

**PENDUGAAN PARAMETER REGRESI POISSON DENGAN  
METODE BAYESIAN *MARKOV CHAIN MONTE CARLO***

**(MCMC)**

**(Studi Kasus DBD di Kota Malang)**

**SKRIPSI**

**Oleh:**

**FEBY INDRIANA YUSUF**

**0710950041 – 95**



**PROGRAM STUDI STATISTIKA  
JURUSAN MATEMATIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS BRAWIJAYA  
MALANG  
2012**

**LEMBAR PENGESAHAN SKRIPSI  
PENDUGAAN PARAMETER REGRESI POISSON DENGAN  
METODE BAYESIAN *MARKOV CHAIN MONTE CARLO*  
(MCMC)  
(Studi Kasus DBD di Kota Malang)**

oleh:

**FEBY INDRIANA YUSUF  
NIM. 0710950041**

**Setelah dipertahankan di depan Majelis Penguji  
pada tanggal 07 Februari 2012  
dan dinyatakan memenuhi syarat untuk memperoleh gelar  
Sarjana Sains dalam bidang Statistika**

**Dosen Pembimbing I**

**Dosen Pembimbing II**

**Eni Sumarminingsih, S.Si., MM.  
NIP. 197705152002122009**

**Prof.Dr.Ir.Waego Hadi Nugroho  
NIP. 195212071979031003**

**Mengetahui,  
Ketua Jurusan Matematika  
Fakultas MIPA Universitas Brawijaya**

**Dr. Abdul Rouf Alghofari, MSc  
NIP. 19670907 199203 1 001**

## LEMBAR PERNYATAAN

Saya yang bertanda tangan di bawah ini:

Nama : Feby Indriana Yusuf  
NIM : 0710950041-95  
Jurusan : Matematika  
Penulisan Skripsi berjudul : **PENDUGAAN PARAMETER  
REGRESI POISSON DENGAN METODE BAYESIAN  
MARKOV CHAIN MONTE CARLO (MCMC)**  
(Studi Kasus DBD di Kota Malang)

Dengan ini menyatakan bahwa:

1. Isi dari Skripsi yang saya buat adalah benar-benar karya sendiri dan tidak menjiplak karya orang lain selain nama-nama yang termaktub di isi dan tertulis di daftar pustaka dalam Skripsi.
2. Apabila di kemudian hari ternyata Skripsi yang saya tulis terbukti hasil jiplakan, maka saya akan bersedia menanggung segala risiko yang akan saya terima.

Demikian pernyataan ini dibuat dengan segala kesadaran.

Malang, 07 Februari 2012  
Yang menyatakan

(Feby Indriana Yusuf)  
NIM. 0710950041

**PENDUGAAN PARAMETER REGRESI POISSON DENGAN  
METODE BAYESIAN *MARKOV CHAIN MONTE CARLO*  
(MCMC)  
(Studi Kasus DBD di Kota Malang)**

**ABSTRAK**

Model Regresi Poisson digunakan untuk memodelkan respon data yang merupakan data *count* (jumlah) yang berdistribusi Poisson. Tujuan penelitian ini yaitu mengetahui variabel-variabel yang berpengaruh terhadap penyebaran penyakit DBD dengan pendekatan *Markov Chain Monte Carlo* (MCMC) pada pendugaan parameter model regresi Poisson menggunakan metode Bayesian. Variabel-variabel yang diindikasikan mempengaruhi jumlah penderita DBD di Kota Malang pada tahun 2010 adalah indeks kepadatan penduduk, banyaknya sarana kesehatan, jumlah rumah tangga miskin, jumlah sekolah masing-masing kelurahan, persentase rumah tangga ber PHBS, rata-rata banyaknya penderita DBD di daerah sekitar dan banyaknya TPS di masing-masing kelurahan. Dalam pembentukan model regresi Poisson ini, pendugaan parameter dilakukan dengan iterasi sebanyak 85000. Dari model yang didapat, pada taraf nyata sebesar 5% ( $\alpha=0,05$ ) diketahui bahwa variabel yang berpengaruh terhadap jumlah penderita DBD di Kota Malang adalah indeks kepadatan penduduk, banyaknya sarana kesehatan, jumlah rumah tangga miskin, jumlah sekolah masing-masing kelurahan dan rata-rata banyaknya penderita DBD di daerah sekitar

Kata kunci : regresi Poisson, metode *Bayes*, *Markov Chain Monte Carlo* (MCMC).

**PARAMETER ESTIMATE OF POISSON REGRESSION  
USING BAYESIAN MARKOV CHAIN MONTE CARLO  
(MCMC) METHOD  
(Case Study of DHF in Malang City)**

**ABSTRACT**

Poisson regression model is used to modeling the response count (number) data which has Poisson distribution. The purpose of this study is to know the variables that influence the spread of dengue fever using Bayesian methods with Markov Chain Monte Carlo (MCMC) approach to estimate parameters of Poisson regression model. Variables that affect the number of DHF cases is indicated in Malang in 2010 is an index of population density, number of health facilities, the number of poor households, the number of schools each village, the percentage of households had PHBS, the average number of DHF cases in the area around and the number of landfills their respective districts. In forming this Poisson regression model, parameter estimation is done as many as 85.000 iterations. At the model signification 5% ( $\alpha = 0.05$ ) the variables affect the number of DHF patients in Malang are the index of population density, number of health facilities, the number of poor households, the number of schools in each village and the average number of DHF cases around the area.

Keywords : Poisson regression, bayes method, Markov Chain Monte Carlo (MCMC).

## KATA PENGANTAR

Puji syukur kehadirat Allah SWT atas berkah, rahmat dan hidayah-Nya sehingga penulisan Skripsi ini dapat terselesaikan dengan baik. Skripsi yang berjudul Pendugaan Parameter Regresi Poisson Dengan Metode Bayesian *Markov Chain Monte Carlo* (MCMC) (Studi Kasus DBD di Kota Malang) ini merupakan salah satu prasyarat kelulusan di Program Studi Statistika, Fakultas MIPA, Universitas Brawijaya.

Dalam penyusunan skripsi ini, penulis telah banyak dibantu oleh beberapa pihak. Oleh karena itu dalam kesempatan ini penulis ingin menyampaikan terima kasih kepada :

1. Ibu Eni Sumarminingsih, S.Si, MM. selaku Dosen Pembimbing I atas segala masukan, nasehat dan bimbingan selama proses penyelesaian Tugas Akhir.
2. Bapak Prof.Dr.Ir.Waego Hadi Nugroho selaku Dosen Pembimbing II atas masukan, nasehat dan bimbingan yang telah diberikan.
3. Ibu Dr.Ir.Ni Wayan Surya Wardhani,MS. selaku Dosen penguji atas saran dan masukan yang telah diberikan.
4. Bapak Dr. Abdul Rouf Alghofari, MSc. selaku Ketua Jurusan Matematika Fakultas MIPA Universitas Brawijaya Malang.
5. Ayah, Ibu, kakak dan adikku tercinta atas doa serta dukungan baik moril maupun materiil.
6. Saudara-saudara Statistika angkatan 2007 atas semangat, informasi dan persahabatan yang terbina.
7. Seluruh staf pengajaran Matematika atas bantuan dan kerjasamanya.
8. Semua pihak yang telah membantu hingga selesainya skripsi ini.

Penulis menyadari keterbatasan dalam penulisan skripsi ini. Oleh karena itu saran ataupun kritik yang membangun akan sangat berguna bagi penulis untuk mengembangkan kemampuan menulis ilmiah. Skripsi ini semoga dapat bermanfaat bagi para pembaca.

Malang, 07 Februari 2012

Penulis

## DAFTAR ISI

	Halaman
<b>HALAMAN JUDUL</b> .....	i
<b>HALAMAN PENGESAHAN</b> .....	ii
<b>HALAMAN PERNYATAAN</b> .....	iii
<b>ABSTRAK</b> .....	iv
<b>ABSTRACT</b> .....	v
<b>KATA PENGANTAR</b> .....	vi
<b>DAFTAR ISI</b> .....	vii
<b>DAFTAR GAMBAR</b> .....	ix
<b>DAFTAR TABEL</b> .....	x
<b>DAFTAR LAMPIRAN</b> .....	xi
<b>BAB I PENDAHULUAN</b>	
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	2
1.3 Batasan Masalah.....	2
1.4 Tujuan.....	2
1.5 Manfaat Penelitian.....	3
<b>BAB II TINJAUAN PUSTAKA</b>	
2.1 Demam Berdarah Dangué (DBD) di Kota Malang.....	5
2.2 Distribusi Poisson.....	6
2.3 Regresi Poisson .....	8
2.4 <i>Over</i> Dispersi dan <i>Under</i> Dispersi .....	10
2.5 Multikolinieritas .....	11
2.6 Metode Bayes .....	12
2.7 Metode <i>Markov Chain Monte Carlo</i> (MCMC).....	14
2.8 Pengujian Hipotesis.....	17
2.8.1 Pengujian Secara Simultan.....	17
2.8.2 Pengujian Secara Parsial .....	18
2.9 Pengujian Kelayakan Model .....	19
<b>BAB III METODOLOGI PENELITIAN</b>	
3.1 Data Penelitian .....	21
3.2 Metode Bayesian MCMC Pada Model Regresi Poisson.....	21
3.3 Proses Simulasi .....	23
3.4 Tahap Analisis.....	23

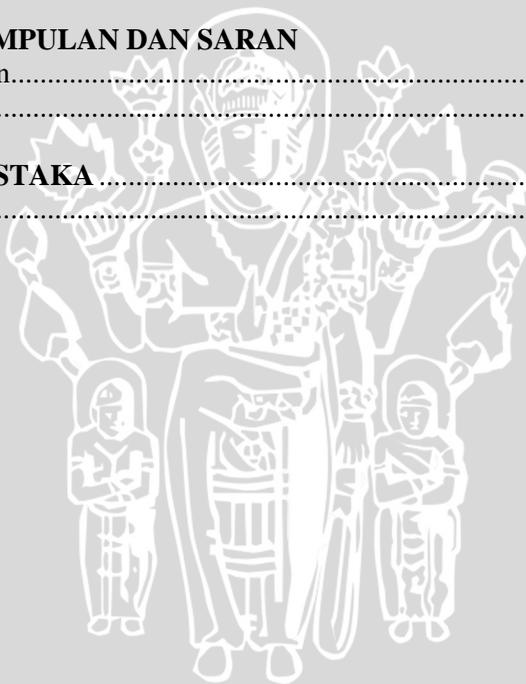
## **BAB IV HASIL DAN PEMBAHASAN**

4.1 Pengujian Distribusi Poisson Terhadap Variabel Y .....	25
4.2 Pemeriksaan <i>Over/under</i> dispersi pada Regresi Poisson.....	25
4.3 Pengujian Multikolinieritas .....	25
4.4 Metode Bayesian MCMC Pada Model Regresi Poisson.....	26
4.5 Pengujian Signifikansi Penduga Parameter Regresi Poisson	27
4.5.1 Pengujian Simultan.....	27
4.5.2 Pengujian Parsial .....	28
4.6 Pengujian Kelayakan Model .....	29
4.7 Interpretasi Model .....	29

## **BAB V KESIMPULAN DAN SARAN**

5.1 Kesimpulan.....	31
5.2 Saran.....	31

<b>DAFTAR PUSTAKA</b> .....	33
<b>LAMPIRAN</b> .....	37



**DAFTAR GAMBAR**

Halaman

Gambar 3.1 Langkah-Langkah Analisis..... 24

**UNIVERSITAS BRAWIJAYA**



## DAFTAR TABEL

	Halaman
Tabel 2.1 Struktur Data Pengamatan.....	9
Tabel 4.1 Hasil Pengujian Multikolinieritas.....	26
Tabel 4.2 Hasil Analisis Pendugaan Parameter Regresi Poisson	26
Tabel 4.3 Hasil Uji Parsial Koefisien Regresi Poisson .....	28
Tabel 4.4 Hasil Uji Parsial Koefisien Regresi Poisson tanpa $\beta_5$ dan $\beta_7$ .....	28



## DAFTAR LAMPIRAN

	Halaman
Lampiran 1. Data Penelitian .....	37
Lampiran 2. Hasil Pengujian Distribusi Poisson .....	40
Lampiran 3. Hasil Pengujian <i>Over/under</i> Dispersi .....	41
Lampiran 4. Hasil Uji Multikolinieritas .....	42
Lampiran 5. Node untuk metode pendugaan Regresi Poisson dengan software WINBUGS 14 .....	43
Lampiran 6. Output WINBUGS 14.....	44
Lampiran 7. Output <i>Likelihood Ratio Chi-Square</i> .....	48
Lampiran 8. Output Statistik Wald.....	49
Lampiran 9. Node untuk metode pendugaan Regresi Poisson tanpa $\beta_5$ dan $\beta_7$ dengan software WINBUGS 14 ..	50
Lampiran 10. Output WINBUGS 14 tanpa $\beta_5$ dan $\beta_7$ .....	51
Lampiran 11. Output Statistik Wald tanpa $\beta_5$ dan $\beta_7$ .....	54
Lampiran 12. Output <i>Likelihood Ratio Chi-Square</i> .....	55
Lampiran 13. Output <i>Goodness of Fit</i> .....	56



# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Semakin bertambahnya penduduk Indonesia menyebabkan kepadatan pada suatu pulau meningkat dan menyebabkan kelebihan proporsi tempat untuk tinggal pada suatu pulau. Dengan semakin banyaknya penduduk yang tinggal berhimpitan pada suatu wilayah akan menyebabkan tingkat kebersihan menurun dan wabah penyakit yang semakin merebak. Berdasarkan BPS (Badan Pusat Statistika) negara Indonesia pada Tahun 2009 memiliki penduduk sebanyak 228,7 juta jiwa, peningkatan yang sangat jauh dari Tahun 2000 yaitu 205,1 juta jiwa. Disamping bertambahnya penduduk Indonesia tiap tahun, jumlah kematian akibat penyakit Demam Berdarah (DBD) dari tahun ke tahun juga meningkat. Kota Malang merupakan kota yang pada Tahun 2010 termasuk dalam wilayah Kejadian Luar Biasa (KLB) DBD yang memiliki angka penderita DBD yang cukup tinggi tiap tahunnya. Pada tahun 2004 – 2010 lebih dari 150 jiwa tiap tahunnya menderita penyakit DBD.

Terjadinya kasus DBD dalam suatu wilayah memiliki hubungan dengan beberapa variabel. Analisis yang mempelajari hubungan di antara variabel-variabel sehingga dari hubungan tersebut dapat menduga variabel yang satu apabila nilai variabel lainnya diketahui, dikenal dengan nama analisis regresi. Dalam analisis regresi terdapat variabel yang dipengaruhi (respon) dan variabel yang mempengaruhi (prediktor).

