



TESIS - KS185411

**PENDEKATAN BAYESIAN PADA MODEL  
REGRESI *MIXTURE* POISSON UNTUK ESTIMASI  
FUNGSI INTENSITAS *SPATIAL POINT PROCESS***

**(Studi Kasus Pola Titik Persebaran Fasilitas Kesehatan Tingkat  
Pertama di Kota Surabaya)**

**TRI MURNIATI  
06211650010018**

Dosen Pembimbing  
Prof. Drs. Nur Iriawan, M.Ikomp, Ph.D  
Dr. rer. pol. Dedy Dwi Prastyo, M.Si

Departemen Statistika  
Fakultas Sains dan Analitika Data  
Institut Teknologi Sepuluh Nopember  
Surabaya 2020



TESIS - KS185411

**PENDEKATAN BAYESIAN PADA MODEL  
REGRESI *MIXTURE* POISSON UNTUK ESTIMASI  
FUNGSI INTENSITAS *SPATIAL POINT PROCESS***

(Studi Kasus Pola Titik Persebaran Fasilitas Kesehatan Tingkat Pertama di Surabaya)

TRI MURNIATI  
06211650010018

DOSEN PEMBIMBING:  
Prof. Drs. Nur Iriawan, M.Ikomp, Ph.D  
Dr. rer. pol. Dedy Dwi Prastyo, M.Si

PROGRAM MAGISTER  
DEPARTEMEN STATISTIKA  
FAKULTAS SAINS DAN ANALITIKA DATA  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA 2020





THESIS - KS185411

**BAYESIAN APPROACH ON THE MODELING OF  
MIXTURE POISSON REGRESSION FOR  
ESTIMATING INTENSITY FUNCTION  
OF SPATIAL POINT PROCESS**

(Case Study of Spatial Point Pattern of Primary Health Centres  
in Surabaya)

TRI MURNIATI  
06211650010018

SUPERVISOR:  
Prof. Drs. Nur Iriawan, M.Ikomp, Ph.D  
Dr. rer. pol. Dedy Dwi Prastyo, M.Si

MAGISTER PROGRAM  
DEPARTEMENT OF STATISTICS  
FAKULTY OF SCIENCE AND DATA ANALITICS  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA 2020



## LEMBAR PENGESAHAN TESIS

Telah disusun untuk memenuhi salah satu syarat memperoleh gelar  
**Magister Statistika (M.Stat)**

di

**Institut Teknologi Sepuluh Nopember**

Oleh:

**TRI MURNIATI**

**NRP.06211650010018**

Tanggal Ujian : 30 Januari 2020

Periode Wisuda : Maret 2020

Disetujui oleh:

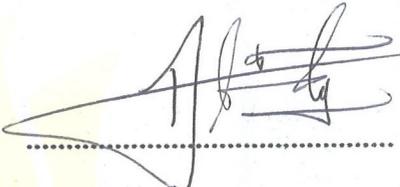
**Pembimbing:**

1. Prof. Drs. Nur Iriawan, M.Ikom, Ph.D.  
NIP. 19621015 198803 1 002



.....

2. Dr. rer. pol. Dedy Dwi Prastyo, M.Si.  
NIP. 19831204 200812 1 002



.....

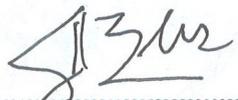
**Penguji:**

3. Dr. Kartika Fithriasari, M.Si.  
NIP. 19691212 199303 2 002



.....

4. Jerry Dwi Trijoyo Purnomo, M.Si, Ph.D.  
NIP. 19810223 200812 1 003



.....



**Kepala Departemen Statistika  
Fakultas Sains dan Analitika Data**

Dr. Kartika Fithriasari, M.Si.  
NIP. 19691212 199303 2 002

7

**PENDEKATAN BAYESIAN PADA  
MODEL REGRESI *MIXTURE* POISSON UNTUK ESTIMASI  
FUNGSI INTENSITAS *SPATIAL POINT PROCESS***

(Studi Kasus Pola Titik Persebaran  
Fasilitas Kesehatan Tingkat Pertama di Surabaya)

Nama Mahasiswa : Tri Murniati  
NRP : 06211650010018  
Dosen Pembimbing : Prof. Drs. Nur Iriawan, M.Ikomp, Ph. D  
Dr. rer.pol. Dedy Dwi Prastyo, S.Si, M.Si

**ABSTRAK**

*Spatial point process* sangat berguna sebagai model statistika untuk menganalisa pengamatan *point pattern* dimana titik (*point*) menunjukkan lokasi objek penelitian. *Point pattern* dapat dilihat sebagai proses Poisson jika diasumsikan antara titik lokasi saling independen yang kemudian disebut sebagai *Poisson Point Process*. Intensitas sebagai momen pertama *Poisson Point Process* didefinisikan sebagai rasio jumlah titik per unit area. Fungsi intensitas dapat dipandang sebagai model loglinier *Poisson Point Process* atau regresi Poisson. Pada kondisi dimana proses Poisson memiliki intensitas yang bervariasi secara spasial (*Nonhomogeneous Poisson Process*), regresi *mixture* Poisson dapat digunakan untuk pemodelan intensitas karena pemodelan dilakukan pada proses Poisson yang dibedakan menjadi  $k$  komponen terbatas yang saling heterogen. Untuk melakukan estimasi parameter regresi *mixture* dibentuk fungsi likelihood yang kompleks sehingga pendekatan Bayesian dengan *Markov Chain Monte Carlo* (MCMC) sering digunakan. Pada penelitian ini pendekatan Bayesian pada model regresi *mixture* Poisson diterapkan untuk memperoleh model intensitas titik lokasi fasilitas kesehatan tingkat pertama (FKTP) di Surabaya. Intensitas FKTP dapat digunakan untuk mengukur pemerataan fasilitas kesehatan yang menjadi perhatian Dinas Kesehatan guna memenuhi kebutuhan dasar kesehatan. Analisis menggunakan *grid* dengan ukuran *irregular tessellation* sesuai dengan batas kecamatan diplotkan pada peta Surabaya. Variabel kovariat yang berkaitan dengan tingkat kebutuhan layanan kesehatan digambarkan dengan kepadatan penduduk dan persentase rumah tangga perilaku hidup bersih sehat serta aksesibilitas yang digambarkan dengan angka keterkaitan dan panjang jalan dalam kondisi baik, mencerminkan karakteristik masing-masing kecamatan. Analisis menunjukkan bahwa persebaran FKTP di Surabaya memiliki intensitas yang *Nonhomogeneous Poisson Process* dengan dua kategori *mixture*. Persebaran FKTP secara signifikan dipengaruhi oleh tingkat kebutuhan layanan kesehatan dan aksesibilitas kecamatan.

**Kata kunci:** *Poisson Point Process, Regresi Mixture Poisson, Analisis Bayesian, Intensitas FKTP*

*(halaman ini sengaja dikosongkan)*

**BAYESIAN APPROACH ON THE MODELING OF  
MIXTURE POISSON REGRESSION FOR ESTIMATING  
INTENSITY FUNCTION OF SPATIAL POINT PROCESS**

(Case Study of Spatial Point Pattern of Primary Health Centres in Surabaya)

Name : Tri Murniati  
Student Identity Number : 06211650010018  
Supervisor : Prof. Drs. Nur Iriawan, M.Ikomp, Ph. D  
Dr. rer.pol. Dedy Dwi Prastyo, S.Si, M.Si

**ABSTRACT**

Spatial point process is statistical model for analyzing point patterns where points indicate the location of research objects. A point pattern can be seen as a Poisson process if it is assumed that between points of location are mutually independent. It is referred as the Poisson Point Process. Intensity as the first moment of the Poisson Point Process is defined as the ratio of the number of points per unit area. The intensity function can be seen as a loglinear Poisson Point Process or Poisson regression model. Under conditions where the Poisson process has a spatially varying intensity (Nonhomogeneous Poisson Process), Poisson mixture regression can be used for intensity modeling because it is done on the Poisson process which is divided into k-th heterogeneous components. To estimate the mixture regression parameters a complex likelihood function is formed so that the Bayesian approach with Markov Chain Monte Carlo (MCMC) is popular. In this study the Bayesian approach to the mixture Poisson regression model was applied to obtain the point intensity model of the location of the primary health centres (PHC) in Surabaya. Intensity of PHC can be used to measure the distribution of health facilities that are the Health Office concern to meet basic health needs. The analysis uses a grid with irregular tessellation size in accordance with the subdistrict boundaries plotted on the Surabaya map. Covariate variables related to the level of need for health services are depicted by population density and percentage of households with healthy hygiene behavior and accessibility as illustrated by the linkage rate and road length in good condition reflecting the characteristics of each district. The analysis shows that PHC distribution in Surabaya follows Nonhomogeneous Poisson Process with two categories of mixture. The distribution of PHC is significantly influenced by the level of health service needs and the accessibility of sub-districts.

**Keywords:** *Poisson Point Process, Mixture Poisson Regression, Bayesian Analysis, PHC Intensity*

*(halaman ini sengaja dikosongkan)*

## KATA PENGANTAR

Alhamdulillah, segala puji bagi Allah SWT yang telah melimpahkan karunia nikmat dan rahmat kepada penulis sehingga tesis yang berjudul “Pendekatan Bayesian Pada Model Regresi *Mixture* Poisson untuk Estimasi Fungsi Intensitas *Spatial Point Process* (Studi Kasus Pola Titik Persebaran Fasilitas Kesehatan Tingkat Pertama di Kota Surabaya)” ini dapat diselesaikan. Tesis ini disusun dalam rangka memenuhi salah satu syarat untuk meraih gelar Magister Sains di Program Pasca Sarjana Statistika Institut Teknologi Sepuluh Nopember (ITS) Surabaya.

Pada kesempatan ini penulis ingin menyampaikan ucapan terima kasih dan penghargaan yang setinggi-tingginya kepada:

1. Orang tua dan seluruh keluarga tercinta yang telah memberikan banyak do'a serta dukungan, sehingga penulis dapat menjalani kuliah sampai sekarang dan dapat menyelesaikan Tesis ini
2. Bapak Prof. Drs. Nur Iriawan, M.Ikom, Ph.D dan Bapak Dr. rer. pol. Dedy Dwi Prastyo, M.Si selaku pembimbing atas segala bimbingan, arahan dan motivasi dalam penyusunan tesis ini.
3. Ibu Dr. Kartika Fithriasari, M.Si, Bapak Dr. Muhammad Mashuri, M.T., dan Bapak Jerry Dwi Trijoyo Purnomo, M.Si, Ph.D, selaku dosen penguji yang telah memberikan saran dan koreksi atas penulisan tesis ini.
4. Bapak dan Ibu dosen Statistika ITS yang telah mencurahkan ilmu dan pengalamannya selama proses studi.
5. Bapak Inspektur II, Inspektorat Jenderal Kementerian Riset, Teknologi, dan Pendidikan Tinggi yang telah memberikan banyak bantuan kepada penulis selama melaksanakan Seminar dan Ujian Tesis.
6. Teman-teman S2-Statistika angkatan 2016 atas bantuan dan dukungannya kepada penulis.
7. Seluruh rekan kerja di Inspektorat II, Inspektorat Jenderal Kementerian Riset, Teknologi, dan Pendidikan Tinggi yang telah memberikan dukungan selama menyelesaikan tesis ini.
8. Semua pihak yang telah membantu penyelesaian tesis ini.

Semoga tesis ini dapat bermanfaat dan menambah wawasan keilmuan. Saran dan kritik yang membangun dari semua pihak, sangat penulis harapkan untuk perbaikan selanjutnya.

Surabaya, Januari 2020

Penulis

## DAFTAR ISI

	Halaman
<b>HALAMAN JUDUL</b> .....	i
<b>LEMBAR PENGESAHAN</b> .....	iii
<b>ABSTRAK</b> .....	v
<b>DAFTAR ISI</b> .....	vii
<b>DAFTAR TABEL</b> .....	ix
<b>DAFTAR GAMBAR</b> .....	xi
<b>BAB I PENDAHULUAN</b> .....	1
1.1 Latar Belakang.....	1
1.2 Perumusan Masalah.....	6
1.3 Tujuan Penelitian.....	6
1.4 Manfaat Penelitian.....	7
1.5 Batasan Masalah Penelitian.....	7
<b>BAB II TINJAUAN PUSTAKA</b> .....	9
2.1 <i>Spatial Point Processes</i> pada $R^d$ .....	9
2.2 <i>Proses Stokastik</i> .....	11
2.3 <i>Poisson Point Processes</i> .....	12
2.4 Uji <i>Goodness of fit</i> Distribusi.....	14
2.5 <i>Finite Mixture Poisson</i> .....	15
2.6 Regresi <i>Mixture Poisson</i> .....	17
2.7 Transformasi Data .....	19
2.7 Metode Bayesian.....	19
2.8.1 Distribusi Prior .....	21
2.8.2 Markov Chain Monte Carlo.....	21
2.8.3 Gibbs Sampling .....	23
2.9 Deviance Information Criterion (DIC) .....	25

2.10 Fasilitas Kesehatan Tingkat Pertama (FKTP) .....	25
<b>BAB III METODE PENELITIAN</b> .....	29
3.1 Sumber Data .....	29
3.2 Variabel Penelitian .....	29
3.3 Struktur Data Penelitian.....	31
3.4 Tahapan Penelitian .....	32
<b>BAB IV ANALISIS DAN PEMBAHASAN</b> .....	37
4.1 Penelitian Sebelumnya .....	37
4.1.1 Pendekatan Bayesian pada Distribusi <i>Mixture</i> Poisson..	37
4.1.2 Pendekatan Bayesian pada Regresi <i>Mixture</i> Poisson .....	53
4.2 Eksplorasi Data FKTP dan Variabel Kovariat .....	58
4.3 Model <i>Spatial Poisson Point Process</i> .....	69
4.3.1 Identifikasi <i>Mixture</i> .....	69
4.3.2 Pemodelan Intensitas FKTP dengan Regresi Poisson dan Regresi <i>Mixture</i> Poisson.....	72
<b>BAB V KESIMPULAN DAN SARAN</b> .....	81
5.1 Kesimpulan.....	81
5.2 Saran .....	82
<b>DAFTAR PUSTAKA</b> .....	83
<b>LAMPIRAN</b>	

## DAFTAR TABEL

	Halaman
Tabel 3.1 Data Lokasi Puskesmas.....	30
Tabel 3.2 Data Lokasi Klinik Pratama.....	30
Tabel 3.3 Variabel Kovariat.....	30
Tabel 3.4 Variabel Penelitian.....	31
Tabel 3.5 Struktur Data Penelitian.....	32
Tabel 4.1 Data <i>point</i> dan <i>mixture</i> .....	44
Tabel 4.2 Estimasi Parameter Regresi Poisson.....	73
Tabel 4.3 Pengujian Distribusi Masing-masing Komponen.....	76
Tabel 4.4 Estimasi Parameter Regresi <i>Mixture</i> Poisson.....	79
Tabel 4.5 Perbandingan DIC Model.....	80

*(halaman ini sengaja dikosongkan)*

## DAFTAR GAMBAR

	Halaman
Gambar 2.1 (a) Objek <i>Tessellations</i> Kota Surabaya dan (b) Voronoi <i>Tessellations</i> .....	9
Gambar 3.1 DAG Regresi Mixture Poisson .....	34
Gambar 3.2 Diagram Alir Penelitian .....	36
Gambar 4.1 Obyek <i>Tessellation</i> Kota Surabaya .....	61
Gambar 4.2 <i>Planar Point Pattern</i> lokasi FKTP.....	61
Gambar 4.3 Intensitas dan <i>Contour</i> dari lokasi FKTP .....	62
Gambar 4.4 Jumlah FKTP masing-masing kecamatan di Surabaya.....	62
Gambar 4.5 Variabel Kovariat Populasi di setiap kecamatan.....	63
Gambar 4.6 Variabel Kovariat Persentase jumlah rumah tangga bersih	65
Gambar 4.7 Variabel Kovariat Angka Keterkaitan di setiap kecamatan	65
Gambar 4.8 Variabel Kovariat Panjang jalan dalam kondisi baik (km).	66
Gambar 4.9 Tranformasi Pixel Variabel Kovariat Populasi di setiap kecamatan .....	66
Gambar 4.10 Tranformasi <i>Pixel</i> Variabel Kovariat Persentase jumlah rumah tangga bersih .....	67
Gambar 4.11 Tranformasi Pixel Variabel Kovariat Angka Keterkaitan di setiap kecamatan .....	67
Gambar 4.12 Tranformasi Pixel Variabel Panjang jalan dalam kondisi baik dalam km.....	69
Gambar 4.13 Histogram FKTP. (a) Histogram dari jumlah FKTP dan (b) Histogram dari 2-komponen dari jumlah FKTP .....	70
Gambar 4.14 Pengujian Distribusi FKTP .....	71
Gambar 4.15 DAG Regresi Poisson .....	73
Gambar 4.16 Regresi Poisson, (a) <i>Density plot</i> , (b) <i>Autocorrelation plot</i> , dan (c) <i>History plot</i> .....	74
Gambar 4.17 Regresi <i>Mixture</i> Poisson DAG.....	77
Gambar 4.18 Regresi <i>Mixture</i> Poisson, (a) <i>Density plot</i> , (b) <i>Autocorrelation</i> <i>plot</i> , dan (c) <i>History plot</i> .....	78

*(halaman ini sengaja dikosongkan)*

# BAB 1

## PENDAHULUAN

### 1.1 Latar Belakang

Data *Spatial point pattern* (SPPt) merupakan kumpulan *dataset* yang diamati berdasarkan titik lokasi spasial yang bersifat random. Contoh paling umum dari *dataset* ini adalah hasil pengamatan titik-titik pohon di hutan, lokasi klaster bintang di angkasa, titik kecelakaan lalu lintas, titik gempa bumi, titik pergerakan binatang buas, atau titik kejadian penyakit. Fokus utama dalam penelitian data SPPt adalah pola persebaran dari titik-titik lokasi tersebut (Baddeley dkk, 2015).

Dalam perkembangannya banyak penelitian tentang data SPPt di bidang kehutanan (Waagepetersen dan Guan, 2009), kependudukan (Chen dan Ge, 2015), kriminologi (Shirota dkk, 2017), kecelakaan lalu lintas (Takbir dkk, 2016), epidemiologi (Quesada dkk, 2017) dan *property sales* (Paci dkk, 2017). Berdasarkan beberapa penelitian yang telah dilakukan, hal pertama yang dilakukan dalam analisis SPPt adalah menyelidiki tentang intensitas titik kejadian. Intensitas adalah ekpektasi banyak titik per unit area. Analisis SPPt dapat berupa estimasi intensitas untuk menentukan apakah intensitas bersifat konstan atau bervariasi secara spasial. Jika terdapat variabel kovariat maka analisis dilakukan untuk mengetahui apakah distribusi intensitas tergantung pada variabel kovariat sebagai contoh apakah kepadatan pohon dalam suatu area bergantung pada keadaan tanahnya (Baddeley dkk, 2015).

Analisis terhadap SPPt termasuk dalam pemodelan stokastik karena berupa pengukuran kejadian pada ruang dua dimensi. Model Stokastik yang menghasilkan SPPt disebut sebagai *spatial point processes* (SPP). Kejadian-kejadian yang diamati dalam sebuah area dihitung berdasarkan pembagian *subarea*, sehingga data yang terkumpul mengikuti proses Poisson. Berdasarkan intensitasnya proses Poisson dibagi menjadi dua jenis yaitu *Homogeneous Poisson Processes* (HPP) dan *Nonhomogeneous Poisson Processes* (NHPP). *Spatial Point Processes* yang tergolong dalam HPP mempunyai ekpektasi intensitas pada masing-masing area konstan. Sedangkan NHPP memiliki ekpektasi intensitas

masing-masing area yang bervariasi secara spasial. Dalam pemodelan SPPT, model yang paling sederhana menggambarkan kondisi yang *Completely Spatial Random* (CSR) atau biasa disebut sebagai HPP. Sedangkan untuk model yang lebih rumit digambarkan oleh NHPP misalnya kluster atau pola reguler (Illian dkk, 2008).

Estimasi intensitas SPP dapat dilakukan dengan cara nonparameterik dan parametrik. Pendekatan dengan nonparametrik banyak digunakan untuk analisis titik persebaran kejadian yang tidak melibatkan variabel kovariat sebagai contoh estimasi intensitas daerah kemiskinan di China yang dilakukan Chen dan Ge (2015) dengan menggunakan fungsi Ripley's K dan estimasi pola spasial desa di masing-masing provinsi menggunakan rasio Average Nearest Neighbor (ANN). Sedangkan pendekatan parametrik dilakukan jika terdapat variabel kovariat yang diduga berpengaruh terhadap persebaran titik kejadian (Baddeley dkk., 2015). Beberapa penelitian yang dilakukan dengan melibatkan variabel kovariat antara lain dilakukan oleh Paci dkk (2017) tentang analisis penjualan properti residen dengan menggunakan *space time point pattern*, Thurman dan Zhu (2014), Choiruddin dkk (2017a), dan Choiruddin dkk (2017b) yang mengembangkan model estimasi intensitas SPP dengan seleksi variabel menggunakan metode regularisasi, seleksi variabel menggunakan metode regularisasi convex dan non-convex dan dengan jumlah pengamatan pada variabel kovariat yang berbeda.

Pada penelitian ini akan dilakukan estimasi fungsi intensitas untuk proses spasial NHPP. Pada sebuah proses spasial NHPP banyaknya titik lokasi SPP yang diamati pada ruang dua dimensi bervariasi secara spasial. Intensitas NHPP yang bervariasi secara spasial bisa dipengaruhi variabel kovariat yang diukur dengan koefisien parameter regresi. Sehingga pada perkembangannya estimasi fungsi intensitas NHPP didekati dengan regresi *mixture* Poisson. *Mixture* model muncul ketika pengukuran variabel random dilakukan berdasarkan dua atau lebih kondisi yang berbeda atau unit yang disampling terdiri dari subpopulasi (Carlin dan Louis, 2008). Kulvanich (2013) pada penelitiannya membahas kasus *mixture Poisson Point Processes* (PPP). *Mixture* PPP timbul dari sebuah model proses Poisson yang menggambarkan variasi antar titik-titik kejadian akibat variabel kovariat

yang bervariasi sehingga populasi dapat membentuk beberapa kluster yang akhirnya membawa pada kasus model *mixture*.

Secara umum pemodelan regresi *mixture* dapat dilakukan dengan dua cara yaitu secara frekuentis dengan menggunakan algoritma *Expectation-Maximization* (EM) dan pendekatan Bayesian. Algoritma EM menduga parameter suatu fungsi dengan menggunakan Maximum Likelihood Estimation (MLE), dimana fungsi mengandung komponen yang *missing* atau tidak diketahui. Dalam regresi *mixture*, parameter peluang yang menentukan pengamatan masuk dalam komponen tertentu dianggap sebagai komponen yang *missing*. Pendekatan dilakukan dengan langkah *Expectation* yaitu mencari nilai ekspektasi untuk fungsi likelihood berdasarkan variabel yang diamati kemudian *Maximization* yaitu mencari MLE dari parameter-parameter dengan memaksimalkan ekspektasi likelihood lokal yang dihasilkan dari langkah *Expectation* (Hogg dkk, 2005). Karena tujuan algoritma adalah mencari nilai maksimum dari fungsi likelihood ada kalanya iterasi membutuhkan waktu yang lama untuk mencapai konvergen dan nilai optimum yang diperoleh cenderung bersifat optimum lokal. Selain itu estimasi parameter yang diperoleh berupa estimasi titik. Pendekatan lain untuk regresi *mixture* adalah Bayesian dengan algoritma *Markov Chain Monte Carlo* (MCMC). Model Bayesian banyak diterapkan untuk SPP karena bertujuan membangun fungsi intensitas untuk prediksi distribusi posterior dari *point pattern*. Dengan menggunakan salah satu metode MCMC yaitu algoritma *Gibbs Sampling* dilakukan estimasi parameter regresi berdasarkan sampel yang dibangkitkan dari distribusi posterior. Keuntungan menggunakan pendekatan Bayesian dalam regresi *mixture* yaitu memperoleh estimasi interval parameter sehingga dapat diuji signifikansi masing-masing parameter.

Beberapa penelitian yang menggunakan pendekatan Bayesian pada kasus SPP dan *mixture* SPP antara lain Leininger dan Gelfand (2017) yang mengembangkan teori analisis Bayesian SPP dan seleksi model untuk proses Poisson, proses Cox, proses Gibbs, dan proses kluster. Pendekatan model Bayesian nonparametrik untuk kasus model *mixture* distribusi dilakukan berdasarkan proses Dirichlet (Kottas dan Sanso, 2007). Walsh dan Raftery (2005) melakukan klasifikasi untuk *mixture* SPP menggunakan *partial bayes factors*.

