menyatakan bahwa penduduk miskin adalah penduduk yang memiliki rata-rata pengeluaran perkapita perbulan dibawah garis kemiskinan (GK). Garis kemiskinan merupakan nilai pengeluaran minimum untuk kebutuhan makanan dan non makanan. Pada tahun 2012 GK di Jawa Timur mencapai Rp 233.202,-. Nilai GK tersebut berada diantara kelompok rata-rata pengeluaran per kapita per bulan yang ditetapkan oleh BPS yaitu sebesar Rp 200.000,- sampai Rp 299.999,- sedangkan BPS hanya menghasilkan data persentase penduduk dengan rata-rata pengeluaran per kapita per bulan sesuai pengelompokan tersebut atau dengan rentang interval Rp 50.000,- atau Rp 100.000,00 sehingga data persentase penduduk miskin diperoleh dengan cara interpolasi.

#### 8. Tingkat kepadatan penduduk ( $X_7$ )

Kepadatan penduduk adalah jumlah penduduk di satu wilayah per- $\text{km}^2$  (kemenkes RI, 2011). Kerr-pontes (2004) dalam penelitiannya menyimpulkan bahwa ada korelasi negatif antara pendidikan, pertumbuhan penduduk, status sosio-ekonomi dan akses menuju layanan kesehatan dengan angka penemuan kusta baru. Semakin tinggi pertumbuhan penduduk maka kepadatan juga akan semakin meningkat. Hal ini berpengaruh terhadap kualitas lingkungan masyarakat.

#### 9. Persentase penduduk umur 10 tahun keatas dengan tingkat pendidikan SD/MI kebawah ( $X_8$ )

Rendahnya tingkat pendidikan seseorang berhubungan dengan pengetahuan terkait penyakit dan penyebarannya. Penelitian Warsini (2007) menyebutkan bahwa tingkat pendidikan penderita kusta yang paling banyak adalah pendidikan SD kebawah yaitu sebanyak (63,9%). Selain itu, Kerr-pontes (2006) juga menyimpulkan bahwa tingkat pendidikan rendah merupakan variabel yang signifikan terhadap angka penemuan kusta.

### 3.3 Langkah Analisis Data

Berdasarkan tujuan penelitian maka langkah analisis yang digunakan adalah

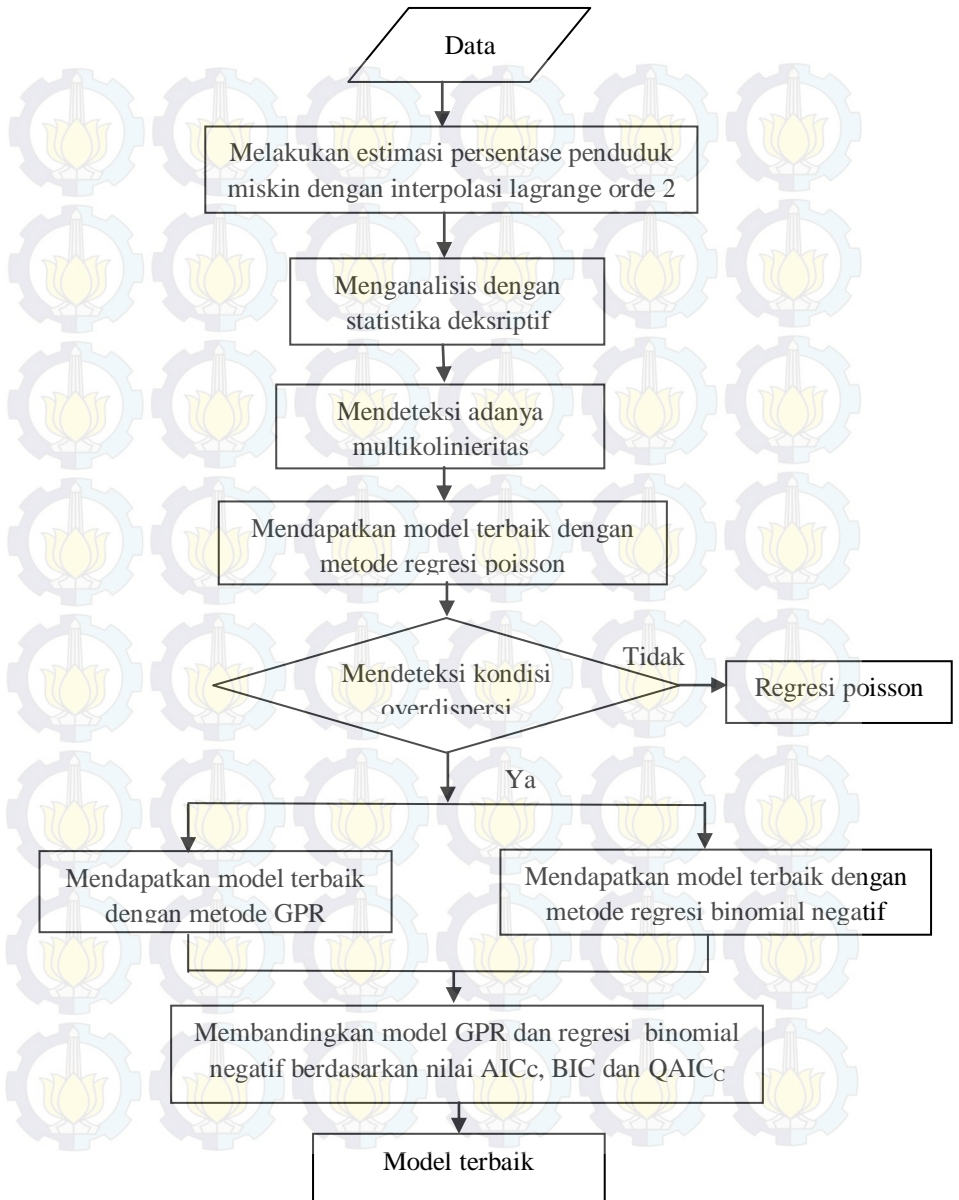
1. Mendeskripsikan karakteristik penderita kusta di Jawa Timur dan faktor-faktor yang diduga mempengaruhi dengan menggunakan analisa statistika deskriptif. Sebelum analisa statistika deskriptif terlebih dahulu dilakukan interpolasi polinomial lagrange orde dua untuk mendapatkan data persentase penduduk miskin di tiap kabupaten/kota di Jawa Timur.
2. Mendapatkan model terbaik untuk jumlah kasus kusta dan faktor-faktor yang diduga mempengaruhi dengan metode regresi poisson, *generalized poisson regression* dan regresi binomial negatif. Langkah-langkahnya adalah sebagai berikut.
  - a. Mendeteksi adanya kasus multikolinieritas pada variabel independen menggunakan nilai VIF (*Variance Inflation*

*Factors*) dan nilai koefisien korelasi *pearson* serta mengatasinya.

- b. Mendapatkan model terbaik menggunakan regresi poisson dengan langkah-langkah sebagai berikut.
    - i. Mengestimasi parameter model regresi poisson
    - ii. Melakukan uji signifikansi secara serentak dan parsial terhadap parameter model regresi poisson.
    - iii. Memperoleh model regresi poisson serta nilai AIC<sub>c</sub>, BIC dan QAIC<sub>c</sub>-nya.
    - iv. Memeriksa adanya keadaan *over/underdispersion* pada model regresi poisson. Jika terdapat keadaan *over/underdispersion* pada model maka dilakukan pemodelan dengan metode *generalized poisson regression* dan regresi binomial negatif.
  - c. Mendapatkan model terbaik menggunakan *generalized poisson regression* (GPR). Berikut adalah langkah-langkah pemodelan dengan GPR.
    - i. Mengestimasi parameter model *generalized poisson regression*
    - ii. Melakukan uji signifikansi secara serentak dan parsial terhadap parameter model *generalized poisson regression*.
    - iii. Memperoleh model *generalized poisson regression* serta nilai AIC<sub>c</sub>, BIC dan QAIC<sub>c</sub> -nya.
  - d. Mendapatkan model terbaik menggunakan regresi binomial negatif. Prosedur untuk mendapatkan model terbaik menggunakan regresi binomial negatif sama seperti pada langkah pemodelan menggunakan GPR.
3. Membandingkan model terbaik hasil *generalized poisson regression* dan regresi binomial negatif berdasarkan nilai AIC<sub>c</sub>, BIC dan QAIC<sub>c</sub> terkecil.

Diagram alir berdasarkan langkah analisis tersebut ditunjukkan pada Gambar 3.1





**Gambar 3.1** Diagram Alir Penelitian

## BAB IV

### ANALISIS DAN PEMBAHASAN

Pada bab ini dibahas beberapa hal untuk menjawab rumusan masalah penelitian, antara lain statistika deskriptif jumlah penderita kusta di Jawa Timur tahun 2012 dan faktor yang diduga mempengaruhi, pemodelan menggunakan regresi poisson dan deteksi *overdispersion*, pemodelan menggunakan analisis *generalized poisson regression* serta analisis regresi binomial negatif. Kemudian dibandingkan antara model GPR dan regresi binomial negatif untuk mendapatkan model yang mampu mengatasi *overdispersion* pada regresi poisson menggunakan kriteria kebaikan AIC<sub>C</sub>, BIC dan QAIC<sub>C</sub>.

Sebelum dilakukan terlebih dahulu dilakukan interpolasi lagrange orde 2 untuk mendapatkan data persentase penduduk miskin di Jawa Timur. Struktur data yang digunakan untuk proses interpolasi adalah sebagai berikut.

**Tabel 4.1** Presentase Penduduk Berdasarkan Golongan Pengeluaran Perkapita Sebulan

No.	Kab./kota	Golongan Pengeluaran Perkapita Sebulan				
		< 99.999	100rb-149.999	150rb-199.999	200rb-299.999	300rb-499.999
1	Pacitan		1,25	11,89	36,81	33,66
2	Ponorogo		1,13	7,76	34,22	34,03
3	Trenggalek	0,09	1,56	9,77	34,46	34,3
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
35	Madiun		0	0,34	15	39,75
36	Surabaya		0,18	0,13	3,89	21,54
37	Batu		0	2,26	12	37,29

Tiga titik data yang digunakan adalah nilai tengah dari golongan pengeluaran perkapita sebulan 150.000-199.999, 200.000-299.999, dan 300.000-499.999 sehingga untuk setiap kabupaten/kota nilai  $x_1$ ,  $x_2$ , dan  $x_3$  adalah

$$x_1 = \frac{150.000 + 199.999}{2} = 174.999$$

$$x_2 = \frac{200.000 + 299.999}{2} = 249.999$$

$$x_3 = \frac{300.000 + 499.999}{2} = 399.999$$

Nilai untuk garis kemiskinan ( $x_0$ ) di Jawa timur adalah 233.202 rupiah. Berikut adalah interpolasi persentase penduduk miskin untuk kabupaten Pacitan yang dihitung berdasarkan persamaan (2.24).

$$L_1(x) = \left( \frac{233.202 - 249.999}{174.999 - 249.999} \right) \left( \frac{233.202 - 399.999}{174.999 - 399.999} \right) = 0,17$$

$$L_2(x) = \left( \frac{233.202 - 174.999}{249.999 - 174.999} \right) \left( \frac{233.202 - 399.999}{249.999 - 399.999} \right) = 0,86$$

$$L_3(x) = \left( \frac{233.202 - 174.999}{399.999 - 174.999} \right) \left( \frac{233.202 - 249.999}{399.999 - 249.999} \right) = -0,03$$

sehingga

$$f_{3-1}(x) = \sum_{i=1}^3 L_i(x) f(x_i)$$

$$\begin{aligned} f_2(x) &= \sum_{i=1}^3 L_i(x) f(x_i) \\ &= (0,17 \times 11,89) + (0,86 \times 36,81) \\ &\quad + (-0,03 \times 33,66) \\ &= 34,92 \% \end{aligned}$$

Persentase penduduk miskin untuk kabupaten/kota lain juga dihitung dengan proses yang sama seperti pada kabupaten Pacitan sehingga didapatkan data persentase penduduk miskin untuk masing-masing kabupaten/kota seperti pada variabel  $X_6$  pada lampiran B halaman 74

#### 4.1 Karakteristik Jumlah Kasus Kusta di Jawa Timur

Provinsi Jawa Timur merupakan wilayah dengan jumlah kasus kusta tertinggi di Indonesia. Dinas kesehatan (2014) menyatakan bahwa pada tahun 2012 terjadi 4.807 kasus kusta baru di Jawa Timur, dengan kata lain sepertiga penderita kusta di Indonesia berada di Jawa Timur. Kasus kusta di Jawa Timur menyebar hampir merata di tiap kabupaten/kota. Banyak faktor yang mempengaruhi kasus kusta di Indonesia diantaranya faktor lingkungan dan faktor perilaku penderita. Untuk mengetahui gambaran awal mengenai kasus kusta maka dilakukan analisis statistika deskriptif terhadap data jumlah penderita kusta dan faktor yang diduga mempengaruhi di Jawa Timur tahun 2012. Statistik deskriptif jumlah kasus kusta dan faktor yang diduga mempengaruhi ditampilkan pada Tabel 4.2.

**Tabel 4.2** Statistika Deskriptif Variabel Penelitian

Variabel	Rata-rata	Varians	Minimum	Maksimum
Y	129,20	22378,5	0,000	589,00
$X_1$	67,69	173,32	38,290	87,17
$X_2$	43,89	223,68	8,500	65,74
$X_3$	25,22	169,45	3,000	62,00
$X_4$	49,13	0,440	47,461	50,19
$X_5$	23,36	5,037	19,611	28,72
$X_6$	21,05	99,15	2,750	41,91
$X_7$	1786,00	4756577	51,000	8463,00
$X_8$	57,21	213,64	31,190	85,76

Tabel 4.2 menunjukkan bahwa rata-rata kasus kusta di Jawa Timur adalah 129 kasus. Kasus terbanyak terjadi di ka-

bupaten Sampang yaitu sebanyak 589 kasus sedangkan kabupaten/kota yang tidak terdapat kasus kusta (0 kasus) adalah kota Batu. Perbedaan jumlah kasus kusta yang tinggi juga ditunjukkan oleh nilai varians yang cukup tinggi yaitu 22378,5. Nilai varians menggambarkan bahwa persebaran kusta di Jawa Timur terkonsentrasi di kabupaten/kota tertentu yaitu Kabupaten/kota yang berada di pesisir pantai utara dan pulau Madura. Daerah tersebut termasuk dalam daerah endemi kusta yang memiliki resiko tinggi terkena kusta. Dinas kesehatan (2012) menyatakan bahwa daerah endemi kusta di Jawa Timur meliputi kabupaten/kota Sumenep, Probolinggo, Jember, Pamekasan, Bangkalan, Tuban, Lumajang, Pasuruan, Sampang, dan Situbondo. Data jumlah kasus kusta di tiap kabupaten/kota di Jawa Timur terdapat pada lampiran A halaman 71.

Nilai varians tertinggi untuk variabel prediktor yang diduga mempengaruhi jumlah kasus kusta di Jawa Timur terdapat pada variabel  $X_7$  (tingkat kepadatan penduduk) yaitu sebesar 4756577 jiwa/ km<sup>2</sup>. Kota Surabaya merupakan kota dengan tingkat kepadatan penduduk tertinggi yaitu mencapai 8463 jiwa/km<sup>2</sup>, sedangkan kota batu merupakan kota dengan tingkat kepadatan penduduk terendah dengan tingkat kepadatan 51 jiwa/km<sup>2</sup>. Artinya, penyebaran penduduk di Jawa Timur belum merata dan mengumpul di kota Surabaya. Hal ini karena Surabaya merupakan kota terbesar di Jawa Timur dan terbesar kedua di Indonesia yang menjadi pusat bisnis, industri, pendidikan dan perdagangan di Jawa Timur sehingga banyak penduduk Jawa Timur yang bertempat tinggal di Surabaya baik untuk keperluan pekerjaan ataupun pendidikan. Kabupaten/kota lain dengan tingkat kepadatan penduduk tinggi yaitu Malang, Madiun dan Mojokerto. Data tingkat kepadatan penduduk masing-masing kabupaten/kota di Jawa Timur ditampilkan pada lampiran B (halaman 74). Faktor lain yang mempengaruhi kasus kusta adalah perilaku dan lingkungan tempat tinggal. Varians yang tinggi untuk variabel prediktor  $X_1$  (persentase rumah tangga yang memiliki rumah sehat) dan

$X_2$  (persentase rumah tangga berperilaku hidup bersih dan sehat) menunjukkan bahwa masih ada kabupaten/kota yang memiliki kesadaran yang rendah akan pentingnya hidup bersih dan sehat serta memiliki rumah sehat. Kabupaten Pamekasan merupakan kabupaten dengan kesadaran berperilaku hidup bersih dan sehat yang cukup dengan persentase rumah tangga yang berperilaku hidup bersih dan sehat sebesar 8,5 % (lampiran B halaman 72). Kesadaran masyarakat tentang pentingnya berperilaku hidup bersih dan sehat tidak lepas dari peran puskesmas dalam mengadakan sosialisasi dan pelatihan hidup bersih dan sehat di setiap wilayah yang menjadi tanggung jawabnya. Oleh karena itu, peran puskesmas di kabupaten/kota di pulau Madura perlu ditingkatkan untuk menekan jumlah kasus kusta di wilayah tersebut.