Regresi Poisson merupakan suatu metode regresi yang tepat untuk memodelkan hubungan kausal antara variabel respon berjenis diskrit yang menyebar Poisson dengan satu atau beberapa variabel penjelas (bisa berjenis diskrit atau kontinyu). Kejadian DBD merupakan kejadian yang menyebar Poisson, sehingga model regresi Poisson tepat digunakan untuk memodelkan hubungan jumlah penderita kasus DBD masing-masing kelurahan di Kota Malang sebagai variabel respon dan beberapa variabel prediktor antara lain kepadatan penduduk, banyaknya sarana kesehatan, jumlah rumah tangga miskin, jumlah sekolah, persentase rumah tangga ber-PHBS (Pola Hidup Bersih dan Sehat), rata-rata banyaknya penderita di

daerah sekitar dan banyaknya TPS (Tempat Pembuangan Sampah) (Roose, 2008).

Bayesian merupakan salah satu metode pendugaan parameter yang merupakan penyederhanaan dari cara klasik yang penuh integral untuk memperoleh model marginal. Di samping itu, metode Bayesian lebih dapat mengakomodir informasi data dari awal (subyektif) ditambah dengan informasi dari sampel. Pendugaan parameter model regresi Poisson dalam penelitian ini menggunakan metode Bayesian dengan algoritma Markov Chain Monte Carlo (MCMC). Menurut Scollnik (1996), *Markov Chain Monte Carlo* (MCMC) merupakan metode simulasi berdasarkan teknik simulasi *Monte Carlo* untuk membangkitkan data dari distribusi tertentu yang cara kerjanya menggunakan sifat rantai Markov. Salah satu simulasi yang terkenal pada MCMC adalah *Gibbs Sampler*, yaitu metode simulasi yang menggunakan distribusi bersyarat penuh yang digabungkan dengan distribusi stasioner ( $\pi(x)$ ). Pendekatan ini akan mempermudah untuk penentuan pendugaan parameter dalam metode Bayesian.

## **1.2 Rumusan Masalah**

Masalah yang terkait dalam penelitian ini adalah variabel-variabel apa saja yang berpengaruh terhadap penyakit DBD pada model regresi poisson dengan metode bayesian untuk memodelkan kasus DBD di Kota Malang ?

## **1.3 Batasan Masalah**

Ruang lingkup penelitian ini adalah pendugaan model regresi Poisson dengan metode Bayesian menggunakan algoritma Markov Chain Monte Carlo (MCMC). Simulasi yang digunakan pada MCMC adalah simulasi *Gibbs Sampler*.

## **1.4 Tujuan**

Tujuan dalam penelitian adalah untuk mengetahui variabel-variabel yang berpengaruh terhadap penyakit DBD dengan pemodelan regresi poisson dan dilakukan pendugaan parameter regresi poisson dengan algoritma Markov Chain Monte Carlo pada metode Bayesian.

## 1.5 Manfaat Penelitian

Penelitian ini diharapkan mampu memberikan penjelasan pendugaan parameter dari regresi poisson dengan metode bayesian yang mengakomodir informasi data dari awal (subyektif) ditambah dengan informasi dari sampel sehingga diperoleh pendugaan parameter menggunakan pendekatan MCMC dengan simulasi *Gibbs Sampler*.

Setelah diperoleh model regresi poisson dan diketahui variabel-variabel yang mempengaruhi penyakit DBD dan diharapkan Dinas Pemerintah Kota Malang dapat lebih mengefektifkan penanggulangan dan pencegahan penyakit DBD. Sehingga penderita dan angka kematian akibat penyakit DBD semakin berkurang.



UNIVERSITAS BRAWIJAYA



## BAB II TINJAUAN PUSTAKA

### 2.1 Demam Berdarah Dengue (DBD) di Kota Malang

Penyakit Demam Berdarah (DBD) disebabkan oleh golongan virus dengue yang ditularkan dari seseorang ke orang lain melalui gigitan nyamuk *Aedes aegypti* dan *Aedes albopictus*. Penyakit DBD pertama kali di Indonesia ditemukan di Surabaya pada tahun 1968 dan pada tahun 1973 mulai merambat ke Jakarta. Sejak itu hampir setiap minggu terdengar ledakan-ledakan sporadik penyakit ini di seluruh wilayah Indonesia (Kristina dan Wulandari, 2004).

Untuk wilayah Jawa Timur, Kota Malang merupakan daerah insiden DBD dengan jumlah penderita meningkat tiap tahunnya. Kota Malang terbagi menjadi 5 kecamatan dengan total 57 kelurahan. Dengan adanya pemodelan penyakit DBD ini diharapkan dapat memberikan informasi variabel-variabel yang merupakan faktor penunjang meningkatnya kasus DBD di wilayah Kota Malang. Pemodelan dilakukan untuk kasus DBD tahun 2010. Variabel-variabel yang digunakan dalam penelitian ini antara lain :

- Banyaknya penderita penyakit DBD di masing-masing kelurahan di Kota Malang tahun 2010.
- Indeks kepadatan penduduk ( $\text{jiwa}/\text{km}^2$ ) yang menunjukkan perbandingan antara banyaknya penduduk dengan luas kecamatan ( $\text{km}^2$ ). Penyebaran virus dengue dapat disebabkan oleh kepadatan penduduk di wilayah tersebut.
- Banyaknya sarana kesehatan yaitu Rumah Sakit Pemerintah. Sarana kesehatan dapat membantu penanganan penyebaran virus dengue yang dapat menyebabkan penyakit DBD. Disamping itu sarana kesehatan yang ada dapat mengoptimalkan penanganan kasus gejala DBD pada masyarakat sekitar.
- Banyaknya rumah tangga miskin perkelurahan di Kota Malang. Keluarga yang termasuk dalam kategori miskin memiliki kecenderungan kurang memperhatikan masalah kesehatan yang disebabkan alasan faktor ekonomi. Selain itu kurangnya pengetahuan tentang pentingnya melakukan vaksinasi yang dapat meningkatkan kekebalan tubuh dapat menyebabkan virus penyakit DBD menyebar dengan pesat dan menambah jumlah penderita penyakit DBD.

- Banyaknya sekolah masing-masing kelurahan meliputi SD/ sederajatnya, SMP/ sederajatnya, SMA/ sederajatnya. Anak sekolah merupakan kelompok umur yang sangat rentan untuk terserang penyakit DBD. Sekolah merupakan salah satu tempat berkumpulnya berbagai orang dari beberapa wilayah sehingga kemungkinan terjadinya pertukaran virus dengue cukup besar.
- Persentase rumah ber-PHBS (Pola Hidup Bersih dan Sehat) merupakan faktor yang sangat mendukung untuk pencegahan berbagai penyakit termasuk DBD. Dalam menentukan rumah tangga yang ber-PHBS, pihak Dinas Kesehatan Kota Malang mengambil sampel per 1000 rumah tangga yang kemudian hasil persentasenya digunakan dalam penelitian ini, dengan perhitungan sebagai berikut :

$$\% \text{ rumah tangga ber - PHBS} = \frac{\text{banyaknya rumah ber - PHBS}}{1000} \times 100\%$$

- Rata-rata penderita DBD di daerah sekitar. Penyakit DBD ditularkan melalui gigitan nyamuk *Aedes aegypti* dan *Aedes albopictus* dimana nyamuk-nyamuk tersebut berpeluang untuk membawa virus ke daerah sekitar, sehingga banyaknya penderita DBD di suatu daerah dapat mempengaruhi banyaknya penderita di daerah laennya (daerah terdekat).
- Tempat Pembuangan Sampah (TPS) merupakan tempat pusat berbagai penyakit, banyaknya TPS di suatu wilayah sangat mempengaruhi peningkatan kasus DBD.

Variabel-variabel diatas digunakan untuk memodelkan kasus DBD. Model yang digunakan yaitu model regresi poisson dengan metode pendugaan bayes.

## 2.2 Distribusi Poisson

Percobaan Poisson merupakan percobaan yang menghasilkan variabel acak (Y) yang bernilai numerik, yaitu banyaknya hasil selama selang waktu tertentu atau dalam daerah tertentu. Suatu percobaan Poisson memiliki sifat :

1. Banyaknya hasil yang terjadi di dalam suatu selang waktu atau daerah tertentu tidak terpengaruh oleh apa yang terjadi pada selang waktu atau daerah lain yang terpisah.
2. Peluang terjadinya suatu hasil dalam selang waktu yang amat

pendek atau dalam daerah yang kecil sebanding dengan panjang selang waktu atau besarnya daerah dan tidak tergantung pada banyaknya hasil yang terjadi di luar selang waktu atau daerah tersebut.

3. Peluang terjadinya lebih dari satu hasil dalam selang waktu yang pendek atau daerah yang sempit tersebut dapat diabaikan (Walpole & Myers, 1995).

Dalam penelitian ini variabel acak Poisson ( $Y$ ) disebut  $Y_i$  yang menyatakan banyaknya penderita DBD di daerah  $i$  yang memiliki distribusi Poisson ( $\lambda_i$ ) dengan  $\lambda_i = \mu_i = \mu(X_{ij}, \beta)$ .

Fungsi peluang distribusi poisson adalah:

$$f(Y_i) = \frac{e^{-(\mu_i)}(\mu_i)^{Y_i}}{Y_i!} \quad (2.1)$$

di mana :

$\mu_i$  = rata-rata banyaknya penderita DBD tiap kelurahan

$e$  = nilai konstan 2,7183

Persamaan di atas disebut juga sebagai fungsi peluang Poisson. Rata-rata dan ragam dari distribusi Poisson adalah sama yaitu (Kutner *et al*, 2004):

$$E\{Y\} = \mu \quad (2.2)$$

dan

$$\sigma^2\{Y\} = \mu \quad (2.3)$$

Hal tersebut dapat dibuktikan dengan penjabaran sebagai berikut:

$$\sum_{y=0}^{\infty} p(y) = 1$$

$$\sum_{y=0}^{\infty} \frac{e^{-\mu} \cdot \mu^y}{y!} = 1$$

$$e^{-\mu} \sum_{y=0}^{\infty} \frac{\mu^y}{y!} = 1$$

$$\sum_{y=0}^{\infty} \frac{\mu^y}{y!} = e^{\mu}$$

kemudian diperoleh nilai harapan:

$$E(Y) = \sum_{y=0}^{\infty} y \cdot p(y)$$

$$= \sum_{y=0}^{\infty} y \frac{e^{-\mu} \mu^y}{y!}$$

$$= e^{-\mu} \sum_{y=0}^{\infty} \frac{y \cdot \mu \cdot \mu^{y-1}}{y(y-1)!}$$

$$= e^{-\mu} \cdot \mu \sum_{y=0}^{\infty} \frac{\mu^{y-1}}{(y-1)!}$$

$$= e^{-\mu} \cdot e^{\mu} \cdot \mu$$

$$= \mu$$

dari nilai harapan selanjutnya dapat dicari ragam:

$$\begin{aligned}
 E(Y^2) &= \sum_{y=0}^{\infty} y^2 \cdot p(y) \\
 &= \sum_{y=0}^{\infty} (y(y-1) + y) \cdot p(y) \\
 &= \sum_{y=0}^{\infty} y(y-1) p(y) + \sum_{y=0}^{\infty} y \cdot p(y) \\
 &= \sum_{y=0}^{\infty} \frac{y(y-1)e^{-\mu}\mu^y}{y!} + \mu \\
 &= e^{-\mu} \sum_{y=0}^{\infty} \frac{y(y-1)\mu^{y-2}\mu^2}{y(y-1)(y-2)!} + \mu \\
 &= e^{-\mu} \mu^2 \sum_{y=2}^{\infty} \frac{\mu^{y-2}}{(y-2)!} + \mu \\
 &= e^{-\mu} \mu^2 e^{\mu} + \mu \\
 &= \mu^2 + \mu \\
 V(Y) &= E(Y^2) - [E(Y)]^2 \\
 &= \mu^2 + \mu - \mu^2 \\
 &= \mu
 \end{aligned}$$

### 2.3 Regresi Poisson

Regresi Poisson merupakan suatu metode regresi yang tepat untuk memodelkan hubungan kausal antara variabel respon berjenis diskrit yang menyebar Poisson dengan satu atau beberapa variabel penjelas (bisa berjenis diskrit atau kontinyu).

Kleinbaum *et al.* (1998) menyatakan bahwa asumsi yang harus dipenuhi dalam regresi Poisson adalah variabel respon Y harus mengikuti distribusi Poisson.

Dalam regresi non linier perlu dilakukan pemeriksaan multikolinieritas, sebagaimana halnya regresi linier. Karena Regresi poisson merupakan regresi non linier, maka variabel-variabel penjelas dalam regresi Poisson juga harus memenuhi asumsi non multikolinieritas, yaitu tidak ada korelasi linier yang tinggi (mendekati sempurna ataupun sempurna) diantara dua atau lebih variabel penjelas.

Pandang  $Y \sim \text{Poisson}(\mu)$ , dengan fungsi peluang yang dinyatakan pada persamaan (2.1), dimana rata-rata dan ragam sama yaitu  $E(Y) = \text{Var}(Y) = \mu$ . Pada model regresi poisson nilai  $\mu$  dari merupakan fungsi dari variabel penjelas.

Jika  $Y_i$  merupakan nilai variabel respon pengamatan ke- $i$ , dimana  $Y_i$  berjenis diskrit dan merupakan variabel acak Poisson, sedangkan  $X_{ij}$  adalah nilai variabel penjelas ke- $j$  ( $j = 1, 2, \dots, p$ ) dari pengamatan ke- $i$  ( $i = 1, 2, \dots, n$ ) dimana  $X_{ij}$  bisa berjenis diskrit atau

kontinyu, maka nilai-nilai pengamatan bagi  $Y_i$  dan  $X_{ij}$  ditunjukkan pada tabel 2.1 berikut ini :

Tabel 2.1 Struktur data pengamatan

$i$	$Y_i$	$X_{i1}$	$X_{i2}$	...	$X_{ip}$
1	$y_1$	$x_{11}$	$x_{12}$	...	$x_{1p}$
2	$y_2$	$x_{21}$	$x_{22}$	...	$x_{2p}$
3	$y_3$	$x_{31}$	$x_{32}$	...	$x_{3p}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	$y_n$	$x_{n1}$	$x_{n2}$	...	$x_{np}$

Secara umum model Regresi Poisson untuk data pada tabel 2.1 dinyatakan dalam bentuk :

$$y_i = \mu_i + \varepsilon_i \quad , i = 1, 2, 3, \dots, n \quad (2.4)$$

dimana :

$y_i$  : nilai pengamatan variabel respon ke- $i$ , yang masing-masing merupakan variabel acak Poisson

$\mu_i$  : rata-rata variabel respon  $Y$  yang dipengaruhi oleh nilai-nilai variabel penjelas pada pengamatan ke- $i$

$\varepsilon_i$  : galat ke- $i$

$\mu_i$  yang biasanya disebut dengan respon rata-rata (*mean respons*) merupakan fungsi dari  $p$  variabel penjelas pada pengamatan ke- $i$ , yaitu  $X_{i1}, X_{i2}, \dots, X_{ip}$ .