Penelitian melakukan pemeriksaan apakah sebuah *point pattern* mengikuti proses Poisson atau *mixture* proses Poisson. Kusumaningrum dkk (2017) melakukan analisis SPP terhadap pola persebaran puskesmas di Surabaya. Pembagian Kota Surabaya berdasarkan grid berukuran 8x4 dengan bentuk persegi sehingga terdapat 32 wilayah dengan variabel kovariat yang digunakan berasal dari karakteristik masing-masing puskesmas. Karakteristik area kecamatan seperti penduduk dan akses belum dapat dimasukkan dalam model karena keterbatasan informasi dari bentuk grid persegi. Pemodelan dilakukan berdasarkan *mixture* model data banyak puskesmas masing-masing *grid*. Pada penelitian ini tidak ada variabel kovariat yang berpengaruh terhadap intensitas Puskesmas. Berdasarkan perkembangannya pendekatan Bayesian banyak digunakan pada *point process* yang memiliki intensitas tidak homogen.

Analisis tentang intensitas NHPP dilakukan terhadap titik-titik lokasi persebaran dari fasilitas kesehatan tingkat pertama (FKTP) di kota Surabaya. FKTP yang dimaksud berdasarkan peraturan menteri kesehatan nomor 71 tahun 2013 merupakan fasilitas kesehatan yang memberikan pelayanan kesehatan tingkat pertama. Masyarakat yang memiliki kartu jaminan Badan Penyelenggara Jaminan Kesehatan (BPJS) harus mendapatkan surat rujukan dari FKTP apabila membutuhkan pelayanan kesehatan tingkat lanjutan (rumah sakit), sehingga keberadaan FKTP menjadi sangat penting guna memberikan pelayanan kesehatan yang mudah dijangkau. Kota Surabaya merupakan kota terbesar kedua di Indonesia. Dengan pertumbuhan penduduk yang cepat diperlukan fasilitas kesehatan yang memadai di seluruh wilayah kota Surabaya. Namun karena perbedaan karakteristik masing-masing kecamatan berakibat pada variasi spasial dari intensitas fasilitas kesehatan di Kota Surabaya. Ada dua dimensi akses perlu diperhatikan dalam penyelenggaraan fasilitas kesehatan yaitu *availability* atau jumlah fasilitas kesehatan dalam suatu wilayah dan *accessibility* yang mengarah pada jarak tempuh pasien ke lokasi fasilitas kesehatan (Nobles dkk, 2014).

Salah satu jenis FKTP yaitu puskesmas (Pusat Kesehatan Masyarakat), harus didirikan minimal satu per kecamatan. Permasalahan pengadaan puskesmas di Kota Surabaya yaitu pemerataan puskesmas yang dapat dilihat dengan ratio jumlah puskesmas terhadap 30.000 penduduk (Kementerian Kesehatan RI, 2015).

Kota Surabaya memiliki 31 kecamatan dengan jumlah penduduk mencapai 2,8 juta penduduk (BPS, 2016). Secara umum jumlah puskesmas yang terletak di Surabaya seharusnya minimal sejumlah 95 puskesmas. Namun saat ini jumlah puskesmas di Surabaya hanya 63 (Dinkes Kota Surabaya, 2017). Hanya 4 kecamatan dari 31 kecamatan di Surabaya yang telah memenuhi jumlah ideal puskesmas yaitu Kecamatan Wonocolo, Kecamatan Lakarsantri, Kecamatan Sambikerep, dan Kecamatan Genteng. Upaya pemerintah untuk mencukupi kebutuhan pelayanan kesehatan salah satunya mengajak pihak swasta untuk turut memberikan pelayanan kesehatan dengan cara mendirikan klinik pratama. Dengan adanya bantuan pihak swasta, kota Surabaya sudah tidak kekurangan fasilitas kesehatan secara jumlah namun pembangunan fasilitas kesehatan cenderung berada di dekat pusat kota sehingga masyarakat yang jauh dari pusat kota harus menempuh perjalanan yang lebih jauh. Berdasarkan data BPS (2016) persentase penduduk Kota Surabaya yang berobat jalan terbesar pada dokter praktek atau klinik pratama sebesar 39,83% dan nomor dua adalah puskesmas dengan 36,97%. Hal ini menunjukkan bahwa antusiasme masyarakat untuk berobat ke puskesmas dan klinik pratama sangat tinggi, sehingga pemerintah khususnya Surabaya sudah seharusnya meningkatkan kualitas serta pemerataan FKTP di Surabaya.

Penyebaran fasilitas kesehatan di daerah perkotaan merupakan inti dari peningkatan utilitas pelayanan kesehatan di kota-kota besar. Eksplorasi hubungan spasial antara lokasi pusat kesehatan dengan karakteristik lingkungan sekitar menjadi faktor penting untuk pengambil keputusan, perencanaan dan sistem pelayanan kesehatan. Keputusan tentang lokasi fasilitas kesehatan sangat penting untuk penyediaan layanan kesehatan sebagai kebutuhan dasar bagi penduduk. Namun variasi cakupan fasilitas kesehatan terjadi akibat karakteristik geografis kependudukan, kegiatan ekonomi, dan aksesibilitas yang berbeda-beda setiap wilayah. *Geographic Information System* (GIS) memiliki peran yang potensial dalam menilai distribusi layanan kesehatan khususnya efektifitas cakupan fasilitas kesehatan yang relevan dengan kepadatan penduduk. Berbagai alat analisis GIS telah digunakan secara luas untuk menganalisis pola distribusi fasilitas kesehatan yang ada dan mencari lokasi baru yang optimal (Mansour, 2016). Perkembangan pemodelan untuk alokasi fasilitas kesehatan melalui GIS memungkinkan peneliti

untuk memodelkan dimensi akses menggunakan model statistika spasial. Dengan GIS dapat dilakukan eksplorasi untuk mengetahui *availability* dari FKTP berdasarkan karakteristik lingkungan sekitar seperti kondisi geografi penduduk, status perekonomian penduduk, dan aksesibilitas (Kwangsoo dan Jun Moon, 2014), sehingga dapat digunakan sebagai informasi dalam menentukan jenis layanan kesehatan yang dibutuhkan masyarakat sekitar dan sebagai pertimbangan untuk menentukan lokasi FKTP yang baru.

Pada penelitian ini akan dilakukan analisis persebaran FKTP dengan batas area atau grid merupakan *tessellation* dari batas-batas kecamatan sehingga jumlah area 31 sesuai banyak kecamatan. Dengan demikian variabel kovariat karakteristik masing-masing kecamatan dapat dimasukkan dalam analisis. Analisis dilakukan terhadap intensitas FKTP pada masing-masing kecamatan berdasarkan variabel kovariat faktor geografi dan kependudukan seperti luas wilayah, kepadatan penduduk, dan kondisi sosial ekonomi masyarakat, faktor aksesibilitas. Faktor-faktor tersebut diduga mempengaruhi perilaku pemanfaatan fasilitas kesehatan yang secara langsung berpengaruh pada intensitas FKTP.

## **1.2 Rumusan Masalah**

Titik lokasi FKTP dapat dipandang sebagai SPpT karena pembangunan puskesmas berdasarkan kebutuhan masyarakat sehingga dapat diasumsikan sebagai pengamatan yang saling independen. Pendekatan proses Poisson sesuai untuk menggambarkan SPP yang bersifat independen. Parameter intensitas dari proses Poisson yang menggambarkan pemerataan FKTP menjadi fokus penelitian. Kondisi FKTP yang tidak merata menyebabkan intensitas proses Poisson bervariasi antar objek kecamatan yang kemudian disebut proses NHPP. Dalam pendirian FKTP faktor-faktor yang berhubungan dengan lokasi, tingkat kebutuhan dan sarana prasarana menjadi pertimbangan penting bagi Dinas Kesehatan dan pihak swasta, sehingga pemodelan intensitas yang nonhomogen memperhatikan adanya variabel kovariat yang dapat dimodelkan dengan model statistika. Berdasarkan uraian tersebut dapat dirumuskan dua permasalahan sebagai berikut:

1. Bagaimana pola persebaran FKTP dan karakteristik masing-masing kecamatan di Kota Surabaya secara spasial?

2. Bagaimana melakukan estimasi fungsi intensitas proses NHPP menggunakan pemodelan regresi *mixture* Poisson dengan pendekatan Bayesian?

### **1.3 Tujuan Penelitian**

Tujuan yang ingin diperoleh dari penelitian ini adalah sebagai berikut.

1. Mendapatkan gambaran secara spasial persebaran FKTP dan Karakteristik masing-masing kecamatan di Kota Surabaya.
2. Mendapatkan estimasi fungsi intensitas lokasi persebaran FKTP menggunakan model regresi *mixture* Poisson dengan pendekatan Bayesian.

### **1.4 Manfaat Penelitian**

Manfaat yang diperoleh dari penelitian ini adalah sebagai berikut.

1. Bagi bidang keilmuan, penelitian ini diharapkan mampu menambah wawasan aplikasi Bayesian terutama model *mixture* untuk kasus SPP.
2. Bagi dinas kesehatan Surabaya dan pihak swasta, hasil penelitian ini diharapkan mampu memberikan masukan tentang lokasi untuk mendirikan FKTP yang baru dan menentukan jenis layanan kesehatan yang sesuai kebutuhan masyarakat sekitar lokasi FKTP.

### **1.5 Batasan Masalah**

Batasan masalah dalam penelitian ini adalah sebagai berikut.

1. FKTP yang dijadikan objek penelitian adalah puskesmas induk dan klinik pratama mitra BPJS yang terletak pada masing-masing kecamatan di Kota Surabaya.
2. Data yang digunakan merupakan data sekunder berupa titik-titik lokasi puskesmas dan titik-titik lokasi klinik pratama yang tercatat di Dinas Kesehatan Surabaya berdasarkan koordinat latitude dan longitude.

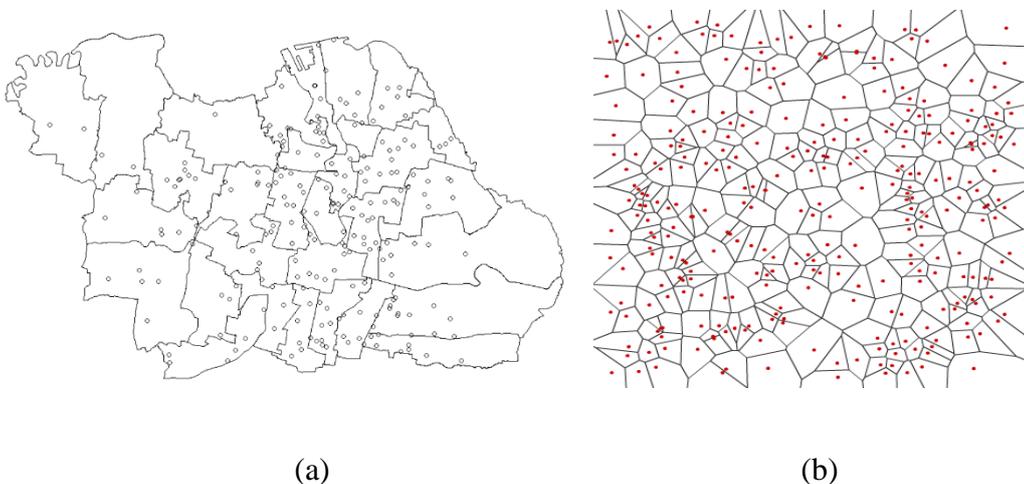
*(halaman ini sengaja dikosongkan)*

## BAB II

### TINJAUAN PUSTAKA

#### 2.1 *Spatial Point Processes* pada $R^d$

*Spatial point processes* digunakan untuk memodelkan *point pattern* dimana *point* merupakan sebuah titik posisi atau pusat objek pada ruang dua atau tiga dimensi. Misalkan dibuat sebuah peta lokasi dari orang yang melakukan panggilan darurat pada suatu hari, maka peta menggambarkan sebuah *random point pattern* dua dimensi, sehingga akan terdapat banyak *point* dan lokasi *point* yang bersifat random. Apabila waktu saat dilakukan panggilan juga dicatat maka akan ada 3 dimensi (ruang  $\times$  waktu). *Point* mungkin dilengkapi dengan *marks* berupa pengukuran atau tipe objek yang kemudian disebut sebagai *marked point processes*. *Spatial point processes*  $Y$  adalah sebuah variabel random diskrit (*countable*) dari sebuah ruang  $B$  dimana diasumsikan bahwa  $B \subseteq R^d$  (Møller dan Waagepetersen, 2004). *Spatial point processes* sangat berguna sebagai model statistik dalam melakukan analisis terhadap *point pattern*, dimana *point* menunjukkan lokasi objek studi (lokasi pohon di hutan, lokasi sarang burung, kasus epidemiologi, atau tindak kejahatan). *Point processes* memiliki peranan yang penting dalam stokastik geometri, sebagai bangunan blok dari model random yang rumit (Model Boolean) dan sebagai contoh sederhana dari *random sets* (Baddeley dkk, 2007).



**Gambar 2.1** (a) Objek *Tessellation* Kota Surabaya dan (b) *Voronoi Tessellation*

Dalam praktiknya, kejadian-kejadian ini dapat diamati sebagai titik yang terbatas pada ruang *window*  $W \subseteq B$  berbentuk persegi panjang atau bentuk tidak beraturan (*irregular*). Objek *window* dengan bentuk *irregular* dapat berupa *tessellation* yaitu pembagian *window* menjadi area tidak saling tumpang tindih. Area selanjutnya disebut sebagai *tiles*. *Tessellation* bisa berupa *list* poligon yang menggambarkan pembagian area administratif sebuah negara menjadi negara bagian/provinsi/pembagian wilayah administrasi yang lebih kecil atau bisa juga dibentuk berdasarkan data titik terluar yang kemudian disebut sebagai Dirichlet atau Voronoi *tessellation* (Baddeley dkk, 2015). Gambar 2.1 (a) menggambarkan kota Surabaya yang diamati sebagai gabungan objek *tessellations* dengan batas kecamatan. Sedangkan Gambar 2.1 (b) menggambarkan pembagian *window* yang dibentuk berdasarkan fungsi Dirichlet atau Voronoi *tessellation*.

Salah satu bentuk analisis *Spatial Point Process* adalah *quadrat counting* yaitu proses menghitung banyak titik yang jatuh pada ruang  $B$  sehingga prosesnya sering disebut sebagai proses Poisson. Pada *quadrat counting*, pengamatan pada *window*  $W$  dibagi menjadi subarea  $B_1, B_2, \dots, B_n$  yang disebut sebagai *quadrat*. Kemudian dilakukan pencacahan banyak titik yang jatuh pada masing-masing *quadrat*,  $n_i = n(\mathbf{x} \cap B_i)$  untuk  $i = 1, 2, \dots, n$ . Karena data *count* merupakan estimator tak bias untuk nilai ekspektasi  $E[n(\mathbf{X} \cap B_i)]$ , maka akan sama dengan nilai rata-rata jika intensitas *homogeneous Poisson process*. Penjelasan lebih lengkap mengenai *homogeneous Poisson process* pada subbab 2.3.

Apabila *dataset point pattern* disertai dengan data kovariat maka analisis dilakukan untuk meneliti apakah intensitas dipengaruhi oleh variabel kovariat dan untuk menghitung besar pengaruhnya. Untuk spasial kovariat numerik  $X(u)$ , dapat diasumsikan secara sederhana hubungan intensitas dengan variabel kovariat sebagai berikut.

$$\lambda(u) = \rho(X(u)) \quad (2.1)$$

dimana  $\rho$  adalah fungsi yang harus diestimasi baik secara parametrik maupun non-parametrik. Proses Poisson dijelaskan oleh fungsi intensitas  $\lambda(u)$ , model statistika untuk *Poisson Point Process* dibentuk untuk memodelkan intensitas yang berarti terdapat hubungan yang erat antara teknik untuk estimasi fungsi

intensitas dengan metode statistika untuk membentuk model *Poisson Point Process* yang sesuai.

Beberapa teknik untuk membentuk model *Poisson Point Process* dengan bentuk intensitas yang spesifik yaitu fungsi loglinier variabel kovariat seperti regresi logistik dimana wilayah studi dibagi menjadi subarea kemudian diamati ada tidaknya titik pada subarea tersebut untuk dijadikan variabel respon, selain itu terdapat regresi Poisson dengan fungsi loglinier intensitas sebagai berikut.

$$\lambda(u) = \exp(\beta_0 + \beta_1 X_1(u) + \beta_2 X_2(u) + \dots + \beta_p X_p(u)) \quad (2.2)$$

dimana  $p$  adalah banyak variabel kovariat dan  $X_p(u)$  adalah variabel kovariat  $p$  pada lokasi  $u$ .

## 2.2 Proses Stokastik

Proses Stokastik adalah himpunan variabel random  $\{Y_i\}_{i \in I}$ , mendapatkan nilai pada sebuah *state space*  $\mathcal{Y}$ , dengan indeks  $I \subset R$ . Apabila himpunan  $I$  dalam  $R$  dikembangkan menjadi ruang yang lebih tinggi dimensinya menjadi  $R^d$ , maka dapat diartikan bahwa proses stokastik melibatkan *state space* berupa ruang bukan waktu. Misalkan  $I \subset R^d$  dan  $\{Y_i\}_{i \in I}$  adalah himpunan random vektor berdimensi  $d$  dengan indeks  $I$  maka  $\{Y_i\}_{i \in I}$  adalah sebuah *random field*  $(N, d)$ . Proses Poisson merupakan sebuah proses stokastik yang sering digunakan untuk menggambarkan variabel random diskrit (*countable*) yang diamati pada ruang  $S$ . Proses Poisson didefinisikan sebagai berikut.

Misalkan  $\mathcal{B}_0(R^d)$  kumpulan himpunan Borel yang dibatasi pada ruang  $R^d$  dan  $\lambda : \mathcal{B}(R^d) \rightarrow [0, \infty]$  adalah ukuran lokal yang terhingga ( $\lambda(B) < \infty$ ) untuk semua  $B \in \mathcal{B}_0(R^d)$ . Maka dapat dikatakan  $\{N_B\}_{B \in \mathcal{B}(R^d)}$  merupakan sebuah ukuran random Poisson dengan intensitas  $\lambda$  jika:

1.  $N_{B_1}, N_{B_2}, \dots$  variabel random independen untuk pasangan *state space* yang saling bebas  $B_1, B_2, \dots \in \mathcal{B}_0(R^d)$ .
2.  $N_B \sim \text{Poisson}(\lambda(B))$  untuk semua  $B \in \mathcal{B}_0(R^d)$ .

Jika  $\lambda$  proporsional terhadap  $v_d$  (sebagai contoh  $\lambda(B) = \lambda v_d(B)$  untuk semua  $B \in \mathcal{B}(R^d)$ ), maka  $\{N_B\}_{B \in \mathcal{B}(R^d)}$  dikatakan sebagai Proses Poisson homogen dengan intensitas  $\lambda$  (Brereton, 2014).

### 2.3 Poisson Point Processes

Proses Poisson dapat digeneralisasi pada ruang dua dimensi, apabila *state* berupa *state space* disebut sebagai *Spatial Poisson point process*. *Spatial Poisson point process*, dengan intensitas  $\lambda > 0$ , adalah *point process* pada  $R^2$  sehingga :

1. Untuk setiap himpunan  $B$  yang dibatasi,  $N(B)$  mengikuti distribusi Poisson dengan mean  $\lambda = \lambda(B)$ ;
2. Jika  $B_1, B_2, \dots, B_n$  sebuah area saling bebas, sehingga  $N_{B_1}, N_{B_2}, \dots, N_{B_n}$  saling independen.

Dapat terlihat bahwa dua sifat tersebut merupakan karakteristik dari Proses Poisson. Konstanta  $\lambda$  adalah ekspektasi banyak *point* (titik lokasi pengamatan) per unit area (Baddeley dkk, 2007). Berdasarkan intensitasnya proses Poisson dibagi menjadi dua jenis yaitu *Homogeneous Poisson Processes* (HPP) dan *Nonhomogeneous Poisson Processes* (NHPP). Beddeley dkk (2015) menguraikan karakteristik dari HPP atau CSR dengan intensitas  $\lambda$  merupakan proses Poisson terhingga dengan sifat-sifat:

1. Distribusi Poisson : jumlah  $n(X \cap B)$  dari random *point* yang terletak di area  $B$  mengikuti distribusi Poisson.
2. Homogen : nilai ekspektasi dari random *point* yang terletak di area  $B$ ,  $E[n(X \cap B)] = \lambda|B|$ .
3. Independen : jika  $B_1, B_2, \dots, B_n$  adalah wilayah yang tidak saling tumpang tindih maka  $n(X \cap B_1)$ ,  $n(X \cap B_2), n(X \cap B_n)$  merupakan variabel random yang saling independen.
4. Sifat kondisional :  $n(X \cap B) = n$  jika  $n$  *point* independen dan berdistribusi secara seragam pada area  $B$ .

Proses Poisson dengan intensitas nonhomogen memiliki ekpektasi intensitas masing-masing area yang bervariasi secara spasial. *Nonhomogeneous Poisson Processes* dengan intensitas  $\lambda(u)$  merupakan proses Poisson terhingga dengan sifat-sifat (Beddeley dkk, 2015):

1. Distribusi Poisson : jumlah  $n(X \cap B)$  dari random *point* yang terletak di area  $B$  mengikuti distribusi Poisson.
2. Fungsi Intensitas : nilai ekspektasi dari random *point* yang terletak di area  $B$  dinyatakan sebagai  $E[n(X \cap B)] = \int_B \lambda(u) du$  dimana  $\lambda(u)$  fungsi intensitas di seluruh area  $B$ .
3. Independen : jika  $B_1, B_2, \dots, B_n$  adalah wilayah yang tidak saling tumpang tindih,  $n(X \cap B_1), n(X \cap B_2), \dots, n(X \cap B_n)$  dapat dinyatakan sebagai variabel random yang saling independen.
4. Identik :  $n(X \cap B) = n$  jika  $n$  *point* adalah independen dan berdistribusi secara identik, dengan Pdf sebagai berikut.

$$f(u) = \frac{\lambda(u)}{\lambda} \quad (2.3)$$

$$\lambda = \int_B \lambda(u) du \quad (2.4)$$

Sifat *superposition* menyatakan bahwa jika  $X$  dan  $Y$  adalah *point process* independen NHPP dengan fungsi intensitas  $\lambda_X(u)$  dan  $\lambda_Y(u)$ , masing-masing superposisi  $X \cup Y$  adalah proses Poisson dengan intensitas  $\lambda_X(u) + \lambda_Y(u)$ .

*Quadrat counting test* digunakan untuk menguji apakah *point process* memiliki intensitas homogen dengan hipotesis sebagai berikut.

$H_0$ : Intensitas *point process* homogen

$H_1$ : Intensitas *point process* nonhomogen

Langkah pertama dalam melakukan pengujian yaitu membagi  $W$  menjadi subarea  $B_1, B_2, \dots, B_n$  dan menghitung banyak titik  $n_1, n_2, \dots, n_n$  pada masing-masing

subarea. Misalkan total titik  $n = \sum_i n_i$  dan total area *window* adalah  $a = \sum_i a_i$  dimana  $a_i$  adalah luas area  $B_i$ , estimasi intensitas adalah  $\bar{\lambda} = n/a$ , dan ekspektasi banyak titik pada subarea  $B_i$  adalah  $e_i = \bar{\lambda}a_i = na_i/a$ . Statistik uji adalah sebagai berikut.

$$\chi^2 = \sum_i \frac{(\text{pengamatan} - \text{ekspektasi})^2}{\text{ekspektasi}}$$

$$\chi^2 = \sum_i \frac{(n_i - e_i)^2}{e_i} = \sum_i \frac{(n_i - \bar{\lambda}a_i)^2}{\bar{\lambda}a_i} \quad (2.5)$$

Statistik uji berdistribusi Chi-square dengan derajat bebas  $(n - 1)$ , sehingga  $H_0$  ditolak jika  $\chi^2 > \chi_{1-\alpha, n-1}^2$  (Beddeley dkk, 2016).