#### **4.2 Pemodelan Jumlah Kasus Kusta di Jawa Timur**

Pemodelan jumlah kasus kusta di Jawa Timur dilakukan menggunakan tiga metode yaitu regresi poisson, *generalized poisson regression* dan regresi binomial negatif. Langkah awal yang dilakukan sebelum pemodelan adalah pemeriksaan multikolinieritas terhadap data.

##### **4.2.1 Pemeriksaan Multikolinieritas**

Pemeriksaan multikolinieritas dilakukan untuk mengetahui adanya hubungan antara variabel prediktor. Kriteria yang dapat digunakan untuk mengidentifikasi adanya kasus multikolinieritas adalah korelasi antar variabel prediktor. Variabel yang memiliki korelasi tinggi adalah variabel  $X_6$  dan  $X_7$  sebesar 0,7034,  $X_8$  dan  $X_6$  sebesar 0,8936 serta  $X_7$  dan  $X_8$  sebesar 0,7659 (lampiran C). Keterangan dan definisi masing-masing variabel prediktor di tampilkan pada halaman 30. Kriteria lain yang dapat digunakan untuk mengidentifikasi kasus multikolinieritas adalah nilai *Variance Inflated Factor* atau VIF. Nilai VIF untuk masing-masing variabel prediktor dapat dilihat pada Tabel 4.3.

**Tabel 4.3** Nilai VIF dari Sembilan Variabel Prediktor.

Variabel	Nilai VIF	Variabel	Nilai VIF
X <sub>1</sub>	2,0617	X <sub>5</sub>	2,0367
X <sub>2</sub>	1,7526	X <sub>6</sub>	5,7829
X <sub>3</sub>	1,4744	X <sub>7</sub>	4,8587
X <sub>4</sub>	1,7904	X <sub>8</sub>	9,7053

Dari hasil analisis Tabel 4.3 dapat disimpulkan bahwa tidak terdapat kasus multikolinieritas antar variabel prediktor karena tidak ada variabel prediktor yang memiliki nilai VIF > 10. Jadi semua variabel tersebut digunakan dalam pemodelan jumlah kasus kusta di Jawa Timur. Selain itu, dalam pemodelan jumlah kasus kusta di Jawa Timur juga ditambahkan variabel *dummy* (Xd). Variabel *dummy* yang digunakan adalah kabupaten/kota yang merupakan daerah endemi kusta. Daerah yang termasuk dalam daerah endemi diberi kode 1 sedangkan daerah lain diberi kode 0. Variabel *dummy* perlu digunakan karena pada daerah yang termasuk endemi kusta terdapat jumlah kasus kusta yang sangat tinggi hingga mencapai nilai ratusan sedangkan ada kabupaten/kota lain yang tidak ditemukan adanya kasus kusta sehingga perlu dimodelkan untuk mengetahui apakah memang ada perbedaan jumlah kasus kusta yang signifikan antara daerah endemi kusta dan yang bukan. Sebelum dilakukan analisis perlu dilakukan standardisasi karena ada variabel prediktor yang memiliki satuan yang berbeda yaitu X<sub>7</sub> (tingkat kepadatan penduduk).

#### 4.2.2 Pemodelan Jumlah Kasus Kusta di Jawa Timur Menggunakan Regresi Poisson

Data jumlah kasus kusta merupakan data *count* yang mengikuti distribusi poisson. Untuk mengetahui faktor-faktor yang mempengaruhi jumlah kasus kusta di Jawa Timur maka dilakukan analisis regresi poisson. Estimasi parameter regresi poisson ditunjukkan pada Tabel 4.4.

**Tabel 4.4** Estimasi Parameter Model Regresi Poisson

Parameter	Estimasi	Standar Error	Z-hit	P-value
$\beta_0$	4,0783	0,02921	140,27	< 0,0001
$\beta_d$	0,7725	0,05680	12,92	< 0,0001
$\beta_1$	-0,0322	0,02244	-1,08	0,2870
$\beta_2$	0,0958	0,02142	3,81	< 0,0005
$\beta_3$	0,1109	0,02465	19,82	0,0013
$\beta_4$	0,4695	0,02313	-17,14	< 0,0001
$\beta_5$	-0,4148	0,02861	18,75	< 0,0001
$\beta_6$	0,5379	0,03688	-9,04	< 0,0001
$\beta_7$	-0,3855	0,04362	-12,51	< 0,0001
$\beta_8$	-0,5705	0,04950	4,89	< 0,0001

Selanjutnya nilai-nilai estimasi tersebut di uji secara serentak dan parsial. Hipotesis untuk pengujian serentak adalah

$$H_0: \beta_d = \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = \beta_8 = 0$$

(tidak ada parameter yang berpengaruh signifikan terhadap model)

$$H_1: \text{paling tidak ada satu } \beta_j \neq 0 ; j = d, 1, 2, 3, 4, 5, 6, 7, 8$$

Statistik uji yang digunakan adalah  $D(\hat{\beta})$  yang diperoleh dari perhitungan dengan persamaan (2.6). Pada lampiran F.1 diketahui bahwa nilai  $D(\hat{\beta})$  adalah sebesar 810,2. Pada tingkat signifikansi ( $\alpha$ ) 10 % nilai  $\chi^2_{(9;0,1)}$  adalah 14,684. Keputusan yang diambil adalah menolak  $H_0$  karena nilai  $D(\hat{\beta}) > \chi^2_{(9;0,1)}$ . Artinya, paling tidak ada satu variabel prediktor yang berpengaruh signifikan terhadap model. Pengaruh yang diberikan setiap variabel prediktor terhadap variabel respon dapat diketahui melalui pengujian parameter secara parsial dengan hipotesis sebagai berikut.

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



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

Jika nilai dari  $|Z_{hit}| > Z_{\alpha/2}$  maka keputusan yang diambil adalah menolak  $H_0$  yang berarti parameter tersebut berpengaruh signifikan terhadap model. Dengan taraf signifikansi ( $\alpha$ ) 10 % maka nilai  $|Z_{hit}|$  untuk semua parameter kecuali  $\beta_1$  lebih besar dari  $Z_{\alpha/2} = 1,64$ . Selain itu, dari hasil analisis Tabel 4.4 diketahui bahwa *p-value* untuk semua parameter kecuali  $\beta_1$  lebih kecil dari 0,1 sehingga parameter yang berpengaruh signifikan positif terhadap model adalah  $\beta_d, \beta_2, \beta_3, \beta_5, \beta_8$  sedangkan parameter yang berpengaruh signifikan negatif adalah  $\beta_4, \beta_6, \beta_7$ .

Variabel yang tidak berpengaruh signifikan terhadap jumlah kasus kusta di Jawa Timur adalah persentase rumah tangga yang memiliki rumah sehat ( $X_1$ ). Variabel tersebut memiliki efek utama yang tidak berpengaruh signifikan terhadap model namun ada kemungkinan bahwa interaksi antara persentase rumah tangga yang memiliki rumah sehat ( $X_1$ ) dengan variabel prediktor lain akan berpengaruh signifikan terhadap jumlah kasus kusta di Jawa Timur. Oleh karena itu, perlu dilakukan pemodelan dengan menambahkan variabel interaksi antara persentase rumah tangga yang memiliki rumah sehat ( $X_1$ ) dengan variabel prediktor lain. Model dengan tambahan interaksi antara persentase rumah tangga yang memiliki rumah sehat ( $X_1$ ) dengan persentase penduduk usia 15-29 tahun ( $X_5$ ) adalah model yang menghasilkan QAIC<sub>C</sub> terkecil yaitu sebesar 64,8256. Nilai QAIC<sub>C</sub> didapatkan melalui perhitungan dengan rumus pada persamaan 2.20 (halaman 25).

Tabel 4.5 merupakan hasil estimasi parameter untuk model regresi poisson dengan Interaksi. Nilai estimasi tersebut kemudian diuji secara serentak dengan hipotesis sebagai berikut.

$H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = \beta_8 = \beta_{15} = 0$   
(tidak ada parameter yang signifikan terhadap model)

$H_1: \text{paling tidak ada satu } \beta_j \neq 0 ; j = d, 1, 2, 3, 4, 5, 6, 7, 8, 15$

**Tabel 4.5** Estimasi Parameter Model Regresi Poisson dengan Interaksi  $X_{15}$ 

Parameter	Estimasi	Standar Error	Z-hit	P-Value
$\beta_0$	4,1156	0,02965	138,82	< 0,0001
$\beta_d$	0,7384	0,05721	12,91	< 0,0001
$\beta_1$	-0,0146	0,02247	-0,65	0,5186
$\beta_2$	0,0889	0,02152	4,13	0,0002
$\beta_3$	0,4959	0,02456	20,19	< 0,0001
$\beta_4$	-0,4224	0,02381	-17,74	< 0,0001
$\beta_5$	0,5087	0,02908	17,49	< 0,0001
$\beta_6$	-0,2900	0,03781	-7,67	< 0,0001
$\beta_7$	-0,5200	0,04394	-11,83	< 0,0001
$\beta_8$	0,1695	0,05171	3,28	0,0023
$\beta_{15}$	-0,0785	0,01896	-4,14	0,0002

Lampiran F.2 menunjukkan nilai statistik uji  $D(\hat{\beta})$  sebesar 793. Pada tingkat signifikansi ( $\alpha$ ) 10% nilai  $\chi^2_{(10;0,1)}$  adalah 15,987 sehingga  $D(\hat{\beta}) > \chi^2_{(10;0,1)}$ . Jadi keputusan yang diambil adalah menolak  $H_0$  yang berarti paling tidak ada satu parameter yang berpengaruh signifikan terhadap model. Selanjutnya dilakukan pengujian parameter secara parsial untuk mengetahui pengaruh yang diberikan masing-masing parameter terhadap model. Hipotesis yang digunakan adalah

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

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

Pada tingkat signifikansi ( $\alpha$ ) 10% hasil pada Tabel 4.5 menunjukkan bahwa semua parameter kecuali  $\beta_1$  memiliki nilai  $|Z_{hit}|$  lebih besar dari  $Z_{\alpha/2} = 1,64$  sehingga dapat disimpulkan bahwa semua parameter kecuali  $\beta_1$  untuk model regresi poisson dengan interaksi antara  $X_1$  dan  $X_5$  berpenga-

ruh signifikan terhadap model. Jadi berdasarkan kriteria QAIC<sub>C</sub> model regresi poisson yang dihasilkan adalah.

$$\hat{\mu} = \exp(4,1156 + 0,7384 X_d + 0,0889 z_2 + 0,4959 z_3 - 0,4224 z_4 + 0,5087 z_5 - 0,2900 z_6 - 0,5200 z_7 + 0,1695 z_8 - 0,0785 z_{15})$$

Model yang dihasilkan setelah dilakukan *unstandardize* adalah

$$\hat{\mu} = \exp(29,58447 + 0,7384 X_d + 0,0059 X_2 + 0,0381 X_3 - 0,6367 X_4 + 0,2267 X_5 - 0,0291 X_6 - 0,00024 X_7 + 0,01159 X_8 - 0,0002 X_{15})$$

Variabel yang berpengaruh positif terhadap jumlah kasus kusta di Jawa Timur adalah variabel *dummy* yaitu daerah endemi kusta, persentase rumah tangga berperilaku hidup bersih dan sehat ( $X_2$ ), persentase keberadaan puskesmas ( $X_3$ ), persentase penduduk usia 15-19 tahun ( $X_5$ ), persentase penduduk umur 10 tahun keatas dengan tingkat pendidikan SD/MI kebawah ( $X_8$ ) sedangkan variabel prediktor yang berpengaruh negatif adalah persentase penduduk laki-laki ( $X_4$ ), persentase penduduk miskin ( $X_6$ ), tingkat kepadatan penduduk ( $X_7$ ) serta interaksi antara persentase antara rumah tangga yang memiliki rumah sehat dan persentase penduduk usia 15-19 tahun ( $X_{15}$ ). Peningkatan maupun penurunan jumlah kasus kusta di Jawa Timur tergantung nilai koefisien masing-masing variabel yang berpengaruh. Untuk variabel prediktor persentase rumah tangga berperilaku hidup bersih dan sehat ( $X_2$ ) dapat dijelaskan bahwa setiap peningkatan 1 % rumah tangga yang berperilaku hidup bersih dan sehat dengan asumsi variabel lain tetap maka jumlah kasus kusta akan meningkat sebesar  $\exp(0,0059) \approx 1$  kasus. Hasil estimasi nilai koefisien untuk variabel prediktor lain yang signifikan menunjukkan bahwa kasus kusta di Jawa Timur akan mengalami peningkatan atau penurunan sebesar 1

kasus untuk setiap peningkatan 1 % masing-masing variabel prediktor dengan asumsi nilai variabel lain tetap kecuali peningkatan persentase penduduk laki-laki yang akan meningkatkan jumlah kasus kusta sebanyak 2 kasus. Variabel *dummy* yang signifikan menunjukkan bahwa ada perbedaan jumlah kasus kusta yang signifikan sebesar  $\exp(0,7384) \approx 2$  kasus antara wilayah endemi dan yang bukan dengan asumsi variabel lain tetap.

Selain itu, hasil analisis menggunakan kriteria  $AIC_C$  dan BIC menunjukkan bahwa model terbaik untuk regresi poisson adalah model dengan interaksi antara persentase rumah tangga yang memiliki rumah sehat dan persentase keberadaan puskesmas ( $X_{13}$ ) karena menghasilkan nilai  $AIC_C$  dan BIC terkecil yaitu 811 dan 818,1. Pengujian secara serentak dan secara parsial menghasilkan kesimpulan bahwa semua variabel yang berpengaruh signifikan terhadap jumlah kasus kusta di Jawa Timur. Nilai estimasi parameter untuk model dengan interaksi antara persentase antara rumah tangga yang memiliki rumah sehat dan persentase keberadaan puskesmas ( $X_{13}$ ) ditampilkan pada lampiran F.3 halaman 83.

#### 4.2.3 Pemeriksaan *Overdispersion*

Dalam regresi poisson terdapat asumsi yang harus dipenuhi yaitu *equidispersion*. Namun kondisi tersebut jarang terpenuhi sehingga muncul adanya kasus *overdispersion*.

Pengujian yang dapat digunakan untuk mendeteksi *overdispersion* adalah dengan menyatakan distribusi poisson termasuk dalam keluarga distribusi Katz. Hipotesis untuk pengujian *overdispersion* adalah

$$H_0: c = 0$$

$$H_1: c > 0$$

Dari nilai estimasi parameter yang telah diperoleh maka bisa didapatkan nilai  $\hat{\mu}_i$  sehingga berdasarkan persamaan (2.8) nilai statistik uji L untuk model dengan interaksi  $X_{15}$  adalah 304,567 sedangkan untuk model dengan interaksi  $X_{13}$

adalah 369,6649.  $\sqrt{2}L$  untuk model dengan interaksi  $X_{15}$  dan model dengan interaksi  $X_{13}$  berturut-turut yaitu 430,7232 dan 522,7851. Nilai tersebut jauh lebih besar daripada  $Z_{\alpha} = 1,29$  pada taraf signifikansi 10 % sehingga menolak  $H_0$  sehingga kesimpulan yang diambil untuk kedua model adalah menolak  $H_0$ .