Dari persamaan (2.2) diperoleh :

$$E(y_i) = E(\mu_i) + E(\varepsilon_i)$$

Jika  $E(\varepsilon_i) = 0$ , maka :

$$\hat{y}_i = \hat{\mu}_i \quad , i = 1, 2, 3, \dots, n \quad (2.5)$$

dimana :

$\hat{y}_i$  : penduga bagi variabel respon ke- $i$

$\hat{\mu}_i$  : penduga bagi respon rata-rata ke- $i$

Persamaan (2.5) merupakan model penduga bagi model umum Regresi Poisson berdasarkan nilai-nilai pengamatan contoh.

Fungsi yang menghubungkan respon rata-rata  $\mu_i$  dengan  $X_i$  dinotasikan dengan  $\mu(\mathbf{X}_{ij}, \boldsymbol{\beta})$ , dan dalam GLM disebut dengan fungsi penghubung (*link function*). Berdasarkan fungsi penghubung ini, maka respon rata-rata  $\mu_i$  dinyatakan dengan :

$$\mu_i = \mu(\mathbf{X}_{ij}, \boldsymbol{\beta})$$

Menurut McCullagh dan Nelder (1989) dalam Agresti (2002), fungsi penghubung yang biasa digunakan dalam regresi Poisson adalah

$\mu(\mathbf{X}_{ij}, \boldsymbol{\beta}) = \exp(\mathbf{X}_{ij}^T \boldsymbol{\beta})$  dimana  $\mathbf{X}_{ij}$  merupakan matriks variabel penjelas berukuran  $n \times (p+1)$ ,  $\boldsymbol{\beta}$  adalah vektor parameter regresi berdimensi  $(p+1) \times 1$ , dan  $\mathbf{X}_{ij}^T \boldsymbol{\beta}$  merupakan fungsi linier. Dengan demikian, penduga bagi  $\mu_i$ , yaitu  $\hat{\mu}_i$ , adalah :

$$\begin{aligned} \hat{\mu}_i &= \hat{\mu}(\mathbf{X}_{ij}, \boldsymbol{\beta}) \\ &= \exp(\mathbf{X}_{ij}^T \boldsymbol{\beta}) \\ &= \exp[\hat{\beta}_0 + \hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2} + \dots + \hat{\beta}_p X_{ip}] \end{aligned} \quad (2.6)$$

Berdasarkan persamaan akhir pada (2.6), yaitu  $\hat{\mu}_i = \exp[\hat{\beta}_0 + \hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2} + \dots + \hat{\beta}_p X_{ip}]$ , maka model umum Regresi Poisson pada persamaan (2.4) dinyatakan sebagai :

$$y_i = \exp[\hat{\beta}_0 + \hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2} + \dots + \hat{\beta}_p X_{ip}] + \varepsilon_i \quad (2.7)$$

yang dapat diduga dengan :

$$\hat{y}_i = \exp[\hat{\beta}_0 + \hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2} + \dots + \hat{\beta}_p X_{ip}]$$

Dari persamaan (2.7) dibentuk fungsi likelihood yang digunakan dalam pendugaan sebagai berikut :

$$\begin{aligned} L(\boldsymbol{\beta}) &= \prod_{i=1}^n f_i(Y_i) \\ L(\boldsymbol{\beta}) &= \prod_{i=1}^n \frac{[\mu(\mathbf{X}_{ij}, \boldsymbol{\beta})]^{Y_i} \exp[-\mu(\mathbf{X}_{ij}, \boldsymbol{\beta})]}{Y_i!} \\ L(\boldsymbol{\beta}) &= \frac{\{\prod_{i=1}^n [\mu(\mathbf{X}_{ij}, \boldsymbol{\beta})]^{Y_i}\} \exp[-\sum_{i=1}^n \mu(\mathbf{X}_{ij}, \boldsymbol{\beta})]}{\prod_{i=1}^n Y_i!} \end{aligned}$$

Dari fungsi likelihood diatas dapat dilakukan pendugaan parameter regresi poisson dengan menggunakan metode bayes.

## 2.4 Over Dispersi dan Under Dispersi

*Over/under* dispersi pada regresi Poisson menunjukkan bahwa tidak terpenuhinya sifat *equal* dispersi. Keadaan di mana ragam variabel Poisson lebih besar daripada rata-rata disebut *over* dispersi, sedangkan *under* dispersi adalah keadaan di mana ragam variabel Poisson lebih kecil dari rata-rata (Højsgaard and Halekøh, 2005).

*Over* dispersi lebih sering terjadi daripada *under* dispersi. Salah satu penyebab *over* dispersi adalah keheterogenan subjek yang diamati (Agresti, 2002).

Ragam dari distribusi Poisson sama dengan rata-ratanya ( $\sigma^2 = \mu$ ) dan parameter dispersi bernilai 1 ( $\omega = 1$ ). Pemeriksaan terjadinya *over* dispersi dan *under* dispersi dapat dilakukan dengan menggunakan nilai statistik Khi-Kuadrat Pearson dibagi dengan

derajat bebas residual yang diperoleh dari hasil analisis Regresi Poisson yang secara matematis ditulis sebagai berikut:

$$\frac{\chi^2_{pearson}}{db} \quad (2.8)$$

dengan  $db = n - p$ , di mana  $n$  adalah banyaknya pengamatan dan  $p$  adalah banyaknya variabel penjelas (Højsgaard and Halekøh, 2005).

*Over* dispersi terjadi jika ( $\omega > 1$ ) sedangkan *under* dispersi terjadi apabila ( $\omega < 1$ ).

Menurut Ismail dan Jemain (2007), statistik uji Khi-Kuadrat Pearson dapat dinyatakan dengan:

$$\sum_{i=1}^n \frac{(y_i - \hat{\mu}_i)^2}{\hat{\sigma}_i^2} \sim \chi^2_{(n-p)} \quad (2.9)$$

di mana:

$\hat{\sigma}_i^2$ : penduga bagi ragam respon ke- $i$

Pada sebaran Poisson, karena  $\hat{\sigma}_i^2 = \hat{\mu}_i$ , maka statistik Uji Khi-Kuadrat Pearson dinyatakan sebagai:

$$\sum_{i=1}^n \frac{(y_i - \hat{\mu}_i)^2}{\hat{\mu}_i} \sim \chi^2_{(n-p)}$$

## 2.5 Multikolinieritas

Salah satu asumsi dalam regresi klasik adalah tidak terjadinya kolinieritas diantara variabel penjelas, baik sempurna atau mendekati sempurna. Yang dimaksud dengan kolinieritas adalah adanya hubungan antara dua variabel penjelas (Gujarati, 1995).

Sembiring (1995) menyatakan jika dalam suatu model regresi terdapat kolinieritas, maka akan mempengaruhi hubungan antara variabel penjelas dengan variabel respon. Taksiran koefisien dari variabel penjelas menjadi tidak tunggal, melainkan tak terhingga banyaknya, sehingga sulit untuk melakukan pendugaan. Selain itu menurut Gujarati (1995), multikolinieritas dalam regresi dapat mengakibatkan besarnya simpangan baku dari penduga koefisien regresi.

Menurut Kutner *et al.* (2004), salah satu cara yang dapat dilakukan untuk mendeteksi adanya multikolinieritas adalah dengan menggunakan VIF (*Variance Inflation Factor*). VIF mengukur seberapa jauh peningkatan ragam penduga bagi koefisien regresi bila terjadi multikolinieritas dalam suatu model regresi. Tingginya nilai

VIF mengindikasikan tingginya multikolinieritas yang sedang terjadi. Idealnya,  $VIF = 1$ , yang berarti hubungan antara variabel-variabel penjelas saling bebas (tidak terjadi multikolinieritas, baik sempurna ataupun tidak sempurna). Beberapa peneliti menetapkan bahwa multikolinieritas terjadi jika  $VIF > 10$ .

Nilai VIF untuk variabel penjelas ke- $j$  dinyatakan dengan :

$$VIF_j = \frac{1}{1-R_j^2} \quad (2.10)$$

di mana :

$$R_j^2 = \frac{\sum_{j=1}^n (\hat{Y}_j - \bar{Y})^2}{\sum_{j=1}^n (Y_j - \bar{Y}_j)^2}$$

$R_j^2$  merupakan nilai  $R^2$  yang diperoleh melalui regresi OLS (*Ordinary Least Square*) antara variabel penjelas ke- $j$  dengan variabel penjelas lainnya yang terdapat dalam model.

Jika hubungan antara variabel penjelas ke- $j$  dan variabel penjelas lainnya mendekati linier, maka  $R_j^2$  akan mendekati 1, sehingga nilai  $VIF_j$  besar. Sedangkan, jika hubungan antara variabel penjelas ke- $j$  dan variabel penjelas lainnya tidak linier, maka  $R_j^2$  akan mendekati 0, sehingga nilai  $VIF_j$  akan mendekati 1 (Kutner *et al.*, 2004).

## 2.6 Metode Bayes

Pendugaan parameter yang digunakan dalam model regresi Poisson ini adalah Bayesian *Markov Chain Monte Carlo* (MCMC). Menurut Martin (2003) Bayesian *Inference* dapat menangani heterogenan data yang menyebabkan *over* dispersi pada regresi Poisson, dikarenakan model hirarki dari metode bayes dapat menangani *over* dispersi pada model. Model hirarki pada regresi Poisson adalah sebagai berikut :

$$y_i \sim \text{Poisson}(\lambda_i)$$

$$\lambda_i \sim \mu_i$$

$$\mu_i = e^{X_{ij}\beta_j}$$

$$\beta_j \sim \text{Normal}(0, \sigma^2)$$

$$\sigma^2 \propto 1/\sigma^2 = 1/100000$$

Pereira (1999) menyatakan bahwa metode bayes memberikan hasil pendugaan yang lebih baik daripada pendugaan titik dalam

metode klasik. Hal ini disebabkan pendugaan parameter klasik hanya berdasarkan informasi dari data sampel, di mana ukuran sampel sangat berpengaruh terhadap hasil pendugaan. Dalam metode bayes selain menggunakan informasi dari data sampel juga dipertimbangkan informasi dari distribusi prior untuk mendapatkan distribusi posterior, sehingga hasil pendugaan dalam metode bayes akan lebih baik. Namun pada kenyataannya, distribusi prior tidak mudah ditentukan dan distribusi posterior menjadi sulit diperoleh sehingga metode bayes cukup sulit diselesaikan secara analitik.

Untuk mengatasi masalah tersebut, maka dikembangkan teknik simulasi sehingga metode bayes mudah diselesaikan. Teknik simulasi yang biasa digunakan dalam metode bayes adalah algoritma *Markov Chain Monte Carlo* (MCMC). Menurut Scollnik (1996), metode MCMC merupakan algoritma untuk mendapatkan data sampel suatu variabel acak dengan teknik sampling berdasarkan sifat rantai markov. Salah satu teknik dalam metode MCMC yang terkenal adalah *Gibbs Sampler*.

Dalam melakukan proses simulasi, *Gibbs Sampler* menggunakan distribusi bersyarat untuk membangkitkan data sampel variabel acak. Metode bayes merupakan metode yang menggabungkan informasi prior dengan pengamatan di dalam percobaan sehingga menghasilkan distribusi posterior. Distribusi posterior kemudian digunakan untuk memperbarui distribusi prior melalui data pengamatan (Pereira, 1999). Berger (1985) menyebutkan bahwa distribusi prior dapat dibedakan menjadi dua berdasarkan bentuk distribusi dan berdasarkan penentuan parameter dalam distribusi prior tersebut. Berdasarkan bentuk distribusi, terdapat dua macam distribusi prior, yaitu:

1. Distribusi prior konjugat, yaitu pemberian bentuk distribusi prior yang sepola dengan distribusi data. Misalnya distribusi beta untuk parameter  $p$  merupakan prior konjugat bagi data yang menyebar binomial dan distribusi normal untuk parameter  $\mu$  merupakan prior konjugat bagi data yang menyebar normal.
2. Distribusi prior non konjugat yaitu pemberian bentuk distribusi prior yang tidak sepola dengan distribusi data.

Distribusi prior terbagi menjadi dua dalam menentukan parameter:

1. Distribusi prior informatif, yaitu penentuan parameter distribusi prior berdasarkan informasi dari data.

2. Distribusi prior non informatif, yaitu penentuan parameter distribusi prior tidak berdasarkan informasi dari data. (Iversen, 1984)

Apabila  $\theta$  merupakan suatu nilai variabel acak dengan distribusi peluang  $f(\theta)$ , maka  $f(\theta)$  sering disebut distribusi awal atau distribusi prior. Peluang yang dikaitkan dengan distribusi prior disebut juga peluang subyektif, karena ingin mengukur derajat keyakinan peneliti mengenai letak parameter yang ingin diduga dengan menggunakan pengalaman dan pengetahuannya. Selanjutnya distribusi prior  $f(\theta)$  digunakan bersama distribusi bersyarat  $f(x|\theta)$  dalam distribusi gabungan sampel.

$$f(x, \theta) = f(x|\theta)f(\theta) \quad (2.11)$$

Dari sini dapat diperoleh distribusi marginal

$$g(x) = \begin{cases} \sum_{\theta} f(x, \theta) & , (bila \text{ diskrit}) \\ \int_{-\infty}^{\infty} f(x, \theta) d\theta & , (bila \text{ kontinyu}) \end{cases} \quad (2.12)$$

Jadi distribusi posterior dapat ditulis sebagai

$$f(\theta|x) = \frac{f(x;\theta)}{g(x)} \quad (2.13)$$

yang menyatakan derajat keyakinan peneliti mengenai letak parameter  $\beta$  setelah sampel diamati (Walpole & Myers, 1995).

Distribusi bersyarat dalam metode bayes juga dapat disebut sebagai fungsi *Likelihood*,  $L(\theta) = f(x|\theta)$ , sehingga metode bayes juga dapat didefinisikan sebagai penggabungan fungsi *Likelihood* dan distribusi prior (Gamerman, 1997).

$$f(\theta|x) \propto f(x|\theta)f(\theta) \quad (2.14)$$

Untuk menentukan nilai duga parameter dari distribusi. posterior  $\theta$  digunakan metode *Markov Chain Monte Carlo* (MCMC). Metode MCMC ini cukup efektif untuk menentukan nilai duga parameter dari distribusi posterior yang sangat kompleks dan cukup sulit jika diselesaikan dengan metode lain (Pereira, 1999).