## 2.4 Uji Goodness of fit Distribusi

Chi-Square *goodness of fit test* (Snedecor dan Cochran, 1989) digunakan untuk menguji apakah data dari sebuah populasi mengikuti pola distribusi tertentu. Chi-Square *goodness of fit test* dapat diaplikasikan ke semua distribusi univariat diskrit dengan menghitung fungsi distribusi kumulatif. Chi-Square *goodness of fit test* adalah alternatif dari Anderson Darling dan Kolmogorov Smirnov *goodness of fit test*. Perbedaannya Chi-Square *goodness of fit test* dapat digunakan untuk menguji data dengan distribusi diskrit sedangkan Anderson Darling dan Kolmogorov Smirnov terbatas pada distribusi kontinyu. Hipotesis pengujian dapat dinyatakan sebagai berikut.

$H_0$  : Data mengikuti pola fungsi distribusi tertentu.

$H_1$  : Data tidak mengikuti pola fungsi distribusi tertentu.

Penolakan hipotesis tersebut dapat dilakukan dengan menggunakan statistik uji sebagai berikut.

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i} \quad (2.6)$$

Dimana  $O_i$  adalah frekuensi data yang diamati,  $E_i$  adalah nilai ekspektasi dari frekuensi data yang diamati berdasarkan distribusi data yang diuji. Statistik uji berdistribusi Chi-Square dengan derajat bebas  $(n - c)$ , dimana  $n$  adalah banyak sel yang tidak kosong dan  $c$  adalah banyak parameter yang diestimasi, sehingga  $H_0$  ditolak jika  $\chi^2 > \chi_{1-\alpha, n-c}^2$ .

## 2.5 Finite Mixture Poisson

Distribusi Poisson adalah salah satu jenis distribusi dari banyaknya kejadian pada interval waktu tertentu atau suatu wilayah tertentu. Kejadian tersebut tergantung pada selang waktu tertentu atau suatu wilayah tertentu, dimana hasil pengamatan berupa data diskrit dan antar kejadian saling independen. Jika variabel random diskrit  $y$  merupakan distribusi Poisson dengan parameter  $\lambda$  maka fungsi peluang dari distribusi Poisson dinyatakan sebagai berikut (McCullagh dan Nelder, 1983).

$$f(y, \lambda) = \frac{e^{-\lambda} \lambda^y}{y!}, y = 0, 1, 2, \dots, n \quad (2.7)$$

Karakteristik distribusi Poisson yaitu bahwa mean dan varians data sama ( $\lambda$ ), atau dapat dituliskan sebagai berikut.

$$E(Y) = Var(Y) = \lambda \quad (2.8)$$

Pada kasus dimana  $\lambda$  berasal dari *nonhomogeneous Poisson process* dapat diperkenalkan sebuah distribusi *k-finite mixture Poisson* dengan fungsi peluang sebagai berikut (Mufudza dan Erol, 2016).

$$f_{mix}(y|\boldsymbol{\lambda}, \boldsymbol{p}) = \sum_{k=1}^K p_k f_k(y|\lambda_k)$$

$$f_{mix}(y|\boldsymbol{\lambda}, \boldsymbol{p}) = \sum_{k=1}^K p_k \frac{e^{-\lambda_k} \lambda_k^y}{y!}$$

$$E_{mix}(y|\boldsymbol{\lambda}, \boldsymbol{p}) = \sum_{k=1}^K p_k \lambda_k$$

$$Var_{mix}(y|\boldsymbol{\lambda}, \boldsymbol{p}) = E_{mix}(y|\boldsymbol{\lambda}, \boldsymbol{p}) + (\sum_{k=1}^K p_k \lambda_k^2 - [E_{mix}(y|\boldsymbol{\lambda}, \boldsymbol{p})]^2) \quad (2.9)$$

dimana

$f_{mix}(y|\boldsymbol{\lambda}, \boldsymbol{p})$  : fungsi densitas dari *finite mixture*

$f_k(y|\lambda_k)$  : fungsi densitas ke- $k$  dari sebanyak  $K$  komponen penyusun model *mixture*

$\lambda_k$  : parameter dengan elemen  $(\lambda_1, \lambda_2, \dots, \lambda_K)$

$p_k$  : parameter proporsi komponen *mixture* dengan elemen

$(p_1, p_2, \dots, p_K), \sum_{k=1}^K p_k = 1, \text{ serta } 0 \leq p_k \leq 1, k = 1, 2, \dots, K.$

Diasumsikan bahwa  $\lambda_1 < \lambda_2 < \dots < \lambda_K$  untuk memastikan identitas *finite mixture* diatas. Pada model *mixture* diperkenalkan formulasi *missing data* dimana observasi  $y_i$  diasumsikan muncul dari komponen *mixture* yang spesifik namun tidak diketahui. Misalkan  $Y_1, Y_2, \dots, Y_n$  terdistribusi sesuai persamaan (2.9), dengan vektor parameter  $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_K)$  dan proporsi  $\boldsymbol{p} = (p_1, p_2, \dots, p_K)$ . Pada kondisi tersebut untuk masing-masing pengamatan  $y_i, i = 1, 2, \dots, n$ , diperkenalkan vektor indikator yang tidak teramati  $\boldsymbol{T} = (\boldsymbol{T}_1, \boldsymbol{T}_2, \dots, \boldsymbol{T}_n)$  yang memperkenalkan  $T_{ik} = 1$  jika  $y_i$  merupakan anggota komponen ke- $k$  atau  $T_{ik} = 0$  jika  $y_i$  merupakan bukan anggota komponen ke- $k$ . Fungsi  $f(y_i | T_{ik} = 1)$  memiliki densitas Poisson ( $\lambda_k$ ) dan  $f(T_{ik} = 1 | p) = p_k$ . Distribusi untuk vektor tidak teramati  $\boldsymbol{T}_i = (T_{i1}, T_{i2}, \dots, T_{iK})$  dengan proporsi  $\boldsymbol{p} = (p_1, p_2, \dots, p_K)$  adalah Multinomial( $1, p_1, p_2, \dots, p_K$ ) (Hamdah dan Riffi, 2015).

Misalkan pengamatan  $\boldsymbol{y} = (y_1, y_2, \dots, y_n)$  berasal dari model *mixture*, dengan vektor parameter  $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_K)$  dan proporsi  $\boldsymbol{p} = (p_1, p_2, \dots, p_K)$ , likelihood densitas *mixture* adalah sebagai berikut.

$$L_{mix}(\boldsymbol{y} | \boldsymbol{\lambda}, \boldsymbol{p}) = \prod_{i=1}^n f_{mix}(y_i | \boldsymbol{\lambda}, \boldsymbol{p}) \quad (2.10)$$

Dengan asumsi  $\lambda_1 < \lambda_2 < \dots < \lambda_K$  dan memperhatikan bahwa dalam distribusi *finite mixture*, pengamatan  $y_i$  masuk pada  $k$  komponen maka dapat dinyatakan  $y_i$  pada komponen 1 sebagai  $(y_1, y_2, \dots, y_{n_1})$ , untuk pengamatan pada komponen 2  $(y_{n_1+1}, y_{n_1+2}, \dots, y_{n_2})$ , dan untuk pengamatan pada komponen ke- $K$  adalah  $(y_{n_{K-1}+1}, y_{n_{K-1}+2}, \dots, y_n)$  maka fungsi likelihood pada persamaan (2.10) dapat ditulis sebagai berikut.

$$\begin{aligned} L_{mix}(\boldsymbol{y} | \boldsymbol{\lambda}, \boldsymbol{p}) &= \prod_{i_1=1}^{n_1} f_1(y_{i_1} | p_1, \lambda_1) \times \prod_{i_2=n_1+1}^{n_2} f_2(y_{i_2} | p_2, \lambda_2) \times \dots \\ &\times \prod_{i_K=n_{K-1}+1}^n f_K(y_{i_K} | p_K, \lambda_K) \end{aligned} \quad (2.11)$$

dimana  $n_k = (n_1, (n_2 - n_1), \dots, (n - n_{K-1}))$  adalah banyak pengamatan pada masing-masing komponen  $k$ , dengan  $n_1 + (n_2 - n_1) + \dots + (n - n_{K-1}) = n$ . Untuk menyederhanakan persamaan (2.11) diperkenalkan indikator *missing data*

$\mathbf{T}_i = (T_{i1}, T_{i2}, \dots, T_{iK})$ , dengan satu dari  $T_{ik}, k = 1, 2, \dots, K$  bernilai 1 untuk setiap  $i$  masuk dalam komponen  $k$ . Likelihood *mixture* dapat ditulis sebagai berikut.

$$L_{mix}(\mathbf{y}, \mathbf{T} | \boldsymbol{\lambda}, \mathbf{p}) = \prod_{i=1}^n \prod_{k=1}^K \left( p_k f_k(y_i | \lambda_k) \right)^{T_{ik}} \quad (2.12)$$

dimana  $f_k(y_i | \lambda_k)$  adalah distribusi Poisson dengan parameter  $\lambda_k$ .

## 2.6 Regresi *Mixture* Poisson

Regresi Poisson merupakan *Generalized linear model* dengan *link function* logaritma yang sering digunakan untuk memodelkan data *count* dimana variabel respon mengikuti distribusi Poisson (Agresti, 2002). Persamaan model regresi Poisson dapat dituliskan sebagai berikut.

$$\log(\lambda) = \boldsymbol{\beta}^T \mathbf{X}$$

$$\lambda_i = \exp(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_p X_{pi}) \quad (2.13)$$

dengan  $\lambda_i$  merupakan rata-rata banyak kejadian pada interval waktu atau area tertentu. Fungsi likelihood dari regresi Poisson adalah sebagai berikut.

$$\begin{aligned} L(\mathbf{y} | \boldsymbol{\lambda}) &= \prod_{i=1}^n \frac{e^{-\lambda_i} \lambda_i^{y_i}}{y_i!} \\ L(\mathbf{y} | \mathbf{x}, \boldsymbol{\beta}) &= \prod_{i=1}^n \frac{\exp(-\exp(\mathbf{x}_i^T \boldsymbol{\beta})) (\exp(\mathbf{x}_i^T \boldsymbol{\beta}))^{y_i}}{y_i!} \\ &= \frac{\exp(-\sum_{i=1}^n \exp(\mathbf{x}_i^T \boldsymbol{\beta})) (\exp \sum_{i=1}^n y_i \mathbf{x}_i^T \boldsymbol{\beta})}{\prod_{i=1}^n y_i!} \end{aligned} \quad (2.14)$$

Dalam pemodelan regresi dikenal model regresi *mixture* yang digunakan untuk mengakomodasi kasus heterogenitas dengan cara membuat grup observasi data yang homogen dengan cara kluster. Pada kasus regresi *mixture* terbatas dipertimbangkan *mixture* dari  $k$  komponen yang setiap komponen mengikuti distribusi tertentu. Setiap komponen diberikan bobot dan probabilitas prior untuk setiap observasi sehingga masuk pada komponen yang diharapkan. Regresi *mixture* Poisson mengasumsikan bahwa distribusi marginal  $y_i$  mengikuti distribusi *mixture* Poisson sebagaimana pada persamaan (2.9). Standar regresi *mixture* dari *Generalized Linier Model* (GLMs) dinyatakan sebagai berikut.

$$f_{mix}(\mathbf{y}|\mathbf{X}, \Phi) = \sum_{k=1}^K p_k f_k(\mathbf{y}|\mathbf{X}, \boldsymbol{\beta}_k)$$

dimana

$$f_k(\mathbf{y}|\boldsymbol{\lambda}) = \frac{e^{-\lambda_k} \lambda_k^y}{y!}$$

$$f_k(\mathbf{y}|\mathbf{X}, \boldsymbol{\beta}_k) = \frac{\exp(-\exp(\mathbf{X}^T \boldsymbol{\beta}_k)) (\exp(\mathbf{X}^T \boldsymbol{\beta}_k))^y}{y!} \quad (2.15)$$

Dengan  $\Phi = \{\boldsymbol{\beta}, \mathbf{p}\}$ ,  $\boldsymbol{\beta}_k = (\beta_{0k}, \beta_{1k}, \beta_{2k}, \dots, \beta_{pk})^T$  dan  $\beta_{0k}$  adalah konstanta model pada komponen ke- $k$  dan  $\boldsymbol{\beta}_k$  adalah vektor koefisien model pada komponen ke- $k$ , peluang pengamatan  $y_i$  merupakan anggota komponen ke- $k$  yaitu  $\mathbf{p}_k = (p_{1k}, p_{2k}, \dots, p_{pk})$ ,  $\lambda_k = \exp(\mathbf{X}^T \boldsymbol{\beta}_k)$ . Variabel  $y$  merupakan variabel respon yang mengikuti distribusi keluarga eksponensial dengan syarat pada komponen  $k$ , ekspektasi bersyarat variabel respon dinyatakan dengan

$$E_{mix}(y|\mathbf{X}, \Phi) = \sum_{k=1}^K p_k \lambda_k \quad (2.16)$$

dimana  $\Phi = \{\boldsymbol{\beta}, \mathbf{p}\}$ ,  $\boldsymbol{\beta}_k = (\beta_{0k}, \beta_{1k}, \beta_{2k}, \dots, \beta_{pk})^T$ , dan  $\lambda_k = \exp(\mathbf{X}^T \boldsymbol{\beta}_k)$ .

Sebagaimana pada *finite mixture* Poisson, pengembangan regresi *mixture* dikenal istilah *missing data*  $\mathbf{T}_i = (T_{i1}, T_{i2}, \dots, T_{iK})$ , dengan satu dari  $T_{ik}, k = 1, 2, \dots, K, i = 1, 2, \dots, n$ , bernilai 1 untuk setiap  $i$  masuk dalam komponen  $k$ . Variabel random  $T_i$  diasumsikan mengikuti distribusi Multinomial ( $T_i | \mathbf{p} \sim \text{Multinomial}(1, p_1, p_2, \dots, p_K)$ ) dengan fungsi densitas sebagai berikut.

$$f(T_{i1}, \dots, T_{iK} | p_1, \dots, p_K) = \frac{1}{T_{i1}! \dots T_{iK}!} \prod_{k=1}^K p_k^{T_{ik}} = \prod_{k=1}^K p_k^{T_{ik}} \quad (2.17)$$

Karena  $\mathbf{T}_i$  saling independen maka *joint density function*  $\mathbf{T}_i, i = 1, 2, \dots, n$  adalah sebagai berikut.

$$f(\mathbf{T}|\mathbf{p}) = \prod_{i=1}^n f(T_{ik}|\mathbf{p}) = \prod_{i=1}^n \prod_{k=1}^K p_k^{T_{ik}} \quad (2.18)$$

Misalkan pengamatan  $\mathbf{y} = (y_1, y_2, \dots, y_n)$  berasal dari model *mixture*, dengan vektor parameter  $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_K)$  dan proporsi  $\mathbf{p} = (p_1, p_2, \dots, p_K)$ , likelihood dari regresi *mixture* Poisson sesuai persamaan (2.15) adalah sebagai berikut.

$$L_{mix}(\mathbf{y}, \mathbf{T}|\mathbf{X}, \Phi) = f(\mathbf{y}, \mathbf{T}|\mathbf{X}, \Phi, \mathbf{p}) f(\mathbf{T}|\mathbf{y}, \Phi, \mathbf{p})$$

$$= f(\mathbf{y}|\mathbf{X}, \mathbf{T}, \Phi) f(\mathbf{T}|\mathbf{p})$$

$$\begin{aligned}
&= \prod_{i=1}^n \prod_{k=1}^K \left( \frac{e^{-\lambda_k} \lambda_k^{y_i}}{y_i!} \right)^{T_{ik}} \prod_{i=1}^n \prod_{k=1}^K p_k^{T_{ik}} \\
&= \prod_{i=1}^n \prod_{k=1}^K \left( p_k \frac{e^{-\lambda_k} \lambda_k^{y_i}}{y_i!} \right)^{T_{ik}} \quad (2.19)
\end{aligned}$$

pada persamaan (2.19)  $\Phi = \{\beta, p\}$ ,  $\lambda_k = \exp(\mathbf{X}^T \beta_k)$ , dimana  $T_i$  adalah indikator vektor nol dan satu dengan  $T_{ik} = (T_k)_i$  bernilai satu untuk komponen ke- $k$  dan nol untuk lainnya (Mufudza dan Erol, 2016). Model *mixture* dengan asumsi terdapat  $p$  variabel kovariat dengan  $K$  komponen maka bentuk persamaan regresi *mixture* Poisson menjadi sebagai berikut.

$$\begin{aligned}
&p_1(\exp(\beta_{01} + \beta_{11}x_{1i_1} + \dots + \beta_{p1}x_{pi_1})) + p_2(\exp(\beta_{02} + \beta_{12}x_{1i_2} + \dots + \beta_{p2}x_{pi_2})) \\
&\quad + \dots + p_K(\exp(\beta_{0K} + \beta_{1K}x_{1i_K} + \dots + \beta_{pK}x_{pi_K})) \quad (2.20)
\end{aligned}$$

## 2.7 Transformasi data

Pada penelitian ini data kovariat ditampilkan dalam bentuk *pixel image*. Metode yang digunakan pada penelitian ini adalah *kernel smoother of mark*. Metode ini digunakan untuk *spatial smoothing* dari nilai *mark* pada *point pattern*, seperti diameter pada tanaman pinus. *Smoothing* dilakukan dengan menghitung nilai rata-rata dari pepohonan di setiap lingkungan. Misalkan data adalah point  $x_1, \dots, x_n$ , dengan nilai *mark*  $m_1, \dots, m_n$ , yang mana merupakan nilai real yang sesuai dengan titik lokasi  $u$ . Fungsi *spatial* yang digunakan adalah The *Nadaraya-Watson smoother* dituliskan sebagai berikut.

$$\tilde{m}(u) = \frac{\sum_i m_i k(u - x_i)}{\sum_i k(u - x_i)} \quad (2.21)$$

Transformasi pixel image dalam package spatstat menggunakan smooth.ppp untuk *point pattern*. Hasil dari smooth.ppp akan mendekati rata-rata *mark* dalam kumpulan data (Baddeley, Rubak, dan Turner, 2016).

## 2.8 Metode Bayesian

Statistika Bayesian berbeda dengan teori statistika klasik karena semua parameter yang tidak diketahui dalam teori Bayesian dianggap sebagai variabel random. Dalam analisis diperlukan inisial distribusi prior yang berasal dari

informasi yang tersedia sebelum peneliti memasukkan data dalam analisis. Analisis dilakukan untuk mendapatkan distribusi posterior berdasarkan data pengamatan (Ntzoufras, 2009). Hubungan antara distribusi posterior dengan distribusi prior dan likelihood dapat dituliskan sebagai berikut.

$$\text{Distribusi Posterior} \propto \text{likelihood} \times \text{Distribusi Prior} \quad (2.22)$$

Apabila terdapat parameter  $\lambda$  diberikan informasi data  $y$ , maka distribusi probabilitas untuk posterior  $\lambda$  pada data  $y$  akan proporsional terhadap perkalian prior  $\lambda$  dengan fungsi likelihood yang diberikan oleh data  $y$ . Berdasarkan teorema Bayes, distribusi posterior diperoleh berdasarkan persamaan berikut.

$$f(\lambda|y) = \frac{f(y|\lambda)f(\lambda)}{f(y)} \propto f(y|\lambda)f(\lambda) \quad (2.23)$$

Distribusi posterior mengandung data dan distribusi prior yang dituliskan sebagai distribusi prior  $f(\lambda)$  dan likelihood data sebagai berikut.

$$L(y|\lambda) = \prod_{i=1}^n f(y_i|\lambda) \quad (2.24)$$

Fungsi likelihood pada model *mixture* dengan  $k$  komponen dapat dituliskan sebagai berikut sesuai persamaan (2.19).

$$L_{mix} = \prod_{i=1}^n \prod_{k=1}^K \left( p_k \frac{e^{-\lambda_k} \lambda_k^{y_i}}{y_i!} \right)^{T_{ik}} \quad (2.25)$$

dengan syarat  $n_1 + n_2 + \dots + n_K = n$ ,  $k = 1, 2, \dots, K$  adalah banyaknya komponen.

Himpunan data diskrit  $y$  akan dilakukan estimasi terhadap parameter  $\lambda$ . Apabila data mengikuti distribusi Poisson maka dituliskan sebagai

$$y_i \sim \text{Poisson}(\lambda) \text{ untuk } i = 1, 2, 3, \dots, n \quad (2.26)$$

Jika diasumsikan distribusi data  $y_i$  adalah *mixture* Poisson maka dapat dituliskan sebagai berikut.

$$\begin{aligned} y_i &\sim \sum_{k=1}^K p_k f_k(\lambda_k) \\ f_{mix}(y_i|\boldsymbol{\lambda}, \mathbf{p}) &= p_1 f_1(y_i|\lambda_1) + p_2 f_2(y_i|\lambda_2) + \dots + p_K f_K(y_i|\lambda_K) \\ \sum_{k=1}^K p_k &= 1 \end{aligned} \quad (2.27)$$

Model *mixture* hampir serupa dengan model hierarki. Letak kemiripan diperlihatkan dengan adanya variabel tidak teramati  $T_{ik}$  dimana memiliki nilai 1 jika pengamatan ke- $i$  berasal dari komponen ke- $k$  dan bernilai 0 untuk kondisi lainnya. Diketahui  $p$  distribusi untuk masing-masing vektor  $\mathbf{T}_i = (T_{i1}, T_{i2}, \dots, T_{iK})$  adalah Multinomial  $(1; p_1, \dots, p_K)$ . Pada kasus *mixture* parameter  $p$  digunakan

sebagai hiperparameter untuk menentukan distribusi dari  $T_i$ . Likelihood *Joint distribution* dari pengamatan  $y_i$  dengan indikator  $T_i$  dapat dituliskan sebagai

$$L_{mix}(\mathbf{y}, \mathbf{T} | \boldsymbol{\lambda}, \mathbf{p}) = L_{mix}(\mathbf{T} | \mathbf{p}) f(\mathbf{y} | \mathbf{T}, \boldsymbol{\lambda}) = \prod_{i=1}^n \prod_{k=1}^K (p_k f(y_i | \lambda_k))^{T_{ik}} \quad (2.28)$$

Distribusi prior untuk model *mixture* dengan parameter  $(\lambda, p)$  banyak diterapkan sebagai perkalian distribusi prior independen pada  $\lambda$  dan  $p$ , sehingga jika indikator *mixture*  $\mathbf{T}_i = (T_{i1}, T_{i2}, \dots, T_{iK})$  dimodelkan sebagai Multinomial dengan parameter  $p$ , maka distribusi prior *conjugate* adalah Dirichlet,  $p \sim \text{Dirichlet}(\alpha_1, \dots, \alpha_K)$ . Sedangkan untuk masing-masing parameter  $\lambda_k$  menggunakan distribusi prior *conjugate* Gamma,  $\lambda_k \sim \text{Gamma}(\alpha_k, \beta_k)$ . Teori Bayes membawa ke bentuk *joint* distribusi posterior sebagai

$$f_{mix}(\boldsymbol{\lambda}, \mathbf{p} | \mathbf{y}, \mathbf{T}) \propto f_{mix}(\mathbf{y}, \mathbf{T} | \boldsymbol{\lambda}) f(\boldsymbol{\lambda}) = f(\mathbf{y} | \boldsymbol{\lambda}, \mathbf{p}) f(\mathbf{T} | \mathbf{p}) f(\mathbf{p}) f(\boldsymbol{\lambda}) \quad (2.29)$$

Untuk menyelesaikan persamaan likelihood lengkap dari (2.28) memerlukan perkalian dari  $n$  data pada masing-masing elemen  $k$ . Sehingga untuk menyelesaikannya digunakan pendekatan *Markov Chain Monte Carlo* (Gelman dkk, 1995).