Jadi dapat disimpulkan bahwa kondisi *equidispersion* tidak terpenuhi atau terjadi *overdispersion* pada model regresi poisson jumlah kasus kusta di Jawa Timur yang diperoleh berdasarkan kriteria  $QAIC_C$  terkecil maupun  $AIC_C$  dan  $BIC$  terkecil. Adanya *overdispersion* akan menyebabkan hasil estimasi parameter menjadi bias dan nilai penduga bagi kesalahan baku yang lebih kecil (*underestimate*). Untuk mengatasi hal tersebut, maka dilakukan pemodelan menggunakan *generalized poisson regression* (GPR) dan regresi binomial negatif.

#### 4.2.4 Pemodelan Jumlah Kasus Kusta di Jawa Timur Menggunakan GPR

Seperti yang telah dijelaskan sebelumnya, untuk mengatasi adanya *overdispersion* pada model regresi poisson salah satu metode yang dapat digunakan adalah pemodelan dengan *generalized poisson regression*. Tabel 4.6 merupakan hasil estimasi parameter dan pengujian untuk model GPR antara Y dengan delapan variabel prediktor. Pengujian yang dilakukan adalah pengujian estimasi parameter secara serentak dan secara parsial.

**Tabel 4.6** Estimasi Parameter Model GPR

Parameter	Estimasi	Standar Error	Z-hit	P-Value
$\beta_0$	4,0219	0,2128	18,90	<,0001
$\beta_d$	1,1847	0,6928	1,71	0,0956
$\beta_1$	-0,1324	0,2066	-0,64	0,5254
$\beta_2$	0,0852	0,1912	0,45	0,6583
$\beta_3$	0,9379	0,3036	3,09	0,0038

**Tabel 4.6** Estimasi Parameter Model GPR (lanjutan)

$\beta_4$	-0,2583	0,2004	-1,29	0,2054
$\beta_5$	0,4690	0,2243	2,09	0,0435
$\beta_6$	0,3163	0,3612	0,88	0,3869
$\beta_7$	-0,3157	0,3952	-0,80	0,4294
$\beta_8$	-0,2803	0,4073	-0,69	0,4956
$\theta$	0,0592	0,01203	4,92	< 0,0001

Hipotesis yang digunakan untuk pengujian secara serentak parameter model *generalized poisson regression* adalah

$$H_0: \beta_d = \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = \beta_8 = \theta = 0$$

(tidak ada parameter yang signifikan terhadap model)

$$H_1: \text{paling tidak ada satu } \beta_j \neq 0 ; j = 1, 2, 3, 4, 5, 6, 7, 8$$

Pengujian parameter secara serentak menggunakan statistik uji  $D(\hat{\beta})$  (persamaan 2.6 halaman 13). Dari lampiran H diperoleh nilai  $D(\hat{\beta})$  sebesar 386,9. Pada tingkat signifikansi ( $\alpha$ ) 10 % nilai  $\chi^2_{(9;0,1)}$  adalah 14,684 sehingga keputusan yang diambil adalah menolak  $H_0$ . Artinya paling tidak ada satu variabel prediktor yang berpengaruh signifikan terhadap model. Selanjutnya dilakukan uji parameter secara parsial menggunakan statistik uji  $Z_{hit}$  yang ditampilkan pada Tabel 4.6 dengan hipotesis sebagai berikut.

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

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

dan

$$H_0: \theta = 0 \text{ (tidak terdapat kasus } \textit{overdispersion})$$

$$H_1: \theta \neq 0 \text{ (tidak terdapat kasus } \textit{overdispersion})$$

Dari hasil pengujian dengan taraf signifikansi ( $\alpha$ ) 10 % dapat disimpulkan bahwa parameter yang berpengaruh signifikan positif terhadap model adalah  $\beta_d, \beta_2, \beta_5$  dan  $\theta$  sedangkan parameter-parameter lainnya tidak signifikan terhadap model. Parameter-parameter tersebut dikatakan signifikan ka-

rena memiliki  $p$ -value yang lebih kecil dari 0,10 dan memiliki nilai  $|Z_{hit}|$  yang lebih besar dari  $Z_{\alpha/2} = 1,64$ .

Variabel prediktor yang signifikan adalah  $X_d$ ,  $X_3$ ,  $X_5$ , sedangkan variabel prediktor yang tidak berpengaruh signifikan terhadap variabel respon adalah  $X_1$ ,  $X_2$ ,  $X_4$ ,  $X_6$ ,  $X_7$  dan  $X_8$ . Nilai  $AIC_C$ ,  $BIC$ , dan  $QAIC_C$  untuk model GPR dengan delapan variabel prediktor berturut-turut adalah 419,4, 426,6 dan 6569,16 Selanjutnya dilakukan pemodelan dengan menambahkan variabel interaksi antara variabel prediktor yang tidak signifikan dengan variabel prediktor lainnya untuk mengetahui adanya pengaruh interaksi. Pada penelitian ini dipilih model terbaik berdasarkan kriteria nilai  $AIC_C$ ,  $BIC$  dan  $QAIC_C$  terkecil. Nilai estimasi parameter dari masing-masing kemungkinan model tersebut ditampilkan pada lampiran J sedangkan ringkasan model terbaik untuk kombinasi interaksi setiap variabel prediktor yang tidak signifikan dengan variabel prediktor lainnya ditunjukkan pada Tabel 4.7. Model yang memiliki nilai  $QAIC_C$  terkecil adalah model antara variabel respon dengan semua variabel prediktor serta interaksi antara  $X_7$  dan  $X_5$ . dengan nilai  $QAIC_C$  sebesar 31388,79.

**Tabel 4.7** Kemungkinan Model GPR dengan Interaksi Berdasarkan Nilai  $QAIC_C$  terkecil

Kemungkinan model (Y dengan 8 variabel prediktor dan interaksi Xi)	$QAIC_C$	Parameter yang signifikan
$X_{48}$	31908,39	$\beta_0, \beta_3, \beta_5, \theta$
$X_{63}$	40444,42	$\beta_0, \beta_d, \beta_3, \beta_{63}, \theta$
$X_{75}$	31388,79	$\beta_0, \beta_3, \beta_5, \theta$
$X_{83}$	37504,73	$\beta_0, \beta_d, \beta_3, \beta_{83}, \theta$

Tabel 4.7 merupakan hasil estimasi parameter model GPR terpilih berdasarkan kriteria nilai  $QAIC_C$ . Selanjutnya nilai-nilai estimasi tersebut di uji secara serentak dan parsial.

Dengan hipotesis untuk pengujian serentak adalah sebagai berikut.

$$H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = \beta_8 = \beta_{75} \\ = \theta = 0$$

(tidak ada parameter yang signifikan terhadap model)

$$H_1: \text{paling tidak ada satu } \beta_j \neq 0 ; j = d, 1, 2, 3, 4, 5, 6, 7, 8, 75$$

Statistik uji yang digunakan adalah  $D(\hat{\beta})$  yang akan dibandingkan dengan nilai *chi-square*. Nilai  $D(\hat{\beta})$  diperoleh melalui perhitungan menggunakan persamaan 2.6. Pada lampiran J.1 diperoleh nilai  $D(\hat{\beta})$  sebesar 385 sedangkan nilai  $\chi^2_{(p;\alpha)}$  dengan tingkat signifikansi ( $\alpha$ ) 10 % adalah 15,987. Hasil analisis menunjukkan bahwa nilai  $D(\beta) > \chi^2_{(10;0,1)}$  sehingga keputusan yang diambil adalah menolak  $H_0$  yang berarti paling tidak ada satu variabel prediktor yang berpengaruh signifikan terhadap model.

**Tabel 4.8** Estimasi Parameter Model GPR dengan Interaksi  $X_{75}$

parameter	Estimasi	Standar Error	Z-Hit	P-Value
$\beta_0$	4,3049	0,3819	11,27	< 0,0001
$\beta_d$	0,9441	0,7908	1,19	0,2401
$\beta_1$	-0,1921	0,2377	-0,81	0,4242
$\beta_2$	0,0832	0,2076	0,40	0,6910
$\beta_3$	1,2090	0,4015	3,01	0,0047
$\beta_4$	-0,2499	0,2333	-1,07	0,2910
$\beta_5$	0,4750	0,2387	1,99	0,0541
$\beta_6$	0,2656	0,3919	0,68	0,5021
$\beta_7$	0,1431	0,5707	0,25	0,8034
$\beta_8$	-0,0091	0,5673	-0,02	0,9873
$\beta_{75}$	-0,2019	0,1597	-1,26	0,2141
$\theta$	0,0616	0,01228	5,01	< 0,0001



Untuk mengetahui pengaruh yang diberikan setiap variabel prediktor terhadap variabel respon maka dilakukan pengujian paramater 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)

dan

$H_0: \theta = 0$  (tidak terdapat kasus *overdispersion*)

$H_1: \theta \neq 0$  (terdapat kasus *overdispersion*)

Jika nilai dari  $|Z_{hit}| > Z_{\alpha/2}$  maka keputusan yang diambil adalah tolak  $H_0$  yang artinya parameter tersebut berpengaruh signifikan terhadap model. Nilai  $|Z_{hit}|$  untuk parameter  $\beta_0, \beta_3, \beta_5, \theta$  lebih besar dari  $Z_{\alpha/2} = 1,64$  pada tingkat signifikansi ( $\alpha$ ) 10 % sehingga parameter  $\beta_0, \beta_3, \beta_5, \theta$  berpengaruh signifikan terhadap model. Selain itu, berdasarkan Tabel 4.8 diketahui bahwa *p-value* untuk parameter  $\beta_0, \beta_3, \beta_5, \theta$  lebih kecil dari 0,1. Parameter  $\theta$  yang signifikan menunjukkan adanya kasus *overdispersion* pada model yang telah terbentuk. Hasil tersebut sesuai dengan hasil pengujian yang dilakukan pada model regresi poisson. Variabel prediktor yang signifikan adalah  $X_3$  dan  $X_5$  sehingga model GPR yang dihasilkan dengan melihat nilai QAIC<sub>C</sub> terkecil adalah

$$\hat{\mu} = \exp(4,3049 + 1,2090 z_3 + 0,4750 z_5)$$

Setelah dilakukan *unstandardize* didapatkan model GPR untuk jumlah kasus kusta di Jawa timur sebagai berikut.

$$\hat{\mu} = \exp(-2,9807 + 0,0929 X_3 + 0,2116 X_5)$$

Dengan demikian faktor yang berpengaruh signifikan positif terhadap jumlah kasus kusta di Jawa Timur adalah persentase keberadaan puskesmas ( $X_3$ ), persentase penduduk usia 15-29 tahun ( $X_5$ ).

Jumlah kasus kusta di Jawa Timur akan mengalami peningkatan sesuai dengan nilai koefisien masing-masing variabel yang berpengaruh. Hasil analisis model *generalized poisson regression* dengan nilai QAIC<sub>C</sub> terkecil menunjukkan bahwa variabel yang menyebabkan peningkatan kasus kusta di Jawa Timur adalah persentase keberadaan puskesmas dan persentase penduduk usia 15-29 tahun. Setiap peningkatan 1% pukesmas dengan asumsi nilai variabel lain tetap akan meningkatkan sebesar  $\exp(0,0929) \approx 1$  kasus kusta di Jawa Timur. Begitu juga dengan kenaikan persentase penduduk usia 15-29 tahun sebesar 1% akan meningkatkan jumlah kasus kusta di Jawa Timur sebesar  $\exp(0,2116) \approx 1$ . Pada model *generalized poisson regression*, variabel *dummy* yaitu daerah endemi kusta tidak signifikan sehingga ada kemungkinan jumlah kasus kusta di daerah endemi tidak berbeda jauh dengan daerah lain yang bukan daerah endemi.

Jika dipilih model terbaik menggunakan kriteria AIC<sub>C</sub> dan BIC terkecil maka model terbaik untuk masing-masing kombinasi variabel prediktor yang tidak berpengaruh adalah sebagai berikut.

**Tabel 4.9** Kemungkinan Model GPR dengan Interaksi Berdasarkan Nilai AIC<sub>C</sub> dan BIC terkecil

Kemungkinan model (Y dengan 8 variabel prediktor dan interaksi Xi)	AIC <sub>C</sub>	BIC	Parameter yang signifikan
X <sub>47</sub>	422	428,4	$\beta_0, \beta_3, \beta_5, \theta$
X <sub>63</sub>	411,9	418,2	$\beta_0, \beta_d, \beta_3, \beta_{63}, \theta$
X <sub>75</sub>	422	428,4	$\beta_0, \beta_3, \beta_5, \theta$
X <sub>83</sub>	414,3	420,6	$\beta_0, \beta_d, \beta_3, \beta_{83}, \theta$

Model terbaik untuk GPR berdasarkan nilai AIC<sub>C</sub> dan BIC terkecil adalah model tambahan interaksi antara persentase penduduk miskin dan persentase keberadaan puskesmas (X<sub>63</sub>). Nilai estimasi untuk parameter model tersebut ditun-

jukan pada lampiran J.2. Kemudian dilakukan pengujian parameter secara serentak dan parsial terhadap parameter model terpilih seperti pada model GPR yang terpilih berdasarkan QAIC<sub>C</sub> terkecil. Setelah dilakukan pengujian diketahui bahwa faktor yang berpengaruh signifikan terhadap jumlah kasus kusta di Jawa Timur adalah daerah endemi, persentase keberadaan puskesmas serta interaksi antara persentase penduduk miskin dan persentase keberadaan puskesmas. Persentase keberadaan puskesmas berpengaruh positif terhadap jumlah kasus kusta sedangkan interaksi antara persentase penduduk miskin dan persentase keberadaan puskesmas memberikan pengaruh negatif. Hal tersebut menunjukkan bahwa keberadaan puskesmas mampu menekan jumlah kasus kusta di Jawa Timur walaupun persentase penduduk miskin meningkat. Selain itu, variabel *dummy* yang signifikan mengindikasikan bahwa ada perbedaan jumlah kasus kusta antara daerah endemi kusta dan daerah bukan endemi. Model GPR dengan nilai AIC<sub>C</sub> dan BIC terkecil adalah sebagai berikut.

$$\hat{\mu} = \exp(4,1096 + 1,9368 X_d + 0,6347 z_3 - 1,0755 z_{63})$$

Setelah dikembalikan kedalam bentuk awal sebagai variabel  $x$ , maka model tersebut menjadi

$$\hat{\mu} = \exp(4,5772 + 1,9368 X_d + 0,0487 X_3 - 0,0031 X_{63})$$

#### 4.2.5 Pemodelan Jumlah Kasus Kusta di Jawa Timur Menggunakan Regresi Binomial Negatif.

Metode lain yang dapat digunakan untuk mengatasi kasus *overdispersion* selain GPR adalah metode regresi binomial negatif. Langkah analisis yang dilakukan hampir sama dengan pemodelan dengan regresi poisson dan GPR. Estimasi parameter untuk model regresi binomial negatif antara variabel respon dengan sembilan variabel prediktor ditunjukkan pada Tabel 4.10. Model dengan delapan variabel prediktor ter-

sebut memiliki nilai QAIC<sub>C</sub> sebesar 1236,408 dengan jumlah parameter yang signifikan sebanyak 5 parameter termasuk variabel *dummy*. Berikut adalah pengujian signifikansi parameter model regresi binomial negatif secara serentak dan parsial. Tabel 4.10 menampilkan hasil estimasi parameter beserta nilai statistik uji  $Z_{hit}$  untuk pengujian signifikansi parameter secara parsial sedangkan pengujian parameter secara serentak menggunakan statistik uji  $D(\hat{\beta})$ . Berikut adalah hipotesis yang digunakan untuk pengujian signifikansi parameter secara serentak.

$$H_0: \beta_d = \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = \beta_8 = \theta = 0$$

(tidak ada parameter yang signifikan terhadap model)

$$H_1: \text{paling tidak ada satu } \beta_j \neq 0 ; j = d, 1, 2, 3, 4, 5, 6, 7, 8,$$

**Tabel 4.10** Estimasi Parameter Model Regresi Binomial Negatif

Parameter	Estimasi	Standar Error	Z-Hit	P-Value
$\beta_0$	4,0079	0,1441	27821	< 2e-16
$\beta_d$	0,8656	0,40121	2157	0,0309
$\beta_1$	-0,0535	0,1371	-390	0,6963
$\beta_2$	0,0351	0,1281	274	0,7840
$\beta_3$	0,6931	0,1182	5863	4,55e-09
$\beta_4$	-0,29945	0,1351	-2217	0,0266
$\beta_5$	0,5226	0,1688	3097	0,00196
$\beta_6$	0,0094	0,2302	41	0,9674
$\beta_7$	-0,5761	0,2139	-2693	0,0071
$\beta_8$	-0,0622	0,3171	-196	0,8444
$\theta$	3,227	0,834	3,8693	< 0,0001

Dari lampiran K.1 diperoleh nilai statistik uji  $D(\hat{\beta})$  sebesar 371,679. Pada taraf signifikansi ( $\alpha$ ) 10 % nilai  $\chi^2_{(10;0,1)}$  adalah 15,987. Kesimpulan yang diambil adalah me-

olak  $H_0$  karena nilai  $D(\hat{\beta}) > \chi^2_{(10;0,1)}$ . Jadi dapat disimpulkan bahwa paling tidak terdapat satu variabel prediktor yang berpengaruh signifikan terhadap model. Selanjutnya dilakukan pengujian parameter secara parsial untuk mengetahui pengaruh yang diberikan setiap variabel prediktor terhadap variabel respon 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)

dan

$H_0: \theta = 0$  (tidak terdapat kasus *overdispersion*)

$H_1: \theta \neq 0$  (terdapat kasus *overdispersion*)

Berdasarkan hasil tabel 4.10 diketahui bahwa parameter yang signifikan adalah  $\beta_d, \beta_3, \beta_4, \beta_5, \beta_7$ , dan  $\theta$  karena memiliki nilai  $|Z_{hit}|$  lebih besar dari nilai  $Z_{0,05}$  yaitu 1,64. Jadi variabel yang tidak signifikan pada tingkat signifikan ( $\alpha$ ) 10% adalah  $X_d, X_3, X_4, X_5$ , dan  $X_7$ . Selanjutnya dilakukan evaluasi pengaruh interaksi antara variabel prediktor yang tidak signifikan tersebut dengan variabel prediktor lain. Variabel tersebut memiliki efek utama yang tidak berpengaruh signifikan terhadap model namun ada kemungkinan bahwa interaksi dengan variabel prediktor lain akan berpengaruh signifikan terhadap jumlah kasus kusta di Jawa Timur.