## 2.7 Metode *Markov Chain Monte Carlo* (MCMC)

Proses stokastik adalah serangkaian variabel acak  $\{X_t\}$ , di mana  $t \in T$  merupakan indeks waktu atau deretan. Sebuah proses stokastik memperlihatkan sifat Markov jika kejadian pada saat  $t+1$  yaitu variabel acak  $X_{t+1}$  hanya dipengaruhi oleh kejadian satu langkah sebelumnya. Secara matematis sifat rantai Markov dapat ditulis sebagai berikut

$$P\{X_{t+1} = j | X_t = i\} = P\{X_{t+1} = j | X_t = i, X_{t-1} = i_1, X_{t-2} = i_2, \dots, X_0 = i_t\}$$

(2.15)

untuk  $t = 0, 1, 2, \dots$ , dan setiap deretan  $j, i, i_1, \dots, i_t$   
Peluang bersyarat,

$$P\{X_{t+1} = j | X_t = i\} = P_{ij} \quad (2.16)$$

disebut peluang transisi satu langkah sedangkan peluang transisi  $n$  tahap adalah

$$P\{X_{t+n} = j | X_t = i\} = P_{ij}^{(n)} \quad (2.17)$$

Rantai Markov (*Markov Chain*) dikatakan berdistribusi stasioner  $\pi(x)$  jika

$$P\{X_{t+1} = j | X_t = i\} = P\{X_1 = j | X_0 = i\} = P_{ij} \quad (2.18)$$

untuk  $t = 0, 1, 2, \dots$ , sehingga peluang transisi yang diulang tidak berubah setiap waktu. Rantai Markov dengan distribusi stasioner  $\pi(x)$  disebut *irreducible* jika setiap keadaan  $j$  dapat dicapai dari setiap keadaan  $i$  setelah sejumlah terbatas transisi, di mana  $i \neq j$  atau  $P_{ij}^{(n)} > 0$  untuk  $1 \leq n < \infty$  (Hines, 1990).

*Markov Chain Monte Carlo* merupakan metode simulasi berdasarkan teknik simulasi *Monte Carlo* untuk membangkitkan data dari distribusi tertentu yang cara kerjanya menggunakan sifat rantai Markov. Simulasi *Monte Carlo* merupakan suatu pendekatan untuk menduga fungsi distribusi dari variabel acak  $\{X_i\}$ .

Salah satu metode simulasi yang terkenal dalam MCMC adalah *Gibbs Sampler*. *Gibbs Sampler* adalah metode yang menggunakan distribusi bersyarat penuh yang dihubungkan dengan distribusi stasioner ( $\pi(x)$ ) (Scollnik, 1996).

Misalkan  $\pi(x_1, x_2, \dots, x_k)$  merupakan distribusi gabungan dari variabel  $X_j$  yang bersesuaian dengan distribusi stasioner ( $\pi(x)$ ), maka  $\pi(x_j)$  merupakan distribusi marginal dan  $\pi(x_j | x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_k)$  merupakan distribusi bersyarat penuh dari variabel  $X_j$  di mana  $j = 1, 2, \dots, k$ . Sehingga proses *Gibbs Samplerr* dilakukan dengan cara membangkitkan sampel dari distribusi bersyarat penuh dengan mengikuti langkah-langkah sebagai berikut:

1. Memilih nilai awal  $\mathbf{x}^{(0)} = (x_1^{(0)}, x_2^{(0)}, \dots, x_k^{(0)})$

Memilih nilai awal untuk  $\mathbf{x}^{(0)}$  adalah sembarang nilai sesuai dengan batas ketentuan masing-masing distribusi. Nilai ini akan digunakan untuk menduga  $\mathbf{x}^{(1)}$  dan  $\mathbf{x}^{(1)}$  akan digunakan untuk menduga  $\mathbf{x}^{(2)}$  dan seterusnya.

2. Dilakukan simulasi pengambilan sampel dari variabel acak berdatarkan distribusi bersyarat penuh

$$\begin{aligned} \mathbf{X}_1^{(i)} &\sim \pi(x_1 | x_2^{(i-1)}, \dots, x_k^{(i-1)}) \\ \mathbf{X}_2^{(i)} &\sim \pi(x_2 | x_1^{(i)}, x_3^{(i-1)}, \dots, x_k^{(i-1)}) \\ \mathbf{X}_3^{(i)} &\sim \pi(x_3 | x_1^{(i)}, x_2^{(i)}, x_4^{(i-1)}, \dots, x_k^{(i-1)}) \\ &\vdots \\ \mathbf{X}_k^{(i)} &\sim \pi(x_k | x_1^{(i)}, x_2^{(i)}, \dots, x_k^{(i-1)}) \end{aligned}$$

Sehingga  $\mathbf{X}^{(i)} = (\mathbf{X}_1^{(i)}, \mathbf{X}_2^{(i)}, \dots, \mathbf{X}_k^{(i)})$

Simulasi ini dilakukan mulai  $i = 1$  hingga  $n$  iterasi yang diinginkan. Ketika  $n \rightarrow \infty$  maka nilai  $\mathbf{X}^{(n)}$  dari masing-masing distribusi bersyarat penuh akan tampak sangat acak. Untuk menduga variabel acak dari distribusi marginal yang dihasilkan dari proses simulasi digunakan rata-rata dari distribusi bersyarat penuh, yaitu

$$\pi(x) = \frac{1}{n} \sum_{i=1}^n \pi(x | \theta^{(i)}) \quad (2.19)$$

di mana :

$n$  : banyaknya iterasi yang diinginkan

$\theta^{(i)}$  : vektor parameter

Nilai awal  $\mathbf{x}^{(0)}$  digunakan untuk menduga nilai  $\mathbf{x}^{(1)}$  dan seterusnya hingga  $\mathbf{x}^{(n)}$ . Hal ini berakibat secara langsung pada  $\mathbf{x}^{(1)}$  yang bergantung pada  $\mathbf{x}^{(0)}$ . Namun pada kondisi  $n \rightarrow \infty$  pengaruh nilai awal  $\mathbf{x}^{(0)}$  akan berkurang terhadap  $\mathbf{x}^{(n)}$ , sehingga dianggap tidak berpengaruh lagi. Pada kondisi ini rantai Markov akan menuju distribusi stasioner ( $\pi(x)$ ) yang tidak bergantung pada nilai awal  $\mathbf{x}^{(0)}$ . Jika iterasi ke- $m$  ( $m < n$ ) sampel random sudah konvergen, maka sampel random ke-1 hingga ke- $m$  dihilangkan dan sampel random yang digunakan untuk menduga rata-rata dari distribusi stasioner ( $\pi(x)$ ) adalah iterasi ke- $(m + 1)$  hingga  $n$ . Kondisi dimana iterasi ke- $m$  yang ditiadakan untuk menghilangkan pengaruh nilai awal disebut kondisi *burn-in*. Karena iterasi ke- $m$  ditiadakan sehingga rata-rata dari distribusi bersyarat penuh dari persamaan (2.15) menjadi

$$\pi(x) = \frac{1}{n-m} \sum_{i=m+1}^n \pi(x | \theta^{(i)}) \quad (2.20)$$

Biasanya nilai 10-100 yang pertama dari iterasi dapat dibuang untuk menghilangkan pengaruh nilai awal.

Setelah pengambilan sampel random untuk menduga rata-rata

$\pi(x)$  berdasarkan kondisi *burn-in*, kemudian melihat nilai koefisien autokorelasi (ACF) apakah pengaruh ketergantungan dari nilai-nilai hasil iterasi ada atau tidak. Jika pada plot ACF terdapat nilai koefisien autokorelasi lag ke- $l$  berada di luar Batas selang  $\pm 2\sqrt{n}$  ( $n$  adalah banyaknya iterasi) maka terdapat pengaruh ketergantungan dari nilai-nilai hasil iterasi, sehingga pengambilan sampel untuk menduga rata-rata dari distribusi stasioner ( $\pi(x)$ ) dilakukan setiap kelipatan ke- $l$  (Scollnik, 1996).

## 2.8 Pengujian Hipotesis

### 2.8.1. Pengujian Secara Simultan

Menurut Hosmer dan Lemeshow (1989), pengujian koefisien regresi secara serentak dilakukan untuk menguji signifikansi dari koefisien regresi secara bersama-sama. Uji yang biasa digunakan adalah uji Nisbah Kemungkinan (*Likelihood Ratio Test*).

Hipotesis yang melandasi uji Nisbah Kemungkinan adalah :

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

versus

$$H_1 : \text{paling sedikit ada satu } \beta_j \neq 0$$

Jika  $H_0$  benar, maka statistik uji yang digunakan dalam uji Nisbah Kemungkinan, yaitu :

$$G = -2(L_R - L_F) \sim \chi^2_{(v)} \quad (2.21)$$

dimana :

$L_R$  = *log-likelihood* model tanpa seluruh variabel  $X_j$   
(Model Intersep)

$L_F$  = *log-likelihood* model dengan intersep dan seluruh variabel  $X_j$  (Model Penuh)

$v$  = selisih antara banyaknya parameter yang diduga dalam model penuh dan model intersep

Pada model regresi Poisson diperoleh persamaan :

$$L_R = \sum_{i=1}^n \left[ -y_i + Y_i \log \left( \frac{y_i}{n} \right) - \log(Y_i!) \right]$$

$$L_F = \sum_{i=1}^n \left[ -\mu_i + Y_i \log(\mu_i) - \log(Y_i!) \right]$$

Sehingga statistik uji G dapat dinyatakan sebagai :

$$G = 2 \left[ \sum_{i=1}^n \left[ y_i - Y_i \log \left( \frac{y_i}{n} \right) \right] - \sum_{i=1}^n [\mu_i + Y_i \log(\mu_i)] \right] \quad (2.22)$$

Jika  $P[\chi^2_{(v)} < W]$  lebih kecil dari taraf  $\alpha$  (taraf nyata yang diinginkan) maka  $H_0$  ditolak. Hal ini mengindikasikan bahwa paling sedikit ada satu  $\beta_j$  yang tidak sama dengan nol, atau dengan kata lain variabel penjelas  $X_j$  berpengaruh nyata terhadap variabel respon  $Y$ .

### 2.8.2. Pengujian Secara Parsial

Menurut Hosmer dan Lemeshow (1989), uji koefisien secara parsial digunakan untuk memeriksa peranan koefisien regresi dari masing-masing variabel penjelas secara individu pada model dengan cara membandingkan penduga dengan ragam penduga. Uji yang biasa yang digunakan untuk menentukan nyata tidaknya setiap koefisien regresi adalah uji Wald.

Hipotesis yang digunakan dalam pengujian ini adalah :

$$H_0 : \beta_j = 0$$

versus

$$H_1 : \beta_j \neq 0$$

Jika  $H_0$  benar, maka statistik uji yang digunakan dalam uji Wald yaitu :

$$W = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \sim Z \quad (2.23)$$

atau,

$$W^2 = \frac{\hat{\beta}_j^2}{[SE(\hat{\beta}_j)]^2} \sim \chi^2_{(p)} \quad (2.24)$$

dimana :

$\hat{\beta}_j$  : koefisien regresi pada variabel ke- $j$

$SE(\hat{\beta}_j)$  : galat baku penduga koefisien regresi

$p$  : banyaknya variabel penjelas

Nilai  $SE(\hat{\beta}_j)$  dapat diperoleh dari :

$$SE(\hat{\beta}_j) = \frac{\sum_{j=1}^n (\beta_j - \bar{\beta})^2}{n-1} \quad (2.25)$$

Menurut Agresti (2002), jika ukuran contoh besar ( $n \geq 30$ ), maka statistik uji Wald akan mengikuti distribusi normal baku, seperti pada persamaan (2.23). Sedangkan jika ukuran contoh kecil ( $n \leq 30$ ) maka statistik uji Wald akan menyebar  $\chi^2$ , seperti pada persamaan (2.24).

Hipotesis nol ( $H_0$ ) ditolak jika statistik  $W > Z_{tabel}$  atau p-value lebih kecil dari taraf  $\alpha$  (taraf nyata yang diinginkan). Artinya, koefisien dari variabel penjelas ke- $j$  berpengaruh signifikan dalam model (Hosmer dan Lemeshow, 1989).

## 2.9 Pengujian Kelayakan Model

Agresti (2002) menyatakan bahwa ada beberapa metode yang dapat digunakan untuk menguji kelayakan suatu model, salah satunya dengan menggunakan Uji Khi-Kuadrat Pearson (*Pearson Chi-Square*). Hipotesis yang digunakan untuk menguji kelayakan model Regresi Poisson berdasarkan Uji Khi-Kuadrat adalah :

$$H_0 : E(Y) = \exp(X^T \beta)$$

versus

$$H_1 : E(Y) \neq \exp(X^T \beta)$$

Statistik Uji Khi-Kuadrat untuk Regresi Poisson yaitu :

$$\chi^2 = \sum_{i=1}^n \frac{(y_i - \hat{\mu}_i)^2}{\hat{\mu}_i} \sim \chi^2_{(n-p)} \quad (2.26)$$

Model yang diperoleh sesuai dan layak digunakan jika nilai  $\chi^2_{Pearson} < \chi^2_{\alpha(n-p)}$ , sehingga menghasilkan keputusan untuk menerima  $H_0$ .

UNIVERSITAS BRAWIJAYA



## BAB III METODE PENELITIAN

### 3.1 Data Penelitian

Dalam penelitian ini data yang digunakan adalah data sekunder yang diperoleh dari Dinas Kesehatan Kota Malang dan Badan Pusat Statistika Kota Malang. Data yang digunakan antara lain :

- Banyaknya penderita penyakit DBD di masing-masing kelurahan Kota Malang tahun 2010 ( $Y$ )
- Indeks kepadatan penduduk perkelurahan ( $X_1$ )
- Banyaknya sarana kesehatan perkelurahan ( $X_2$ )
- Banyaknya rumah tangga miskin perkelurahan ( $X_3$ )
- Banyaknya sekolah masing-masing kelurahan ( $X_4$ )
- Persentase rumah yang ber-PHBS ( $X_5$ )
- Rata-rata penderita di daerah sekitar ( $X_6$ )
- Banyaknya TPS masing-masing kelurahan ( $X_7$ )

Data-data diatas terdapat pada lampiran.

### 3.2 Metode Bayesian MCMC Pada Model Regresi Poisson

Untuk menduga parameter model regresi Poisson yang digunakan untuk mengetahui variabel-variabel yang berpengaruh terhadap banyaknya kasus DBD di Kota Malang dilakukan dengan simulasi metode Bayesian MCMC.

Proses simulasi ini digunakan untuk membangkitkan parameter  $\beta_j$  untuk menduga parameter  $\beta_1 - \beta_7$ . Langkah-langkah dalam proses simulasi adalah sebagai berikut :

1. Membentuk model Bayesian berdasarkan fungsi Likelihood  $\beta$  atau distribusi bersyarat  $Y_i, X_{ij}$  Poisson ( $\mu(\mathbf{X}_{ij}^T, \boldsymbol{\beta})$ ) dan distribusi prior  $\beta_j$  (Normal ( $\mu_j, \sigma_j^2$ )).
2. Model Bayesian berdasarkan fungsi Likelihood Poisson dan distribusi prior akhirnya dapat membentuk distribusi posterior sebagai berikut :

$$f(\beta_j | y_i, x_{ij}) = \exp \left( - \sum_{j=1}^p \frac{1}{2\sigma_j^2} \beta_j^2 + \sum_{j=1}^p \left( \frac{\mu_j}{\sigma_j^2} + \sum_{i=1}^n Y_i X_{ij} \right) - \sum_{i=1}^n \exp \left( \sum_{j=1}^p X_{ij} \beta_j \right) \right)$$

(Doss dan Narasimhan, 1994).