Diberikan bahwa  $f_k(y_{ik} | \mathbf{x}_{ik}, \boldsymbol{\beta}_k)$  sebagaimana pada persamaan (2.15),  $i = 1, 2, \dots, n$ ,  $k = 1, 2, \dots, K$ ,  $i_k = 1, 2, \dots, n_k$ ,  $n_1 + n_2 + \dots + n_K = n$ , dan  $n_k$  adalah banyak pengamatan pada subpopulasi ke- $k$ , dan  $k = 1, 2, \dots, K$  adalah banyak komponen. Untuk 2 komponen regresi *mixture* Poisson, distribusi posterior dapat dituliskan sebagaimana persamaan (2.30).

$$\begin{aligned} f(\boldsymbol{\Phi} | Y_i) &\propto f(y_i | \mathbf{p}, \boldsymbol{\beta}) \times f(\mathbf{p}) \times f(\boldsymbol{\beta}) \\ &= \prod_{i=1}^n \sum_{k=1}^2 p_k f_k(y_{ik} | \mathbf{x}_{ik}(u_{ik}), \boldsymbol{\beta}_k) \times f(\mathbf{p}) \times f(\boldsymbol{\beta}) \\ &= \left\{ \prod_{i=1}^n \left[ p_1 \left[ \exp(\beta_{01} + \beta_{11} x_{1i_1}(u_{i_1}) + \dots + \beta_{p_1} x_{pi_1}(u_{i_1})) \right] \right. \right. \\ &\quad \left. \left. + p_2 \left[ \exp(\beta_{02} + \beta_{12} x_{1i_2}(u_{i_2}) + \dots + \beta_{p_2} x_{pi_2}(u_{i_2})) \right] \right] \right\} \times f(\mathbf{p}) \times f(\boldsymbol{\beta}) \end{aligned} \quad (2.30)$$

### 2.8.1 Distribusi Prior

Penentuan prior menjadi sangat penting dari metode Bayesian karena berpengaruh pada bentuk distribusi posterior. Berikut ini macam-macam distribusi prior yang digunakan di metode Bayesian.

1. *Conjugate prior* dan *non conjugate prior*, yaitu *prior* ditentukan sesuai dengan pola *likelihood* data (Box dan Tiao, 1973).
2. *Proper prior* atau *improper prior* (*Jeffreys prior*), yaitu *prior* yang terkait dengan pemberian bobot atau densitas di setiap titik sehingga terdistribusi secara *uniform* atau tidak (Ntzoufras, 2009).
3. *Informative prior* atau *non informative prior*, yaitu *prior* yang berkaitan dengan ketersediaan pengetahuan atau informasi sebelumnya mengenai pola distribusi data (Box dan Tiao, 1973).
4. *Pseudo prior*, yaitu *prior* ditentukan berdasarkan hasil elaborasi dari metode klasik. Misalnya *prior* ditentukan berdasarkan hasil dari estimasi parameter model regresi dengan *Ordinary Least Squares/OLS* (Carlin dan Chib, 1995).

### 2.8.2 Markov Chain Monte Carlo

Markov chain diperkenalkan oleh Andrei Andreyevich Markov pada tahun 1906. Markov Chain Monte Carlo (MCMC) digunakan untuk membangkitkan data sampel yang berasal dari distribusi posterior. Teori bayes biasanya diuraikan tanpa fungsi normalisasi atau berupa bentuk proporsional terhadap prior kali *likelihood*, sehingga formula ini tidak menunjukkan fungsi densitas distribusi posterior secara utuh tetapi hanya memberikan sebuah bentuk distribusi posterior. Dengan kata lain, dapat ditentukan keberadaan modus dan nilai relatif diantara dua titik. Namun formula ini tidak memberikan faktor skala yang membentuk sebuah densitas. Hal ini berarti tidak dapat dilakukan inferensi terhadap parameter distribusi posterior. Fungsi densitas posterior diperoleh dengan melakukan *scaling* konstanta normalisasi sehingga nilai integral denominator sama dengan satu.

$$f(\lambda|y) = \frac{f(y|\lambda)f(\lambda)}{\int_{-\infty}^{\infty} f(y|\lambda)f(\lambda)d\lambda} \quad (2.31)$$

Bentuk *close form* dari integral denominator hanya terpenuhi di beberapa kasus. Pada kasus lainnya diperlukan penyelesaian secara numerik. Sebagai contoh pada kasus *finite mixture* Poisson persamaan (2.9).

$$f(\boldsymbol{\lambda}|y, \mathbf{p}) = \frac{f_{mix}(y|\boldsymbol{\lambda}, \mathbf{p})f(\boldsymbol{\lambda})f(\mathbf{p})}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{mix}(y|\boldsymbol{\lambda}, \mathbf{p})f(\boldsymbol{\lambda})f(\mathbf{p})d\boldsymbol{\lambda}d\mathbf{p}} \quad (2.32)$$

dimana

$$f_{mix}(y|\boldsymbol{\lambda}, \mathbf{p}) = \sum_{k=1}^K p_k f_k(y|\lambda_k)$$

$$f(\mathbf{p}) = \frac{\Gamma(\sum_{k=1}^K \delta_k)}{\prod_{k=1}^K \Gamma(\delta_k)} \prod_{k=1}^K p_k^{\delta_k - 1}$$

$$f(\boldsymbol{\lambda}) = \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta}$$

Hasil integral denominator sulit untuk diperoleh bentuk *close form*. Pendekatan komputasi terhadap statistika Bayesian memungkinkan posterior didekati dengan sampel Monte Carlo yang berasal dari posterior. Random sampel Monte Carlo dari posterior akan mendekati nilai asli posterior jika sampel yang digunakan cukup. Ide dari MCMC yaitu membangkitkan data sampel yang berasal dari posterior sesuai dengan proses markov chain dengan menggunakan simulasi Monte Carlo secara iteratif hingga diperoleh kondisi yang konvergen (Hamdah, 2015).

### 2.8.3 Gibbs Sampling

Algoritma Gibbs Sampling dikembangkan oleh Geman (1984). Prosedur yang dilakukan pada Gibbs Sampling untuk mixture  $k$  komponen yaitu (Hamdah, 2015).

1. Pada saat  $t = 0$ , memberi nilai state awal  $:(\lambda_1^0, \dots, \lambda_K^0, \mathbf{p}^0, \mathbf{T}_1^0, \dots, \mathbf{T}_n^0)$  dengan memperhatikan karakteristik datanya.
2. Untuk  $t = 1, 2, \dots, N$ , memperbarui nilai masing-masing variabel pada iterasi ke- $t$  :
  - a. *Update* Gibbs untuk  $\lambda_k, k = 1, \dots, K$  : membangkitkan sampel  $\lambda_k^t$  dari  $gamma(\alpha_k + \sum_{i=1}^n y_i T_{ik}, \sum_{i=1}^n T_{ik} + \beta_k)$  dengan menggunakan nilai  $T_{ik}$  paling baru.

b. *Update* Gibbs untuk proporsi  $\mathbf{p}$  : membangkitkan sampel  $\mathbf{p}^t$  dari Dirichlet  $(1 + \sum_{i=1}^n T_{i1}, \dots, 1 + \sum_{i=1}^n T_{iK})$  dengan nilai  $T_{i1}, \dots, T_{iK}$  paling baru.

c. *Update* Gibbs untuk indikator  $\mathbf{T}_i$ : membangkitkan sampel  $T_i^t$  dari Multinomial  $(1, p_{i1}, \dots, p_{iK})$   $i = 1, \dots, n, k = 1, \dots, K$ , dimana

$$p_{ik} = \frac{f(y_i|\lambda_k)p_k}{f(y_i)} \quad k = 1, \dots, K \quad (2.31)$$

menggunakan nilai  $\lambda_k$  dan  $\mathbf{p}$ .

d. Sehingga diperoleh *state* Markov Chain  $(\lambda_1^t, \dots, \lambda_K^t, \mathbf{p}^t, \mathbf{T}_1^t, \dots, \mathbf{T}_n^t)$  yang digunakan untuk membangkitkan sampel pada iterasi  $t + 1$ .

3. Kembali ke langkah ke-2 sampai sebanyak  $N - 1$  kali iterasi untuk memperoleh Markov Chain dengan ukuran  $N$ .

Iriawan (2001) menguraikan prosedur yang dilakukan pada Gibbs Sampling untuk pemodelan estimasi parameter regresi yaitu.

1. Menetapkan nilai awal untuk parameter regresi *mixture* Poisson:

$\Phi^{(t)} = (\mathbf{p}, \boldsymbol{\beta})^{(t)}$ , untuk parameter  $\mathbf{p} = (p_1, p_2, \dots, p_K)$ ,  $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_K)$ ,  $\boldsymbol{\beta}_1 = (\beta_{01}, \beta_{11}, \dots, \beta_{p1})$ ,  $\boldsymbol{\beta}_2 = (\beta_{02}, \beta_{12}, \dots, \beta_{p2})$ ,  $\dots$ ,  $\boldsymbol{\beta}_K = (\beta_{0K}, \beta_{1K}, \dots, \beta_{pK})$ , pada iterasi  $t = 0$ .

2. Membangkitkan komponen parameter untuk masing-masing *mixture*

a. Membangkitkan  $\mathbf{p}^{t+1}$  dari *full conditional posterior distribution* pada persamaan (2.29).

$$f(\mathbf{p} | y, \mathbf{T}_i, \boldsymbol{\beta}_1^{(t)}, \boldsymbol{\beta}_2^{(t)}, \dots, \boldsymbol{\beta}_K^{(t)})$$

b. Membangkitkan  $\boldsymbol{\beta}_1^{(t+1)}$  dari *full conditional posterior distribution* pada persamaan (2.30)

$$f(\boldsymbol{\beta}_1 | y, \mathbf{p}^{t+1}, \boldsymbol{\beta}_2^{(t)}, \boldsymbol{\beta}_3^{(t)}, \dots, \boldsymbol{\beta}_K^{(t)})$$

c. Membangkitkan  $\boldsymbol{\beta}_2^{(t+1)}$  dari *full conditional posterior distribution*

$$f(\boldsymbol{\beta}_2 | y, \mathbf{p}^{t+1}, \boldsymbol{\beta}_1^{(t+1)}, \boldsymbol{\beta}_3^{(t)}, \dots, \boldsymbol{\beta}_K^{(t)})$$

⋮

d. Membangkitkan  $\boldsymbol{\beta}_k^{(t+1)}$  dari *full conditional posterior distribution*

$$f(\boldsymbol{\beta}_k | y, \mathbf{p}^{t+1}, \boldsymbol{\beta}_1^{(t+1)}, \boldsymbol{\beta}_2^{(t+1)}, \dots, \boldsymbol{\beta}_{k-1}^{(t+1)}, \boldsymbol{\beta}_{k+1}^{(t)}, \dots, \boldsymbol{\beta}_K^{(t)})$$

⋮

e. Membangkitkan  $\boldsymbol{\beta}_K^{(t+1)}$  dari *full conditional posterior distribution*

$$f(\boldsymbol{\beta}_K | y, \mathbf{p}^{t+1}, \boldsymbol{\beta}_1^{(t+1)}, \boldsymbol{\beta}_2^{(t)}, \dots, \boldsymbol{\beta}_{K-1}^{(t+1)}).$$

3. Meningkatkan iterasi  $t$ , dimana iterasi baru  $t = t + 1$
4. Mengulangi langkah 2 dan langkah 3 hingga sebanyak  $M$  kali, dimana  $M \rightarrow \infty$  atau semua parameter telah mencapai konvergen.

## 2.9 Deviance Information Criterion (DIC)

*Deviance Information Criterion* (DIC) diperkenalkan oleh Spiegelhalter dkk (2002) sebagai ukuran perbandingan model dan kecukupan model. Persamaan yang digunakan untuk menghitung DIC dinyatakan sebagai berikut.

$$DIC(m) = 2\overline{D(\lambda_m, m)} - D(\bar{\lambda}_m, m) = D(\bar{\lambda}_m, m) + 2p_m$$

dengan  $D(\lambda_m, m)$  ukuran *deviance* biasa, yang nilainya sama dengan dua kali loglikelihood

$$D(\lambda_m, m) = -2 \log f(y | \lambda_m, m) \quad (2.32)$$

dan  $\overline{D(\lambda_m, m)}$  adalah nilai *mean* posterior,  $p_m$  dapat dinyatakan sebagai jumlah efektif untuk parameter model  $m$  dengan rumus

$$p_m = \overline{D(\lambda_m, m)} - D(\bar{\lambda}_m, m) \quad (2.33)$$

$\bar{\lambda}_m$  merupakan mean posterior dari parameter yang dimasukkan dalam model  $m$ . DIC yang semakin kecil menunjukkan model yang semakin baik (Ntzoufras, 2009).

## 2.10 Fasilitas Kesehatan Tingkat Pertama (FKTP)

Berdasarkan peraturan presiden nomor 19 tahun 2016, bahwa pelayanan kesehatan pada FKTP mencakup beberapa poin, diantaranya adalah :

1. Administrasi pelayanan
2. Pelayanan promotif dan preventif
3. Pemeriksaan, pengobatan, dan konsultasi medis
4. Tindakan medis non spesialistik, baik operatif maupun non operatif
5. Pelayanan obat dan bahan medis habis pakai

6. Pemeriksaan penunjang diagnostik laboratorium tingkat pratama dan
7. Rawat inap tingkat pertama sesuai indikasi medis.

Fasilitas kesehatan yang termasuk FKTP adalah:

1. Puskesmas

Puskesmas adalah Unit Pelaksana Teknis Dinas (UPTD) kesehatan kabupaten/kota yang bertanggung jawab menyelenggarakan pembangunan kesehatan di suatu wilayah kerja. Puskesmas sebagai salah satu jenis fasilitas pelayanan kesehatan tingkat pertama memiliki peranan penting dalam sistem kesehatan nasional, khususnya subsistem upaya kesehatan. Dalam penyelenggaraan puskesmas perlu ditata ulang untuk meningkatkan aksesibilitas, keterjangkauan, dan kualitas kesehatan dalam rangka meningkatkan derajat masyarakat serta menyukseskan program jaminan sosial nasional (Permenkes-No-75, 2014). Pembangunan kesehatan yang diselenggarakan di puskesmas bertujuan untuk mewujudkan masyarakat yang:

- a. Memiliki perilaku sehat yang meliputi kesadaran, kemauan, dan kemampuan hidup sehat;
- b. Mampu menjangkau pelayanan kesehatan bermutu
- c. Hidup di lingkungan sehat; dan
- d. Memiliki derajat kesehatan optimal, baik individu, keluarga, kelompok dan masyarakat.

Prinsip penyelenggaraan puskesmas meliputi:

- a. Paradigma sehat;
- b. Pertanggungjawaban wilayah;
- c. Kemandirian masyarakat;
- d. Pemerataan;
- e. Teknologi tepat guna; dan
- f. Keterpaduan dan kesinambungan.

Berdasarkan prinsip pemerataan, puskesmas menyelenggarakan pelayanan kesehatan yang dapat diakses dan dijangkau oleh seluruh masyarakat di wilayah kerjanya secara adil tanpa membedakan status sosial, ekonomi, agama, budaya, dan kepercayaan. Puskesmas harus

didirikan pada setiap kecamatan dengan mempertimbangkan kebutuhan pelayanan, jumlah penduduk, aksesibilitas, persyaratan lokasi, bangunan, sarana prasarana, peralatan kesehatan, ketenagaan, kefarmasian dan laboratorium. Rasio ideal dokter menurut WHO adalah 1 dokter untuk 2.500 penduduk. Sebanyak 17.507 dokter bekerja di Puskesmas sehingga setiap puskesmas rata-rata memiliki 1,8 dokter. Akan tetapi data kementerian kesehatan menunjukkan 938 dari 9.599 puskesmas di Indonesia masih kekurangan bahwa tidak memiliki dokter akibat distribusi yang tidak merata.

#### 2. Praktek Dokter Umum dan Dokter Gigi

Praktik dokter umum adalah rangkaian kegiatan yang dilakukan oleh dokter umum terhadap pasien dalam melaksanakan upaya kesehatan. Praktik dokter gigi adalah rangkaian kegiatan yang dilakukan oleh dokter gigi terhadap pasien dalam melaksanakan upaya kesehatan. (UU No. 29 Tahun 2004)

#### 3. Klinik Pratama / Umum

Klinik umum adalah fasilitas pelayanan kesehatan yang menyelenggarakan pelayanan kesehatan perorangan yang menyediakan pelayanan medis dasar, diselenggarakan oleh lebih dari satu jenis tenaga kesehatan dan dipimpin oleh seorang tenaga medis. (Permenkes No. 28 tahun 2011).

#### 4. RS Kelas D Pratama

RS Pratama adalah rumah sakit umum yang mempunyai fasilitas dan kemampuan pelayanan kesehatan dasar yang tidak membedakan kelas perawatan dalam upaya menjamin peningkatan akses bagi masyarakat dalam rangka penyelenggaraan kegiatan upaya kesehatan perorangan yang memberikan pelayanan gawat darurat selama 24 jam, pelayanan rawat jalan, dan rawat inap.

*(halaman ini sengaja dikosongkan)*

## **BAB III**

### **METODE PENELITIAN**

#### **3.1 Sumber Data**

Data yang digunakan dalam penelitian merupakan data sekunder yang diperoleh berdasarkan data Dinas Kesehatan Surabaya mengenai data puskesmas dan alamat lengkap untuk menentukan titik latitude dan longitude masing-masing puskesmas dan data klinik pratama. Data puskesmas dan klinik pratama merupakan tempat rujukan untuk memperoleh layanan kesehatan BPJS. Data variabel kovariat diperoleh dari publikasi Badan Pusat Statistik Kota Surabaya dengan judul Kota Surabaya dalam Angka tahun 2017, sedangkan data tingkat aksesibilitas diperoleh dari publikasi tugas akhir Mursalim (2018) dengan judul “Pengukuran Aksesibilitas Kecamatan di Wilayah Pemerintah Kota Surabaya”.

#### **3.2 Variabel Penelitian**

Variabel yang akan diteliti merupakan pola titik-titik lokasi puskesmas dan klinik pratama di Kota Surabaya beserta variabel kovariat yang diduga berpengaruh terhadap persebaran FKTP. Variabel kovariat yang digunakan merupakan karakteristik masing-masing kecamatan.

##### **1. Variabel respon penelitian**

Variabel respon yang digunakan yaitu banyak puskesmas dan klinik pratama yang diamati pada objek *tessellations* berupa batas-batas masing-masing kecamatan, sehingga sebanyak 63 puskesmas dan 155 klinik pratama yang berada di Kota Surabaya berdasarkan titik lokasinya dimasukkan ke dalam objek kecamatan. Selanjutnya banyaknya titik puskesmas dan klinik pratama yang jatuh pada objek *tessellations* diamati sebagai variabel random dari sebuah proses Poisson. Masing-masing titik lokasi puskesmas dan klinik pratama pada ruang dua dimensi dinyatakan dalam garis latitude dan longitude. Berikut ini merupakan data lokasi puskesmas yang diperoleh dari Dinas Kesehatan Surabaya.

**Tabel 3.1** Data lokasi puskesmas

No	Nama Puskesmas	Latitude	Longitude
1	Gayungan	-7,338074	112,718704
2	Kedurus	-7,319671	112,709634
3	Gununganyar	-7,340858	112,783992
⋮	⋮	⋮	⋮
62	Tambak Wedi	-7,217434	112,771583
63	Sememi	-7,248419	112,635390

Sedangkan untuk data lokasi klinik pratama ditampilkan pada Tabel 3.2 berikut ini.

**Tabel 3.2** Data lokasi klinik pratama

No	Nama Klinik Pratama	Alamat
1	Klinik Nurani Jaya 83 (JST)	K.H. Abdul Karim No 17 SBY
2	BP Widya Mandiri II (108)	Gubeng Kertajaya V C No 24
3	BP Klinik Kebangkitan (JST)	Manukan Madya 141 Tandes SBY
⋮	⋮	⋮
154	Klinik Rahap Bersalin Al-Azhar	Dupak Bandarejo No.23
155	Putri Rahayu	Jl. Mastrip IX No.9 Karang Pilang

## 2. Variabel kovariat penelitian

Variabel kovariat penelitian merupakan karakteristik yang diambil pada masing-masing kecamatan. Dengan menggunakan pusat administrasi kecamatan sebagai titik lokasi variabel kovariat dilakukan transformasi *pixel* ke dalam bentuk *pixel image* sesuai dengan batasan objek *tessellation*. Hasil pengukuran variabel kovariat pada lokasi pusat administrasi kecamatan adalah sebagai berikut.

**Tabel 3.3** Variabel Kovariat

No	Nama	Lintang	Bujur	$X_1$	$X_2$	$X_3$	$X_4$
1	Gayungan	-7,33821	112,71666	72,24	92,08	60	73,43
2	Karangpilang	-7,33335	112,69942	76,68	94,58	75	40,85
3	Gununganyar	-7,34094	112,78310	54,42	98,75	71	47,69
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
30	Kenjeran	-7,22651	112,77555	177,45	75,77	75	28,45
31	Benowo	-7,24885	112,63539	24,22	91,16	77	29,47

Deskripsi masing-masing variabel penelitian akan dijelaskan pada Tabel 3.4 sebagai berikut.

**Tabel 3.4** Variabel Penelitian

<b>Variabel</b>	<b>Deskripsi</b>	<b>Tipe Data</b>
Banyak FKTP (Y)	Banyak Puskesmas dan Klinik Pratama yang terletak pada masing-masing objek <i>tessellations</i>	<i>Count</i>
X <sub>1</sub> (Kepadatan Penduduk)	Rasio jumlah penduduk masing-masing kecamatan terhadap luas area kecamatan	Rasio
X <sub>2</sub> (Persentase Rumah Tangga Bersih dan Sehat)	Rumah Tangga Bersih adalah rumah tangga yang dapat memelihara, meningkatkan, dan melindungi kesehatan setiap anggota rumah tangga dari penyakit dan kondisi lingkungan yang tidak mendukung kehidupan yang sehat.	Rasio
X <sub>3</sub> (Angka Keterkaitan)	Indeks Aksesibilitas diukur dengan menggunakan parameter waktu tempuh sebagai sistem pengukuran tidak langsung. Data direkam berdasarkan perangkat lunak Google Maps, dan metode pengukuran aksesibilitas dan penggambaran jaringan jalan didasarkan pada wilayah kecamatan, yang terdiri dari 31 kecamatan. Kantor kecamatan ditetapkan sebagai pusat kecamatan dan jalan yang menghubungkannya ditetapkan sebagai rute.	Rasio
X <sub>4</sub> (Kondisi Jalan Baik dalam km)	Jalan nasional / provinsi / kota yang melewati kecamatan dalam kondisi baik	Rasio

### 3.3 Struktur Data Penelitian

Struktur data penelitian dibangun berdasarkan variabel respon dan variabel kovariat ditunjukkan pada Tabel 3.5 sebagai berikut.

**Tabel 3.5** Struktur Data Penelitian

Kecamatan ( <i>tessellations</i> )	(Y)	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>
1	Y <sub>1</sub>	X <sub>1.1</sub>	X <sub>2.1</sub>	X <sub>3.1</sub>	X <sub>4.1</sub>
2	Y <sub>2</sub>	X <sub>1.2</sub>	X <sub>2.2</sub>	X <sub>3.2</sub>	X <sub>4.2</sub>
3	Y <sub>3</sub>	X <sub>1.3</sub>	X <sub>2.3</sub>	X <sub>3.3</sub>	X <sub>4.3</sub>
⋮	⋮	⋮	⋮	⋮	⋮
30	Y <sub>30</sub>	X <sub>1.30</sub>	X <sub>2.30</sub>	X <sub>3.30</sub>	X <sub>4.30</sub>
31	Y <sub>31</sub>	X <sub>1.31</sub>	X <sub>2.31</sub>	X <sub>3.31</sub>	X <sub>4.31</sub>

### 3.4 Tahapan Penelitian

Metode dan tahapan penelitian yang akan dilakukan untuk mencapai tujuan penelitian diuraikan sebagai berikut.

1. Melakukan eksplorasi data titik lokasi FKTP di Kota Surabaya

Titik-titik lokasi puskesmas dan klinik pratama yang ditunjukkan berdasarkan latitude dan longitude dipetakan pada peta Kota Surabaya. Membentuk objek *tessellations* berdasarkan *polygon* kecamatan sebagai *window* untuk mengamati pola persebaran *point pattern*. Identifikasi persebaran dengan metode *quadrat count* sehingga diperoleh data banyak FKTP pada masing-masing objek *tessellations*. Kemudian menghitung intensitas FKTP berdasarkan luas area kecamatan.

2. Melakukan Eksplorasi Variabel Kovariat

Identifikasi pola variabel kovariat dan hubungan terhadap persebaran FKTP di Surabaya. Melakukan transformasi pixel untuk masing-masing variabel kovariat.

3. Melakukan Identifikasi SPP

Pola *point pattern* dianalisis untuk menentukan proses yang sesuai untuk menggambarkan kondisi persebaran FKTP di Kota Surabaya. Apabila titik FKTP cenderung menyebar menjauhi satu sama lain maka proses yang sesuai untuk menggambarkan kondisi tersebut adalah proses Gibbs. Untuk proses yang saling independen digambarkan melalui proses poisson dan untuk pola Klaster digambarkan dengan proses Cox. Apabila intensitas konstan maka proses disebut HPP dan NHPP jika intensitas bervariasi secara spasial. Identifikasi homogenitas proses dilakukan dengan menggunakan pengujian *quadrat counting test*.