**Tabel 4.11** Kemungkinan Model Regresi Binomial Negatif Berdasarkan nilai QAIC<sub>C</sub>

Kemungkinan model (Y dengan 8 variabel prediktor dan interaksi Xi)	QAIC <sub>C</sub>	Parameter yang signifikan
$X_{15}$	1277,28	$\beta_0, \beta_d, \beta_4, \beta_5, \beta_7, \theta$
$X_{25}$	1242,82	$\beta_0, \beta_d, \beta_4, \beta_5, \beta_7, \theta$
$X_{65}$	1237,13	$\beta_0, \beta_d, \beta_3, \beta_4, \beta_5, \beta_7, \theta$
$X_{85}$	1253	$\beta_0, \beta_d, \beta_3, \beta_4, \beta_5, \beta_7, \theta$

. Tabel 4.11 adalah hasil pemilihan model dengan interaksi yang menghasilkan nilai QAIC<sub>C</sub> terkecil untuk masing-masing kombinasi variabel prediktor yang tidak signifikan. Tabel 4.11 menunjukkan bahwa model yang memiliki parameter signifikan paling banyak dan nilai QAIC<sub>C</sub> terkecil adalah model dengan tambahan variabel interaksi X<sub>65</sub>. Hasil estimasi parameter model regresi binomial dengan tambahan variabel interaksi X<sub>65</sub> ditampilkan pada Tabel 4.12 sebagai berikut.

**Tabel 4.12** Estimasi Parameter Model Regresi Binomial Negatif dengan Interaksi X<sub>65</sub>

Parameter	Estimasi	Standar Error	Z-hit	P-Value
$\beta_0$	4,0059	0,1483	27,017	< 2e-16
$\beta_d$	0,8625	0,4015	2,148	0,0317
$\beta_1$	-0,0512	0,1397	-0,367	0,7138
$\beta_2$	0,0335	0,1301	0,258	0,7967
$\beta_3$	0,6909	0,1188	5,817	6e-09
$\beta_4$	-0,3016	0,1368	-2,204	0,0275
$\beta_5$	0,5277	0,1732	3,047	0,0023
$\beta_6$	0,0116	0,2344	0,050	0,9604
$\beta_7$	-0,5839	0,2329	-2,508	0,0122
$\beta_8$	-0,0607	0,3179	-0,191	0,8487
$\beta_{65}$	-0,0109	0,1156	-0,094	0,9252
$\theta$	3,229	0,835	3,8671	< 0,0001

Nilai estimasi parameter tersebut kemudian diuji secara serentak dan parsial untuk mengetahui signifikansinya terhadap model. Hipotesis untuk pengujian serentak adalah

$$H_0: \beta_d = \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = \beta_8 = \theta = \beta_{65} = 0$$

(tidak ada parameter yang berpengaruh signifikan terhadap model)

$H_1$ : paling tidak ada satu  $\beta_j \neq 0 ; j = d, 1, 2, 3, 4, 5, 6, 7, 8$

Statistik uji yang digunakan adalah  $D(\hat{\beta})$  yang akan dibandingkan dengan nilai *chi-square*. Pada lampiran K.2 diperoleh nilai  $D(\hat{\beta})$  adalah 371,587. Pada taraf signifikansi ( $\alpha$ ) 10 % nilai  $\chi^2_{(11;0,1)}$  adalah 17,275 sehingga kesimpulan yang diambil adalah menolak  $H_0$  karena nilai  $D(\hat{\beta}) > \chi^2_{(11;0,1)}$ . Artinya, paling tidak ada satu variabel prediktor yang berpengaruh signifikan terhadap model. Pengaruh yang diberikan setiap variabel prediktor terhadap variabel respon dapat diketahui melalui pengujian parameter 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)

dan

$H_0: \theta = 0$  (tidak terdapat kasus *overdispersion*)

$H_1: \theta \neq 0$  (terdapat kasus *overdispersion*)

Jika nilai dari  $|Z_{hit}| > Z_{\alpha/2}$  maka keputusan yang diambil adalah menolak  $H_0$  yang berarti parameter tersebut berpengaruh signifikan terhadap model. Dari hasil analisis Tabel 4.12 diketahui bahwa parameter yang signifikan adalah  $\beta_a, \beta_3, \beta_4, \beta_5, \beta_7$  dan  $\theta$ . Hal ini disebabkan nilai  $|Z_{hit}|$  untuk parameter tersebut lebih besar dari nilai  $Z_{0,05}$  yaitu 1,64 sehingga variabel prediktor yang berpengaruh signifikan positif adalah  $X_d, X_3, X_5$  sedangkan variabel prediktor yang berpengaruh signifikan negatif antara lain  $X_4, X_7$ . Jadi model regresi binomial negatif untuk jumlah kasus kusta di Jawa Timur adalah sebagai berikut.

$$\hat{\mu} = \exp(4,0059 + 0,8625 X_d + 0,6909 z_3 - 0,3016 z_4 + 0,5278 z_5 - 0,5839 z_7)$$

Model yang dihasilkan setelah dilakukan *unstandardize* adalah sebagai berikut

$$\hat{\mu} = \exp(19,9911 + 0,8625 X_d + 0,05308 X_3 - 0,4546 X_4 + 0,2351 X_5 - 0,00027 X_7)$$

Hasil pengujian parsial parameter  $\theta$  memperoleh kesimpulan yang sama dengan hasil pengujian *overdispersion* pada regresi poisson dan GPR. Hal tersebut semakin menegaskan bahwa terdapat kasus *overdispersion* pada pemodelan jumlah kasus kusta di Jawa Timur. Selain itu, dapat dikatakan bahwa metode GPR dan regresi binomial negatif lebih baik dalam memodelkan jumlah kasus kusta di Jawa Timur karena mampu menaksir parameter dispersi.

Jadi faktor-faktor yang mempengaruhi jumlah kasus kusta di Jawa Timur pada taraf signifikansi 10 % adalah variabel *dummy* yaitu daerah endemi kusta, persentase keberadaan puskesmas ( $X_3$ ), persentase penduduk laki-laki ( $X_4$ ), persentase penduduk usia 15-29 tahun ( $X_5$ ), tingkat kepadatan penduduk ( $X_7$ ). Variabel *dummy* yang signifikan positif menunjukkan bahwa jumlah kasus kusta di daerah endemi berbeda dengan daerah yang bukan endemi kusta. Dengan kata lain, kabupaten/kota yang termasuk dalam daerah endemi kusta akan memiliki jumlah kasus kusta yang lebih tinggi dibandingkan kabupaten/kota lain. Jumlah kasus kusta yang lebih tinggi di daerah endemi tersebut secara tidak langsung akan meningkatkan jumlah kasus kusta di Jawa Timur.

Peningkatan atau penurunan jumlah kasus kusta di Jawa Timur tergantung dari nilai koefisien masing-masing variabel yang berpengaruh. Dari hasil model regresi binomial negatif yang diperoleh dapat disimpulkan bahwa Jumlah kasus kusta di Jawa Timur akan meningkat sebesar 1 kasus untuk setiap peningkatan masing-masing variabel prediktor sebesar satu satuan dengan asumsi nilai variabel lain tetap. Selain itu, adanya daerah endemi kusta juga akan meningkatkan jumlah kasus kusta sebesar  $\exp(0,8625) \approx 2$  kasus di Jawa Timur.



Selain menggunakan kriteria QAIC<sub>C</sub>, model terbaik juga dapat dipilih berdasarkan kriteria AIC<sub>C</sub> dan BIC terkecil. Model yang terpilih akan berbeda dengan model yang terpilih berdasarkan kriteria QAIC<sub>C</sub>. Hal ini karena kriteria QAIC<sub>C</sub> juga mempertimbangkan adanya parameter dispersi dalam model seperti yang ditunjukkan pada persamaan (2.20). Ringkasan model terbaik untuk masing-masing interaksi antara parameter yang tidak signifikan dengan variabel prediktor lain ditampilkan pada Tabel 4.13 berikut.

**Tabel 4.13** Kemungkinan Model Regresi Binomial Negatif dengan Interaksi Berdasarkan Nilai AIC<sub>C</sub> dan BIC terkecil

Kemungkinan model (Y dengan 8 variabel prediktor dan interaksi Xi)	AIC <sub>C</sub>	BIC	Parameter yang signifikan
X <sub>13</sub>	405,5	411,8	$\beta_0, \beta_d, \beta_3, \beta_4, \beta_7, \beta_{13}, \theta$
X <sub>25</sub>	408,5	414,8	$\beta_0, \beta_d, \beta_4, \beta_5, \beta_7, \theta$
X <sub>63</sub>	405	412,3	$\beta_0, \beta_d, \beta_3, \beta_4, \beta_5, \beta_7, \beta_{63}, \theta$
X <sub>83</sub>	405,5	411,9	$\beta_0, \beta_d, \beta_3, \beta_4, \beta_5, \beta_7, \beta_{83}, \theta$

Model yang terpilih adalah model dengan tambahan variabel interaksi antara persentase penduduk umur 10 tahun keatas dengan tingkat pendidikan SD/MI kebawah dan persentase keberadaan puskesmas (X<sub>83</sub>) dengan nilai AIC<sub>C</sub> sebesar 405,5 dan BIC sebesar 411,9. Model tersebut dipilih karena memiliki jumlah parameter signifikan yang lebih banyak dan memiliki selisih nilai AIC<sub>C</sub> dan BIC yang tidak terlalu besar dengan model yang menghasilkan AIC<sub>C</sub> dan BIC terkecil. Selanjutnya dilakukan pengujian secara serentak dan parsial terhadap parameter model yang ditampilkan lampiran K.3 halaman 86. Hasil analisis menyimpulkan bahwa parameter yang signifikan terhadap model adalah  $\beta_0, \beta_d, \beta_3, \beta_4, \beta_5, \beta_7, \beta_{83}, \theta$  sehingga model regresi binomial negatif yang memiliki nilai AIC<sub>C</sub> dan BIC terkecil adalah

$$\hat{\mu} = \exp(3,9905 + 1,1001 X_d + 0,5091 z_3 - 0,2555 z_4 + 0,4212 z_5 - 0,5718 z_7 - 0,2203 z_{83})$$

sehingga setelah dikembalikan dalam bentuk variabel  $x$  model tersebut menjadi

$$\hat{\mu} = \exp(18,4186 + 1,1001 X_d + 0,0391 - 0,3851 X_4 + 0,1877 X_5 - 0,8190 X_7 - 0,00027 X_{83})$$

Faktor yang mempengaruhi jumlah kasus kusta di Jawa Timur adalah variabel *dummy* ( $X_d$ ), persentase keberadaan puskesmas ( $X_3$ ), persentase penduduk laki-laki ( $X_4$ ), persentase penduduk usia 15-29 tahun ( $X_5$ ), tingkat kepadatan penduduk ( $X_7$ ) serta interaksi antara persentase penduduk umur 10 tahun keatas dengan tingkat pendidikan SD/MI kebawah dan persentase keberadaan puskesmas ( $X_{83}$ ). Faktor yang menyebabkan penambahan jumlah kasus kusta adalah variabel *dummy*, persentase keberadaan puskesmas dan Persentase penduduk usia 15-29 tahun sedangkan penambahan 1% variabel lain yang signifikan akan menyebabkan penurunan jumlah penderita kusta di Jawa Timur.

### 4.3 Pemilihan Model Terbaik

Perbandingan model *generalized poisson regression* dan regresi binomial negatif dilakukan untuk mengetahui model yang lebih baik dalam pemodelan jumlah kasus kusta di Jawa Timur. Kriteria pemilihan model terbaik yang digunakan adalah  $AIC_C$ ,  $BIC$  dan  $QAIC_C$  yang dihitung berdasarkan persamaan (2.18), (2.17) dan (2.20).

**Tabel 4.14** Pemilihan Model Terbaik Berdasarkan  $AIC_C$  dan  $BIC$

Model	Variabel Yang Signifikan	$AIC_C$	$BIC$
GPR	$X_d, X_3, X_{63}$	411,9	418,2
Regresi binomial negatif	$X_d, X_3, X_4, X_5, X_7, X_{83}$	405,5	411,9

Model yang terpilih menggunakan kriteria  $AIC_C$  sama dengan model yang terpilih menggunakan kriteria BIC terkecil. Hal tersebut karena kriteria  $AIC_C$  dan BIC belum memperhitungkan adanya nilai parameter dispersi dalam perhitungan nilainya. Nilai  $AIC_C$  dan BIC tergantung dari jumlah parameter model dan nilai *In-likelihood* dari model sehingga nilai  $AIC_C$  dan BIC akan sebanding dan menghasilkan model yang sama jika digunakan sebagai kriteria pemilihan model. Model terbaik adalah model dengan nilai  $AIC_C$  dan BIC terkecil. Tabel 4.14 menunjukkan bahwa model regresi binomial negatif memiliki nilai  $AIC_C$  dan BIC lebih kecil daripada model GPR sehingga dapat disimpulkan bahwa model regresi binomial negatif lebih baik dalam memodelkan data yang mengalami *overdispersion* daripada GPR.

**Tabel 4.15** Pemilihan Model Terbaik Berdasarkan Nilai  $QAIC_C$

Model	Variabel Yang Signifikan	$QAIC_C$
GPR	$X_3, X_5$	31388,79
Regresi binomial negatif	$X_d, X_3, X_4, X_5, X_7$	1237,13

Model terbaik adalah model yang memiliki nilai  $QAIC_C$  terkecil. Nilai  $QAIC_C$  pada Tabel 4.15 menunjukkan bahwa model dengan nilai  $QAIC_C$  terkecil adalah model regresi binomial negatif. Hal tersebut sesuai dengan analisis sebelumnya yang menyatakan bahwa model regresi binomial negatif. Dengan demikian model terbaik yang lebih sesuai dalam memodelkan kasus *overdispersion* pada regresi poisson untuk jumlah kasus kusta di Jawa Timur adalah model regresi binomial negatif.