3. Berdasarkan distribusi bersyarat Poisson ( $\mu_i$ ) dan distribusi posterior dilakukan pengambilan sampel dengan *Gibbs Sampler* yaitu sebagai berikut :

- a. Menentukan nilai awal  $\beta_j^{(0)}$  yang selanjutnya nilai  $\beta_j^{(0)}$  digunakan untuk menentukan  $\beta_j^{(1)}$  dan seterusnya nilai  $\beta_j^{(n-1)}$  yang menentukan nilai  $\beta_j^{(n)}$  .
- b. Dilakukan iterasi dari  $i=1$  hingga  $n$  (dalam penelitian ini  $n=85000$ ).

$$\begin{aligned}
 & \beta_j^{(1)} \sim \pi(\beta_j^{(0)} | Y_i, X_{ij}) \propto \\
 i = 1 \rightarrow & \exp\left(-\sum_{j=1}^p \frac{1}{2\sigma_j^2} \beta_j^{2(0)} + \sum_{j=1}^p \left(\frac{\mu_j}{\sigma_j^2} + \sum_{i=1}^n Y_i, X_{ij}\right) - \sum_{i=1}^n \exp\left(\sum_{j=1}^p X_{ij} \beta_j^{(0)}\right)\right)
 \end{aligned}$$

$$\begin{aligned}
 & \beta_j^{(2)} \sim \pi(\beta_j^{(1)} | Y_i, X_{ij}) \propto \\
 i = 2 \rightarrow & \exp\left(-\sum_{j=1}^p \frac{1}{2\sigma_j^2} \beta_j^{2(1)} + \sum_{j=1}^p \left(\frac{\mu_j}{\sigma_j^2} + \sum_{i=1}^n Y_i, X_{ij}\right) - \sum_{i=1}^n \exp\left(\sum_{j=1}^p X_{ij} \beta_j^{(1)}\right)\right)
 \end{aligned}$$

⋮

$$\begin{aligned}
 & \beta_j^{(n)} \sim \pi(\beta_j^{(n-1)} | Y_i, X_{ij}) \propto \\
 i = n \rightarrow & \exp\left(-\sum_{j=1}^p \frac{1}{2\sigma_j^2} \beta_j^{2(n-1)} + \sum_{j=1}^p \left(\frac{\mu_j}{\sigma_j^2} + \sum_{i=1}^n Y_i, X_{ij}\right) - \sum_{i=1}^n \exp\left(\sum_{j=1}^p X_{ij} \beta_j^{(n-1)}\right)\right)
 \end{aligned}$$

4. Untuk menduga parameter Regresi Poisson dilakukan langkah-langkah sebagai berikut :

- a. Membuang 5000 nilai awal ( $m=5000$ ) dari hasil langkah 2 untuk menghilangkan pengaruh nilai awal
- b. Membuat plot ACF untuk mengetahui apakah  $\beta_j^{(n)}$  yang diperoleh sudah saling bebas
  - Jika sudah bebas maka dapat dilakukan pengujian parameter
  - Jika tidak bebas maka dihitung nilai  $l$  yaitu jumlah ACF yang berada diluar batas dan dilakukan pengambilan sampel setiap kelipatan  $ke-l$ , kemudian dilakukan pengujian parameter  $\beta_j$

### 3.3 Proses Simulasi

Proses simulasi pada pendugaan parameter Bayesian MCMC dilakukan dengan bantuan paket program WinBUGS (*Windows Bayesian inference Using Gibbs Sampling*) versi 1.4 yang dilakukan dengan langkah-langkah sebagai berikut:

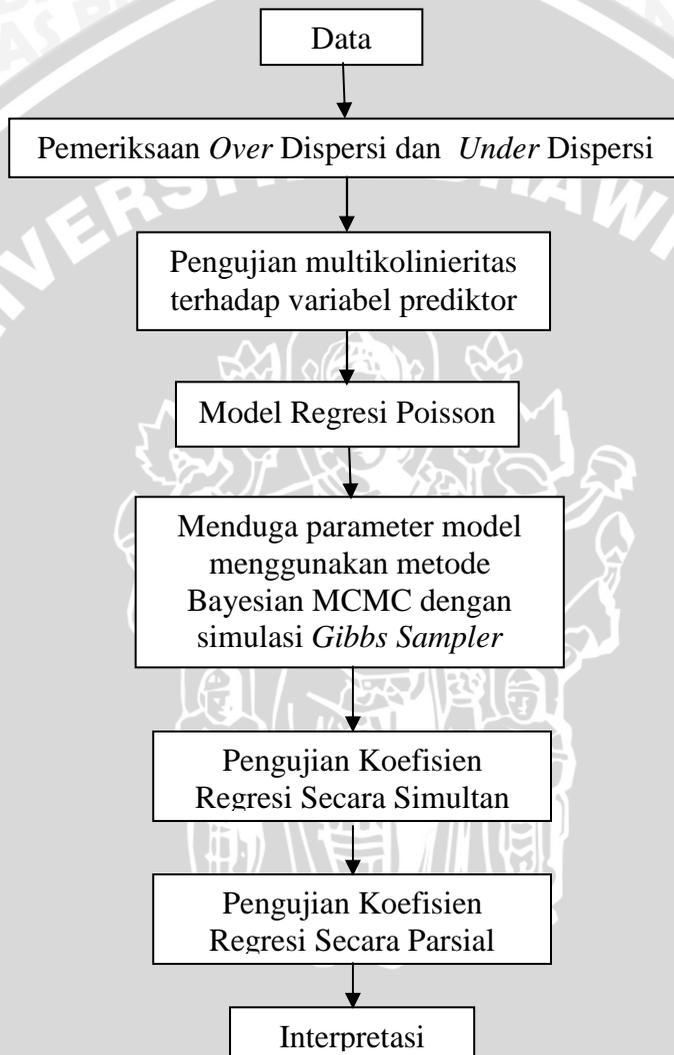
- a. Membuat model dalam menu *Model Specification*. Menu ini berisi pengecekan program pada *Model*, *Data* dan *Iniits*.
- b. Memasukkan parameter yang akan diduga dalam menu *Inference/Sample*.
- c. Melakukan iterasi dengan memilih menu *Model/Update* dan memasukkan nilai banyaknya iterasi yang akan dilakukan. Iterasi berakhir apabila telah konvergen yang ditandai dengan pola acak pada *dynamic trace* yang terbentuk. Cara lain yang dapat dilakukan adalah melihat nilai *MC error*. Iterasi dihentikan apabila *MC error* bernilai kurang dari 5% simpangan baku.
- d. Menggunakan menu *Sample*, akan didapatkan output berupa pendugaan parameter, gambar plot ACF, *kernel density* dan *dynamic trace*.

### 3.4 Tahap Analisis

Tahapan dalam menduga parameter model regresi poisson, yaitu :

- a. Menentukan model umum regresi poisson, seperti pada persamaan (2.7)
- b. Menguji overdispersi / underdispersi pada data
- c. Menduga parameter regresi Poisson dengan metode Bayesianian Markov Chain Monte Carlo (MCMC) dengan bantuan software WinBUGS 1.4.
- d. Melakukan pengujian asumsi multikolinieritas masing-masing variabel prediktor pada model yang terbentuk
- e. Melakukan pengujian signifikansi model melalui uji parsial dan uji simultan pada model yang terbentuk
- f. Menguji Kelayakan model
- g. Menyatakan model akhir
- h. Melakukan interpretasi peubah-peubah prediktor yang mempengaruhi peubah respon pada data.

Bagan alur metode analisis disajikan pada gambar 3.1 sebagai berikut :



Gambar 3.1 Langkah-langkah analisis

## BAB IV HASIL DAN PEMBAHASAN

### 4.1 Pengujian Distribusi Poisson Terhadap Variabel Y

Pada regresi Poisson variabel dependen ( $y$ ) harus berdistribusi Poisson, sehingga dilakukan pengujian distribusi. Pengujian dilakukan dengan uji *Kolmogorov Smirnov* dengan hipotesis :

$H_0$  : Data pada variabel Y berdistribusi Poisson

$H_1$  : Data pada variabel Y tidak berdistribusi Poisson

Dari hasil pengujian diperoleh nilai  $p$ -value sebesar 0,199 ( $> \alpha=0,05$ ) sehingga dapat disimpulkan bahwa variabel banyaknya penderita DBD di masing-masing kelurahan di Kota Malang memiliki distribusi Poisson.

### 4.2 Pemeriksaan *Over/under* dispersi Regresi Poisson

*Over* dispersi terjadi jika ( $\omega > 1$ ) sedangkan *under* dispersi terjadi apabila ( $\omega < 1$ ). Pemeriksaan terjadinya *over* dispersi dan *under* dispersi dilakukan dengan menggunakan nilai statistik Khi-Kuadrat Pearson dibagi dengan derajat bebas residual yang diperoleh dari hasil analisis Regresi Poisson (hasil perhitungan dapat dilihat pada Lampiran 3). Dari output pada Lampiran 3 didapatkan rasio nilai statistik Khi-Kuadrat Pearson dengan derajat bebas residual sebesar 1,640. Nilai 1,640 menunjukkan bahwa data banyaknya penderita DBD pada masing-masing kelurahan di Kota Malang mengalami *over* dispersi karena rasionya lebih dari 1.

### 4.3 Pengujian Non Multikolinieritas

Seerti halnya pada regresi linier, pada regresi Poisson juga dilakukan pengujian multikolinieritas untuk mengetahui korelasi antara variabel prediktor. Setelah dilakukan pengujian dengan VIF seperti pada persamaan (2.10) diperoleh hasil yang ditunjukkan pada Tabel 4.1

Tabel 4.1 Hasil pengujian multikolinieritas

Variabel	VIF
X <sub>1</sub>	1,256
X <sub>2</sub>	1,132
X <sub>3</sub>	1,400
X <sub>4</sub>	1,227
X <sub>5</sub>	1,445
X <sub>6</sub>	1,115
X <sub>7</sub>	1,344

Tabel 4.1 menunjukkan bahwa X<sub>1</sub> - X<sub>7</sub> memiliki nilai VIF < 10, sehingga dapat disimpulkan bahwa tidak terjadi multikolinieritas antara variabel prediktor.

#### 4.4 Metode Bayes MCMC pada Model Regresi Poisson

Penaksiran model regresi Poisson dilakukan dengan menggunakan metode *Bayes MCMC* dengan pembangkitan sampel menggunakan *Gibbs Sampler*. Proses pendugaan dilakukan dengan bantuan software WINBUGS 14. Proses simulasi diawali dengan membentuk model Bayes berdasarkan fungsi Likelihood  $L(\beta)$  dan distribusi prior  $\beta_j(\text{Normal}(\mu_j, \sigma_j^2))$ . Selanjutnya melakukan iterasi *Gibbs Sampler* berdasarkan distribusi Poisson ( $\mu_i$ ) dan distribusi Posterior ( $f(\beta_j|Y_i, X_{ij})$ ).

Iterasi dilakukan sebanyak 85000 sampel dengan menghilangkan 5000 nilai pertama guna menghilangkan pengaruh nilai awal (burn-in). Karena 5000 nilai pertama dihilangkan sehingga sampel yang digunakan sebagai pendugaan parameter adalah sebanyak 80000 sampel. Hasil pendugaan parameter Regresi Poisson ditunjukkan pada Tabel 4.2.

Tabel 4.2 Hasil analisis pendugaan parameter Regresi Poisson

Parameter	Mean ( $\beta$ )	sd	MC error
alpha ( $\alpha$ )	1,773	0,2055	0,005366
beta [1] ( $\beta_1$ )	4,5E-6	5,32E-6	4,333E-6
beta [2] ( $\beta_2$ )	0,04617	0,02352	1,985E-4
beta [3] ( $\beta_3$ )	0,001053	1,44E-4	1,383E-6
beta [4] ( $\beta_4$ )	-1,365E-4	0,001604	1,766E-5
beta [5] ( $\beta_5$ )	3,283E-4	0,005357	1,4E-3
beta [6] ( $\beta_6$ )	0,006168	0,006824	9,906E-5
beta [7] ( $\beta_7$ )	0,01398	0,04618	6,992E-2

Tabel 4.2 menunjukkan bahwa mean adalah rata-rata dari 80000 sampel iterasi pada masing-masing parameter. Nilai rata-rata itu digunakan sebagai parameter model Regresi Poisson, sehingga terbentuk model sebagai berikut :

$$Y_i = \exp [1,773 + (4,5E - 6) X_{i1} + 0,04617X_{i2} + 0,001053 X_{i3} - (1,365E - 4) X_{i4} + (3,283E - 4) X_{i5} + 0,006168X_{i6} + 0,013X_{i7}]$$

Nilai *sd* (*standard deviation*) adalah simpangan baku dari 80000 sampel iterasi. Simpangan baku yang kecil (mendekati nilai nol) menunjukkan bahwa tingkat fluktuasi dari 80000 sampel iterasi adalah kecil atau dapat dikatakan bahwa nilai dari 80000 sampel iterasi mengelompok di sekitar nilai rata-rata.

MC *error* merupakan salah baku dari proses *Markov Chain* atau tingkat kesalahan dari proses *Markov*. Menurut Spiegelhalter, dkk (2003) nilai MC *error* harus kurang dari 5% dari simpangan baku. Dan semua nilai MC *error* pada Tabel 4.2 menunjukkan kurang dari 5% dari simpangan baku.

## 4.5 Pengujian Signifikansi Penduga Parameter Regresi

### 4.5.1 Pengujian Simultan

Pengujian signifikansi parameter regresi secara simultan didasarkan pada statistik uji G. Berdasarkan hasil analisis diperoleh nilai G sebesar 104,219 (ouput selengkapnya pada Lampiran 7) dengan  $\chi_{0,05;1}^2 = 3,841$ .

Hal ini menunjukkan bahwa kepadatan penduduk, banyaknya sarana kesehatan, banyaknya rumah tangga miskin, banyaknya sekolah, persentase rumah tangga ber-PHBS (Pola Hidup Bersih dan Sehat), rata-rata jumlah penderita di daerah sekitar dan banyaknya TPS (Tempat Pembuangan Sampah) memberikan pengaruh yang signifikan terhadap banyaknya penderita DBD di masing-masing kelurahan Kota Malang. Hal ini juga berarti bahwa model sesuai pada taraf nyata 95%.

### 4.5.2 Pengujian Parsial

Pengujian signifikansi parameter regresi secara parsial dilakukan untuk mengetahui sejauh mana tingkat signifikansi dari pendugaan parameter regresi. Pengujian signifikansi parameter regresi secara parsial berdasarkan statistik uji Wald yang hasilnya secara ringkas disajikan pada tabel 4.3.