4. Melakukan Identifikasi Model *Mixture*

Dengan menggunakan histogram intensitas FKTP dilakukan identifikasi apakah pola persebaran FKTP terdiri dari satu proses poisson atau lebih yang kemudian disebut sebagai *mixture* proses Poisson.

5. Membentuk Fungsi untuk Estimasi Intensitas

Seperti telah dijelaskan sebelumnya bahwa dalam analisis SPP intensitas akan menjadi fokus utama. Sehingga pada langkah ini dilakukan pemodelan terhadap intensitas FKTP di Kota Surabaya dengan menggunakan *mixture* proses Poisson dengan menggunakan WinBUGS.

- a. Menentukan banyak komponen penyusun *mixture*
- b. Menentukan anggota *mixture*  $k$  komponen

Menentukan anggota *mixture*  $k$  komponen dengan menggunakan *Gibbs sampling*.

i) Pada saat  $t = 0$ , memberi nilai state awal  $(\lambda_1^0, \dots, \lambda_K^0, \mathbf{p}^0, \mathbf{T}_1^0, \dots, \mathbf{T}_n^0)$  dengan memperhatikan karakteristik datanya.

ii) Untuk  $t = 1, 2, \dots, N$ , memperbarui nilai masing-masing variabel pada iterasi ke- $t$  :

a) *Update* Gibbs untuk  $\lambda_k, k = 1, \dots, K$ : membangkitkan sampel  $\lambda_k^t$  dari  $gamma(\alpha_k + \sum_{i=1}^n y_i T_{ik}, \sum_{i=1}^n T_{ik} + \beta_k)$  dengan menggunakan nilai  $T_{ik}$  paling baru.

b) *Update* Gibbs untuk proporsi  $\mathbf{p}$ : membangkitkan sampel  $p^t$  dari  $Dirichlet(1 + \sum_{i=1}^n T_{i1}, \dots, 1 + \sum_{i=1}^n T_{iK})$  dengan nilai  $T_{i1}, \dots, T_{iK}$  paling baru.

c) *Update* Gibbs untuk indikator  $\mathbf{T}_i$ : membangkitkan sampel  $\mathbf{T}_i^t$  dari *multinomial*  $(1, p_{i1}, \dots, p_{iK})$   $i = 1, \dots, n, k = 1, \dots, K$ , dimana

$$p_{ik} = \frac{f(y_i|\lambda_k)p_k}{f(y_i)} \quad k = 1, \dots, K$$

menggunakan nilai  $\lambda_k$  dan  $\mathbf{p}$ .

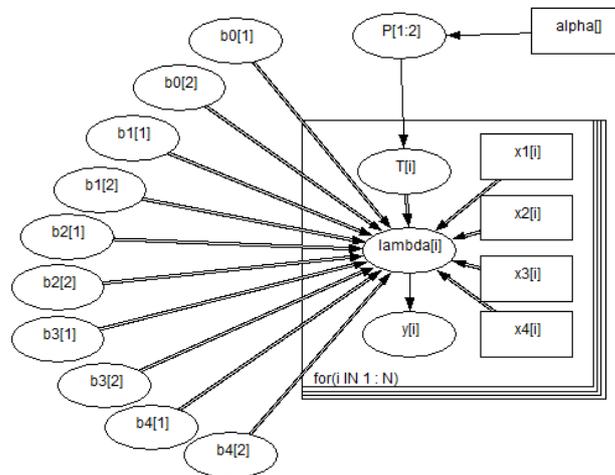
d) Sehingga diperoleh *state* Markov Chain  $(\lambda_1^t, \dots, \lambda_K^t, \mathbf{p}^t, \mathbf{T}_1^t, \dots, \mathbf{T}_n^t)$  yang digunakan untuk membangkitkan sampel pada iterasi  $t + 1$ .

iii) Kembali ke langkah ke-ii sampai sebanyak  $N - 1$  kali iterasi untuk memperoleh Markov Chain dengan ukuran  $N$ .

c. Menentukan distribusi *prior* dan *hyperprior* dari parameter dan hyperparameter yang akan diestimasi.

d. Menyusun *Directed Acyclic Graph* (DAG) pada WinBUGS model *mixture* Poisson

Gambar 3.1 merupakan bentuk *Directed Acyclic Graph* (DAG) pada WinBUGS untuk model *mixture* Poisson. Menentukan besarnya sampel yang akan dibangkitkan. Sampel dibangkitkan sampai dapat memenuhi sifat *irreducible, recurrent, aperiodic*.



**Gambar 3.1** DAG Regresi *Mixture* Poisson

e. Melakukan proses iterasi penaksiran parameter

Estimasi parameter regresi *mixture* poisson dilakukan dengan Gibbs Sampling.

i. Menetapkan nilai awal untuk parameter regresi *mixture* Poisson:

$$\Phi^{(t)} = (\mathbf{p}, \boldsymbol{\beta})^{(t)}, \text{ untuk parameter } \mathbf{p} = (p_1, p_2, \dots, p_K),$$

$$\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_K), \text{ pada iterasi } t = 0.$$

ii. Membangkitkan komponen parameter untuk masing-masing *mixture*

a) Membangkitkan  $\mathbf{p}^{t+1}$  dari *full conditional posterior distribution* pada persamaan (2.29)

$$f(\mathbf{p} | y, \boldsymbol{\beta}_1^{(t)}, \boldsymbol{\beta}_2^{(t)}, \dots, \boldsymbol{\beta}_K^{(t)})$$

b) Membangkitkan  $\boldsymbol{\beta}_1^{(t+1)}$  dari *full conditional posterior distribution* pada persamaan (2.30)

$$f(\boldsymbol{\beta}_1 | y, \mathbf{p}^{t+1}, \boldsymbol{\beta}_2^{(t)}, \boldsymbol{\beta}_3^{(t)}, \dots, \boldsymbol{\beta}_K^{(t)})$$

c) Membangkitkan  $\boldsymbol{\beta}_2^{(t+1)}$  dari *full conditional posterior distribution*

$$f(\boldsymbol{\beta}_2 | y, \mathbf{p}^{t+1}, \boldsymbol{\beta}_1^{(t+1)}, \boldsymbol{\beta}_3^{(t)}, \dots, \boldsymbol{\beta}_K^{(t)})$$

⋮

d) Membangkitkan  $\boldsymbol{\beta}_k^{(t+1)}$  dari *full conditional posterior distribution*

$$f(\boldsymbol{\beta}_k | y, \mathbf{p}^{t+1}, \boldsymbol{\beta}_1^{(t+1)}, \boldsymbol{\beta}_2^{(t+1)}, \dots, \boldsymbol{\beta}_{k-1}^{(t+1)}, \boldsymbol{\beta}_{k+1}^{(t)}, \dots, \boldsymbol{\beta}_K^{(t)})$$

⋮

e) Membangkitkan  $\boldsymbol{\beta}_K^{(t+1)}$  dari *full conditional posterior distribution*

$$f(\boldsymbol{\beta}_K | y, \mathbf{p}^{t+1}, \boldsymbol{\beta}_1^{(t+1)}, \boldsymbol{\beta}_2^{(t)}, \dots, \boldsymbol{\beta}_{K-1}^{(t+1)}).$$

iii. Meningkatkan iterasi  $t$ , dimana iterasi baru  $t = t + 1$

iv. Mengulangi langkah 2 dan langkah 3 hingga sebanyak  $M$  kali, dimana  $M \rightarrow \infty$  atau semua parameter telah mencapai konvergen.

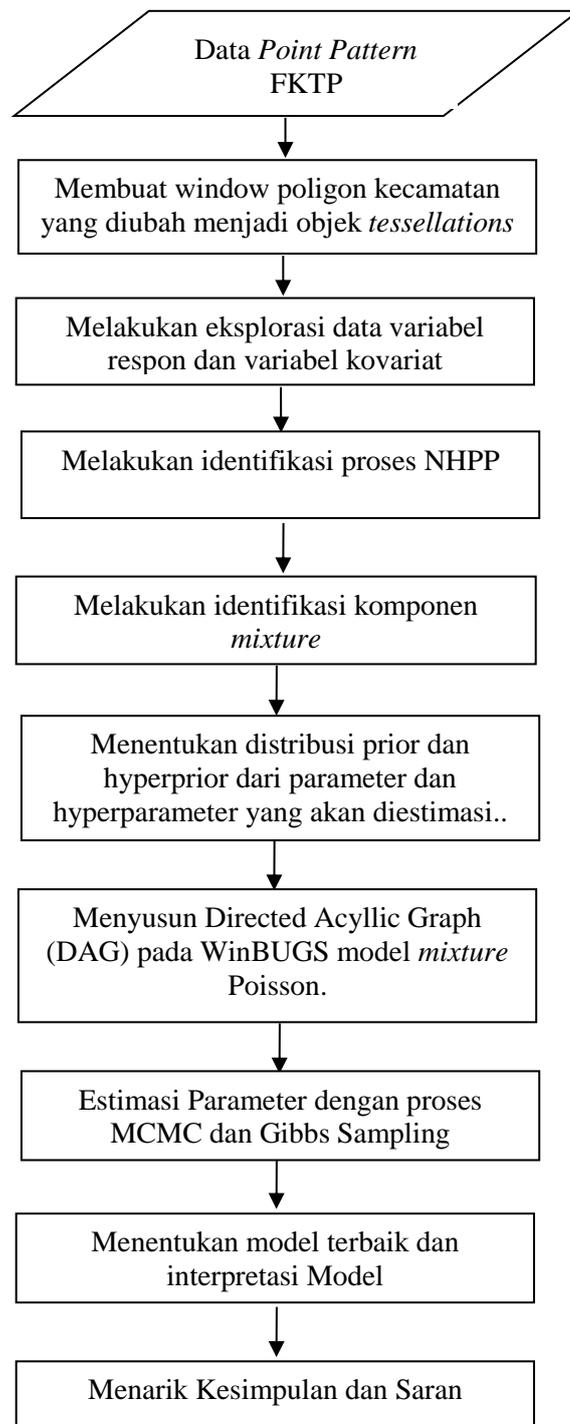
6. Menentukan model terbaik

Model terbaik ditentukan berdasarkan nilai DIC terkecil. Berdasarkan model terbaik dilakukan intepretasi model.

7. Membuat kesimpulan dan saran

Kesimpulan berupa variabel yang berpengaruh terhadap intensitas FKTP sehingga bisa menjadi pertimbangan Dinas Kesehatan Surabaya dalam upaya pemerataan FKTP.

Berdasarkan tahapan proses yang telah diuraikan, maka dapat digambarkan diagram alir seperti pada Gambar 3.2 sebagai berikut.



**Gambar 3.2** Diagram Alir Penelitian

## BAB IV

### ANALISIS DAN PEMBAHASAN

#### 4.1 Penelitian Sebelumnya

Pada penelitian ini menggunakan dari berbagai sumber referensi. Referensi pertama yaitu Tesis dengan judul Bayesian Inference on Finite *Mixtures* of Poisson Distributions oleh Hamdah (2015). Dalam penelitian tersebut menjelaskan mengenai *Finite Mixtures of Poisson Distributions* dengan pendekatan Bayesian dengan mengambil contoh dua komponen *mixture* serta dilakukan analisis untuk menentukan distribusi prior yang sesuai untuk masing-masing parameter *Finite Mixtures of Poisson Distributions*.

##### 4.1.1 Pendekatan Bayesian pada Distribusi *Mixture Poisson*

Dasar inferensi Bayesian diturunkan dari teorema Bayes, yang menggambarkan hubungan antara peluang bersyarat dari dua himpunan kejadian  $A_i = \{A_1, A_2, \dots, A_n\}$  dan  $B$  sebagai berikut.

$$\begin{aligned} P(A_i|B) &= \frac{P(B|A_i)P(A_i)}{P(B)} \\ &= \frac{P(B|A_i)P(A_i)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + \dots + P(B|A_n)P(A_n)} \\ &= \frac{P(B|A_i)P(A_i)}{\sum_i P(B|A_i)P(A_i)} \end{aligned}$$

Dalam statistika inferensi, dua hal penting adalah data dan parameter. Konsep Bayesian menyatakan bahwa parameter adalah random variabel, jadi misalkan  $A_i = \{A_1, A_2, \dots, A_n\}$  adalah suatu himpunan dari parameter-parameter yang tidak teramati (namun memiliki distribusi) sehingga akan diganti notasinya menjadi  $\theta_i = \theta_1, \theta_2, \dots, \theta_n$  dengan  $\theta_i$  mempartisi semesta. Kemudian untuk setiap  $\theta_i, i = 1, 2, \dots, n$ , akan memiliki distribusi dengan fungsi probabilitas yaitu  $f(\theta_i)$  yang berasal dari informasi yang telah tersedia sebelumnya atau secara subjektif ditentukan peneliti, yang disebut dengan probabilitas priordan distribusinya disebut sebagai distribusi prior. Sedangkan  $B$  dimisalkan sebagai data observasi yang baru diperoleh dari pengamatan, kemudian dimisalkan sebagai  $x$ . Jadi secara singkat

dapat digambarkan bahwa dalam suatu pengamatan, akan diperoleh satu set data pengamatan  $x$  dan satu keyakinan tentang karakteristik data ( $\theta$ ). Dengan mengganti  $B$  dengan pengamatan  $x = (x_1, x_2, \dots, x_n)$ ,  $A$  dengan parameter  $\theta$ , dan probabilitas  $P$  dengan fungsi  $f$ , menghasilkan persamaan sebagai berikut.

$$f(\theta|x) = \frac{f(x|\theta)f(\theta)}{f(x)} \quad (4.1)$$

$f(\theta)$  : distribusi prior untuk parameter  $\theta$  sebelum  $y$  diamati;

$f(x|\theta)$  : likelihood  $x$  dibawah sebuah model;

$f(\theta|x)$  : distribusi posterior dari parameter  $\theta$ ;

$f(x)$  : sebuah konstanta karena merupakan total peluang dari  $f(x|\theta)f(\theta)$ , dimana

$f(x) = \sum_i f(x|\theta_i) f(\theta_i)$ , untuk diskrit

$f(x) = \int f(x|\theta)f(\theta)d\theta$ , untuk kontinyu

Misalkan  $X_1, X_2, \dots, X_n$  merupakan sampel i.i.d. dari distribusi  $f(x|\theta)$ ,  $f(\theta)$  merupakan distribusi prior dan  $f(x|\theta)$  merupakan distribusi posterior. Maka likelihood marginal atau prediksi distribusi prior diperoleh berdasarkan:

$$f(x) = \int_{-\infty}^{\infty} f(x|\theta)f(\theta) d\theta \quad (4.2)$$

Jadi penyebut  $f(x) = \int_{-\infty}^{\infty} f(x|\theta)f(\theta) d\theta$  mendefinisikan likelihood marginal  $x$ , atau distribusi prediktif prior  $x$ , dan dapat diatur sebagai konstanta  $c$ . Distribusi prediktif prior menunjukkan data  $x$  yang seharusnya atau yang diberikan oleh model sebelum  $x$  diobservasi. Hanya fungsi probabilitas prior dan fungsi likelihood model yang digunakan untuk likelihood marginal  $x$ .

Dalam teori probabilitas, konstanta normalisasi merupakan sebuah konstanta dimana sebuah fungsi non-negatif harus dikalikan agar luas area dalam grafik sama dengan 1, misal digunakan untuk membuat sebuah fungsi menjadi fungsi densitas peluang. Misalnya, jika didefinisikan sebuah fungsi  $f(x)$  sebagai berikut.

$$f(x) = e^{-x^2/2}, x \in (-\infty, \infty)$$

$$\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{\infty} e^{-x^2/2} dx = \sqrt{2\pi} \quad (4.3)$$

Jika didefinisikan fungsi  $\varphi(x)$  sebagai

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \quad (4.4)$$

sehingga

$$\int_{-\infty}^{\infty} \varphi(x) dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = 1 \quad (4.5)$$

Fungsi  $\varphi(x)$  merupakan fungsi densitas peluang distribusi normal standar. Dan konstanta  $\frac{1}{\sqrt{2\pi}}$  merupakan konstanta normalisasi fungsi  $f(x)$ . Dengan melihat kembali pada persamaan (4.1) dan (4.2), dapat diketahui bahwa, likelihood marginal  $x$  menormalisasi *joint* distribusi posterior,  $f(\theta|x)$ , memastikan agar fungsi tersebut menjadi sebuah distribusi probabilitas dan bernilai integral sama dengan satu.

Dengan mengganti  $f(x)$  dengan  $c$  sesuai dengan persamaan 4.1 dimana  $f(x)$  adalah sebuah konstanta karena merupakan total peluang dari  $f(x|\theta)f(\theta)$  atau merupakan simbol dari konstanta proporsionalitas yang biasanya disebut fungsi konstanta normalisasi  $f(x)$ , model berdasarkan teorema bayes menjadi sebagai berikut.

$$f(\theta|x) = \frac{f(x|\theta)f(\theta)}{c} \quad (4.6)$$

Dengan menghilangkan konstanta  $c$  pada model maka persamaan “=” berubah menjadi proporsional “ $\propto$ ”

$$f(\theta|x) \propto f(x|\theta)f(\theta) \quad (4.7)$$

Dengan kata lain posterior yang tidak ternormalisasi proporsional terhadap perkalian likelihood dengan prior.

Posterior  $\propto$  Likelihood x Prior

Likelihood dari parameter  $\theta$  merupakan suatu probabilitas bersyarat yaitu data pengamatan  $x$  yang diperoleh berada dibawah kondisi parameter  $\theta$ . Likelihood  $f(x|\theta_i)$  tersebut menjadi pembobot bagi setiap prior  $f(\theta_i)$  dari data observasi  $x$ .  $f(x|\theta_i)$  diberi nama fungsi likelihood karena parameter  $\theta$  karena semakin besar fungsi  $f(x|\theta)$  akan “*more likely*” (lebih mirip) terhadap nilai parameter  $\theta$  yang sebenarnya dibanding fungsi  $f(x|\theta_i)$  yang lebih kecil.

Sedangkan  $f(\theta_i|x)$  ) adalah fungsi distribusi peluang posterior dari parameter  $\theta$  dengan syarat data observasi  $x$ . Secara sederhana, prior adalah fungsi parameter  $\theta_i$  sebelum observasi dilakukan yang diperoleh berdasarkan pengetahuan peneliti secara subjektif, dan posterior adalah fungsi parameter  $\theta$

yang dilakukan *adjust* dari prior setelah memperoleh data pengamatan. Posterior yang diperoleh ini kelak bisa menjadi prior pada pengamatan berikutnya. Dan posterior inilah yang menjadi target dalam analisis statistik Bayesian.

Misalkan  $X_1, X_2, \dots, X_n$  merupakan sampel i.i.d. berdistribusi Poisson ( $\theta$ ) dengan densitas prior diberikan oleh:

$$f(\theta) = e^{-\theta}, \theta > 0 \quad (4.8)$$

Fungsi likelihood adalah sebagai berikut.

$$\begin{aligned} L(x|\theta) &= \prod_{i=1}^n f(x_i|\theta) \\ &= \prod_{i=1}^n \frac{1}{x_i!} e^{-\theta} \theta^{x_i} \\ &= \left( \frac{1}{x_1!} e^{-\theta} \theta^{x_1} \right) \left( \frac{1}{x_2!} e^{-\theta} \theta^{x_2} \right) \dots \left( \frac{1}{x_n!} e^{-\theta} \theta^{x_n} \right) \\ &= \frac{1}{x_1! x_2! \dots x_n!} e^{-\theta} e^{-\theta} \dots e^{-\theta} \theta^{x_1} \dots \theta^{x_n} \\ &= \frac{1}{x_1! x_2! \dots x_n!} e^{-n\theta} \theta^{x_1 + x_2 + \dots + x_n} \\ &= \frac{1}{x_1! x_2! \dots x_n!} e^{-n\theta} \theta^{\sum_{i=1}^n x_i} \end{aligned} \quad (4.9)$$

Densitas posterior adalah sebagai berikut.

$$\begin{aligned} f(\theta|x) &\propto L(x|\theta) f(\theta) \\ &= \frac{1}{x_1! x_2! \dots x_n!} e^{-n\theta} \theta^{\sum_{i=1}^n x_i} e^{-\theta} \\ &= \frac{1}{x_1! x_2! \dots x_n!} e^{-n\theta - \theta} \theta^{\sum_{i=1}^n x_i} \\ &= \frac{1}{x_1! x_2! \dots x_n!} e^{-(n+1)\theta} \theta^{\sum_{i=1}^n x_i} \end{aligned} \quad (4.10)$$

Untuk konstanta yang tidak memuat  $\theta$  tidak ditulis sehingga persamaan menjadi

$$f(\theta|x) \propto e^{-(n+1)\theta} \theta^{\sum_{i=1}^n x_i} \quad (4.11)$$

Maka terlihat bahwa persamaan tersebut merupakan distribusi Gamma dengan parameter  $1 + \sum_{i=1}^n x_i$  dan  $n + 1$  sehingga

$$(\theta|x) \sim \text{Gamma}(1 + \sum_{i=1}^n x_i, n + 1) \quad (4.12)$$

### **Conjugate Prior**

Keluarga distribusi probabilitas dinamakan keluarga konjugat prior untuk sampel i.i.d. dari likelihood  $f(x|\theta)$ , jika distribusi prior untuk  $\theta$  adalah anggota keluarga distribusi, maka distribusi posterior untuk  $\theta$  juga merupakan anggota

keluarga distribusi tersebut, untuk berbagai ukuran sampel dan nilai sampel. Misalkan  $X_1, X_2, \dots, X_n$  merupakan sampel i.i.d. berdistribusi Poisson ( $\theta$ ) dengan densitas prior diberikan oleh distribusi Gamma ( $\alpha, \beta$ ) dengan fungsi densitas prior sebagai berikut:

$$f(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta}, \theta > 0, \alpha > 0, \beta > 0. \quad (4.13)$$

Fungsi Likelihood dapat dituliskan sebagai berikut:

$$\begin{aligned} L(x|\theta) &= \prod_{i=1}^n f(x_i|\theta) \\ &= \prod_{i=1}^n \frac{1}{x_i!} e^{-\theta} \theta^{x_i} \\ &= \left( \frac{1}{x_1!} e^{-\theta} \theta^{x_1} \right) \left( \frac{1}{x_2!} e^{-\theta} \theta^{x_2} \right) \dots \left( \frac{1}{x_n!} e^{-\theta} \theta^{x_n} \right) \\ &= \frac{1}{x_1! x_2! \dots x_n!} e^{-\theta} e^{-\theta} \dots e^{-\theta} \theta^{x_1} \dots \theta^{x_n} \\ &= \frac{1}{x_1! x_2! \dots x_n!} e^{-n\theta} \theta^{x_1 + x_2 + \dots + x_n} \\ &= \frac{1}{x_1! x_2! \dots x_n!} e^{-n\theta} \theta^{\sum_{i=1}^n x_i} \end{aligned} \quad (4.14)$$

Densitas Posterior dapat dituliskan sebagai berikut:

$$\begin{aligned} f(\theta|x) &\propto L(x|\theta) f(\theta) \\ &= \frac{1}{x_1! x_2! \dots x_n!} e^{-n\theta} \theta^{\sum_{i=1}^n x_i} \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta} \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{1}{x_1! x_2! \dots x_n!} \theta^{\sum_{i=1}^n x_i + \alpha - 1} e^{-n\theta - \beta\theta} \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{1}{x_1! x_2! \dots x_n!} \theta^{\sum_{i=1}^n x_i + \alpha - 1} e^{-(n+\beta)\theta} \\ f(\theta|x) &\propto \theta^{\sum_{i=1}^n x_i + \alpha - 1} e^{-(n+\beta)\theta} \end{aligned} \quad (4.15)$$

Bentuk akhir pada persamaan 4.15 merupakan densitas distribusi Gamma dengan parameter,  $\sum_{i=1}^n x_i + \alpha$ , dan  $(n + \beta)$  sehingga

$$(\theta|x) \sim \text{Gamma} \left( \sum_{i=1}^n x_i + \alpha, n + \beta \right)$$

Terlihat  $f(\theta|x)$  merupakan satu keluarga distribusi dengan distribusi prior  $f(\theta)$  dengan parameter yang berbeda sehingga  $f(\theta)$  merupakan konjugat prior untuk  $\theta$ .