Persentase keberadaan puskesmas, persentase penduduk laki-laki, persentase penduduk usia 15-29 tahun, tingkat kepadatan penduduk, dan interaksi antara persentase penduduk umur 10 tahun keatas dengan tingkat pendidikan SD/MI kebawah dan persentase keberadaan puskesmas adalah faktor-

faktor yang mempengaruhi jumlah kasus kusta di Jawa Timur berdasarkan pemilihan model menggunakan kriteria AIC<sub>C</sub> dan BIC. Setiap penambahan persentase keberadaan puskesmas maka akan meningkatkan jumlah kasus kusta di Jawa Timur. Dengan kata lain, persentase keberadaan puskesmas memiliki hubungan searah dengan jumlah kasus kusta di Jawa Timur. Keberadaan puskesmas secara tidak langsung meningkatkan jumlah penderita kusta. Hal ini karena dengan meningkatnya persentase puskesmas maka jumlah penderita kusta yang terdeteksi menjadi bertambah. Penderita kusta yang sebelumnya tidak tercatat atau tidak terdeteksi akan diketahui sehingga dengan adanya peningkatan persentase puskesmas jumlah penderita seolah-olah meningkat. Selain itu, persentase penduduk usia 15-29 tahun juga memiliki hubungan searah dengan jumlah kasus kusta di Jawa Timur. Setiap peningkatan 1% keberadaan puskesmas maka akan meningkatkan 1 kasus kusta di Jawa Timur. Begitu juga dengan peningkatan 1% penduduk usia 15-29 tahun maka kasus kusta di Jawa Timur akan meningkat sebesar 1 kasus. Peningkatan jumlah kasus kusta di Jawa Timur juga dipengaruhi oleh kabupaten/kota yang merupakan daerah endemi kusta karena di daerah-daerah endemi tersebut jumlah kasus kusta yang ditemukan lebih banyak dibandingkan kabupaten/kota lain yang bukan merupakan daerah endemi kusta.

Interaksi antara persentase penduduk umur 10 tahun keatas dengan tingkat pendidikan SD/MI kebawah dan persentase keberadaan puskesmas juga berpengaruh terhadap jumlah kasus kusta di Jawa Timur. Setiap penambahan persentase penduduk umur 10 tahun keatas dengan tingkat pendidikan SD/MI kebawah dan keberadaan puskesmas maka jumlah kasus kusta akan menurun. Dengan kata lain, interaksi antara Persentase penduduk umur 10 tahun keatas dengan tingkat pendidikan SD/MI kebawah dan persentase keberadaan puskesmas memiliki hubungan berlawanan arah dengan jumlah kasus kusta di Jawa Timur. Hal ini menunjukkan bahwa

walaupun persentase penduduk umur 10 tahun keatas dengan tingkat pendidikan SD/MI kebawah meningkat, jumlah penderita kusta bisa menurun jika persentase puskesmas meningkat. Penduduk dengan tingkat pendidikan SD/MI kebawah merupakan kelompok masyarakat yang beresiko terkena kusta. Pendidikan yang rendah memungkinkan mereka tidak memiliki pengathuan yang cukup terkait penyakit kusta baik itu penyebab, tanda-tanda ataupun pencegahannya sehingga rawan terkena kusta. Namun, jika mereka mendapat sosialisasi dan pelayanan kesehatan yang memadai maka kemungkinan terjangkit kusta akan berkurang karena pada dasarnya setiap orang memiliki kekebalan terhadap penyakit kusta. Selain itu, masa inkubasi bakteri penyebab kusta cenderung lama yaitu 3-5 tahun sehingga jika penyakit kusta terdeteksi sejak dini dan dilakukan pengobatan secara rutin orang yang diduga terkena kusta akan dapat disembuhkan sebelum terjadi kecacatan.

Faktor-faktor yang berpengaruh terhadap jumlah penderita kusta di Jawa Timur berdasarkan model yang memiliki nilai QAIC<sub>C</sub> terkecil sama seperti model yang terpilih berdasarkan kriteria AIC<sub>C</sub> dan BIC tetapi faktor interaksi antara persentase penduduk umur 10 tahun keatas dengan tingkat pendidikan SD/MI kebawah dan persentase penduduk usia 15-29 tahun tidak berpengaruh terhadap jumlah kasus kusta di Jawa Timur.

## BAB V

### PENUTUP

#### 5.1 Kesimpulan

Berdasarkan analisis dan pembahasan yang telah dilakukan maka didapatkan kesimpulan sebagai berikut.

1. Jumlah kasus kusta di Jawa Timur pada tahun 2012 menunjukkan tingkat kasus yang masih tinggi dengan rata-rata sebesar 1129 kasus. Kabupaten Sampang merupakan kabupaten dengan jumlah kasus kusta tertinggi yaitu sebesar 589 kasus sedangkan kota batu merupakan kota yang tidak terdapat kasus kusta. Selain itu, nilai varians jumlah kasus kusta sangat tinggi yaitu 22378,5. Hal ini menunjukkan bahwa persebaran kasus kusta masih tinggi di daerah pulau Madura.

2. Hasil pemodelan dari faktor-faktor yang diduga mempengaruhi jumlah kasus kusta di Jawa Timur pada tahun 2012 menggunakan regresi poisson ternyata memberikan hasil adanya pengaruh *overdispersion* sehingga metode *Generalized poisson regression* dan regresi binomial negatif digunakan untuk mengatasi kasus tersebut. Hasil pemodelan untuk masing-masing metode adalah sebagai berikut.

a. Hasil Pemodelan Menggunakan Regresi Poisson.

Model regresi poisson yang diperoleh menggunakan kriteria QAIC<sub>c</sub> adalah

$$\hat{\mu} = \exp(29,58447 + 0,7384 X_d + 0,0059 X_2 + 0,0381 X_3 - 0,6367 X_4 + 0,2267 X_5 - 0,0291 X_6 - 0,00024 X_7 + 0,01159 X_8 - 0,0002 X_{15})$$

Faktor yang berpengaruh terhadap jumlah kasus kusta di Jawa Timur adalah daerah endemi kusta, persentase rumah tangga berperilaku hidup bersih dan sehat, persentase keberadaan puskesmas, persentase penduduk usia 15-19 tahun, persentase penduduk umur 10 tahun keatas dengan tingkat pendidikan SD/MI kebawah, persentase penduduk laki-laki, persentase

penduduk miskin, tingkat kepadatan penduduk ( $X_7$ ) serta interaksi antara persentase antara rumah tangga yang memiliki rumah sehat dan persentase penduduk usia 15-19 tahun.

Jika menggunakan kriteria  $AIC_C$  dan  $BIC$  model yang diperoleh adalah sebagai berikut.

$$\hat{\mu} = \exp(28,3136 + 0,7848 X_d - 0,005 X_1 + 0,0066 X_2 + 0,0314 X_3 - 0,6132 X_4 + 0,2285 X_5 - 0,0316 X_6 - 0,00026 X_7 + 0,01246 X_8 + 0,00031 X_{13})$$

b. Hasil Pemodelan Menggunakan GPR

Faktor yang berpengaruh signifikan terhadap jumlah kasus kusta di Jawa Timur pada tahun 2012 adalah persentase keberadaan puskesmas ( $X_3$ ), persentase penduduk usia 15-29 tahun. Jadi model GPR yang diperoleh menggunakan kriteria  $QAIC_C$  adalah

$$\hat{\mu} = \exp(-2,9807 + 0,0929 X_3 + 0,2116 X_5)$$

Sedangkan model GPR dengan nilai  $AIC_C$  dan  $BIC$  terkecil adalah sebagai berikut.

$$\hat{\mu} = \exp(4,5772 + 1,9368 X_d + 0,0487 X_3 - 0,0031 X_{63})$$

dengan faktor yang signifikan adalah persentase keberadaan puskesmas serta interaksi antara persentase penduduk miskin dan persentase keberadaan puskesmas

c. Hasil Pemodelan Menggunakan Regresi Binomial Negatif

Analisis regresi binomial negatif menghasilkan model terbaik dengan  $QAIC_C$  terkecil sebagai berikut.

$$\hat{\mu} = \exp(19,9911 + 0,8625 X_d + 0,05308 X_3 - 0,4546 X_4 + 0,2351 X_5 - 0,00027 X_7)$$

Berdasarkan model tersebut maka daerah endemi kusta, persentase keberadaan puskesmas, persentase penduduk laki-laki,

persentase penduduk usia 15-29 tahun, tingkat kepadatan penduduk, adalah faktor-faktor yang berpengaruh terhadap peningkatan ataupun penurunan jumlah kasus kusta di Jawa Timur tahun 2012.

Sedangkan faktor yang mempengaruhi jumlah kasus kusta di Jawa Timur berdasarkan kriteria  $AIC_C$  dan  $BIC$  terkecil adalah daerah endemi, persentase keberadaan puskesmas, persentase penduduk laki-laki, persentase penduduk usia 15-29 tahun, tingkat kepadatan penduduk serta interaksi antara persentase penduduk umur 10 tahun keatas dengan tingkat pendidikan SD/MI kebawah dan persentase keberadaan puskesmas. model regresi binomial negatif yang memiliki nilai  $AIC_C$  dan  $BIC$  terkecil adalah

$$\hat{\mu} = \exp(18,4186 + 1,1001 X_d + 0,0391 X_3 - 0,3851 X_4 + 0,1877 X_5 - 0,8190 X_7 - 0,00027 X_{83})$$

3. Model regresi binomial negatif merupakan model terbaik untuk mengatasi kasus *overdispersion* pada regresi poisson karena menghasilkan nilai  $AIC_C$ ,  $BIC$  dan  $QAIC_C$  yang lebih kecil daripada model GPR.

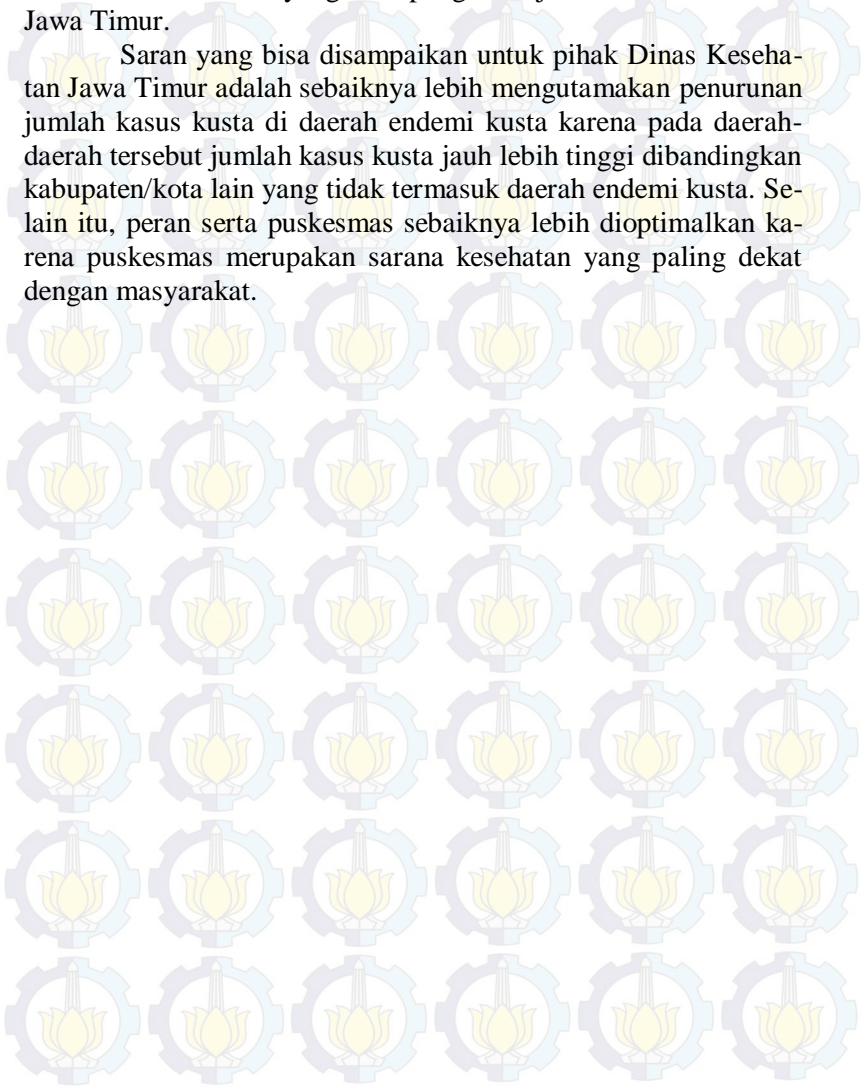
## 5.2 Saran

pada penelitian ini masih banyak permasalahan dan berbagai aspek lain yang belum dikaji secara mendalam salah satunya adalah analisis yang lebih detail terhadap nilai dispersi pada model GPR dan regresi binomial negatif. Selain itu, perlu dikaji lebih mendalam terkait pengaruh aspek geografis suatu daerah, karena ada kemungkinan bahwa kasus kusta dipengaruhi dipengaruhi oleh letak dan kondisi suatu daerah. Oleh karena itu, saran yang bisa direkomendasikan untuk penelitian selanjutnya adalah penggunaan metode yang bisa mengatasi kasus *overdispersion* dengan memperhatikan faktor spasial sehingga dapat memberikan penjelasan yang lebih mendalam terkait faktor yang mempengaruhi jumlah kasus kusta di Jawa Timur. Interaksi untuk lebih dari



dua variabel prediktor juga perlu menjadi pertimbangan dalam analisis faktor-faktor yang mempengaruhi jumlah kasus kusta di Jawa Timur.

Saran yang bisa disampaikan untuk pihak Dinas Kesehatan Jawa Timur adalah sebaiknya lebih mengutamakan penurunan jumlah kasus kusta di daerah endemi kusta karena pada daerah-daerah tersebut jumlah kasus kusta jauh lebih tinggi dibandingkan kabupaten/kota lain yang tidak termasuk daerah endemi kusta. Selain itu, peran serta puskesmas sebaiknya lebih dioptimalkan karena puskesmas merupakan sarana kesehatan yang paling dekat dengan masyarakat.



## DAFTAR PUSTAKA

- Bakker, M. I., *et al.* (2006). *Risk Factor For Developing Leprosy- a Population-Based Cohort Study In Indonesia*. Netherland: KIT Biomedical Research.
- Bernadus, Vanny. (2010). *Informasi kusta dan gejalanya*. Diunduh dari alamat [www.doktersehat.com](http://www.doktersehat.com), Pada Senin, 03 Februari 2014
- BPS Provinsi Jawa Timur. (2012). *Berita Resmi Statistika NO. 45/07/35/Th.X*. Diunduh dari alamat [jatim.bps.go.id](http://jatim.bps.go.id), Pada Selasa 06 Mei 2014
- Burnham, K.P., Anderson, D.R. (2002). *Model Selection and Multimodel Inference: A Prectical Information Theoritic Approach, Second Edition*. USA: Springer
- Cameron, A.C., Trivedi, P.K. (1998). *Regression Analysis of Count Data*. United Kingdom: Cambridge University Press.
- Chamidah, Nur. (2008). *Pemodelan Regresi Binomial Negatif Untuk Mengatasi Overdispersi Pada Regresi Poisson pada Kasus Kematian Ibu Melahirkan di Jawa Timur*. Tugas Akhir, Jurusan Statistika, Institut Teknologi Sepuluh Nopember Surabaya
- Chapra, Steven.C. (2008). *Applied Numerical Methods With Matlab For Engineers and Scientist*. New york: McGraw-Hill.
- Dean, C. (1992). Testing For Detecting Overdispersion In Poisson and Binomial Regression Models. *Journal Of The American Statistical Association. Theory and Methods* 87 (418), 451-457
- Dinas Kesehatan Provinsi Jawa Timur. (2012). *Jatim Berupaya Menekan Kusta*. Diunduh dari alamat <http://kominfo.jatimprov.go.id>, Pada Selasa 20 Mei 2014

- Dinas Kesehatan Provinsi Jawa Timur. (2014). *Peringatan Hari Kusta Sedunia 2014*. Diunduh dari alamat [www.depkes.go.id](http://www.depkes.go.id), Pada Kamis 30 Januari 2014
- Famoye, F., Wulu, J.T. & Singh, K.P. (2004). On the generalized poisson regression model with an application to accident data. *Journal of Data Science*, 2 , 287-295.
- Ghahfarokhi, M.A.B., Iravani, H. & Sepehri, M.R. (2008). Application Of Katz Family Of Distribution For Detecting And Testing Overdispersion In Poisson Regression Models. *World Academy of Science, Engineering and Technology*, 42,514-519
- Hansen, J.A. & Chaignat, C.L. (2013). *Neglected Tropical Disease: Equity and Social Determinant (Bb 8)*. World Health Organization
- Hocking, R. (1996). *Methods and Application of Linear Models*. New York :John Wiley and Sons, Inc.
- Kemntrian Kesehatan RI. (2011). *Petunjuk Teknis Penyusunan Profil Kesehatan Kabupaten/Kota*. Jakarta: Kemntrian Kesehatan RI
- Kerr-Pontes, L.R., *et al.* (2004). Inequality and leprosy in northeast brazil: an ecological study. *International Journal of Epidemiology* 33 (2), 262-269
- Kerr-Pontes, L.R., *et al.* (2006). Socioeconomic, enviromental, and behavioural risk factors for leprosy in North-east Brazil: result of a case-control study. *International Journal of Epidemiology* 35, 994-1000
- 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.