Tabel 4.3 Hasil Uji Parsial Koefisien Regresi *Poisson*

Variabel	Koefisien	Std.Error	Statistik Wald	$Z_{\alpha} = Z_{0,025}$
Konstanta	1,773	0,005366	322,96	1,96
X <sub>1</sub>	4,5E-6	4,333E-6	103,854	
X <sub>2</sub>	0,04617	1,985E-4	232,595	
X <sub>3</sub>	0,001053	1,383E-6	761,39	
X <sub>4</sub>	-1,365E-4	1,766E-5	-7,729	
X <sub>5</sub>	3,283E-4	1,4E-3	0,234	
X <sub>6</sub>	0,006168	9,906E-5	62,265	
X <sub>7</sub>	0,01398	6,992E-2	0,199	

Tabel 4.3 menunjukkan bahwa statistik Wald untuk pendugaan parameter  $\beta_1$  sebesar 103,854  $> Z_{tabel} = 1,96$ . Dari statistik Wald tersebut dapat dikatakan bahwa  $X_1$  berpengaruh signifikan terhadap model yang terbentuk. Begitu pula dengan statistik Wald dari  $\beta_2, \beta_3, \beta_4$  dan parameter  $\beta_6$  berturut-turut sebesar 232,595; 761,39; -7,729 dan 62,265, masing-masing |statistik Wald| menunjukkan nilai yang lebih dari  $Z_{tabel} = 1,96$  yang berarti  $X_2, X_3, X_4$  dan  $X_6$  berpengaruh signifikan terhadap model.

Sedangkan statistik Wald dari  $\beta_5$  dan  $\beta_7$  sebesar 0,234 dan 0,199 ( $< Z_{tabel} = 1,96$ ). Dari nilai tersebut dapat disimpulkan bahwa variabel  $X_5$  dan  $X_7$  tidak berpengaruh signifikan terhadap model. Parameter yang tidak signifikan pada model Regresi Poisson ini dikeluarkan dan kemudian dilakukan pengujian parsial terhadap parameter yang dihasilkan. Melalui uji parsial tanpa variabel  $X_5$  dan  $X_7$  diperoleh hasil statistik Wald seperti pada Tabel 4.4.

Tabel 4.4 Hasil Uji Parsial Koefisien Regresi *Poisson* tanpa  $\beta_5$  dan  $\beta_7$

Variabel	Koefisien	Std.Error	Statistik Wald	$Z_{\alpha} = Z_{0,025}$
Konstanta	1,762	0,001696	1038,915	1,96
X <sub>1</sub>	4,586E-6	4,234E-8	108,314	
X <sub>2</sub>	0,04646	2,191E-4	212,049	
X <sub>3</sub>	0,001064	1,348E-6	789,318	
X <sub>4</sub>	3,544E-5	1,095E-5	3,236	
X <sub>6</sub>	0,005758	8,268E-5	69,642	

Tabel 4.4 menunjukkan bahwa seluruh variabel berpengaruh nyata terhadap model regresi Poisson yang terbentuk. Sehingga diperoleh model akhir yang signifikan sebagai berikut :

$$Y_i = \exp [1,773 + (4,586E - 6) X_{i1} + 0,04646X_{i2} + 0,001064 X_{i3} + (3,544E - 4) X_{i4} + 0,005758 X_{i6}]$$

Pengujian signifikansi parameter regresi secara simultan didasarkan pada statistik uji G. Statistik uji G pada model Regresi Poisson. Berdasarkan output program dengan model tanpa  $X_5$  dan  $X_7$  diperoleh nilai G sebesar 103,917 (ouput selengkapnya pada Lampiran 12) dengan  $\chi^2_{0,05;1} = 3,841$ . Karena nilai statistik uji G > titik kritis maka tolak  $H_0$  model sesuai pada taraf nyata 95% yang berarti secara simultan model berpegaruh signifikan.

#### 4.6 Pengujian Kelayakan Model

Pengujian kelayakan model menggunakan Uji Khi-Kuadrat Pearson (*Pearson Chi-Square*). Hipotesis yang digunakan untuk menguji kelayakan model Regresi Poisson berdasarkan Uji Khi-Kuadrat adalah :

$$H_0 : E(Y) = \text{eXp}(X^T \beta)$$

versus

$$H_1 : E(Y) \neq \text{eXp}(X^T \beta)$$

Dari hasil output diperoleh nilai  $\chi^2 = 56,565$  dan  $\chi^2_{0,05;51} = 68,658$  dan dapat diputuskan bahwa  $\chi^2_{\text{Pearson}} < \chi^2_{\alpha(n-p)}$  (Terima  $H_0$ ). Sehingga dapat disimpulkan bahwa model layak untuk digunakan.

#### 4.7 Interpretasi Model

Dari pendugaan parameter Regresi Poisson dengan metode Bayesian dan telah dilakukan pengujian parameter diperoleh model sebagai berikut :

$$Y_i = \exp [1,773 + (4,586E - 6) X_{i1} + 0,04646X_{i2} + 0,001064 X_{i3} + (3,544E - 4) X_{i4} + 0,005758 X_{i6}]$$

Dari model diatas diketahui nilai alpha ( $\alpha$ ) yang merupakan konstanta dari model yang menunjukkan persentase rata-rata jumlah penderita DBD di Kota Malang jika tidak terjadi peningkatan atau

penurunan pada kepadatan penduduk ( $X_1$ ), banyaknya sarana kesehatan ( $X_2$ ), banyaknya rumah tangga miskin ( $X_3$ ), banyaknya sekolah ( $X_4$ ) dan banyaknya penderita DBD di daerah sekitar ( $X_6$ ) yaitu sebesar 1,773% penderita DBD di masing-masing kelurahan di Kota Malang.

Jika kepadatan penduduk ( $X_1$ ) meningkat sebanyak satu sedangkan variabel lain dianggap konstan maka  $Y_i$  akan meningkat sebesar 4,586E-6% dari banyaknya penderita sebelumnya, dikarenakan semakin padat penduduk maka membuat tempat berkembangbiaknya nyamuk pembawa virus dangue akan bertambah. Penambahan 1 nilai pada variabel  $X_2$  jika variabel lain konstan maka akan meningkatkan banyaknya penderita DBD sebesar 0,04646%, bertambahnya sarana kesehatan juga menyebabkan peningkatan banyaknya penderita DBD dikarenakan pada RS khususnya merupakan tempat yang memiliki kemungkinan besar dalam penularan suatu penyakit.

Jika pada suatu wilayah di Kota Malang terjadi peningkatan satu rumah tangga miskin maka akan meningkatkan banyaknya penderita DBD sebesar 0,001064%. Begitu pula dengan peningkatan yang terjadi pada jumlah sekolah ( $X_4$ ) dan rata-rata banyaknya penderita di daerah sekitar ( $X_6$ ) akan meningkatkan jumlah penderita DBD berturut-turut sebesar 3,544E-4% jika bertambah satu sekolah dan sebesar 0,005758% jika bertambahnya penderita di daerah sekitar sebanyak satu orang karena anak-anak di bangku sekolah merupakan usia yang rawan terserang DBD dan wilayah yang berdekatan juga merupakan faktor yang memungkinkan penyebaran nyamuk pembawa virus DBD.

## BAB V KESIMPULAN DAN SARAN

### 5.1 Kesimpulan

Kesimpulan yang bisa diperoleh dari penelitian ini adalah sebagai berikut:

1. Model Regresi Poisson yang terbentuk setelah dilakukan pendugaan parameter dengan metode Bayes MCMC dan pengujian hipotesis serta kelayakan model adalah

$$Y_i = \exp [1,773 + (4,586E - 6) X_{i1} + 0,04646X_{i2} + 0,001064 X_{i3} + (3,544E - 4) X_{i4} + 0,005758 X_{i6}]$$

2. Berdasarkan model diatas dapat disimpulkan bahwa variabel yang berpengaruh terhadap jumlah penderita DBD di Kota Malang adalah indeks kepadatan penduduk perkelurahan ( $X_1$ ), banyaknya sarana kesehatan perkelurahan ( $X_2$ ), banyaknya rumah tangga miskin ( $X_3$ ), banyaknya sekolah masing-masing kelurahan ( $X_4$ ) dan rata-rata jumlah penderita DBD di daerah sekitar ( $X_6$ ).

### 5.2 Saran

Saran yang bisa diberikan dalam penelitian selanjutnya sebagai berikut:

1. Sebaiknya menambahkan lebih banyak variabel prediktor seperti lokasi sekolah, persentase ABJ, banyaknya penyelenggaraan penanggulangan DBD (Sosialisasi 3M) dan variabel lain yang berpengaruh terhadap banyaknya penderita DBD di suatu wilayah.
2. Penelitian selanjutnya disarankan dapat memperhatikan faktor spasial yang sangat berperan terhadap penyebaran wabah penyakit dan dilakukan pemetaan untuk mengetahui wilayah endemik dan non endemik DBD.
3. Untuk Dinas Pemerintahan Kota Malang supaya lebih efektif terhadap penanggulangan kasus DBD di kota Malang, karena setelah dilakukan pemodelan 5 dari 7 variabel yang berpengaruh nyata terhadap peningkatan kasus DBD di Kota Malang. Pananggulangan yang efektif misalnya dengan penyuluan 3M (Menguras, Menimbun dan Menutup) tempat-tempat jentik nyamuk dapat berkembang,

menerapkan pola hidup bersih dan sehat, memberikan vaksinasi DBD untuk balita atau anak-anak di Kota Malang.

UNIVERSITAS BRAWIJAYA



## DAFTAR PUSTAKA

- Agresti, A. 2002. *An Introduction to Categorical Data Analysis*. John Wiley & Sons. New York.
- Berger, J.O. 1985. *Statistical Decision Theory and Bayesian Analysis, Second Edition*. Springer-Verlag New York, Inc. New York.
- Doss, H. Dan B. Narasimhan. 1994. *Bayesian Poisson Regression using the Gibbs Sampler : Sensitivity Analysis through Dynamic Graphics*. Department of Statistics, Florida State University.
- Gelman, A., John B. Carlin, Hal S. Stern dan Donald B. Rubin. 2004. *Bayesian Data Analysis Second Edition*. Chapman and Hall/ CRC Press Company, New York.
- Gujarati, D. 1995. *Ekonometrika Dasar Edisi Keempat. Terjemahan Sumarno Zain*. Erlangga. Jakarta.
- Hines, W.W. dan D.C Montgomery. 1990. *Probabilita dan Statistik Dalam Ilmu Rekayasa dan Manajemen (edisi kedua)*. Penerjemah : Ruidiansyah. Pendamping : Adler Haymans Manurung. UI-Press.
- Hosmer, D. W dan H. Lemeshow. 1989. *Applied Logistic Regression*. John Wiley and Sons, New York.
- Iversen, G. R. 1984. *Bayesian Statistical Inference*. The International Professional Publishers.
- Kecamatan Klojen Dalam Angka 2010*. Badan Pusat Statistika (BPS) Kota Malang.

*Kecamatan Sukun Dalam Angka 2010.* Badan Pusat Statistika (BPS) Kota Malang.

*Kecamatan Kedungkandang Dalam Angka 2010.* Badan Pusat Statistika (BPS) Kota Malang.

*Kecamatan Lowokwaru Dalam Angka 2010.* Badan Pusat Statistika (BPS) Kota Malang.

*Kecamatan Blimbing Dalam Angka 2010.* Badan Pusat Statistika (BPS) Kota Malang.

Kleinbaum, D.G., L.L. Kupper, K.E. Muller dan A.Nizam. 1998. *Applied Regression Analysis and Other Multivariable Third Edition.* Duxbury Press, Pacific Grove.

Kristina dan L. Wulandari. 2004. *Demam Berdarah Dangué.* Departemen Kesehatan Jakarta (Online), (<http://www.litbang.depkes.go.id/maskes/052004/DEMAMBERDARAH1.pdf>, diakses pada tanggal 18 April 2011)

Kutner, M. H, C. J. Nachtsheim, dan J. Neter. 2004. *Applied Linear Regression Models. 4th Edition.* Mc. Graw-Hill Companies, Inc. New York.

Martin, A. D. 2003. *Bayesian Inference for Heterogeneous Event Count.* Sage Publication, Washington.

Ntzoufras, I. 2009. *Bayesian Modeling Using WinBUGS.* John Wiley & Sons, Inc. New Jersey.

Pereira, F. 1999. *Practical “Modern” Bayesian Statistics In Actuarial Science.* General Insurance Convention.

Roose. 2008. *Hubungan Sosiomografi dan Lingkungan Dengan Kejadian Penyakit DBD di Kecamatan Bukit Raya Kota Pekanbaru*. USU Pers.

Sembiring, R.K. 1995. *Analisis Regresi*. Penerbit ITB, Bandung.

Scollnik, D.P.M. 1996. *An Introduction To Markov Chain Monte Carlo Methods and Their Actual Applications*. Department of Mathematics and Statistics University of Calgary.

Walpole, R.E. dan R. H. Myers. 1995. *Ilmu Peluang dan Statistika Untuk Insinyur dan Ilmuwan (edisi keempat)*. Penerbit ITB Bandung.