### Distribusi *Finite Mixture*

Distribusi *mixture* banyak digunakan untuk memodelkan data di mana setiap pengamatan diasumsikan berasal dari distribusi dengan parameter yang berbeda. Dengan kata lain, dapat dikatakan model *mixture* digunakan ketika populasi sampling terdiri dari sejumlah subpopulasi yang memiliki parameter yang berbeda.

Sebuah variable random  $X$  dikatakan memiliki distribusi *finite mixture* jika fungsi densitas  $f$  dapat diekspresikan dalam bentuk sebagai berikut.

$$f_{mix}(x|\boldsymbol{\theta}, \mathbf{p}) = \sum_{k=1}^K p_k f_k(x|\theta_k) \quad (4.16)$$

Dimana  $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_k)$ ,  $\mathbf{p} = (p_1, p_2, \dots, p_K)$  untuk nilai peluang  $p_k > 0, k = 1, \dots, K, k > 1$ , dengan  $\sum_{k=1}^K p_k = 1$ . Variabel random tersebut dikatakan memiliki densitas *k-finite*.

Keterangan:

1.  $\theta_k$  adalah vektor parameter atau skalar yang mengacu pada komponen *mixture* ke- $k$  dan  $\theta_k$  disebut dengan parameter *mixture* atau parameter komponen ke- $k$ .
2. Biasanya  $p_k$  disebut sebagai proporsi *mixture* atau bobot dan sering kali tidak diketahui, dan satu  $p_k$  sama dengan proporsi  $x$  bagian dari komponen ke-  $k$ .
3. Banyak  $k$  komponen tidak diketahui sehingga dianggap variabel random.

### *Missing Data*

Misalkan  $X_1, X_2, \dots, X_n$  adalah sampel i.i.d. dengan distribusi sesuai fungsi sesuai dengan persamaan (2.9), dengan vektor parameter  $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_K)$  dan proporsi  $\mathbf{p} = (p_1, p_2, \dots, p_K)$ . Dalam situasi tersebut untuk masing-masing pengamatan  $x_i, i = 1, 2, \dots, n$  dengan parameter indikator  $\mathbf{T}_i$  diperkenalkan sehingga  $T_{ik} = 1$  mengindikasikan bahwa pengamatan  $x_i$  bagian dari komponen *mixture* ke- $k$  dan  $T_{ik} = 0$  mengindikasikan bahwa pengamatan  $x_i$  bukan bagian dari komponen *mixture* ke- $k$

$$T_{ik} = \begin{cases} 1 \\ 0 \end{cases}$$

Untuk setiap  $T_i$  merupakan hasil percobaan tunggal yang sama dengan satu pada  $k$  kemungkinan komponen *mixture*, dengan peluang  $p_1, p_2, \dots, p_K \in (0,1)$  untuk  $k = 1, 2, \dots, K$  dan  $\sum_{k=1}^K p_k = 1$ . Fungsi densitas  $f(x_i|T_{ik} = 1)$  merupakan fungsi densitas dari komponen *mixture* ke- $k$  yang mengikuti distribusi Poisson ( $\lambda_k$ ). Besar peluang  $x_i$  merupakan anggota komponen *mixture* ke- $k$  adalah  $f(T_{ik} = 1|p) = p_k$ . Distribusi untuk masing-masing vektor tidak teramati  $T_i = (T_{i1}, \dots, T_{iK})$  dan  $T = (T_1, \dots, T_n)$  dengan nilai peluang  $p = (p_1, p_2, \dots, p_K)$ , adalah distribusi Multinomial  $(1, p_1, p_2, \dots, p_K)$ .

Pada kasus dari  $n$  i.i.d. pengamatan  $(x_1, x_2, \dots, x_n)$  dari model *mixture*, didefinisikan nilai peluang  $x_i$  bagian dari komponen *mixture* ke- $k$  dengan  $p = (p_1, p_2, \dots, p_K)$  dan parameter  $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_K)$ , fungsi likelihood dari densitas *mixture* sesuai fungsi *mixture* adalah sebagai berikut.

$$L_{mix}(\mathbf{x}|\boldsymbol{\lambda}, \mathbf{p}) = \prod_{i=1}^n f_{mix}(x_i|\boldsymbol{\lambda}, \mathbf{p}) \quad (4.17)$$

Perhitungan lengkap dari distribusi posterior dan khususnya perhitungan nilai ekspektasi distribusi posterior melibatkan perkalian likelihood *mixture* ke- $K^n$ , yang secara komputasi terlalu sulit untuk digunakan apabila pengamatan dalam jumlah banyak. Jadi representasi *missing data* dari distribusi *mixture* dapat dimanfaatkan sebagai salah satu cara untuk melakukan perhitungan numerik distribusi posterior. Pada persamaan (4.17) dapat digunakan indikator *missing data*  $T_i$  dimana tepat satu  $T_{ik}$  sama dengan 1 untuk setiap  $i$ , dan dapat ditulis likelihood *mixture* seperti pada persamaan (2.12).

$$L_{mix}(\mathbf{x}, \mathbf{T}|\boldsymbol{\lambda}, \mathbf{p}) = \prod_{i=1}^n \sum_{k=1}^K T_{ik} p_k f_k(x_i|\lambda_k) \quad (4.18)$$

$$L_{mix}(\mathbf{x}, \mathbf{T}|\boldsymbol{\lambda}, \mathbf{p}) = \prod_{i=1}^n \prod_{k=1}^K (p_k f_k(x_i|\lambda_k))^{T_{ik}} \quad (4.19)$$

Dimana  $f_k(x_i|\lambda_k)$  pada penelitian ini adalah distribusi Poisson dengan parameter  $\lambda_k$ . Pembuktian persamaan 4.19 dari persamaan 4.18 dapat dilihat dengan permisalan sebagai berikut. Misalkan data berasal dari distribusi *finite mixture* dengan  $k = 2$  dan  $n = 4$ . Parameter masing-masing komponen *mixture* adalah  $\lambda_1$  dan  $\lambda_2$ . Data diasumsikan berasal dari distribusi Poisson.

**Tabel 4.1** Data *point* dan *mixture*

	$x_i$	$k$
1	$x_1$	1
2	$x_2$	1
3	$x_3$	2
4	$x_4$	2

Likelihood data *finite mixture* Poisson adalah sebagai berikut.

$$\begin{aligned}
 L_{mix}(\mathbf{x}, \mathbf{T} | \boldsymbol{\lambda}, \mathbf{p}) &= \prod_{i=1}^n \sum_{k=1}^K T_{ik} p_k f_k(x_i | \lambda_k) \\
 &= [T_{11}f(x_1 | \lambda_1)p_1 + T_{12}f(x_1 | \lambda_2)p_2] \cdot \\
 &\quad [T_{21}f(x_2 | \lambda_1)p_1 + T_{22}f(x_2 | \lambda_2)p_2] \cdot \\
 &\quad [T_{31}f(x_3 | \lambda_1)p_1 + T_{32}f(x_3 | \lambda_2)p_2] \cdot \\
 &\quad [T_{41}f(x_4 | \lambda_1)p_1 + T_{42}f(x_4 | \lambda_2)p_2] \\
 &= [T_{11}f(x_1 | \lambda_1)p_1 + 0] \cdot [T_{21}f(x_2 | \lambda_1)p_1 + 0] \cdot [0 + T_{32}f(x_3 | \lambda_2)p_2] \cdot \\
 &\quad [0 + T_{42}f(x_4 | \lambda_2)p_2] \\
 &= \frac{e^{-\lambda_1} \lambda_1^{x_1}}{x_1!} p_1 \frac{e^{-\lambda_1} \lambda_1^{x_2}}{x_2!} p_1 \frac{e^{-\lambda_2} \lambda_2^{x_3}}{x_3!} p_2 \frac{e^{-\lambda_2} \lambda_2^{x_4}}{x_4!} p_2 \\
 &\quad (T_{11} = T_{21} = T_{32} = T_{42} = 1) \\
 &= \prod_{i=1}^n \prod_{k=1}^K T_{ik} p_k f_k(x_i | \lambda_k) \\
 &= \prod_{i=1}^n \prod_{k=1}^K (p_k f_k(x_i | \lambda_k))^{T_{ik}} \tag{4.20}
 \end{aligned}$$

Persamaan (4.10) identik dengan persamaan (4.19) yang kemudian disebut sebagai fungsi likelihood lengkap.

### **Finite Mixture Poisson**

Distribusi *mixture* Poisson banyak digunakan dalam berbagai disiplin ilmu untuk memodelkan data dimana setiap pengamatan diasumsikan berasal dari

sejumlah distribusi Poisson dengan parameter yang berbeda. Pada bagian ini disajikan model *mixture* Poisson menggunakan formulasi *missing data*.

Misalkan  $X$  adalah data banyak kejadian dalam interval waktu atau area, dengan  $\lambda$  rata-rata banyak kejadian per interval waktu atau area. Fungsi probabilitas hasil pengamatan  $x$  merupakan fungsi densitas Poisson dinyatakan sebagai berikut.

$$P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0, 1, 2, \dots, n$$

Fungsi peluang  $k$ -finite mixture Poisson diberikan sesuai persamaan (2.9), yaitu

$$f_{mix}(x|\boldsymbol{\lambda}, \mathbf{p}) = \sum_{k=1}^K p_k \frac{e^{-\lambda_k} \lambda_k^x}{x!}$$

dimana  $\mathbf{p} = (p_1, p_2, \dots, p_K)$ ,  $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_K)$  untuk nilai peluang  $p_k > 0$ ,  $k = 1, \dots, K$ , dengan  $\sum_{k=1}^K p_k = 1$ . Pada saat melakukan identifikasi keanggotaan *finite mixture*, diasumsikan bahwa parameter  $\lambda_1 < \lambda_2 < \dots < \lambda_K$ . Komponen pertama memiliki nilai  $\lambda$  yang lebih kecil, sehingga pada proses identifikasi keanggotaan *finite mixture* perlu dilakukan langkah mengurutkan data pengamatan dari nilai yang paling kecil hingga terbesar. *Finite mixture* yang dinyatakan sebagai fungsi berikut  $f_{mix}(x) = \sum_{k=1}^K p_k f_k(x|\theta_k)$  dapat diidentifikasi jika memiliki dua representasi berikut.

$$f_{mix}(x) = \sum_{k=1}^K p_k f_k(x|\theta_k) \text{ dan } f_{mix}'(x) = \sum_{k=1}^{K'} p'_k f_k(x|\theta'_k),$$

Kemudian  $f \equiv f'$  jika dan hanya jika  $k \equiv k'$  dan terdapat permutasi  $\pi$  dari indeks  $(1, 2, \dots, K)$  sehingga  $p_k = p'_{\pi_k}$  dan  $\theta_k = \theta'_{\pi_k}$  dan juga untuk setiap konstan  $i$  dan  $k = 1, 2, \dots, K$ , karena  $T_{ik}$  terdapat dua kemungkinan nilai yaitu 0 atau 1, maka  $f(T_{ik} = 0|p) = 1 - f(T_{ik} = 1|p) = 1 - p_k$ . untuk masing-masing pengamatan  $\mathbf{T}_i = (T_{ik})_{k=1}^K$ , diketahui

$$\mathbf{T}_i | p \sim \text{Multinomial}(1, p_1, p_2, \dots, p_K)$$

sehingga fungsi densitas indikator  $\mathbf{T}_i = (T_{ik})_{k=1}^K$  adalah

$$f(T_{i1}, \dots, T_{iK} | p_1, \dots, p_K) = \frac{1!}{T_{i1}! \dots T_{iK}!} \prod_{k=1}^K p_k^{T_{ik}} = \prod_{k=1}^K p_k^{T_{ik}} \quad (4.21)$$

pada persamaan 4.21 kemungkinan nilai  $T_{ik}$  adalah 0 atau 1 maka elemen  $\frac{1!}{T_{i1}! \dots T_{iK}!}$  sama dengan satu dan masing-masing  $T_i$  saling independen satu sama lain, sehingga *joint* densitas adalah

$$f(\mathbf{T}|\mathbf{p}) = \prod_{i=1}^n \prod_{k=1}^K p_k^{T_{ik}} \quad (4.22)$$

Misalkan  $X = X_1, X_2, \dots, X_n$  merupakan random sampel i.i.d. dari densitas *Mixture* Poisson. Likelihood densitas *mixture* adalah sebagai berikut.

$$L_{mix}(\mathbf{x}|\boldsymbol{\lambda}, \mathbf{p}) = \prod_{i=1}^n f_{mix}(x_i|\boldsymbol{\lambda}, \mathbf{p}) \quad (4.23)$$

Dan dengan menggunakan indikator  $T_i$  likelihood dapat dituliskan sebagai berikut.

$$\begin{aligned} L_{mix}(\mathbf{x}, \mathbf{T} | \boldsymbol{\lambda}, \mathbf{p}) &= f_{mix}(\mathbf{x}|\mathbf{T}, \boldsymbol{\lambda}, \mathbf{p}) f(\mathbf{T} | \mathbf{p}, \boldsymbol{\lambda}, \mathbf{x}) \\ &= f_{mix}(\mathbf{x}|\mathbf{T}, \boldsymbol{\lambda}) f(\mathbf{T} | \mathbf{p}) \\ &= \prod_{i=1}^n \prod_{k=1}^K \left( \frac{e^{-\lambda_k} \lambda_k^{x_i}}{x_i!} \right)^{T_{ik}} \prod_{i=1}^n \prod_{k=1}^K (p_k)^{T_{ik}} \\ &= \prod_{i=1}^n \prod_{k=1}^K \left( p_k \frac{e^{-\lambda_k} \lambda_k^{x_i}}{x_i!} \right)^{T_{ik}} \end{aligned} \quad (4.24)$$

### Densitas Prior

Sekarang akan dibuktikan bahwa *conjugate prior* untuk parameter Multinomial  $p = (p_1, \dots, p_K)$  adalah Dirichlet ( $\delta$ ). Misalkan vektor indikator tidak teramati  $\mathbf{T}_i = (T_{i1}, T_{i2}, \dots, T_{iK}), k = 1, 2, \dots, K$  adalah sampel i.i.d. Multinomial  $(1, p_1, \dots, p_K)$ , kemudian dengan fungsi densitas masing-masing  $\mathbf{T}_i, i = 1, \dots, n$  adalah sebagi berikut.

$$f(T_{i1}, \dots, T_{iK} | p_1, \dots, p_K) = \frac{\Gamma(\sum_{k=1}^K T_{ik} + 1)}{\prod_{k=1}^K \Gamma(T_{ik} + 1)} \prod_{k=1}^K p_k^{T_{ik}} \quad (4.25)$$

dimana  $T_{ik} \in \{0, 1\}$ , dan  $\sum_{k=1}^K T_{ik} = 1$ . Dapat diketahui bahwa terdapat satu hasil percobaan tunggal yang sama persis dengan hasil identifikasi keanggotaan dari *fixed finite*, dengan peluang  $p_1, \dots, p_K$  ( sehingga  $p_k \in (0, 1)$  untuk  $k = 1, \dots, K$  dan  $\sum_{k=1}^K p_k = 1$ ). Misalkan distribusi prior adalah Dirichlet ( $\delta$ ), dengan

parameter  $\delta_k = 1$ ,  $p \sim \text{Dir}(p_1, \dots, p_K, \delta_1, \dots, \delta_K)$ , maka densitas prior adalah sebagai berikut.

$$\begin{aligned}
 f(p_1, \dots, p_K | \delta_1, \dots, \delta_K) &= \frac{\Gamma(\sum_{k=1}^K \delta_k)}{\prod_{k=1}^K \Gamma(\delta_k)} \prod_{k=1}^K p_k^{\delta_k - 1} \\
 &= \frac{\Gamma(\sum_{k=1}^K 1)}{\prod_{k=1}^K \Gamma(1)} \prod_{k=1}^K p_k^0 \\
 &= \frac{\Gamma(K)}{\prod_{k=1}^K \Gamma(1)} \\
 &= \frac{(K-1)!}{\prod_{k=1}^K 1} = (K-1)!
 \end{aligned} \tag{4.26}$$

dimana  $p_k \in (0,1)$  dan  $\sum_{k=1}^K p_k = 1$ ,  $k \geq 2$ , dan  $\delta_k > 0$ .

Misalkan  $\mathbf{T}_i = (T_{i1}, T_{i2}, \dots, T_{iK})$ , kemudian fungsi likelihood adalah sebagai berikut.

$$L(\mathbf{T} | \mathbf{p}) = \prod_{i=1}^n \frac{\Gamma(\sum_{k=1}^K T_{ik} + 1)}{\prod_{k=1}^K \Gamma(T_{ik} + 1)} \prod_{k=1}^K p_k^{T_{ik}} \tag{4.27}$$

Densitas posterior adalah:

$$\begin{aligned}
 f(\mathbf{p} | \mathbf{T}) &\propto L(\mathbf{T} | \mathbf{p}) f(\mathbf{p}) \\
 &= \prod_{i=1}^n \frac{\Gamma(\sum_{k=1}^K T_{ik} + 1)}{\prod_{k=1}^K \Gamma(T_{ik} + 1)} \prod_{k=1}^K p_k^{T_{ik}} \frac{\Gamma(\sum_{k=1}^K \delta_k)}{\prod_{k=1}^K \Gamma(\delta_k)} \prod_{k=1}^K p_k^{\delta_k - 1} \\
 &\propto \prod_{i=1}^n \prod_{k=1}^K p_k^{T_{ik}} \prod_{k=1}^K p_k^{\delta_k - 1} \\
 &= \prod_{k=1}^K p_k^{\sum_{i=1}^n T_{ik}} \prod_{k=1}^K p_k^{\delta_k - 1} \\
 &= \prod_{k=1}^K p_k^{\sum_{i=1}^n T_{ik} + \delta_k - 1}
 \end{aligned} \tag{4.28}$$

Terlihat bahwa bentuk tersebut merupakan fungsi densitas Dirichlet ( $\sum_{i=1}^n T_{i1} + \delta_1, \dots, \sum_{i=1}^n T_{iK} + \delta_K$ ). Terlihat bahwa densitas posterior  $f(\mathbf{p} | \mathbf{T})$  merupakan satu keluarga distribusi dengan densitas prior  $f(\mathbf{p})$  dengan parameter berbeda, sehingga  $f(\mathbf{p})$  merupakan *conjugate* prior untuk vektor parameter  $\mathbf{p}$ . Untuk

parameter  $\lambda_k$ ,  $k = 1, \dots, K$ , densitas Gamma yang sering digunakan sebagai distribusi prior. Misalkan  $f(\lambda_k)$  menunjukkan densitas prior dari  $k$  parameter *mixture* Poisson kemudian

$$\lambda_k \sim \text{Gamma}(\alpha, \beta)$$

Kemudian akan dibuktikan untuk Poisson ( $\theta$ ) dengan parameter  $\theta$ . Misalkan  $X = X_1, X_2, \dots, X_n$  adalah sampel iid Poisson ( $\theta$ ), dan densitas prior Gamma ( $\alpha, \beta$ ) dengan densitas prior yaitu

$$f(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta}, \theta > 0, \alpha > 0, \beta > 0$$

Densitas likelihood adalah sebagai berikut.

$$\begin{aligned} L(\mathbf{x}|\theta) &= \prod_{i=1}^n f(x_i|\theta) \\ &= \prod_{i=1}^n \frac{1}{x_i!} e^{-\theta} \theta^{x_i} \\ &= \left( \frac{1}{x_1!} e^{-\theta} \theta^{x_1} \right) \left( \frac{1}{x_2!} e^{-\theta} \theta^{x_2} \right) \dots \left( \frac{1}{x_n!} e^{-\theta} \theta^{x_n} \right) \\ L(\mathbf{x}|\theta) &= \frac{1}{x_1! x_2! \dots x_n!} e^{-\theta} e^{-\theta} \dots e^{-\theta} \theta^{x_1} \dots \theta^{x_n} \\ &= \frac{1}{x_1! x_2! \dots x_n!} e^{-n\theta} \theta^{x_1+x_2+\dots+x_n} \\ &= \frac{1}{x_1! x_2! \dots x_n!} e^{-n\theta} \theta^{\sum_{i=1}^n x_i} \end{aligned} \quad (4.29)$$

Densitas posterior adalah

$$\begin{aligned} f(\theta|\mathbf{x}) &\propto L(\mathbf{x}|\theta) f(\theta) \\ &= \frac{1}{x_1! x_2! \dots x_n!} e^{-n\theta} \theta^{\sum_{i=1}^n x_i} \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta} \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{1}{x_1! x_2! \dots x_n!} \theta^{\sum_{i=1}^n x_i + \alpha - 1} e^{-n\theta - \beta\theta} \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{1}{x_1! x_2! \dots x_n!} \theta^{\sum_{i=1}^n x_i + \alpha - 1} e^{-(n+\beta)\theta} \\ f(\theta|\mathbf{x}) &\propto \theta^{\sum_{i=1}^n x_i + \alpha - 1} e^{-(n+\beta)\theta} \end{aligned} \quad (4.30)$$

Terbukti bahwa densitas posterior merupakan densitas distribusi Gamma dengan parameter  $\sum_{i=1}^n x_i + \alpha$  dan  $n + \beta$  sehingga

$$(\theta|\mathbf{x}) \sim \text{Gamma}\left(\sum_{i=1}^n x_i + \alpha, n + \beta\right)$$

Terlihat bahwa densitas posterior  $f(\theta|\mathbf{x})$  merupakan satu keluarga distribusi dengan densitas prior  $f(\theta)$  dengan parameter berbeda. Sehingga  $f(\theta)$  merupakan conjugate prior untuk  $\theta$ .

### Densitas Posterior

Dengan menggunakan independensi bersyarat pada seluruh variabel, dimana semua parameter saling independen maka *joint density function* dari seluruh variabel dapat ditulis sebagai berikut.

$$\begin{aligned} f_{mix}(\boldsymbol{\lambda}, \mathbf{p}, \mathbf{T}, \mathbf{x}) &= f_{mix}(\mathbf{x}, \mathbf{T} | \boldsymbol{\lambda}, \mathbf{p}) f(\boldsymbol{\lambda}, \mathbf{p}) \\ &= f_{mix}(\mathbf{x}, \mathbf{T}_i | \boldsymbol{\lambda}, \mathbf{p}) g(\boldsymbol{\lambda}) \pi(\mathbf{p}) \end{aligned} \quad (4.31)$$

( $\boldsymbol{\lambda}$  dan  $\mathbf{p}$  saling independen)

Dimana  $\mathbf{p} = (p_k)_{k=1}^K$ ,  $\mathbf{T}_i = (T_{ik})_{k=1}^K$ ,  $\boldsymbol{\lambda} = (\lambda_k)_{k=1}^K$ ,  $\mathbf{x} = (x_i)_{i=1}^n$ .

Dengan catatan bahwa parameter  $\lambda_k$  adalah saling independen, sehingga *joint density prior* untuk  $\boldsymbol{\lambda}$  diberikan oleh

$$g(\boldsymbol{\lambda}) = g(\lambda_1) \dots g(\lambda_K).$$

Dengan menggunakan teorema Bayes maka *joint density posterior* diberikan oleh

$$\begin{aligned} f_{mix}(\boldsymbol{\lambda}, \mathbf{p}, \mathbf{T} | \mathbf{x}) &= \frac{f(\boldsymbol{\lambda}, \mathbf{p}, \mathbf{T}, \mathbf{x})}{f(\mathbf{x})} \\ &\propto f(\boldsymbol{\lambda}, \mathbf{p}, \mathbf{T}, \mathbf{x}) \\ &= f(\mathbf{x}, \mathbf{T} | \boldsymbol{\lambda}, \mathbf{p}) g(\boldsymbol{\lambda}) \pi(\mathbf{p}) \\ &= f(\mathbf{x} | \boldsymbol{\lambda}, \mathbf{T}) f(\mathbf{T} | \mathbf{p}) g(\boldsymbol{\lambda}) \pi(\mathbf{p}) \end{aligned} \quad (4.32)$$

dengan melihat persamaan (4.24) dan persamaan (4.26) maka persamaan (4.32) dapat dituliskan sebagai berikut.

$$f_{mix}(\boldsymbol{\lambda}, \mathbf{p}, \mathbf{T} | \mathbf{x}) = \prod_{i=1}^n \prod_{k=1}^K \left( \frac{e^{-\lambda_k} \lambda_k^{x_i}}{x_i!} \right)^{T_{ik}} \prod_{i=1}^n \prod_{k=1}^K (p_k)^{T_{ik}} \times g(\lambda_1) \times \dots \times g(\lambda_K) \times (K-1)!$$

dimana  $g(\lambda_k) \sim \text{Gamma}(\alpha, \beta)$ , dengan parameter  $\alpha, \beta$  untuk masing-masing  $\lambda_k$ ,  $k = 1, \dots, K$ .