Kusumaningsih, Dian Rahayu. (2012). *Pemodelan Pengaruh Iklim Terhadap Angka Kejadian Demam Berdarah Dengue di Surabaya*. Tugas Akhir, Jurusan Statistika, Institut Teknologi Sepuluh Nopember Surabaya

Myers, R.H.(1990). *Classical and Modern Regression with Application*, Second Edition. Boston: PWS-KENT Publishing Company.

Naik SS, Thakar UH, Phrande AM & Ganapati R. (1999). Survey of leprosy in unapproachable and uncovered areas. *Indian Journal of Leprosy* 71, 333–335.

Norlatifah, Sutomo, A. H., Solikhah. (2010). Hubungan kondisi fisik, sarana air bersih dan karakteristik masyarakat dengan kejadian kusta di kabupaten Tapin Kalimantan Selatan. *Jurnal KES MAS* 4,(3) ,144-239

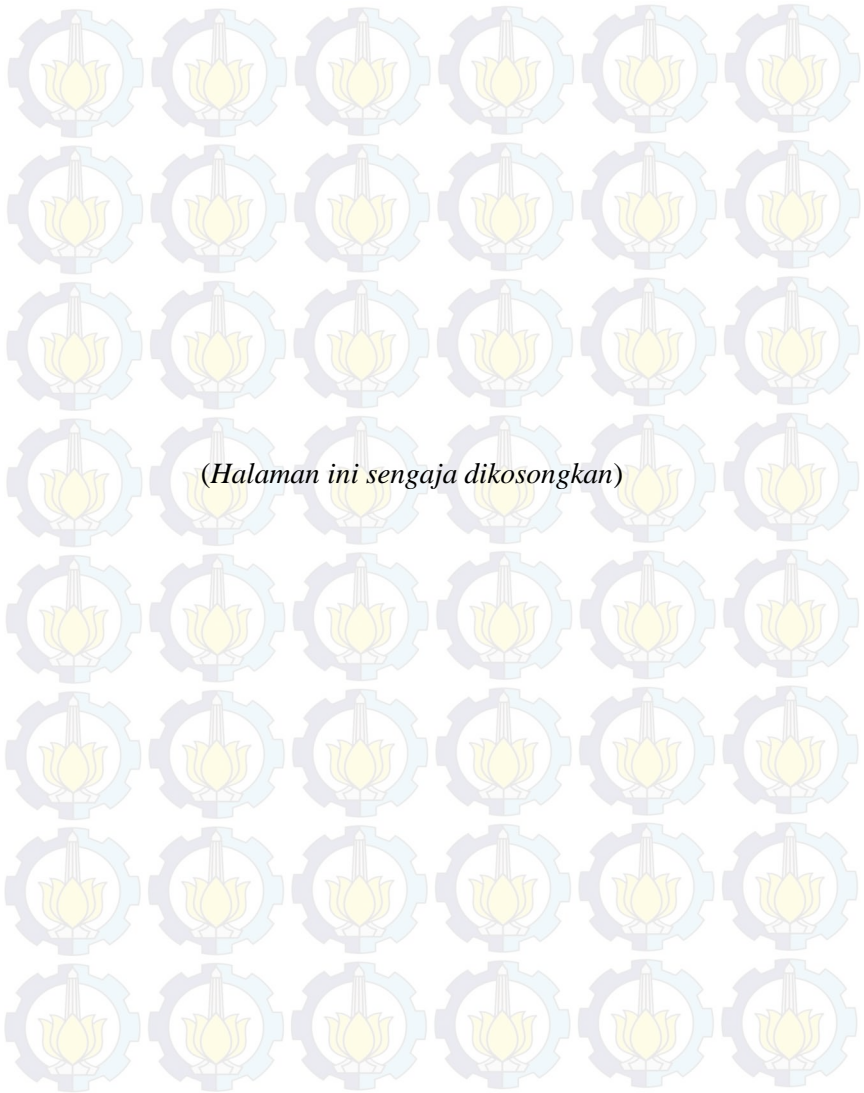
Susanto, Nugroho. (2006). *Faktor-Faktor yang Berhubungan dengan Tingkat Kecacatan Penderita Kusta*. Tesis, Jurusan Ilmu-iIlmu Kesehatan, Universitas Gadjah Mada

Warsini, S. (2007). *Faktor-faktor yang Mempengaruhi Penularan Kusta pada Kontak Serumah di Propinsi DIY dan Kabupaten Klaten*. Tesis, Program Studi Ilmu Kedokteran Tropis, Universitas Gadjah Mada

Winkelmann, R. and Zimmermann, K. F. (1994). Count data models for demographic data .*Mathematical Population Studies*, 4, 205-221.

World Health Organization (WHO). (2013). *Weekly Epidemiological Report*. 83(33): 293-300.

Zucchini, Walter. (2000). An introduction to model selection. *Journal of Mathematical Psychology* 44 : 41-61



*(Halaman ini sengaja dikosongkan)*

## BIOGRAFI PENULIS



Terlahir di Lamongan pada tanggal 29 Juni, Sari Putri merupakan anak kedua dari dua bersaudara dengan kakak perempuan bernama Erlin Dian Saptarini. Putri dari pasangan Bapak Karjiman dan Ibu Rupi'ah ini menempuh jenjang pendidikan formal mulai dari TK Dharma Wanita Ngarum, MI Islamiyah Ngarum, SMPN 3 Babat, SMAN 1 Babat dan pada tahun 2010 ia diterima menjadi mahasiswa Jurusan Statistika ITS. Selain menjalani aktifitas akademik, Sari juga mengasah *softskill* dengan bergabung dalam beberapa organisasi dan kegiatan kemahasiswaan diantaranya Tim Media FORSIS-ITS 2011/2012 dan 2012/2013, Tim Sekretaris BEM FMIPA-ITS 2011/2012, Sekretaris Umum BEM FMIPA-ITS 2012/2013, Tim Soal *Data Analysis Competition* (DAC) 2012, dan kepanitiaan kegiatan kampus lainnya. Sari juga memiliki pengalaman menjadi asisten dosen mata kuliah Pengendalian Kualitas Statistik, menjadi ketua tim Program Kreatifitas Mahasiswa Bidang Penelitian yang didanai Dikti dan menjalani Kerja Praktek di PT Angkasa Pura 1 cabang Bandar Udara Juanda. Apabila pembaca memiliki saran, kritik, atau ingin berdiskusi dengan penulis, silahkan kirim email ke [sariputri01@gmail.com](mailto:sariputri01@gmail.com).

## DAFTAR LAMPIRAN

	Halaman
<b>Lampiran A</b> Jumlah Kasus Kusta Tiap Kabupaten/Kota di Jawa Timur Tahun 2012 .....	69
<b>Lampiran B</b> Variabel Prediktor .....	70
<b>Lampiran C</b> Matriks Korelasi Sembilan Variabel Prediktor.....	74
<b>Lampiran D</b> Nilai VIF Sembilan Variabel Prediktor .....	74
<b>Lampiran E</b> Program SAS Regresi Poisson dan Nilai Devians.....	76
<b>Lampiran F</b> Output SAS untuk Regresi Poisson .....	78
<b>Lampiran G</b> Program SAS untuk GPR.....	80
<b>Lampiran H</b> Output SAS untuk GPR .....	81
<b>Lampiran I</b> Program SAS GPR dengan Interaksi .....	82
<b>Lampiran J</b> Output SAS untuk GPR dengan Interaksi .....	86
<b>Lampiran K</b> Regresi Binomial Negatif.....	89





## LAMPIRAN

### LAMPIRAN A. Jumlah Kasus Kusta Tiap Kabupaten/Kota di Jawa Timur Tahun 2012

No	Kabupaten/Kota	Y	No	Kabupaten/Kota	Y
1	Pacitan	23	20	Ngawi	62
2	Ponorogo	65	21	Bojonegoro	130
3	Trenggalek	19	22	Tuban	243
4	Tulung Agung	27	23	Lamongan	144
5	Blitar	15	24	Gresik	158
6	Kediri	36	25	Bangkalan	447
7	Malang	71	26	Sampang	589
8	Lumajang	190	27	Pamekasan	220
9	Jember	371	28	Sumenep	489
10	Banyuwangi	55	29	Kediri (Ko)	8
11	Bondowoso	20	30	Blitar (Ko)	4
12	Situbondo	291	31	Malang (Ko)	15
13	Probolinggo	288	32	Probolinggo (Ko)	24
14	Pasuruan	215	33	Pasuruan (Ko)	18
15	Sidoarjo	76	34	Mojokerto (Ko)	9
16	Jombang	126	35	Madiun (Ko)	7
17	Nganjuk	110	36	Surabaya (Ko)	152
18	Madiun	33	37	Batu (Ko)	0
19	Magetan	31			

**Lampiran B. Variabel Prediktor**

No	Kabupaten/Kota	$X_d$	$X_1$	$X_2$
1	Pacitan	0	48,89	54,81
2	Ponorogo	0	61,17	35,09
3	Trenggalek	0	60,54	34,35
4	Tulung Agung	0	57,8	35,32
5	Blitar	0	65,18	49,63
6	Kediri	0	56,46	64,89
7	Malang	0	70,55	57,25
8	Lumajang	1	76,55	40,52
9	Jember	1	80,02	63,8
10	Banyuwangi	0	73,76	38,63
11	Bondowoso	0	38,93	14,55
12	Situbondo	1	53,07	18,86
13	Probolinggo	1	38,29	20,05
14	Pasuruan	1	64,05	38,59
15	Sidoarjo	0	68,63	56,93
16	Jombang	0	77,69	45,31
17	Nganjuk	0	50,53	30,91
18	Madiun	0	69,27	46,92
19	Magetan	0	73,93	64,57
20	Ngawi	0	62,7	62,32
21	Bojonegoro	0	77,37	43,49
22	Tuban	1	64,96	53,67
23	Lamongan	0	84,46	45,54
24	Gresik	0	87,17	54,84
25	Bangkalan	1	81,18	39,69
26	Sampang	1	51,92	29,09
27	Pamekasan	1	51,24	8,5
28	Sumenep	1	56,29	59,99
29	Kediri (Ko)	0	83,3	65,74
30	Blitar (Ko)	0	81,93	30,44
31	Malang (Ko)	0	86,84	36,07
32	Probolinggo (Ko)	0	78,62	52,19
33	Pasuruan (Ko)	0	68,79	38,52
34	Mojokerto (Ko)	0	81,12	53,9
35	Madiun (Ko)	0	72,41	44,46
36	Surabaya (Ko)	0	80,99	62,97
37	Batu (Ko)	0	67,8	31,48

**LAMPIRAN B. Variabel prediktor (lanjutan)**

No	Kabupaten/Kota	$X_3$	$X_4$	$X_5$
1	Pacitan	24,00	48,69	20,19
2	Ponorogo	31,00	49,86	20,67
3	Trenggalek	22,00	49,56	21,47
4	Tulung Agung	31,00	48,62	21,45
5	Blitar	24,00	49,96	21,01
6	Kediri	37,00	50,06	22,44
7	Malang	39,00	50,15	23,18
8	Lumajang	25,00	48,71	22,23
9	Jember	49,00	49,04	23,46
10	Banyuwangi	45,00	49,65	21,52
11	Bondowoso	25,00	48,60	21,70
12	Situbondo	17,00	48,68	24,54
13	Probolinggo	33,00	48,69	23,42
14	Pasuruan	33,00	49,43	25,57
15	Sidoarjo	26,00	50,13	24,92
16	Jombang	34,00	49,63	23,74
17	Nganjuk	20,00	49,59	22,02
18	Madiun	26,00	49,23	19,61
19	Magetan	22,00	48,57	19,74
20	Ngawi	24,00	48,58	20,46
21	Bojonegoro	36,00	49,32	21,88
22	Tuban	33,00	49,28	23,44
23	Lamongan	33,00	48,46	22,36
24	Gresik	32,00	49,47	24,56
25	Bangkalan	22,00	47,65	26,28
26	Sampang	21,00	48,62	27,64
27	Pamekasan	20,00	48,49	26,56
28	Sumenep	30,00	47,46	22,35
29	Kediri (Ko)	9,00	49,74	26,12
30	Blitar (Ko)	3,00	49,45	23,64
31	Malang (Ko)	15,00	49,21	28,72
32	Probolinggo (Ko)	6,00	49,13	24,53
33	Pasuruan (Ko)	8,00	49,46	25,52
34	Mojokerto (Ko)	5,00	49,06	24,06
35	Madiun (Ko)	6,00	48,25	22,68
36	Surabaya (Ko)	62,00	49,33	26,54
37	Batu (Ko)	5,00	50,19	24,02

**LAMPIRAN B. Variabel prediktor (lanjutan)**

No	Kabupaten/Kota	$X_6$	$X_7$	$X_8$
1	Pacitan	32,76	383,21	66,82
2	Ponorogo	29,83	576,75	60,30
3	Trenggalek	30,37	545,28	61,84
4	Tulung Agung	21,93	870,65	53,84
5	Blitar	21,54	624,64	61,51
6	Kediri	23,93	997,45	56,03
7	Malang	21,43	719,44	62,52
8	Lumajang	31,99	561,78	72,67
9	Jember	29,17	705,34	70,10
10	Banyuwangi	17,94	435,93	60,12
11	Bondowoso	25,25	474,52	73,29
12	Situbondo	24,23	397,03	70,82
13	Probolinggo	26,34	672,25	73,53
14	Pasuruan	23,86	1037,55	65,24
15	Sidoarjo	3,47	2815,96	33,66
16	Jombang	18,21	1092,96	51,73
17	Nganjuk	25,11	798,69	57,85
18	Madiun	18,82	659,12	57,82
19	Magetan	19,87	881,24	55,38
20	Ngawi	28,20	588,27	61,82
21	Bojonegoro	22,03	526,56	63,72
22	Tuban	24,30	572,53	65,76
23	Lamongan	15,74	678,64	55,87
24	Gresik	11,16	980,17	42,28
25	Bangkalan	41,91	711,77	77,52
26	Sampang	32,15	735,21	85,76
27	Pamekasan	38,99	1027,18	70,61
28	Sumenep	34,71	503,65	79,11
29	Kediri (Ko)	11,17	3966,36	36,10
30	Blitar (Ko)	13,60	4131,23	39,03
31	Malang (Ko)	3,16	7587,52	31,84
32	Probolinggo (Ko)	14,62	4118,76	47,48
33	Pasuruan (Ko)	11,41	5001,18	45,27
34	Mojokerto (Ko)	5,32	6127,50	32,25
35	Madiun (Ko)	11,85	5071,21	31,19
36	Surabaya (Ko)	2,75	8463,47	35,11
37	Batu (Ko)	9,65	51,07	51,07

**Keterangan :**

Y : Jumlah penderita kusta

X<sub>d</sub> : Variabel *dummy* yaitu daerah endemi kusta

X<sub>1</sub> : Persentase rumah tangga yang memiliki rumah sehat

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

X<sub>3</sub> : Persentase keberadaan puskesmas

X<sub>4</sub> : Persentase penduduk laki-laki

X<sub>5</sub> : Persentase penduduk usia 15-29 tahun

X<sub>6</sub> : Persentase penduduk miskin

X<sub>7</sub> : Tingkat kepadatan penduduk

X<sub>8</sub> : Persentase penduduk umur 10 tahun keatas dengan tingkat pendidikan SD/MI kebawah