UNIVERSITAS BRAWIJAYA



## Lampiran 1. Data Penelitian

No	Kecamatan	Kelurahan	Y	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>
1	Klojen	Kasin	19	16024	5	486	19	38,70%	17	2
2		Sukoharjo	10	21804	0	300	20	42,80%	13	1
3		Kiduldalem	8	13418	2	234	10	21,00%	21	1
4		Kauman	9	17128	6	418	23	36,50%	10	2
5		Bareng	8	17418	2	312	18	30,00%	11	3
6		Gadingkasri	15	14021	3	280	14	38,90%	16	3
7		Oro oro dowo	13	10375	2	357	19	60,00%	9	4
8		Klojen	4	9870	3	213	22	65,00%	15	4
9		Rampal Celaket	12	13420	4	137	12	45,20%	18	2
10		Samaan	11	20794	5	486	11	41,00%	20	1
11		Penanggungan	15	21344	2	425	17	40,20%	9	1
12	Sukun	Kebonsari	3	5439	2	315	13	50,30%	9	0
13		Gadang	13	9588	2	297	17	10,60%	11	1
14		Ciptomulyo	12	19766	4	793	14	20,70%	14	0
15		Sukun	8	14531	3	302	29	40,60%	27	2
16		Bandungrejosari	27	10144	5	987	30	32,50%	18	0
17		Bakalan Krajan	6	4157	1	156	10	30,40%	6	1
18		Mulyorejo	19	4289	7	603	13	39,00%	21	1
19		Bandulan	10	6096	3	351	11	30,10%	22	1
20		Tanjungrejo	30	27607	1	1035	20	29,90%	29	1
21		Pisangcandi	19	9996	2	458	24	40,60%	13	0
22		Karangbesuki	6	5869	1	126	22	41,10%	9	2
23	Kedung Kandang	Arjowinangun	9	3032	1	251	42	42,00%	6	1
24		Tlogowaru	5	1145	2	148	46	35,00%	8	0
25		Wonokoyo	18	896	1	452	37	36,50%	12	1
26		Bumiayu	13	3504	2	326	50	35,00%	17	1
27		Buring	8	1536	1	247	56	45,00%	9	2
28		Mergosono	16	31473	2	645	71	52,00%	19	1
29		Kotalama	12	34222	2	410	158	53,10%	12	2

## Lampiran 1. Lanjutan

No	Kecamatan	Kelurahan	Y	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>
30		Kedungkandang	10	1961	3	480	45	52,50%	11	2
31		Sawojajar	48	16382	4	1537	97	60,00%	17	4
32		Madyopuro	11	4941	1	341	97	32,00%	22	1
33		Lesanpuro	9	4845	3	351	90	30,60%	7	1
34		Cemorokandang	0	3143	1	703	37	36,00%	15	1
35	Lowokwaru	Merjosari	7	4666	2	289	15	39,80%	9	1
36		Dinoyo	14	13197	5	205	18	37,50%	7	2
37		Sumbersari	8	11173	1	222	16	36,50%	29	2
38		Ketawanggede	16	9973	0	204	7	50,00%	18	1
39		Jatimulyo	15	7735	3	382	18	29,80%	12	1
40		Lowokwaru	16	14976	1	414	14	36,50%	7	3
41		Tulusrejo	13	12386	2	504	16	32,50%	26	2
42		Mojolangu	17	8618	3	326	21	36,40%	21	1
43		Tanjungsekar	12	7149	1	420	11	30,00%	19	1
44		Tasikmadu	11	2047	1	296	9	35,10%	3	1
45		Tunggulwulung	10	3148	1	204	7	30,20%	25	1
46		Tlogomas	9	7135	2	354	17	36,50%	10	2
47	Blimbing	Jodipan	8	26916	0	606	12	34,00%	18	1
48		Polehan	13	17145	1	554	10	35,00%	4	2
49		Kesatrian	14	7394	4	288	19	32,50%	14	1
50		Bunulrejo	12	13627	4	625	25	36,50%	17	2
51		Purwanto	10	12829	4	614	24	33,50%	19	1
52		Pandanwangi	12	6362	3	573	23	33,80%	11	1
53		Blimbing	8	9467	1	223	15	32,00%	14	1
54		Purwodadi	16	11563	6	739	22	36,50%	11	1
55		Polowijen	11	7401	3	121	5	39,00%	11	0
56		Arjosari	8	6483	1	216	21	38,00%	19	0
57		Balearjosari	3	4681	1	210	4	37,50%	11	2

Keterangan tabel :

Y : banyaknya penderita DBD

$X_1$  : Indeks Kepadatan Penduduk (Jiwa/ $\text{Km}^2$ )

$X_2$  : Banyaknya sarana kesehatan

$X_3$  : Jumlah Rumah Tangga miskin

$X_4$  : Jumlah sekolah (SD/ sederajatnya, SMP/ Sederajatnya, SMA/ sederajatnya)

$X_5$  : Persentase Rumah Tangga ber-PHBS (%)

$X_6$  : Rata-rata banyaknya penderita di daerah sekitar

$X_7$  : Banyaknya TPS



## Lampiran 2. Hasil Pengujian Distribusi Poisson

### One-Sample Kolmogorov-Smirnov Test

		Y
N		57
Poisson Parameter <sup>a</sup>	Mean	12.26
Most Extreme Differences	Absolute	.142
	Positive	.142
	Negative	-.061
Kolmogorov-Smirnov Z		1.073
Asymp. Sig. (2-tailed)		.199
a. Test distribution is Poisson.		



### Lampiran 3. Hasil Pengujian untuk *Over/under* dispersi

#### Goodness of Fit<sup>P</sup>

	Value	df	Value/df
Deviance	95.214	49	1.943
Scaled Deviance	95.214	49	
Pearson Chi-Square	80.365	49	1.640
Scaled Pearson Chi-Square	80.365	49	
Log Likelihood <sup>a</sup>	-166.847		
Akaike's Information Criterion (AIC)	349.695		
Finite Sample Corrected AIC (AICC)	352.695		
Bayesian Information Criterion (BIC)	366.039		
Consistent AIC (CAIC)	374.039		

Dependent Variable: Y

Model: (Intercept), X1, X2, X3, X4, X5, X6, X7

a. The full log likelihood function is displayed and used in computing information criteria.

b. Information criteria are in small-is-better form.



### Lampiran 4. Hasil Uji Non-Multikolinieritas

#### Coefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	.000	3.491		.000	1.000		
	X1	-1.004E-5	.000	-.011	-.102	.919	.796	1.256
	X2	.322	.443	.073	.727	.471	.883	1.132
	X3	.019	.003	.684	6.150	.000	.714	1.400
	X4	.007	.027	.027	.261	.795	.815	1.227
	X5	3.143	8.613	.041	.365	.717	.692	1.445
	X6	.090	.114	.078	.787	.435	.896	1.115
	X7	.632	.810	.085	.781	.439	.744	1.344

a. Dependent Variable: Y



## Lampiran 5. Node untuk metode pendugaan Regresi Poisson dengan software WINBUGS 14

```

model
{
  for (i in 1 : N) {
    y[i] ~ dpois(mu[i])
    log(mu[i]) <- alpha+ beta[1]*x1[i] + beta[2]*x2[i] + beta[3]*x3[i] +
beta[4]*x4[i] + beta[5]*x5[i]
      + beta[6]*x6[i] + beta[7]*x7[i]
  }
  # prior distribution
  alpha ~ dflat ( )
  for (j in 1:7) {
    beta [j] ~ dnorm(0.0, 1.0E-5)
  }
}

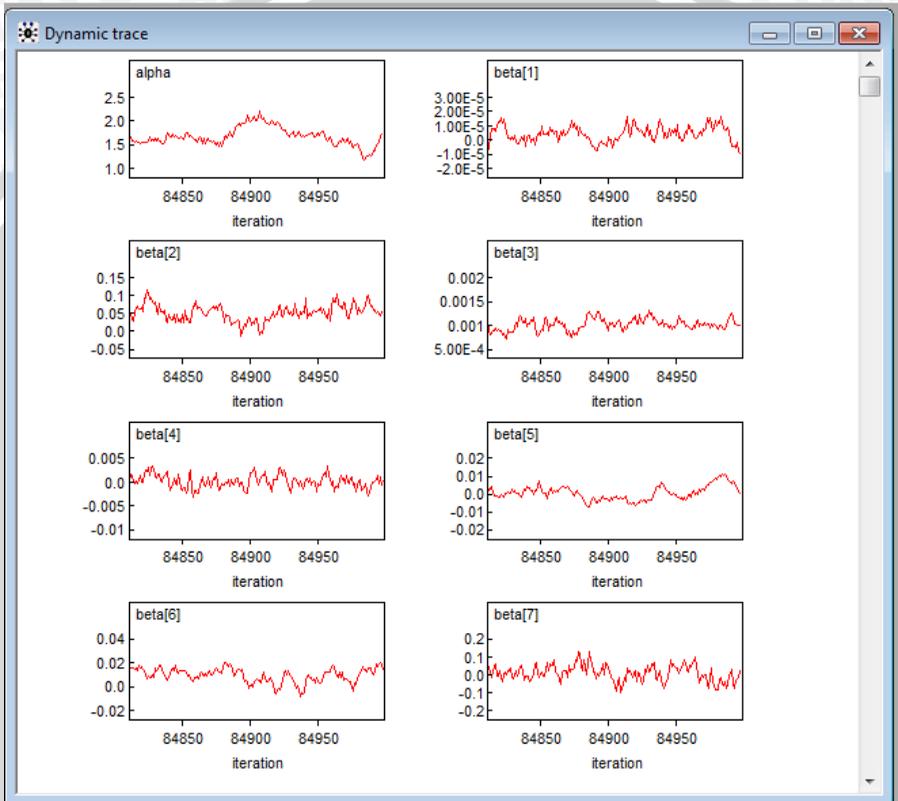
DATA
list( N=57,
y=c(19,10,8,9,8,15,13,4,12,11,15,3,13,12,8,27,6,19,10,30,19,6,9,5,18,13,8,16,12,10,4
8,11,9,0,7,14,8,16,15,16,13,17,12,11,10,9,8,13,14,12,10,12,8,16,11,8,3),
x1=c(16017,21804,13418,17128,17418,13871,10325,9870,13420,20794,21344,5439,
9588,19766,14531,10144,4157,4289,6096,27607,9996,5869,3032,1145,896,3504,15
36,31473,34222,1961,16382,4941,4845,3143,4666,13197,11173,9973,7735,14976,1
2386,8618,7149,2047,3148,7135,26916,17145,7394,13627,12829,6362,9467,11563,
7401,6483,4681),
x2=c(5,0,2,6,2,3,2,3,4,5,2,2,2,4,3,5,1,7,3,1,2,1,1,2,1,2,1,2,2,3,4,1,3,1,2,5,1,0,3,1,2,3,1
,1,1,2,0,1,4,4,4,3,1,6,3,1,1),
x3=c(486,300,234,418,312,280,357,213,137,486,425,315,297,793,302,987,156,603,3
51,1035,458,126,251,148,452,326,247,645,410,480,1537,341,351,703,289,205,222,2
04,382,414,504,326,420,296,204,354,606,554,288,625,614,573,223,428,121,116,210
),
x4=c(19,20,10,23,18,14,19,22,12,11,17,13,17,14,29,30,10,13,11,20,24,22,42,46,37,5
0,56,71,158,45,97,97,90,37,15,18,16,7,18,14,16,21,11,9,7,17,12,10,19,25,24,23,15,2
2,5,21,4),
x5=c(38.7,42.8,21.0,36.5,30.0,38.9,60.0,65.0,45.2,41.0,40.2,50.3,10.6,20.7,40.6,32.5,
30.4,39.0,30.1,29.9,40.6,41.1,42.0,35.0,36.5,35.0,45.0,52.0,53.1,52.5,60.0,32.0,30.6,
36.0,39.8,37.5,36.5,50.0,29.8,36.5,32.5,36.4,30.0,35.1,30.2,36.5,34.0,35.0,32.5,36.5,
33.5,33.8,32.0,36.5,39.0,38.0,37.5),
x6=c(17,13,21,10,11,16,9,15,18,20,9,9,11,14,27,18,6,21,22,29,13,9,6,8,12,17,9,19,12
,11,17,22,7,15,9,7,29,18,12,7,26,21,19,3,25,10,18,4,14,17,19,11,14,11,11,19,11),
x7=c(2,1,1,2,3,3,4,4,2,1,1,0,1,0,2,0,1,1,1,1,0,2,1,0,1,1,2,1,2,2,4,1,1,1,1,2,2,1,1,3,2,1,1
,1,1,2,1,2,1,2,1,1,1,1,0,0,2))

INITS
list(alpha=1; beta=c(1,1,1,1,1,1,1))

```

## Lampiran 6. Output WINBUGS 14

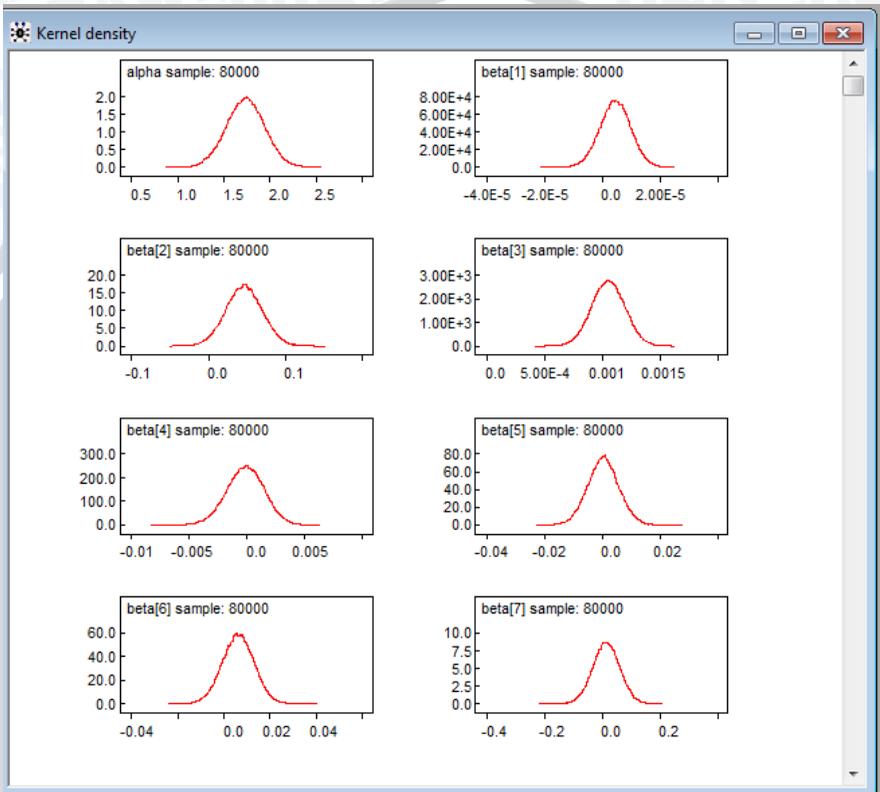
- Dynamic trace plot



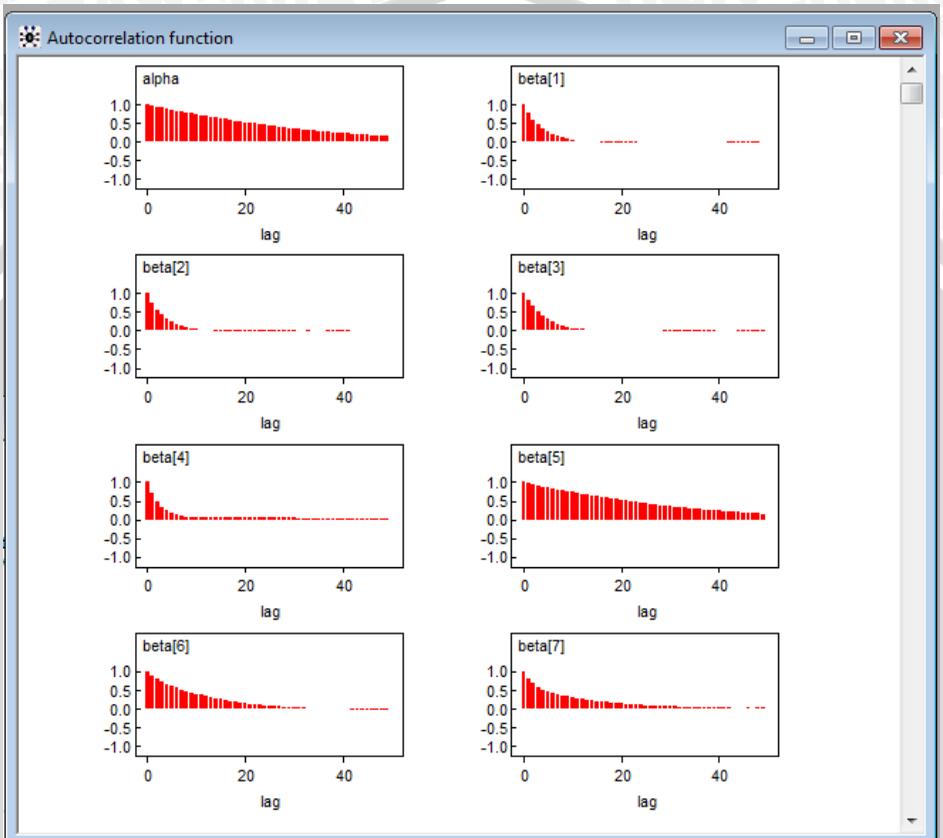
- Node Statistics

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
alpha	1.733	0.2055	0.005366	1.325	1.735	2.133	5001	80000
beta[1]	4.5E-6	5.32E-6	4.333E-8	-5.966E-6	4.532E-6	1.484E-5	5001	80000
beta[2]	0.04617	0.02352	1.985E-4	7.358E-5	0.04614	0.09232	5001	80000
beta[3]	0.001053	1.44E-4	1.383E-6	7.715E-4	0.001052	0.001336	5001	80000
beta[4]	-1.365E-4	0.001604	1.766E-5	-0.003343	-1.145E-4	0.002937	5001	80000
beta[5]	3.283E-4	0.005357	1.4E-3	-0.01009	3.293E-4	0.0109	5001	80000
beta[6]	0.006168	0.006824	9.906E-5	-0.00725	0.006186	0.0194	5001	80000
beta[7]	0.01398	0.04618	6.992E-2	-0.07683	0.01389	0.104	5001	80000