### Full Conditional Posterior Distributions

*Gibbs Sampler* adalah salah satu dari metode *Markov Chain Monte Carlo* (MCMC), di mana *full conditional posterior distribution* dari seluruh parameter diperlukan. Menggunakan likelihood, prior, dan *joint density posterior* diperoleh semua *full conditional posterior distribution* dengan mengabaikan semua konstanta yang tidak mengandung parameter. Perhatikan bahwa untuk *finite mixture Poisson* memiliki likelihood distribusi sebagai berikut.

$$L_{mix}(\mathbf{x}, \mathbf{T} | \boldsymbol{\lambda}, \mathbf{p}) = \prod_{i=1}^n \prod_{k=1}^K \left( p_k \frac{e^{-\lambda_k} \lambda_k^{x_i}}{x_i!} \right)^{T_{ik}}$$

dan distribusi prior adalah untuk masing-masing parameter  $\lambda_k$  dan  $p$  adalah sebagai berikut.

untuk  $\lambda_k$ ,  $k = 1, \dots, K$ ,

$$\lambda_k \sim \text{Gamma}(\alpha, \beta)$$

untuk proporsi  $p$ ,

$$p_k \sim \text{Dirichlet}(\delta, \dots, \delta), \delta = 1.$$

### Posterior $\lambda_k$

*Full conditional posterior distribution* untuk  $\lambda_k$  adalah

$$\begin{aligned} f(\lambda_k | \lambda_1, \dots, \lambda_{k-1}, \lambda_{k+1}, \dots, \lambda_K, \mathbf{p}, \mathbf{T}, \mathbf{x}) &\propto L(\mathbf{x} | \boldsymbol{\lambda}, \mathbf{T}) g(\lambda_k) \\ &= \left( \prod_{i=1}^n \prod_{k=1}^K \left( \frac{e^{-\lambda_k} \lambda_k^{x_i}}{x_i!} \right)^{T_{ik}} \right) \times \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda_k^{\alpha-1} e^{-\beta \lambda_k} \\ &= \prod_{i=1}^n \left( \frac{e^{-\lambda_k} \lambda_k^{x_i}}{x_i!} \right)^{T_{ik}} \times \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda_k^{\alpha-1} e^{-\beta \lambda_k} \\ &= \prod_{i=1}^n \left( \frac{1}{x_i!} \right)^{T_{ik}} (e^{-\lambda_k} \lambda_k^{x_i})^{T_{ik}} \times \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda_k^{\alpha-1} e^{-\beta \lambda_k} \\ &\propto \prod_{i=1}^n (e^{-\lambda_k} \lambda_k^{x_i})^{T_{ik}} \times \lambda_k^{\alpha-1} e^{-\beta \lambda_k} \\ &= e^{-\lambda_k \sum_{i=1}^n T_{ik}} \times \lambda_k^{\sum_{i=1}^n x_i T_{ik}} \times \lambda_k^{\alpha-1} e^{-\beta \lambda_k} \\ &= (e^{-\lambda_k \sum_{i=1}^n T_{ik}} \times e^{-\beta \lambda_k}) \times \left( \lambda_k^{\sum_{i=1}^n x_i T_{ik}} \times \lambda_k^{\alpha-1} \right) \end{aligned}$$

$$= (e^{-\lambda_k(\sum_{i=1}^n T_{ik} + \beta)}) \left( \lambda_k^{\sum_{i=1}^n x_i T_{ik} + \alpha - 1} \right) \quad (4.33)$$

Maka  $\lambda_k \sim \text{Gamma}(\sum_{i=1}^n x_i T_{ik} + \alpha, \sum_{i=1}^n T_{ik} + \beta)$ .

Sebagai catatan pada baris kedua proporsional karena

$$\prod_{k=1}^K \left( \frac{e^{-\lambda_k} \lambda_k^{x_i}}{x_i!} \right)^{T_{ik}} = \left( \frac{e^{-\lambda_1} \lambda_1^{x_i}}{x_i!} \right)^{T_{i1}} \times \dots \times \left( \frac{e^{-\lambda_{k-1}} \lambda_{k-1}^{x_i}}{x_i!} \right)^{T_{i(k-1)}} \times \left( \frac{e^{-\lambda_k} \lambda_k^{x_i}}{x_i!} \right)^{T_{ik}} \times$$

$$\left( \frac{e^{-\lambda_{k+1}} \lambda_{k+1}^{x_i}}{x_i!} \right)^{T_{i(k+1)}} \times \dots \times \left( \frac{e^{-\lambda_K} \lambda_K^{x_i}}{x_i!} \right)^{T_{iK}} \quad (\text{tidak mengandung } \lambda_k)$$

### Posterior $\mathbf{p}$

Full conditional posterior distribution untuk  $\mathbf{p}$  adalah

$$f(\mathbf{p}|\mathbf{T}, \boldsymbol{\lambda}, \mathbf{x}) \propto f(\mathbf{T}|\mathbf{p})\pi(\mathbf{p})$$

$$= \prod_{i=1}^n \prod_{k=1}^K (p_k)^{T_{ik}} (K-1)!$$

$$= \prod_{k=1}^K p_k^{\sum_{i=1}^n T_{ik}} (K-1)!$$

$$f(\mathbf{p}|\mathbf{T}, \boldsymbol{\lambda}, \mathbf{x}) \propto \prod_{k=1}^K p_k^{\sum_{i=1}^n T_{ik}}$$

$$= \prod_{k=1}^K p_k^{(\sum_{i=1}^n T_{ik} + 1) - 1} \quad (4.34)$$

Jelas bahwa

$$\mathbf{p} \sim \text{Dirichlet} \left( 1 + \sum_{i=1}^n T_{i1}, \dots, 1 + \sum_{i=1}^n T_{iK} \right).$$

### Posterior $\mathbf{T}_i$

Untuk masing-masing pengamatan  $x_i$ ,  $i = 1, \dots, n$  terdapat indikator  $\mathbf{T}_i$  sebagai berikut.

$$\mathbf{T}_i = (T_{i1}, T_{i2}, \dots, T_{iK}) = (T_{ik})_{k=1}^K$$

dimana untuk masing-masing  $T_{ik}$  memiliki dua nilai yaitu 1 atau 0, dan masing-masing  $\mathbf{T}_i$  hanya satu  $T_{ik}$  yang sama dengan 1 dan lainnya bernilai 0. Kemudian untuk semua  $i$ ,  $\sum_{k=1}^K T_{ik} = 1$ . Dengan menggunakan teorema Bayes untuk semua  $i = 1, \dots, n$ , dan untuk  $k = 1, \dots, K$ ,

$$\begin{aligned}
f(T_{ik} = 1|x_i, \boldsymbol{\lambda}, \mathbf{p}) &= \frac{f(x_i|\boldsymbol{\lambda}, \mathbf{p}, T_{ik} = 1)f(T_{ik} = 1|\boldsymbol{\lambda}, \mathbf{p}, x_i)}{\sum_{k=1}^K f(x_i|\boldsymbol{\lambda}, \mathbf{p}, T_{ik} = 1)f(T_{ik} = 1|\boldsymbol{\lambda}, \mathbf{p}, x_i)} \\
&= \frac{f(x_i|\boldsymbol{\lambda}, T_{ik} = 1)f(T_{ik} = 1|\mathbf{p})}{\sum_{k=1}^K f(x_i|\boldsymbol{\lambda}, T_{ik} = 1)f(T_{ik} = 1|\mathbf{p})} \\
&= \frac{f(x_i|\lambda_k)p_k}{\sum_{k=1}^K f(x_i|\lambda_k)p_k} \\
&= \frac{f(x_i|\lambda_k)p_k}{f(x_i)} \tag{4.35}
\end{aligned}$$

Karena untuk masing-masing  $T_{ik}$  memiliki dua nilai yaitu 1 atau 0 kemudian

$$f(T_{ik} = 0|x_i, \boldsymbol{\lambda}, \mathbf{p}) = 1 - f(T_{ik} = 1|x_i, \boldsymbol{\lambda}, \mathbf{p}) = 1 - \frac{f(x_i|\lambda_k)p_k}{f(x_i)}$$

maka distribusi indikator  $\mathbf{T}_i = (T_{ik})_{k=1}^K \sim \text{Multinomial}(1, p_{i1}, \dots, p_{iK})$ , untuk  $i = 1, \dots, n$ ,  $k = 1, \dots, K$ , dimana

$$p_{ik} = \frac{f(x_i|\lambda_k)p_k}{f(x_i)}, k = 1, \dots, K$$

### Update Gibbs untuk *fixed k*

Misalkan *Mixture Poissons* dimana terdapat  $k$  komponen *mixture*. Semua parameter *mixture Poisson* memiliki *full conditional densities* yang telah diketahui dan mudah untuk mengambil dilakukan pengambilan sampel. Oleh karena itu dapat dilakukan *update Gibbs* pada proses sampling di mana sampel berasal dari *full conditional densities*. Algoritma Gibbs untuk  $k$  komponen adalah

1. Pada saat  $t = 0$ , memberi nilai state awal  $(\lambda_1^0, \dots, \lambda_K^0, \mathbf{p}^0, \mathbf{T}_1^0, \dots, \mathbf{T}_n^0)$  dengan memperhatikan karakteristik datanya.
2. Untuk  $t = 1, 2, \dots, N$ , memperbarui nilai masing-masing variabel pada iterasi ke- $t$  :
  - a. *Update Gibbs* untuk  $\lambda_k, k = 1, \dots, K$ : membangkitkan sampel  $\lambda_k^t$  dari  $\text{gamma}(\alpha_k + \sum_{i=1}^n y_i T_{ik}, \sum_{i=1}^n T_{ik} + \beta_k)$  dengan menggunakan nilai  $T_{ik}$  paling baru.
  - b. *Update Gibbs* untuk proporsi  $\mathbf{p}$ : membangkitkan sampel  $\mathbf{p}^t$  dari  $\text{Dirichlet}(1 + \sum_{i=1}^n T_{i1}, \dots, 1 + \sum_{i=1}^n T_{iK})$  dengan nilai  $T_{i1}, \dots, T_{iK}$  paling baru.

- c. *Update* Gibbs untuk indikator  $T_i$ : membangkitkan sampel  $T_i^t$  dari Multinomial  $(1, p_{i1}, \dots, p_{iK})$   $i = 1, \dots, n, k = 1, \dots, K$ , dimana

$$p_{ik} = \frac{f(y_i|\lambda_k)p_k}{f(y_i)} \quad k = 1, \dots, K \quad (2.31)$$

menggunakan nilai  $\lambda_k$  dan  $\mathbf{p}$ .

- d. Sehingga diperoleh *state* Markov Chain  $(\lambda_1^t, \dots, \lambda_K^t, \mathbf{p}^t, T_1^t, \dots, T_n^t)$  yang digunakan untuk membangkitkan sampel pada iterasi  $t + 1$ .

3. Kembali ke langkah ke-2 sampai sebanyak  $N - 1$  kali iterasi untuk memperoleh Markov Chain dengan ukuran  $N$ .

#### 4.1.2 Pendekatan Bayesien pada Regresi *Mixture Poisson*

Regresi Poisson merupakan contoh GLM dengan distribusi dari respon  $Y$  dengan vektor kovariat  $\mathbf{x}$  memiliki pdf Poisson sebagai berikut.

$$(y, \lambda) = \frac{e^{-\lambda}\lambda^y}{y!}, \quad y = 0, 1, 2, \dots, n$$

dimana  $E(Y) = Var(Y) = \lambda$ , fungsi link  $g(\lambda) = \log(\lambda) = \boldsymbol{\beta}^T \mathbf{x}$ . *Finite mixture model* memungkinkan pemodelan yang fleksibel dari data heterogen karena menggabungkan kombinasi populasi heterogenitas baik diskrit maupun kontinyu. Vektor random  $\mathbf{y} = (y_1, y_2, \dots, y_N)'$  masuk kategori *finite mixture distribution* jika pdf  $p(\mathbf{y})$  dari distribusi mengikuti fungsi berikut.

$$f_{mix}(y_i) = \sum_{k=1}^K p_k f_k(y_i)$$

$$f_{mix}(\mathbf{y}|\boldsymbol{\lambda}) = p_1 f_1(\mathbf{y}|\lambda_1) + p_2 f_2(\mathbf{y}|\lambda_2) + \dots + p_K f_K(\mathbf{y}|\lambda_K) \quad (4.36)$$

dimana  $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_K)'$  parameter distribusi poisson dan  $\mathbf{p} = (p_1, p_2, \dots, p_K)'$  merupakan bobot masing-masing distribusi,  $k = 1, 2, \dots, n, K$  merupakan banyak komponen *mixture* dan untuk masing-masing  $k$ ,  $f_k(y_i)$  adalah pdf dan  $p_k$  adalah kuantitas non negatif dengan jumlah sama dengan satu,  $\sum_{k=1}^K p_k = 1$ . Fungsi densitas  $f(\cdot|\lambda_k)$  disebut sebagai fungsi distribusi untuk komponen  $k$ . Pada sebagian besar aplikasi, diasumsikan bahwa semua distribusi komponen muncul dari distribusi yang sama,  $f(\cdot|\lambda_k)$  dalam kasus ini adalah distribusi Poisson. Dengan formula *finite mixture model*, heterogenitas dalam data dapat dihitung

dengan dua acara. Pertama dengan menghitung heterogenitas populasi dengan memilih banyak indikator komponen  $T_i$  yang tidak teramati, masing-masing komponen dapat disebut sebagai sub-populasi. Indikator komponen  $T_i$  adalah representasi diskrit dari heterogenitas data, karena rata-rata tingkat kejadian diperkirakan dengan sejumlah komponen. Dalam hal ini, model *finite mixture* mengasumsikan bahwa ada lebih dari satu komponen dalam *dataset*. Jika fungsi densitas probabilitas tidak mengambil komponen *mixture*, estimasi parameter yang dihasilkan akan sangat tidak stabil dan tidak akurat. Dalam kasus tertentu, dimungkinkan untuk memilih model regresi tradisional (model regresi Poisson atau Binomial Negatif) yang tidak menjelaskan heterogenitas karena adanya subpopulasi yang berbeda. Ini dapat dilakukan dengan menetapkan banyak komponen  $k = 1$ . Cara kedua tergantung pada pemilihan komponen distribusi,  $f(\cdot|\lambda_k)$ , yang dapat mengakomodasikan heterogenitas pada masing-masing komponen. Untuk model Regresi *Mixture* Poisson, heterogenitas pada masing-masing komponen dihitung dengan memasukkan variabel kovariat pada fungsi rata-rata tingkat kejadian. *Finite Mixture* dari regresi Poisson sangat cocok untuk kasus populasi yang heterogen yang terdiri dari dua atau lebih komponen. Model Regresi *mixture* Poisson sebagai contoh GLMs dengan distribusi respon  $y$  dapat dinyatakan sebagai berikut.

$$f_{mix}(\mathbf{y}|\mathbf{X}, \Phi) = \sum_{k=1}^K p_k f_k(\mathbf{y}|\mathbf{X}, \boldsymbol{\beta}_k)$$

dimana

$$\begin{aligned} f_k(\mathbf{y}|\mathbf{X}, \boldsymbol{\beta}_k) &= \frac{e^{-\lambda_k} \lambda_k^y}{y!} \\ &= \frac{\exp(-\exp(\mathbf{X}^T \boldsymbol{\beta}_k)) (\exp(\mathbf{X}^T \boldsymbol{\beta}_k))^y}{y!} \end{aligned} \quad (4.37)$$

dimana  $\Phi = \{\boldsymbol{\beta}, \mathbf{p}\}$ ,  $\boldsymbol{\beta}_k = (\beta_{0k}, \beta_{1k}, \beta_{2k}, \dots, \beta_{pk})^T$ , dan  $\mathbf{p} = (p_1, p_2, \dots, p_K)$  adalah proporsi komponen *mixture*,  $k$  adalah banyak komponen *mixture*.  $\beta_{0k}$  adalah konstanta model pada komponen ke- $k$  dan  $\boldsymbol{\beta}_k$  adalah vektor koefisien model pada komponen ke- $k$  dengan ukuran  $n \times Q$ , dimana  $Q$  adalah banyak variabel kovariat.  $\mathbf{p} = (p_1, p_2, \dots, p_K)$ ,  $\lambda_k = \exp(\mathbf{X}^T \boldsymbol{\beta}_k)$ . Variabel  $y$  merupakan variabel respon yang

mengikuti distribusi keluarga eksponensial dengan syarat pada komponen  $k$ , ekspektasi bersyarat variabel respon dinyatakan dengan

$$E(y|\mathbf{X}, \Phi) = \sum_{k=1}^K p_k \lambda_k$$

Sehingga model *mixture* dengan asumsi terdapat dua komponen Poisson dapat dituliskan sebagai berikut.

$$\begin{aligned} \Pr(Y = y) &= p \frac{e^{-\lambda_1} \lambda_1^y}{y!} + (1 - p) \frac{e^{-\lambda_2} \lambda_2^y}{y!} \quad y = 0, 1, 2, \dots \\ \lambda_i &= p_1 \left[ \exp(\beta_{01} + \beta_{11}x_{1i_1} + \dots + \beta_{p_1}x_{pi_1}) \right] + \\ &\quad p_2 \left[ \exp(\beta_{02} + \beta_{12}x_{1i_2} + \dots + \beta_{p_2}x_{pi_2}) \right] \end{aligned} \quad (4.38)$$

Dalam regresi, heterogenitas terjadi jika variabel kovariat yang penting dimasukkan dalam kumpulan data atau dalam beberapa hal terdapat komponen penting yang tidak diperhitungkan dalam analisis, yang mengarah pada bias estimasi parameter. Model Regresi *Mixture* mengakomodasi perbedaan akibat masalah heterogenitas dengan memasukkan pengamatan yang homogen ke dalam kelompok yang sama. Perbedaan di antara model-model regresi *mixture* terutama tergantung pada bagaimana probabilitas posterior dihitung untuk variabel indikator. Telah ditunjukkan bahwa regresi *mixture* dengan variabel concomitant ditentukan join distribusi dengan aturan jenis regresi logistik dan menentukan conditional distribusi dari respon untuk masing-masing model.

Melonggarkan asumsi normalitas dan homogenitas dalam distribusi parameter sangat mungkin dengan menggunakan model regresi *mixture* terutama menggunakan formulasi GLMs seperti yang dipelajari oleh Grun dan Leisch (2008). Pada model finite *mixture* regresi, terdapat  $k$  komponen dimana masing-masing komponen mengikuti distribusi parameter tertentu. Setiap komponen memiliki prior bobot dan probabilitas untuk pengamatan yang berasal dari nilai komponen awal dan *mixture* distribusi diberikan oleh jumlah terboboti  $k$  komponen. Bobot dapat tergantung pada variabel lain yang tidak diamati yang disebut sebagai variabel concomitant.

Model Regresi *Mixture* Poisson dikembangkan berdasarkan *Mixture of Generalized Linear Model* yang disebut *Mixture of Generalized Linear*

*Regression Model.* Dalam kerangka kerja *Generalized Linier Model*, variabel acak  $Y_i$  yang dinamai sebagai variabel dependen, memiliki hubungan linier dengan kovariat  $X_1, X_2, \dots, X_p$  sebagai berikut.

$$\eta_i = g(\mu_i) = g(E(Y_i|X)) = \sum_{j=1}^p \beta_j X_{ij},$$

dimana  $\eta$  adalah prediktor linier,  $g(\cdot)$  adalah fungsi link,  $\mu_i$  adalah nilai ekspektasi dari variabel acak  $Y_i$  dan  $\beta$  adalah parameter regresi. Hubungan linear pada persamaan di atas memungkinkan distribusi variabel dependen diasumsikan menjadi bentuk distribusi keluarga eksponensial (yaitu Gaussian, Poisson, Gamma, atau Bernoulli). Fungsi link yang berbeda dapat digunakan untuk menggambarkan hubungan tersebut. Fungsi link kanonik adalah salah satu fungsi penghubung yang ditentukan oleh fungsi densitas eksponensial. Untuk distribusi Poisson, fungsi link kanonik adalah fungsi logit yang dapat didefinisikan sebagai

$$g(\mu_i) = \log(\mu_i) = \sum_{j=1}^p \beta_j X_{ij} = \mathbf{x}_i^T \boldsymbol{\beta}$$

sehingga fungsi *mixture* dapat ditulis sebagai berikut.

$$f(\mathbf{Y}|\mathbf{T}, \mathbf{p}, \mathbf{X}, \boldsymbol{\beta}) = \sum_{k=1}^K p_k f_k(\mathbf{Y}|\mathbf{X}, \boldsymbol{\beta}_k), \quad (4.39)$$

dimana

$$f_k(\mathbf{Y}|\mathbf{X}, \boldsymbol{\beta}_k) \sim \text{Poisson}(\log_k(\boldsymbol{\mu})),$$

Rata-rata  $f_k(\mathbf{Y}|\mathbf{X}, \boldsymbol{\beta}_k)$  merupakan distribusi Poisson dengan parameter  $\logit_k(\boldsymbol{\mu})$ ,

$\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_n)'$ ,  $\mathbf{X}' = (\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n)'$ , dan  $\boldsymbol{\beta}_k = (\beta_1, \beta_2, \dots, \beta_p)'$ .

### Densitas Prior

Sebagaimana pada *finite mixture Poisson conjugate prior* untuk parameter Multinomial  $\mathbf{p} = (p_1, \dots, p_K)$  adalah Dirichlet ( $\delta$ ). Misalkan vektor indikator  $\mathbf{T}_i = T_{i1}, T_{i2}, \dots, T_{iK}, k = 1, 2, \dots, K$  adalah sampel i.i.d. Multinomial  $(1, p_1, \dots, p_K)$ , kemudian dengan fungsi densitas masing-masing  $\mathbf{T}_i, i = 1, \dots, n$  adalah sebagai berikut.

$$f(T_{i1}, \dots, T_{iK} | p_1, \dots, p_K) = \frac{\Gamma(\sum_{k=1}^K T_{ij} + 1)}{\prod_{k=1}^K \Gamma(T_{ik} + 1)} \prod_{k=1}^K p_k^{T_{ik}} \quad (4.40)$$

dimana  $T_{ik} \in \{0,1\}$ , dan  $\sum_{k=1}^K T_{ik} = 1$ .

Prior untuk vektor parameter  $\boldsymbol{\beta}$  dinotasikan sebagai  $f(\beta_{pk})$  dimana  $p = 1,2,3,4$ ,  $k = 1,2$  sebanyak parameter regresi Poisson.  $f(\beta_{pk})$  berdistribusi normal yang merupakan *pseudo prior*, dengan  $\beta_{pk} \sim N(\mu_{\beta_{pk}}, \sigma_{\beta_{pk}}^2)$  memiliki fungsi densitas sebagai berikut.

$$f(\beta_{pk}) = \frac{1}{\sqrt{2\pi\sigma_{\beta_{pk}}^2}} \exp\left[-\frac{1}{2\sigma_{\beta_{pk}}^2} (\beta_{pk} - \mu_{\beta_{pk}})^2\right] \quad (4.41)$$

untuk semua parameter  $\beta_{pk}$ ,  $p = 1,2,3,4$ ,  $k = 1,2$  saling independen.

### Bayesian MCMC

Misalkan  $\boldsymbol{\Theta} = (\Theta_1, \Theta_2, \dots, \Theta_d)' = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_K, \mathbf{p})'$  menyatakan semua parameter pada regresi *mixture* Poisson. Dengan menggunakan teorema Bayes, distribusi probabilitas  $f(\boldsymbol{\Theta} | \mathbf{Y}, \mathbf{T}, \mathbf{X})$  dapat dinyatakan sebagai berikut.

$$\begin{aligned} f(\boldsymbol{\Theta} | \mathbf{Y}, \mathbf{T}, \mathbf{X}) &= \frac{f(\mathbf{Y} | \boldsymbol{\Theta}, \mathbf{T}, \mathbf{X}) p(\boldsymbol{\Theta})}{f(\mathbf{Y})} \\ &\propto f(\mathbf{Y} | \boldsymbol{\Theta}, \mathbf{T}, \mathbf{X}) p(\boldsymbol{\Theta}) \end{aligned}$$

dimana  $p(\boldsymbol{\Theta})$  adalah distribusi prior dari  $\boldsymbol{\Theta}$  dan  $f(\mathbf{Y} | \boldsymbol{\Theta}, \mathbf{T}, \mathbf{X})$  adalah likelihood *mixture*. Dalam pendekatan inferensi Bayesian, proses estimasi parameter dilakukan dengan mengintegrasikan distribusi posterior. Integrasi dapat dilakukan secara numerik dengan prosedur simulasi yang umumnya dikenal sebagai metode *Markov Chain Monte Carlo* (MCMC). Umumnya metode *Markov Chain Monte Carlo* bekerja dengan langkah-langkah berikut.