### LAMPIRAN C. Matrik Korelasi Delapan Variabel Prediktor

```
MTB > name m1 "CORR1"
MTB > Correlation 'x1'-'x8' 'CORR1'
MTB > Print CORR1
```

#### Data Display

Matrix CORR1

1,00000	0,48780	-0,02628	0,16972
0,48780	1,00000	0,27023	0,15632
-0,02628	0,27023	1,00000	0,07808
0,16972	0,15632	0,07808	1,00000
0,26006	-0,17567	-0,14801	0,00406
-0,53187	-0,27335	0,16816	-0,47850
0,47172	0,19026	-0,22433	0,08375
-0,58139	-0,32467	0,28902	-0,41691
0,26006	-0,53187	0,47172	-0,58139
-0,17567	-0,27335	0,19026	-0,32467
-0,14801	0,16816	-0,22433	0,28902
0,00406	-0,47850	0,08375	-0,41691
1,00000	-0,25873	0,52839	-0,20917
-0,25873	1,00000	-0,70340	0,89363
0,52839	-0,70340	1,00000	-0,76597
-0,20917	0,89363	-0,76597	1,00000

### LAMPIRAN D. Nilai VIF Sembilan Variabel Prediktor

```
MTB > name m2 "VIF1"
MTB > invert 'corr1' 'VIF1'
MTB > Print 'VIF1'
```

#### Data Display

Matrix VIF1

2,06173	-0,79936	-0,12828	0,34196
-0,79936	1,75262	-0,53524	-0,03657
-0,12828	-0,53524	1,47439	-0,38565
0,34196	-0,03657	-0,38565	1,79043
-0,69665	0,59219	0,18285	-0,41920
0,12078	-0,26736	0,51965	0,77498
0,57554	-0,26567	-0,41589	1,39979
1,30599	0,40302	-1,57995	1,33682
-0,69665	0,12078	0,57554	1,30599
0,59219	-0,26736	-0,26567	0,40302
0,18285	0,51965	-0,41589	-1,57995
-0,41920	0,77498	1,39979	1,33682
2,03663	0,31625	-1,90932	-1,75947
0,31625	5,78289	0,38285	-4,65202
-1,90932	0,38285	4,85874	3,93232
-1,75947	-4,65202	3,93232	9,70531

**LAMPIRAN E. Program SAS Regresi Poisson dan Nilai Devians**  
**E.1 Macro SAS Regresi Poisson Y dengan  $X_1 X_2 X_3 X_4 X_5 X_6 X_7 X_8$**

```

data poisson;
input y xd x1 x2 x3 x4 x5 x6 x7 x8;
datalines;
23      0      -1.4278  0.7302  -0.0934  -0.6692  -1.4108  1.1765
        -0.6433  0.6573
65      0      -0.4950  -0.5883  0.4443   1.0911  -1.1961  0.8821
        -0.5546  0.2112
19      0      -0.5428  -0.6378  -0.2471  0.6363  -0.8408  0.9357
        -0.5690  0.3166
.
.
.
7       0       0.3588  0.0382  -1.4762  -1.3307  -0.3016  -0.9239
        1.5062  -1.7803
152    0       1.0105  1.2758  2.8257   0.2999  1.4168  -1.8372
        3.0616  -1.5122
0      0       0.0086  -0.8297  -1.5530  1.5890  0.2937  -1.1447
        -0.7956  -0.4202
;
run;
title 'poisson model';
proc nlmixed data=poisson start hess tech=newrap alpha=0.1;
parms a0=0 ad=0 a1=0 a2=0 a3=0 a4=0 a5=0 a6=0 a7=0 a8=0;
eta =a0+ad*xd+a1*x1+a2*x2+a3*x3+a4*x4+a5*x5+a6*x6+a7*x7
+a8*x8;
lambda=exp(eta);
model y ~ poisson(lambda);
run;

```

**E.2 Macro SAS Regresi Poisson Y dengan  $X_1 X_2 X_3 X_4 X_5 X_6 X_7 X_8 X_9 X_{15}$**

```

data poisson;
input y xd x1 x2 x3 x4 x5 x6 x7 x8 x9 x15;
datalines;
23      0      -1.4278  0.7302  -0.0934  -0.6692  -1.4108  1.1765
        -0.6433  0.6573  2.0143

```



```

65      0      -0.4950 -0.5883 0.4443  1.0911 -1.1961 0.8821
      -0.5546 0.2112 0.5921
19      0      -0.5428 -0.6378 -0.2471 0.6363 -0.8408 0.9357
      -0.5690 0.3166 0.4564
.
.
.
152    0      1.0105 1.2758 2.8257 0.2999 1.4168 -1.8372
      3.0616 -1.5122 1.4317
0      0      0.0086 -0.8297 -1.5530 1.5890 0.2937 -1.1447
      -0.7956 -0.4202 0.0025;

run;
title 'poisson model';
proc nlmixed data=poisson start hess tech=newrap alpha=0.1;
parms a0=0 ad=0 a1=0 a2=0 a3=0 a4=0 a5=0 a6=0 a7=0 a8=0 a9=0
a15=0;
eta =a0+ad*xd+a1*x1+a2*x2+a3*x3+a4*x4+a5*x5+a6*x6+a7*x7
+a8*x8+a9*x9+a15*x15;
lambda=exp(eta);
model y ~ poisson(lambda);
run;

```

### E.3 Macro SAS Regresi Poisson Y dengan $X_1$ $X_2$ $X_3$ $X_4$ $X_5$ $X_6$ $X_7$ $X_8$ $X_9$ $X_{13}$

```

data poisson;
input y xd x1 x2 x3 x4 x5 x6 x7 x8 x9 x13;
datalines;
23      0      -1.4278 0.7302 -0.0934 -0.6692 -1.4108 1.1765
      -0.6433 0.6573 0.1334
65      0      -0.4950 -0.5883 0.4443  1.0911 -1.1961 0.8821
      -0.5546 0.2112 -0.2199
19      0      -0.5428 -0.6378 -0.2471 0.6363 -0.8408 0.9357
      -0.5690 0.3166 0.1341
.
.
.
152    0      1.0105 1.2758 2.8257 0.2999 1.4168 -1.8372
      3.0616 -1.5122 2.8555
0      0      0.0086 -0.8297 -1.5530 1.5890 0.2937 -1.1447

```

```

-0.7956 -0.4202 -0.0134;
run;
title 'poisson model';
proc nlmixed data=poisson start hess tech=newwrap alpha=0.1;
parms a0=0 ad=0 a1=0 a2=0 a3=0 a4=0 a5=0 a6=0 a7=0 a8=0 a9=0
a13=0;
eta=a0+ad*xd+a1*x1+a2*x2+a3*x3+a4*x4+a5*x5+a6*x6+a7*x7
+a8*x8+a9*x9+a13*x13;
lambda=exp(eta);
model y ~ poisson(lambda);
run;

```

#### E.4 Macro SAS Nilai Devians Y dengan $X_1 X_2 X_3 X_4 X_5 X_6 X_7 X_8 X_{15}$

```

data deviance;
input y xd x1 x2 x3 x4 x5 x6 x7 x8 x15;
cards;
23 0 -1.4278 0.7302 -0.0934 -0.6692 -1.4108 1.1765
-0.6433 0.6573 2.0143
65 0 -0.4950 -0.5883 0.4443 1.0911 -1.1961 0.8821
-0.5546 0.2112 0.5921
19 0 -0.5428 -0.6378 -0.2471 0.6363 -0.8408 0.9357
-0.5690 0.3166 0.4564
.
.
.
152 0 1.0105 1.2758 2.8257 0.2999 1.4168 -1.8372
3.0616 -1.5122 1.4317
0 0 0.0086 -0.8297 -1.5530 1.5890 0.2937 -1.1447
-0.7956 -0.4202 0.0025;
run;
proc genmod data=deviance;
model y = xd x1 x2 x3 x4 x5 x6 x7 x8 x15/dist = poisson
link = log
type1
type3 wald
scale=deviance;
run;

```

### E.5 Macro SAS Nilai Devians Y dengan $X_1$ $X_2$ $X_3$ $X_4$ $X_5$ $X_6$ $X_7$ $X_8$ $X_9$ $X_{13}$

```

data deviance;
input y xd x1 x2 x3 x4 x5 x6 x7 x8 x9 x13;
cards;
23    0    -1.4278 0.7302 -0.0934 -0.6692 -1.4108 1.1765
      -0.6433 0.6573 0.1334
65    0    -0.4950 -0.5883 0.4443 1.0911 -1.1961 0.8821
      -0.5546 0.2112 -0.2199
19    0    -0.5428 -0.6378 -0.2471 0.6363 -0.8408 0.9357
      -0.5690 0.3166 0.1341
.
.
.
152   0    1.0105 1.2758 2.8257 0.2999 1.4168 -1.8372
      3.0616 -1.5122 2.8555
0     0    0.0086 -0.8297 -1.5530 1.5890 0.2937 -1.1447 -
0.7956 -0.4202 -0.0134;
run;
proc genmod data=deviance;
  model y = xd x1 x2 x3 x4 x5 x6 x7 x8 x9 x13/dist = poisson
  link = log
  type1
  type3 wald
  scale=deviance;
run;

```

## LAMPIRAN F. Output SAS untuk Regresi Poisson

### F.1 Regresi Poisson Y dengan $X_1$ $X_2$ $X_3$ $X_4$ $X_5$ $X_6$ $X_7$ $X_8$

```

The NL MIXED Procedure
Fit Statistics
-2 Log Likelihood           810.2
AIC (smaller is better)    830.2
AICC (smaller is better)   838.6
BIC (smaller is better)    846.3

poisson model              14:57 Thursday, July 10, 2014 3

The NL MIXED Procedure

Parameter Estimates

```

Parameter	Estimate	Standard Error	DF	t Value	Pr> t	Alpha	Lower	Upper	Gradient
a0	4.0971	0.02921	37	140.27	<.0001	0.1	4.0478	4.1464	4.174E-9
ad	0.7338	0.05680	37	12.92	<.0001	0.1	0.6380	0.8296	2.333E-9
a1	-0.02425	0.02244	37	-1.08	0.2870	0.1	-0.06211	0.01362	-326E-12
a2	0.0815	0.02142	37	3.81	0.0005	0.1	0.04541	0.1177	-437E-15
a3	0.4884	0.02465	37	19.82	<.0001	0.1	0.4469	0.5300	1.36E-9
a4	-0.3965	0.02313	37	-17.14	<.0001	0.1	-0.4355	-0.3574	-1.14E-9
a5	0.5362	0.02861	37	18.75	<.0001	0.1	0.4880	0.5845	6.12E-10
a6	-0.3336	0.03688	37	-9.04	<.0001	0.1	-0.3958	-0.2713	1.891E-9
a7	-0.5459	0.04362	37	-12.51	<.0001	0.1	-0.6195	-0.4723	-1.19E-9
a8	0.2420	0.04950	37	4.89	<.0001	0.1	0.1584	0.3255	2.239E-9

## F.2 Regresi Poisson Y dengan $X_1 X_2 X_3 X_4 X_5 X_6 X_7 X_8 X_9 X_{15}$

Parameter	Estimate	Standard Error	DF	t Value	Pr> t	Alpha	Lower	Upper	Gradient
a0	4.1156	0.02965	37	138.82	<.0001	0.1	4.0656	4.1656	0.000012
ad	0.7384	0.05721	37	12.91	<.0001	0.1	0.6419	0.8349	0.000011
a1	-0.01464	0.02247	37	-0.65	0.5186	0.1	-0.05255	0.02326	-0.000001
a2	0.08893	0.02152	37	4.13	0.0002	0.1	0.05262	0.1252	-0.00001
a3	0.4959	0.02456	37	20.19	<.0001	0.1	0.4544	0.5373	-3.37E-6
a4	-0.4224	0.02381	37	-17.74	<.0001	0.1	-0.4625	-0.3822	-8.36E-6
a5	0.5087	0.02908	37	17.49	<.0001	0.1	0.4597	0.5578	0.000021
a6	-0.2900	0.03781	37	-7.67	<.0001	0.1	-0.3538	-0.2262	0.000012
a7	-0.5200	0.04394	37	-11.83	<.0001	0.1	-0.5941	-0.4458	-5.48E-6
a8	0.1695	0.05171	37	3.28	0.0023	0.1	0.08231	0.2568	0.000022
a15	-0.07849	0.01896	37	-4.14	0.0002	0.1	-0.1105	-0.04649	-0.00002

## F.3 Regresi Poisson Y dengan $X_1 X_2 X_3 X_4 X_5 X_6 X_7 X_8 X_9 X_{13}$

Parameter	Estimate	Standard Error	DF	t Value	Pr> t	Alpha	Lower	Upper	Gradient
a0	4.1156	0.02965	37	138.82	<.0001	0.1	4.0656	4.1656	0.000012
ad	0.7384	0.05721	37	12.91	<.0001	0.1	0.6419	0.8349	0.000011
a1	-0.01464	0.02247	37	-0.65	0.5186	0.1	-0.05255	0.02326	-0.000001
a2	0.08893	0.02152	37	4.13	0.0002	0.1	0.05262	0.1252	-0.00001
a3	0.4959	0.02456	37	20.19	<.0001	0.1	0.4544	0.5373	-3.37E-6
a4	-0.4224	0.02381	37	-17.74	<.0001	0.1	-0.4625	-0.3822	-8.36E-6
a5	0.5087	0.02908	37	17.49	<.0001	0.1	0.4597	0.5578	0.000021
a6	-0.2900	0.03781	37	-7.67	<.0001	0.1	-0.3538	-0.2262	0.000012
a7	-0.5200	0.04394	37	-11.83	<.0001	0.1	-0.5941	-0.4458	-5.48E-6
a8	0.1695	0.05171	37	3.28	0.0023	0.1	0.08231	0.2568	0.000022
a13	-0.07849	0.01896	37	-4.14	0.0002	0.1	-0.1105	-0.04649	-0.00002

Parameter	Estimate	Error	DF	t Value	Pr> t	Alpha	Lower	Upper	Gradient
a0	4.0791	0.02945	37	138.50	<.0001	0.1	4.0294	4.1288	3.973E-9
ad	0.7848	0.05738	37	13.68	<.0001	0.1	0.6880	0.8817	2.159E-9
a1	-0.06646	0.02385	37	-2.79	0.0084	0.1	-0.1067	-0.02622	-213E-12
a2	0.09875	0.02184	37	4.52	<.0001	0.1	0.06191	0.1356	3.84E-13
a3	0.4087	0.02860	37	14.29	<.0001	0.1	0.3604	0.4569	9.44E-10
a4	-0.4068	0.02330	37	-17.46	<.0001	0.1	-0.4461	-0.3675	-1.33E-9
a5	0.5128	0.02891	37	17.74	<.0001	0.1	0.4641	0.5616	6.01E-10
a6	-0.3147	0.03627	37	-8.68	<.0001	0.1	-0.3759	-0.2536	1.879E-9
a7	-0.5650	0.04375	37	-12.91	<.0001	0.1	-0.6388	-0.4911	-1.14E-9
a8	0.1821	0.05008	37	3.64	0.0008	0.1	0.09766	0.2666	2.099E-9
a13	0.1359	0.02420	37	5.61	<.0001	0.1	0.09505	0.1767	6.03E-10

#### F.4 Nilai Devians Y dengan $X_1 X_2 X_3 X_4 X_5 X_6 X_7 X_8 X_{15}$

The SAS System		13:29 Thursday, July 13, 2014		7
The GENMOD Procedure				
Model Information				
Data Set	WORK.DEVIANCE			
Distribution	Poisson			
Link Function	Log			
Dependent Variable	y			
Number of Observations Read				37
Number of Observations Used				37
Criteria For Assessing Goodness Of Fit				
Criterion	DF	Value	Value/DF	
Deviance	26	576.8286	22.1857	
Scaled Deviance	26	26.0000	1.0000	
Pearson Chi-Square	26	639.0091	24.5773	
Scaled Pearson X2	26	28.8027	1.1078	
Log Likelihood	942.2239			