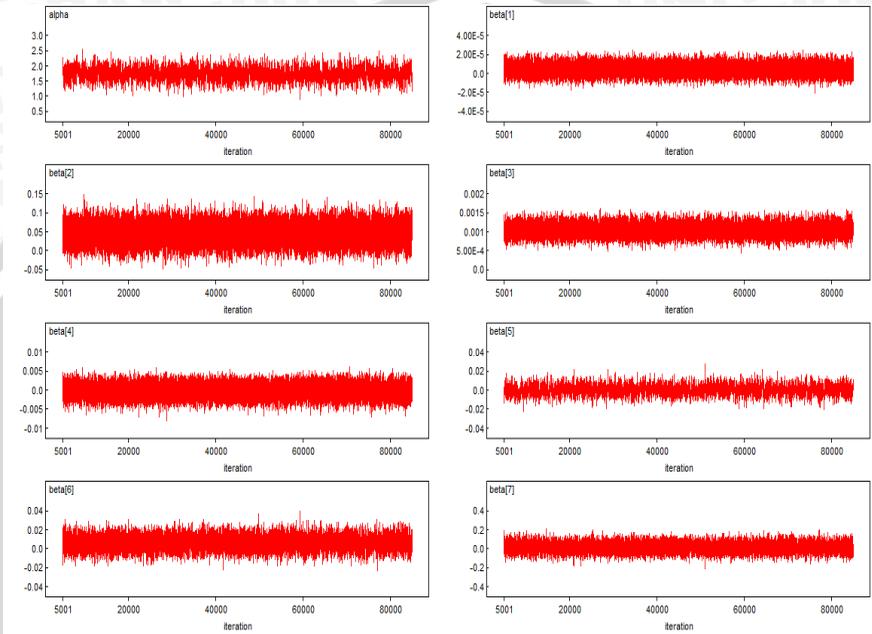
- Kernel Density plot



- Autocorrelation plot



- Time Series plot



## Lampiran 7. Output *Likelihood Ratio Chi-Square*

### Omnibus Test<sup>a</sup>

Likelihood Ratio Chi-Square	df	Sig.
104.219	7	.000

Dependent Variable: Y

Model: (Intercept), X1, X2, X3, X4, X5, X6, X7

a. Compares the fitted model against the intercept-only model.



## Lampiran 8. Output Statistik Wald

Tests of Model Effects

Source	Statistic	
	Wald Chi-Square	df
(Intercept)	322.96	1
X1	103.854	1
X2	232.595	1
X3	761.39	1
X4	-7.729	1
X5	0.234	1
X6	62.265	1
X7	0.199	1

Dependent Variable: Y

Model: (Intercept), X1, X2, X3, X4, X5, X6, X7



**Lampiran 9.** Node untuk metode pendugaan Regresi  
*Generalized Poisson* tanpa  $X_5$  dan  $X_7$  dengan  
 software WINBUGS 14

```

model
{
  for (i in 1 : N) {
    y[i] ~ dpois(mu[i])
    log(mu[i]) <- alpha+ beta[1]*x1[i] + beta[2]*x2[i] + beta[3]*x3[i] +
    beta[4]*x4[i] + beta[5]*x6[i]
  }
  # prior distribution
  alpha ~ dflat ( )
  for (j in 1:5) {
    beta [j] ~ dnorm(0.0, 1.0E-5)
  }
}

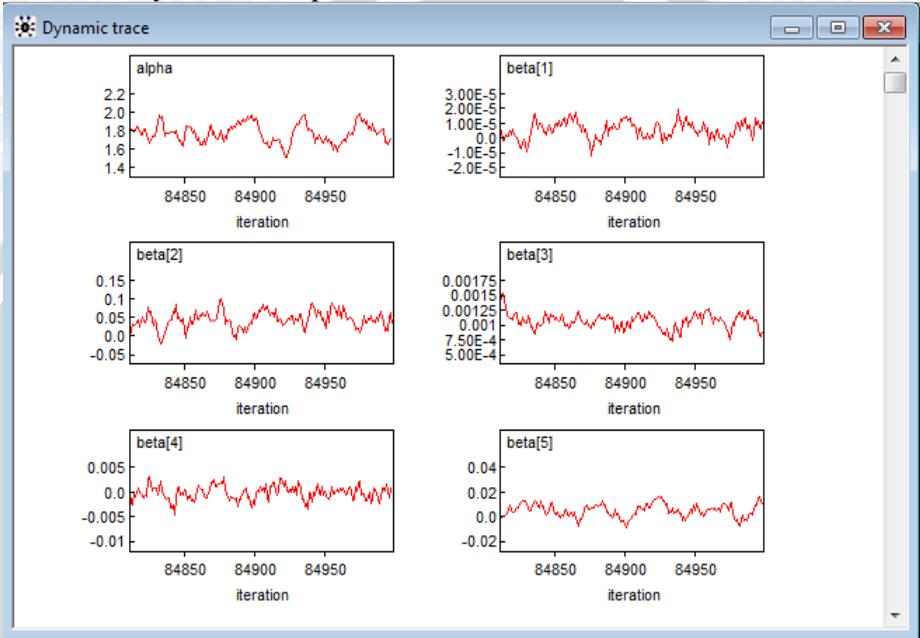
DATA
list( N=57,
y=c(19,10,8,9,8,15,13,4,12,11,15,3,13,12,8,27,6,19,10,30,19,6,9,5,18,13,8,16,12,10
,48,11,9,0,7,14,8,16,15,16,13,17,12,11,10,9,8,13,14,12,10,12,8,16,11,8,3),
x1=c(16017,21804,13418,17128,17418,13871,10325,9870,13420,20794,21344,543
,9,9588,19766,14531,10144,4157,4289,6096,27607,9996,5869,3032,1145,896,3504
,1536,31473,34222,1961,16382,4941,4845,3143,4666,13197,11173,9973,7735,149
76,12386,8618,7149,2047,3148,7135,26916,17145,7394,13627,12829,6362,9467,1
1563,7401,6483,4681),
x2=c(5,0,2,6,2,3,2,3,4,5,2,2,2,4,3,5,1,7,3,1,2,1,1,2,1,2,1,2,2,3,4,1,3,1,2,5,1,0,3,1,2,3
,1,1,1,2,0,1,4,4,4,3,1,6,3,1,1),
x3=c(486,300,234,418,312,280,357,213,137,486,425,315,297,793,302,987,156,603
,351,1035,458,126,251,148,452,326,247,645,410,480,1537,341,351,703,289,205,2
22,204,382,414,504,326,420,296,204,354,606,554,288,625,614,573,223,428,121,1
16,210),
x4=c(19,20,10,23,18,14,19,22,12,11,17,13,17,14,29,30,10,13,11,20,24,22,42,46,37,
50,56,71,158,45,97,97,90,37,15,18,16,7,18,14,16,21,11,9,7,17,12,10,19,25,24,23,1
5,22,5,21,4),
x6=c(17,13,21,10,11,16,9,15,18,20,9,9,11,14,27,18,6,21,22,29,13,9,6,8,12,17,9,19,
12,11,17,22,7,15,9,7,29,18,12,7,26,21,19,3,25,10,18,4,14,17,19,11,14,11,11,19,11))

INITS
list(alpha=1; beta=c(1,1,1,1,1))

```

## Lampiran 10. Output WINBUGS 14

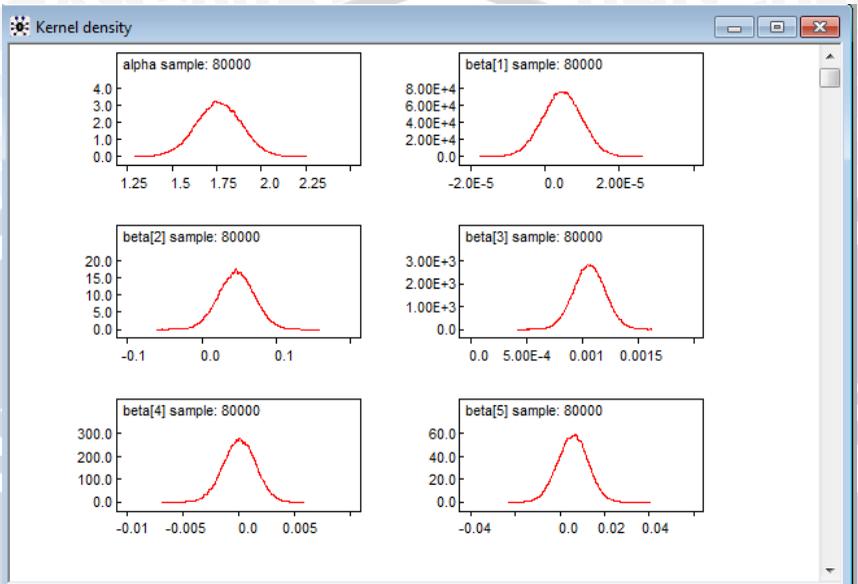
- Dynamic trace plot



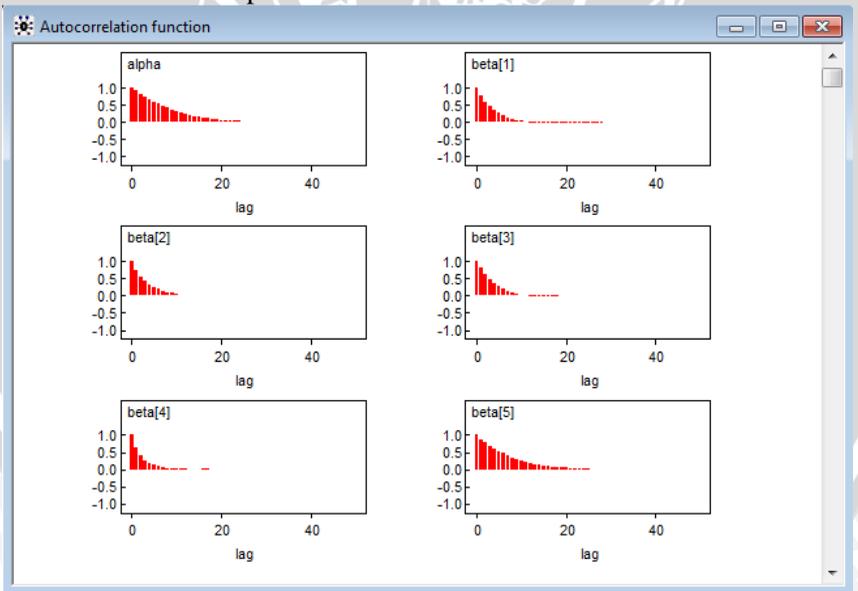
- Node Statistics

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
alpha	1.762	0.1257	0.001696	1.512	1.762	2.006	5001	80000
beta[1]	4.586E-6	5.294E-6	4.234E-8	-5.829E-6	4.606E-6	1.489E-5	5001	80000
beta[2]	0.04646	0.02377	2.191E-4	-1.565E-4	0.04655	0.09295	5001	80000
beta[3]	0.001064	1.409E-4	1.348E-6	7.873E-4	0.001063	0.00134	5001	80000
beta[4]	3.544E-5	0.001451	1.095E-5	-0.002866	5.322E-5	0.00283	5001	80000
beta[5]	0.005758	0.006764	8.268E-5	-0.007531	0.005833	0.01888	5001	80000

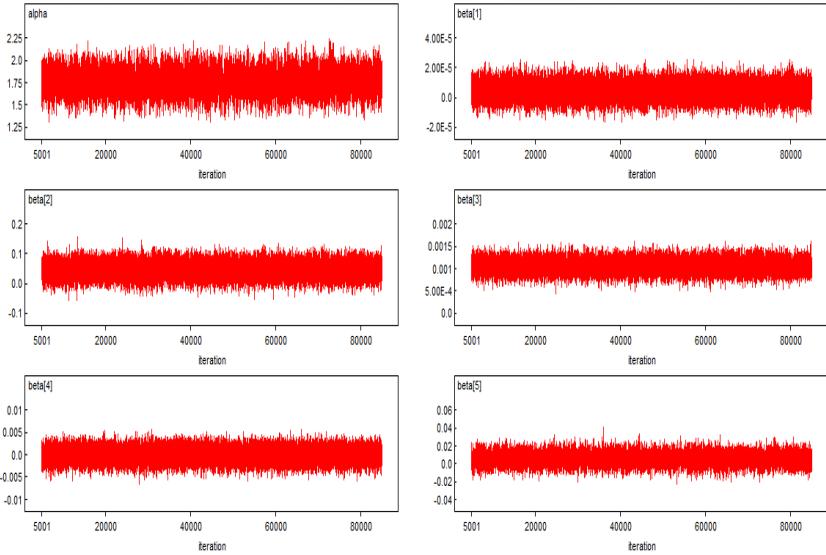
- Kernel Density plot



- Autocorellation plot



- Time Series plot



## Lampiran 11. Output Statistik Wald

### Tests of Model Effects

Source	Statistic	
	Wald Chi-Square	df
(Intercept)	1038.915	1
X1	108.314	1
X2	212.049	1
X3	789.318	1
X4	3.236	1
X6	69,642	1

Dependent Variable: Y

Model: (Intercept), X1, X2, X3, X4, X6



## Lampiran 12. Output *Likelihood Ratio Chi-Square*

Omnibus Test <sup>a</sup>		
Likelihood Ratio Chi-Square	df	Sig.
103.917	5	.000

Dependent Variable: Y

Model: (Intercept), X1, X2, X3, X4, X6

a. Compares the fitted model against the intercept-only model.



### Lampiran 13. Output Goodness of Fit

#### Pearson Chi-Square

Pearson Chi-Square	df
56.565	5

Dependent Variable: Y

Model: (Intercept), X1, X2, X3, X4, X6