#### F.5 Nilai Devians Y dengan $X_1 X_2 X_3 X_4 X_5 X_6 X_7 X_8 X_9 X_{13}$

The SAS System		13:29 Thursday, July 13, 2014		3
The GENMOD Procedure				
Model Information				
Data Set	WORK.DEVIANCE			
Distribution	Poisson			
Link Function	Log			
Dependent Variable	y			
Number of Observations Read				37
Number of Observations Used				37
Criteria For Assessing Goodness Of Fit				
Criterion	DF	Value	Value/DF	
Deviance	26	562.2468	21.6249	

Scaled Deviance	26	26.0000	1.0000
Pearson Chi-Square	26	573.8920	22.0728
Scaled Pearson X2	26	26.5385	1.0207
Log Likelihood		966.9975	

### LAMPIRAN G. Program SAS untuk (GPR)

```

data GPR;
input y xd x1 x2 x3 x4 x5 x6 x7 x8;
datalines;
23 0 -1.4278 0.7302 -0.0934 -0.6692 -1.4108 1.1765
-0.6433 0.6573
65 0 -0.4950 -0.5883 0.4443 1.0911 -1.1961 0.8821
-0.5546 0.2112
19 0 -0.5428 -0.6378 -0.2471 0.6363 -0.8408 0.9357
-0.5690 0.3166
.
.
.
7 0 0.3588 0.0382 -1.4762 -1.3307 -0.3016 -0.9239
1.5062 -1.7803
152 0 1.0105 1.2758 2.8257 0.2999 1.4168 -1.8372
3.0616 -1.5122
0 0 0.0086 -0.8297 -1.5530 1.5890 0.2937 -1.1447
-0.7956 -0.4202
;
run;
/*Generalized Poisson Regression Model*/
title 'Model Generalized Poisson Regression';
proc nlmixed data=GPR tech=newwrap maxiter=500 maxfunc=5000
alpha=0.1;
parms a0=0 ad=0 a1=0 a2=0 a3=0 a4=0 a5=0 a6=0 a7=0 a8=0 teta=0;
lambda= exp(a0+ad*xd+a1*x1+a2*x2+a3*x3+a4*x4+a5*x5+a6*x6
+a7*x7+a8*x8);
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 H. Output SAS untuk GPR

Model Generalized Poisson Regression  
20:08 Thursday, July 10, 2014

The NL MIXED Procedure

NOTE: GCONV convergence criterion satisfied.

Fit Statistics

-2 Log Likelihood	386.9
AIC (smaller is better)	408.9
AICC (smaller is better)	419.4
BIC (smaller is better)	426.6

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr> t	Alpha	Lower	Upper	Gradient
a0	4.0219	0.2128	37	18.90	<.0001	0.1	3.6628	4.3809	-9E-6
ad	1.1847	0.6928	37	1.71	0.0956	0.1	0.01595	2.3535	0.000037
a1	-0.1324	0.2066	37	-0.64	0.5254	0.1	-0.4810	0.2161	3.799E-6
a2	0.08523	0.1912	37	0.45	0.6583	0.1	-0.2373	0.4077	1.776E-7
a3	0.9379	0.3036	37	3.09	0.0038	0.1	0.4257	1.4501	-1.64E-6
a4	-0.2583	0.2004	37	-1.29	0.2054	0.1	-0.5964	0.07980	5.864E-6
a5	0.4690	0.2243	37	2.09	0.0435	0.1	0.09058	0.8474	-6.87E-6
a6	0.3163	0.3612	37	0.88	0.3869	0.1	-0.2931	0.9257	2.694E-6
a7	-0.3157	0.3952	37	-0.80	0.4294	0.1	-0.9824	0.3510	-7.52E-8
a8	-0.2803	0.4073	37	-0.69	0.4956	0.1	-0.9675	0.4069	-8.89E-6
teta	0.05923	0.01203	37	4.92	<.0001	0.1	0.03893	0.07953	-2.96E-6

## LAMPIRAN I. Program SAS GPR dengan Interaksi

### I.1 Program SAS GPR Y dengan $X_1$ $X_2$ $X_3$ $X_4$ $X_5$ $X_6$ $X_7$ $X_8$ $X_9$ $X_{75}$

```

data GPR;
input y xd x1 x2 x3 x4 x5 x6 x7 x8 x9 x75;
datalines;
23 0 -1.4278 0.7302 -0.0934 -0.6692 -1.4108 1.1765
-0.6433 0.6573 0.9076
65 0 -0.4950 -0.5883 0.4443 1.0911 -1.1961 0.8821
-0.5546 0.2112 0.6634
19 0 -0.5428 -0.6378 -0.2471 0.6363 -0.8408 0.9357
-0.5690 0.3166 0.4784
.
.
.
152 0 1.0105 1.2758 2.8257 0.2999 1.4168 -1.8372
3.0616 -1.5122 4.3375
0 0 0.0086 -0.8297 -1.5530 1.5890 0.2937 -1.1447
-0.7956 -0.4202 -0.2337;
run;

```

```

/*Generalized Poisson Regression Model*/
title 'Model Generalized Poisson Regression';
proc nlmixed data=GPR tech=newwrap maxiter=500 maxfunc=5000
alpha=0.1;
parms a0=0 ad=0 a1=0 a2=0 a3=0 a4=0 a5=0 a6=0 a7=0 a8=0 a9=0
a75=0 teta=0;
lambda= exp(a0+ad*xd+a1*x1+a2*x2+a3*x3+a4*x4+a5*x5+a6*x6
+a7*x7+a8*x8+a9*x9+a75*x75);
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;

```

## I.2 Program SAS GPR Y dengan $X_1 X_2 X_3 X_4 X_5 X_6 X_7 X_8 X_9 X_{63}$

```

data GPR;
input y xd x1 x2 x3 x4 x5 x6 x7 x8 x9 x63;
datalines;
23 0 -1.4278 0.7302 -0.0934 -0.6692 -1.4108 1.1765
-0.6433 0.6573 -0.1099
65 0 -0.4950 -0.5883 0.4443 1.0911 -1.1961 0.8821
-0.5546 0.2112 0.3919
19 0 -0.5428 -0.6378 -0.2471 0.6363 -0.8408 0.9357
-0.5690 0.3166 -0.2312
.
.
.
0 0 0.0086 -0.8297 -1.5530 1.5890 0.2937 -1.1447
-0.7956 -0.4202 1.7777
;
run;
/*Generalized Poisson Regression Model*/
title 'Model Generalized Poisson Regression';
proc nlmixed data=GPR tech=newwrap maxiter=500 maxfunc=5000
alpha=0.1;
parms a0=0 ad=0 a1=0 a2=0 a3=0 a4=0 a5=0 a6=0 a7=0 a8=0 a9=0
a63=0 teta=0;
lambda= exp(a0+ad*xd+a1*x1+a2*x2+a3*x3+a4*x4+a5*x5+a6*x6
+a7*x7+a8*x8+a9*x9+a63*x63);

```



```
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 J. Output SAS untuk GPR dengan Interaksi J.1 GPR Y dengan X<sub>1</sub> X<sub>2</sub> X<sub>3</sub> X<sub>4</sub> X<sub>5</sub> X<sub>6</sub> X<sub>7</sub> X<sub>8</sub> X<sub>9</sub> X<sub>75</sub>

```
Model Generalized Poisson Regression
10:57 Thursday, July 10, 2014

The NL MIXED Procedure

NOTE: GCONV convergence criterion satisfied.
```

Fit Statistics

-2 Log Likelihood	385.0
AIC (smaller is better)	409.0
AICC (smaller is better)	422.0
BIC (smaller is better)	428.4

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr> t	Alpha	Lower	Upper	Gradient
a0	4.3049	0.3819	37	11.27	<.0001	0.1	3.6606	4.9492	-7.04E-6
ad	0.9441	0.7908	37	1.19	0.2401	0.1	-0.3901	2.2783	0.000011
a1	-0.1921	0.2377	37	-0.81	0.4242	0.1	-0.5932	0.2090	2.3E-6
a2	0.08317	0.2076	37	0.40	0.6910	0.1	-0.2671	0.4334	-3.58E-7
a3	1.2090	0.4015	37	3.01	0.0047	0.1	0.5317	1.8863	-5.15E-6
a4	-0.2499	0.2333	37	-1.07	0.2910	0.1	-0.6436	0.1437	6.42E-7
a5	0.4750	0.2387	37	1.99	0.0541	0.1	0.07222	0.8777	-1.06E-6
a6	0.2656	0.3919	37	0.68	0.5021	0.1	-0.3956	0.9268	3.60E-6
a7	0.1431	0.5707	37	0.25	0.8034	0.1	-0.8198	1.1059	-7.89E-6
a8	-0.00908	0.5673	37	-0.02	0.9873	0.1	-0.9661	0.9480	-9.8E-6
a75	-0.2019	0.1597	37	-1.26	0.2141	0.1	-0.4714	0.06756	2.183E-6
teta	0.06156	0.01228	37	5.01	<.0001	0.1	0.04083	0.08228	-1.08E-6

## J.2 GPR Y dengan X<sub>1</sub> X<sub>2</sub> X<sub>3</sub> X<sub>4</sub> X<sub>5</sub> X<sub>6</sub> X<sub>7</sub> X<sub>8</sub> X<sub>9</sub> X<sub>63</sub>

```
Model Generalized Poisson Regression
10:57 Thursday, July 10, 2014

The NL MIXED Procedure

NOTE: GCONV convergence criterion satisfied.
```

Fit Statistics

-2 Log Likelihood	374.9
AIC (smaller is better)	398.9
AICC (smaller is better)	411.9

		BIC (smaller is better)				418.2			
								Parameter Estimates	
Parameter	Estimate	Standard Error	DF	t Value	Pr> t	Alpha	Lower	Upper	Gradient
a0	4.1096	0.1760	37	23.34	<.0001	0.1	3.8126	4.4066	5.508E-7
ad	1.9368	0.7021	37	2.76	0.0090	0.1	0.7523	3.1212	-5.75E-6
a1	0.1086	0.1851	37	0.59	0.5609	0.1	-0.2037	0.4209	-6.27E-7
a2	0.04305	0.1520	37	0.28	0.7785	0.1	-0.2133	0.2994	1.263E-7
a3	0.6347	0.2159	37	2.94	0.0056	0.1	0.2705	0.9989	4.291E-7
a4	0.009729	0.1855	37	0.05	0.9585	0.1	-0.3033	0.3228	-1.11E-6
a5	0.04276	0.2834	37	0.15	0.8809	0.1	-0.4353	0.5209	1.983E-6
a6	0.2174	0.3501	37	0.62	0.5384	0.1	-0.3732	0.8081	5.19E-7
a7	0.6121	0.4360	37	1.40	0.1687	0.1	-0.1234	1.3477	-2.29E-6
a8	0.05172	0.3853	37	0.13	0.8939	0.1	-0.5984	0.7018	-5.17E-7
a63	-1.0755	0.3114	37	-3.45	0.0014	0.1	-1.6008	-0.5501	1.819E-6
teta	0.05218	0.009278	37	5.62	<.0001	0.1	0.03653	0.06784	-2.17E-7

## LAMPIRAN K .Regresi Binomial Negatif

### K.1 Y dengan $X_1 X_2 X_3 X_4 X_5 X_6 X_7 X_8$

```
> modelnegbin=glm.nb(y~xd+x1+x2+x3+x4+x5+x6+x7+x8,
data=binom)
> summary(modelnegbin)
```

Call:

```
glm.nb(formula = y ~ xd + x1 + x2 + x3 + x4 + x5 + x6 + x7 +
x8, data = binom, init.theta = 3.227170972, link = log)
```

Deviance Residuals:

```
Min      1Q  Median      3Q      Max
-3.6138 -0.9073 -0.1372  0.4552  2.5630
```

Coefficients:

```
Estimate Std. Error z value Pr(>|z|)
(Intercept) 4.007934  0.144063  27.821 < 2e-16 ***
xd          0.865580  0.401201   2.157 0.03097 *
x1         -0.053534  0.137150  -0.390 0.69629
x2          0.035115  0.128120   0.274 0.78402
x3          0.693181  0.118233   5.863 4.55e-09 ***
x4         -0.299488  0.135083  -2.217 0.02662 *
x5          0.522623  0.168756   3.097 0.00196 **
x6          0.009416  0.230223   0.041 0.96738
x7         -0.576109  0.213946  -2.693 0.00709 **
```

```
x8      -0.062255  0.317101 -0.196 0.84436
```

```
---
```

```
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
(Dispersion parameter for Negative Binomial(3.2272) family taken to be 1)
```

```
Null deviance: 183.960 on 36 degrees of freedom
```

```
Residual deviance: 42.627 on 27 degrees of freedom
```

```
AIC: 393.68
```

```
Number of Fisher Scoring iterations: 1
```

```
Theta: 3.227
```

```
Std. Err.: 0.834
```

```
2 x log-likelihood: -371.679
```

## K.2 Y dengan $X_1$ $X_2$ $X_3$ $X_4$ $X_5$ $X_6$ $X_7$ $X_8$ $X_9$ $X_{65}$

```
> modelnegbin=glm.nb(y~xd+x1+x2+x3+x4+x5+x6+x7+x8+x65,
data=binom)
```

```
> summary(modelnegbin)
```

Call:

```
glm.nb(formula = y ~ xd + x1 + x2 + x3 + x4 + x5 + x6 + x7 +
x8 + x65, data = binom, init.theta = 3.22887955, link = log)
```

Deviance Residuals:

```
Min      1Q  Median      3Q      Max
-3.6202 -0.9193 -0.1233  0.4565  2.5639
```

Coefficients:

```
Estimate Std. Error z value Pr(>|z|)
(Intercept) 4.00588  0.14827 27.017 < 2e-16 ***
xd           0.86248  0.40147  2.148 0.03169 *
x1          -0.05124  0.13971 -0.367 0.71381
x2           0.03353  0.13013  0.258 0.79667
x3           0.69097  0.11879  5.817 6e-09 ***
x4          -0.30161  0.13682 -2.204 0.02749 *
x5           0.52772  0.17320  3.047 0.00231 **
x6           0.01164  0.23435  0.050 0.96040
x7          -0.58394  0.23286 -2.508 0.01215 *
x8          -0.06067  0.31796 -0.191 0.84867
x65         -0.01086  0.11561 -0.094 0.92515
```

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Negative Binomial(3.2289) family taken to be 1)

Null deviance: 184.049 on 36 degrees of freedom

Residual deviance: 42.638 on 26 degrees of freedom

AIC: 395.67

Number of Fisher Scoring iterations: 1

Theta: 3.229

Std. Err.: 0.835

2 x log-likelihood: -371.672

**K.3 Y dengan  $X_1$   $X_2$   $X_3$   $X_4$   $X_5$   $X_6$   $X_7$   $X_8$   $X_9$   $X_{83}$** 

```
> modelnegbin = glm.nb(y ~ xd + x1 + x2 + x3 + x4 + x5 + x6 + x7 + x8 + x83,
data = binom)
```

```
> summary(modelnegbin)
```

Call:

```
glm.nb(formula = y ~ xd + x1 + x2 + x3 + x4 + x5 + x6 + x7 +
x8 + x83, data = binom, init.theta = 3.57436632, link = log)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-3.8564	-0.8050	-0.1611	0.5403	2.5575

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	3.990520	0.137600	29.001	< 2e-16 ***
xd	1.100142	0.405444	2.713	0.006659 **
x1	0.006485	0.133206	0.049	0.961169
x2	0.054725	0.122633	0.446	0.655413
x3	0.509118	0.143396	3.550	0.000385 ***
x4	-0.255491	0.130572	-1.957	0.050382 .
x5	0.421171	0.168943	2.493	0.012667 *
x6	-0.002167	0.219279	-0.010	0.992117
x7	-0.571865	0.204232	-2.800	0.005109 **

```
x8      -0.082759  0.304771 -0.272 0.785972
x83     -0.220301  0.114609 -1.922 0.054582 .
```

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

(Dispersion parameter for Negative Binomial(3.5744) family taken to be 1)

Null deviance: 202.049 on 36 degrees of freedom

Residual deviance: 42.994 on 26 degrees of freedom

AIC: 392.53

Number of Fisher Scoring iterations: 1

Theta: 3.574

Std. Err.: 0.944

2 x log-likelihood: -368.532