1. Menetapkan nilai awal untuk parameter regresi *mixture* Poisson:

$\Phi^{(t)} = (\mathbf{p}, \boldsymbol{\beta})^{(t)}$ , untuk parameter  $\mathbf{p} = (p_1, p_2, \dots, p_K)$ ,  $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_K)$ ,  
 $\boldsymbol{\beta}_1 = (\beta_{01}, \beta_{11}, \dots, \beta_{p1})$ ,  $\boldsymbol{\beta}_2 = (\beta_{02}, \beta_{12}, \dots, \beta_{p2})$ , ...,  $\boldsymbol{\beta}_K = (\beta_{0K}, \beta_{1K}, \dots, \beta_{pK})$ ,  
pada iterasi  $t = 0$ .

2. Membangkitkan komponen parameter untuk masing-masing *mixture*

- a. Membangkitkan  $\mathbf{p}^{t+1}$  dari *full conditional posterior distribution* pada persamaan (2.29).

$$f(\mathbf{p} | y, \boldsymbol{\beta}_1^{(t)}, \boldsymbol{\beta}_2^{(t)}, \dots, \boldsymbol{\beta}_K^{(t)})$$

- b. Membangkitkan  $\boldsymbol{\beta}_1^{(t+1)}$  dari *full conditional posterior distribution* pada persamaan (2.30).

$$f(\boldsymbol{\beta}_1 | y, \mathbf{p}^{t+1}, \boldsymbol{\beta}_2^{(t)}, \boldsymbol{\beta}_3^{(t)}, \dots, \boldsymbol{\beta}_K^{(t)})$$

- c. Membangkitkan  $\boldsymbol{\beta}_2^{(t+1)}$  dari *full conditional posterior distribution*

$$f(\boldsymbol{\beta}_2 | y, \mathbf{p}^{t+1}, \boldsymbol{\beta}_1^{(t+1)}, \boldsymbol{\beta}_3^{(t)}, \dots, \boldsymbol{\beta}_K^{(t)})$$

⋮

- d. Membangkitkan  $\boldsymbol{\beta}_k^{(t+1)}$  dari *full conditional posterior distribution*

$$f(\boldsymbol{\beta}_k | y, \mathbf{p}^{t+1}, \boldsymbol{\beta}_1^{(t+1)}, \boldsymbol{\beta}_2^{(t+1)}, \dots, \boldsymbol{\beta}_{k-1}^{(t+1)}, \boldsymbol{\beta}_{k+1}^{(t)}, \dots, \boldsymbol{\beta}_K^{(t)})$$

⋮

- e. Membangkitkan  $\boldsymbol{\beta}_K^{(t+1)}$  dari *full conditional posterior distribution*

$$f(\boldsymbol{\beta}_K | y, \mathbf{p}^{t+1}, \boldsymbol{\beta}_1^{(t+1)}, \boldsymbol{\beta}_2^{(t)}, \dots, \boldsymbol{\beta}_{K-1}^{(t+1)}).$$

3. Meningkatkan iterasi  $t$ , dimana iterasi baru  $t = t + 1$
4. Mengulangi langkah 2 dan langkah 3 hingga sebanyak  $M$  kali, dimana  $M \rightarrow \infty$  atau semua parameter telah mencapai konvergen.

Untuk regresi *Mixture Poisson*, digunakan distribusi Normal sebagai *conjugate prior* untuk semua parameter  $\boldsymbol{\beta}$  dan distribusi Dirichlet sebagai *conjugate prior* untuk semua parameter proporsi  $\mathbf{p}$ . Parameter hyper-prior dari  $\boldsymbol{\beta}$  dan  $\mathbf{p}$  ditetapkan dari pseudo-prior setelah data dimodelkan menggunakan GLM frekuentif.

## 4.2 Eksplorasi Data FKTP dan Variabel Kovariat

FKTP adalah fasilitas kesehatan tingkat pertama termasuk Puskesmas dan Klinik Pratama. Layanan kesehatan yang disediakan oleh FKTP termasuk dokter umum, layanan kesehatan ibu dan anak, layanan gigi, dll. Karena fungsinya sebagai layanan kesehatan tingkat pertama, FKTP idealnya harus dibangun di setiap kecamatan. Jika pasien memerlukan perawatan lanjutan, dokter di

Puskesmas atau Klinik Pratama harus merujuk pasien ke rumah sakit yang lebih besar.

Distribusi lokasi di Surabaya dapat dianggap sebagai *point pattern*. Analisis *point pattern* ini dapat berguna untuk mengetahui bagaimana lokasi FKTP tersebar di Surabaya. Karena distribusi FKTP tidak terdistribusi secara merata di semua wilayah kecamatan, maka dilakukan analisis spasial *point pattern* yang berfokus pada intensitas pola titik di wilayah tersebut telah ditentukan. Intensitas menunjukkan rasio jumlah titik di suatu daerah dengan ukuran populasi. Dalam analisis yang melibatkan wilayah administratif dengan batas regional tertentu, objek *tessellation* sering digunakan. Alasan utama untuk menggunakan objek *tessellation* adalah mekanisme tampilan komputasi yang dapat mengubah area menjadi grid yang memiliki ukuran yang sama dengan wilayah. Sehingga variabel yang diamati dapat menggambarkan kondisi masing-masing daerah (kecamatan). Manfaat utama dari analisis ini adalah bahwa hasilnya memberikan informasi tentang kecamatan yang membutuhkan lebih banyak FKTP sehingga pemerintah sebagai pembuat kebijakan memiliki informasi yang cukup untuk mengambil tindakan dalam menentukan lokasi FKTP baru atau mendorong penambahan FKTP di suatu kecamatan ke depannya.

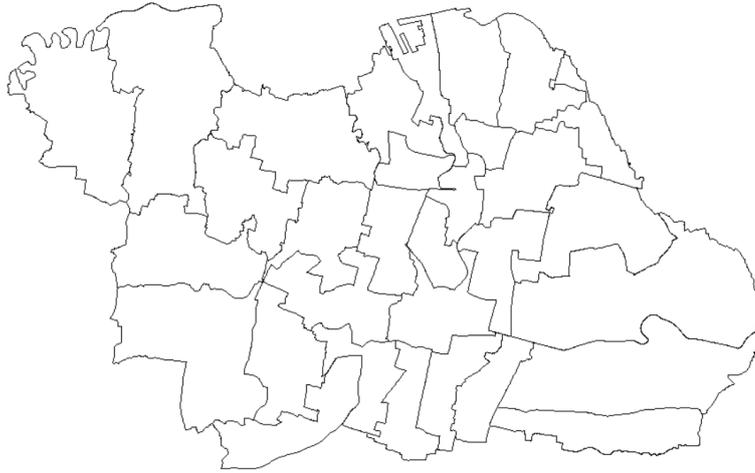
Berdasarkan data Dinas Kesehatan Kota Surabaya pada tahun 2108, jumlah puskesmas di Surabaya berjumlah 63 puskesmas dan 155 Klinik Pratama yang tersebar pada 31 kecamatan di Kota Surabaya. Analisis SPpT tentang FKTP di Surabaya dimulai dengan mentransformasikan data FKTP ke dalam bentuk *point pattern* berdasarkan lokasinya. Wilayah studi dibagi menjadi 31 objek *tessellation* sesuai dengan jumlah kecamatan di Surabaya. Karakteristik dari data lokasi FKTP dapat dilihat secara visual melalui *spatial point pattern*.

Dengan menggunakan bantuan software R dilakukan eksplorasi data *spatial point pattern* dari titik lokasi FKTP di Surabaya. Untuk memetakan titik lokasi FKTP di program R diperlukan peta Kota Surabaya dalam bentuk Data Shapefile dengan package *maptools* dan *sp*. *Shapefile* mewakili objek spasial dalam bentuk daftar garis, daftar titik, atau daftar wilayah *polygonal*, dan setiap objek dalam daftar dapat memiliki variabel tambahan yang melekat. *Dataset* yang disimpan dalam format file disimpan dalam kumpulan file teks, misalnya Surabaya.shp,

Surabaya.prj, Surabaya.sbn, Surabaya.dbf, yang semuanya memiliki nama file yang sama “Surabaya” tetapi ekstensi file yang berbeda. Sedangkan untuk melakukan analisis *spatial point patterns* didominasi dengan menggunakan package spatstat. Package spatstat sendiri dependen dengan package spatstat.data, nlme, rpart, foreign, dan rgdal. Package spatstat menyediakan tipe data (kelas) khusus untuk data spasial. Seperti *point pattern*, *spatial windows*, gambar pixel, pola garis, *tessellation*, dan jaringan linear. Fasilitas dalam package spatstat termasuk menangani *dataset* spasial meliputi pembuatan *dataset*, *query data* dan modifikasi (subset, pemisahan/penggabungan, dll), pencetakan, dan meringkas data. Banyak perhitungan geometris dan matematika dapat dilakukan, termasuk pengukuran jarak, transformasi jarak, *nearest neighbours*, pengukuran (pengukuran luas dan Panjang), dll. Operasi geometris yang tersedia termasuk *shift*, rotasi, *rescaling*, *smoothing transformation*, *union*, *convex hull*, dan operasi morfologi (pelebaran, erosi, penutupan, pembukaan). Untuk setiap tipe data spasial, spatstat menyediakan metode plot yang dirancang untuk menghasilkan tampilan grafis yang sesuai. Berbagai prosedur alternative plot, disesuaikan dengan tujuan statistik yang berbeda. Sebagai tambahan gambar pixel dapat diplot sebagai gambar warna, plot kontur, tampilan perspektif permukaan, histogram, atau fungsi distribusi kumulatif. Paket spatstat mendukung berbagai teknik untuk analisis data eksplorasi dan analisis nonparametrik pola titik spasial. Teknik-teknik tersebut termasuk metode klasik seperti penataan ruang dan *quadrat counting* (Baddeley, 2016).

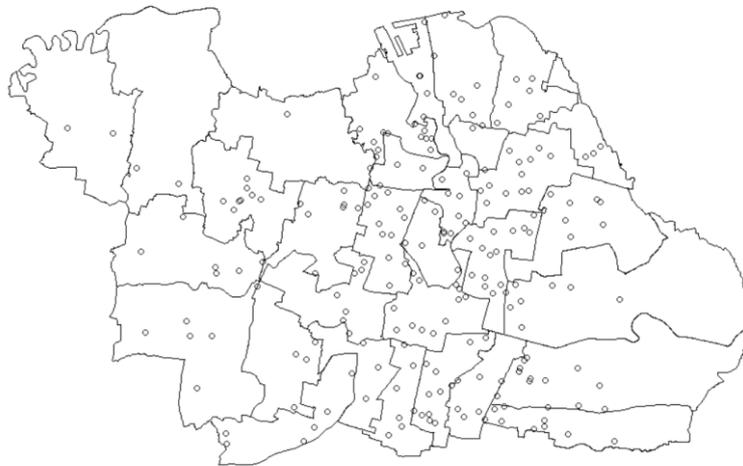
Langkah untuk melakukan eksplorasi data menggunakan *package spatstat* dilakukan dengan *syntax* terlampir pada **Lampiran 4**. Setiap wilayah yang digambarkan pada map (menggunakan grafik vektor) dapat ditampilkan sebagai *polygonal window*. Jendela-jendela semacam itu biasanya digunakan untuk mewakili batas-batas nasional atau wilayah administrasi, seperti wilayah Kota Surabaya. Gambar 4.1 menunjukkan objek *tessellation* kota Surabaya. Untuk membuat obyek *tessellation* dari *list windows* pada Gambar 4.1, dapat dilakukan dengan *tess(tiles=windows)* dimana *windows* adalah *list* obyek dengan kelas “*owin*”. Jendela tidak boleh tumpang tindih. Perintah ini biasanya digunakan ketika wilayah studi dibagi menjadi wilayah administratif (negara bagian, area

kodepos, provinsi/kabupaten/kecamatan) dan batas-batas masing-masing sub-wilayah disediakan oleh file data GIS.

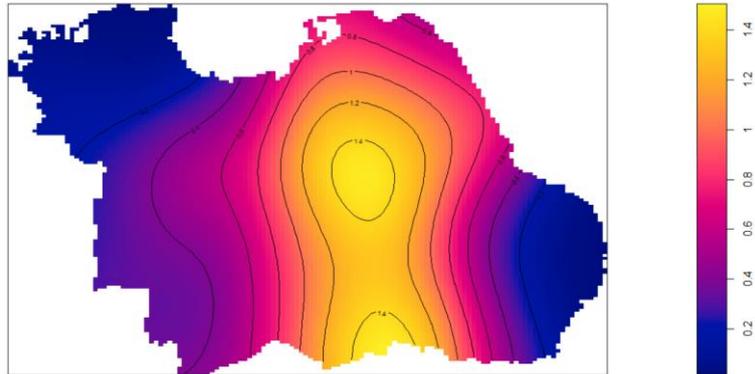


**Gambar 4.1** Obyek *Tessellation* Kota Surabaya

Gambar 4.1 menunjukkan *window* yang akan digunakan untuk melakukan pengamatan *point pattern* berupa lokasi FKTP di Surabaya. Objek *window* telah dibagi menjadi objek *tessellation* sesuai batas masing-masing kecamatan. Dengan menggunakan metode plot untuk pola titik, `plot.ppp`, menampilkan jendela pengamatan untuk pola dan menentukan lokasi spasial titik-titik FKTP pada peta Surabaya menggunakan skala yang sama untuk koordinat x dan y.

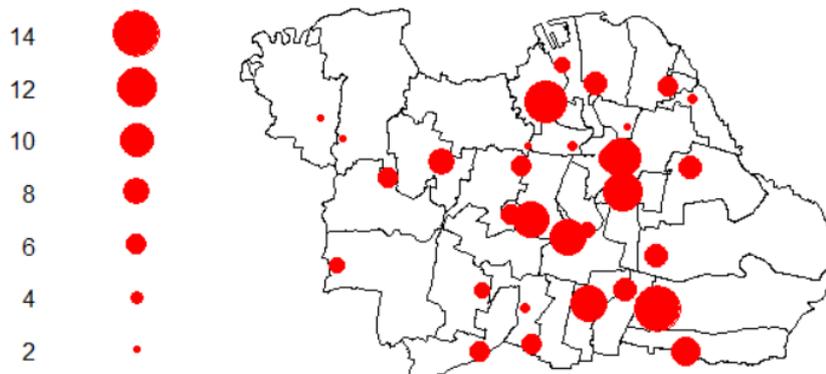


**Gambar 4.2:** *Planar Point Pattern* lokasi FKTP



**Gambar 4.3:** Intensitas dan *Contour* dari lokasi FKTP

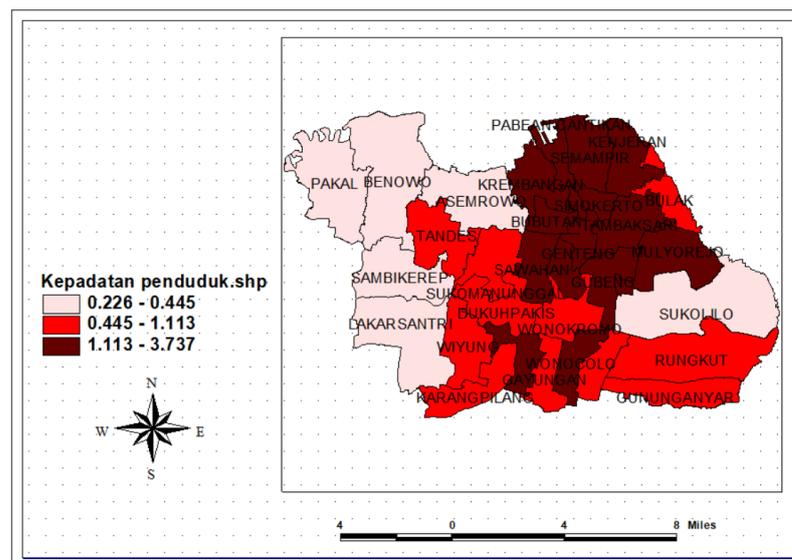
Gambar 4.2 merupakan visualisasi dari data lokasi FKTP dengan menggunakan *spatial point pattern* di Surabaya. Intensitas FKTP yang menunjukkan rasio FKTP per area pada Gambar 4.3 menjelaskan bahwa pengembangan FKTP cenderung bergerak ke arah pusat Kota Surabaya. *Contour* yang terbentuk menunjukkan terdapat 2 (dua) puncak kepadatan dari persebaran lokasi FKTP, hal ini merupakan indikasi tidak homogenya persebaran lokasi FKTP. Secara sederhana banyak titik pada masing-masing *list tessellation* dalam hal ini kecamatan dapat dihitung dengan *quadrat counting*. Pada *quadrat counting*, pengamatan pada window  $W$  dibagi menjadi subregion yang disebut *quadrat*. Kemudian dihitung banyak titik yang jatuh pada masing-masing subregion. Hasil *output quadrat counting* menunjukkan banyak FKTP pada masing-masing kecamatan. Proses penghitungan ini merupakan sebuah *Poisson Point Process*.



**Gambar 4.4:** Jumlah FKTP masing-masing kecamatan di Surabaya

Gambar 4.4 menunjukkan bahwa titik merah menampilkan jumlah FKTP di setiap lokasi. Semakin besar titik lingkaran merah, semakin tinggi jumlah FKTP di suatu area. Lokasi FKTP di Surabaya masih banyak tersebar di Surabaya Pusat. Ini berarti bahwa kecamatan di Surabaya Pusat cenderung memiliki lebih banyak FKTP daripada daerah lain. Meskipun, jumlah FKTP telah memenuhi jumlah minimum fasilitas kesehatan di Surabaya, lokasi FKTP cenderung berkembang menuju kedekatan wilayah Surabaya Pusat. Hal ini menyebabkan jarak yang lebih jauh bagi orang yang tinggal di pinggiran kota Surabaya untuk mencapai FKTP.

Variabel-variabel yang mewakili setiap kecamatan menggambarkan kondisi sosial demografi, populasi, dan aksesibilitas kecamatan yang digambarkan oleh Gambar 4.5. Variabel total populasi menunjukkan jumlah orang yang harus mendapatkan layanan kesehatan. Oleh karena itu, semakin padat penduduk suatu daerah, semakin besar keperluan akan ketersediaan fasilitas kesehatan. Gambar 4.5 menggambarkan kondisi distribusi penduduk di kota Surabaya. Sebagian besar penduduk Surabaya tinggal di daerah yang dekat dengan Surabaya Pusat karena warnanya merah dan merah bata yang menunjukkan jumlah penduduk yang lebih tinggi. Semakin jauh dari pusat kota penduduk Surabaya semakin sedikit.

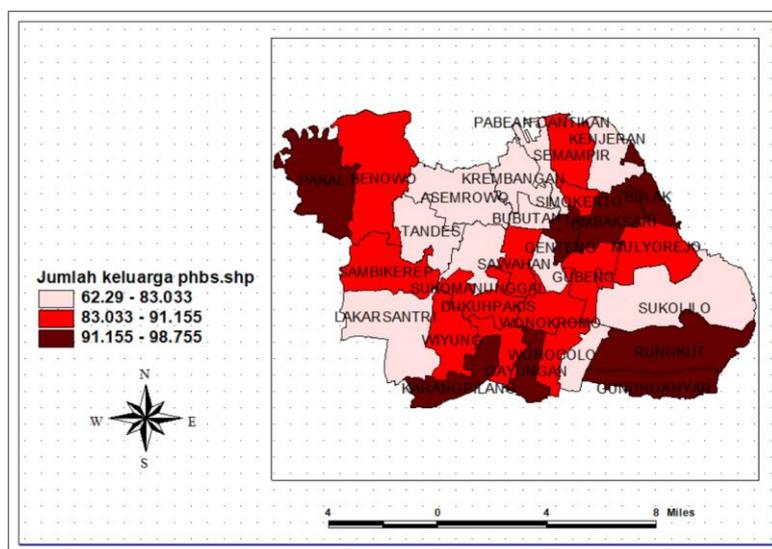


**Gambar 4.5:** Variabel Kovariat, Populasi di setiap kecamatan

Selanjutnya untuk variabel kovariat kedua yaitu variabel persentase rumah tangga yang berperilaku hidup bersih dan sehat merupakan hasil bagi antara jumlah rumah tangga yang berperilaku hidup bersih dan sehat terhadap jumlah rumah

tangga yang dipantau di tiap kecamatan di Surabaya dikalikan dengan 100%. Kebiasaan rumah tangga hidup bersih dan sehat menunjukkan sikap dalam menjaga kebersihan dan kesehatan diri sendiri dan lingkungan sekitar. Dengan persentase jumlah rumah tangga yang berperilaku hidup bersih dan sehat yang tinggi diasumsikan masyarakat memiliki kesadaran untuk melakukan pemeriksaan rutin ke fasilitas kesehatan dalam rangka menjaga kesehatan sehingga tingkat kebutuhan akan fasilitas kesehatan semakin tinggi.

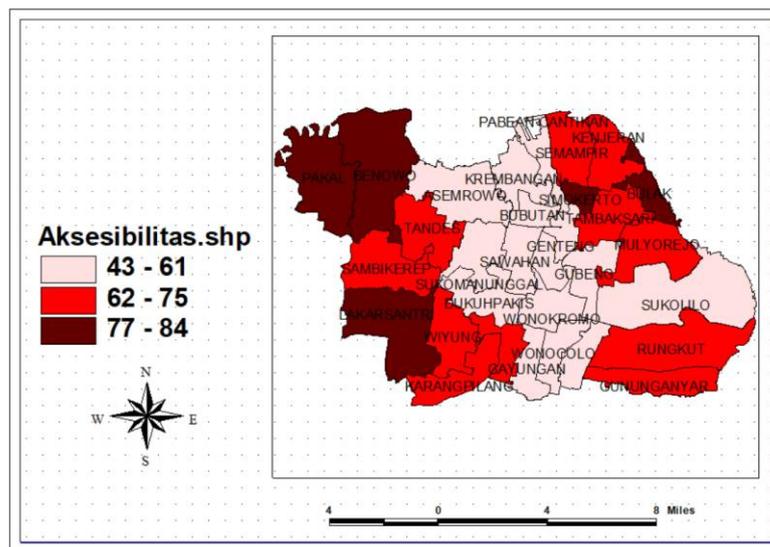
Gambar 4.6 menunjukkan distribusi rumah tangga bersih. Rumah tangga bersih mencerminkan tingkat kesadaran masyarakat akan kesehatan. Semakin tinggi tingkat kesadaran kesehatan masyarakat, semakin banyak orang melakukan pemeriksaan kesehatan secara teratur sehingga jumlah fasilitas kesehatan meningkat. Distribusi rumah tangga bersih relatif lebih merata di Surabaya.



**Gambar 4.6:** Variabel Kovariat Persentase jumlah rumah tangga bersih sehat

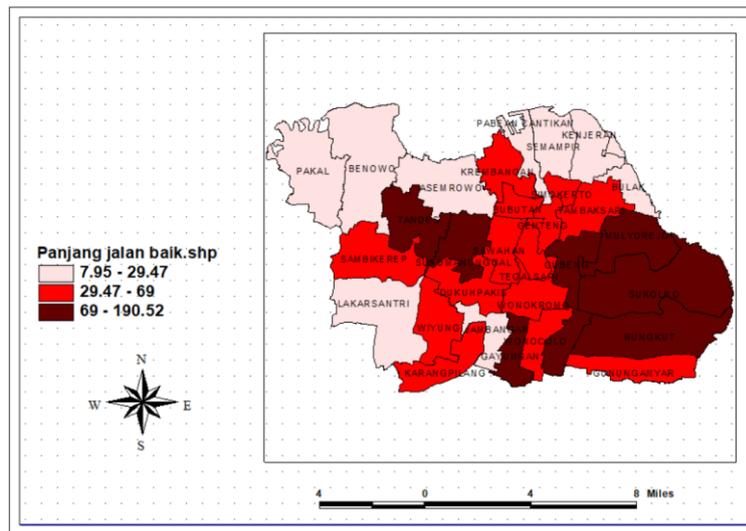
Tingkat akses di kecamatan ditunjukkan oleh indeks aksesibilitas yang digambarkan oleh angka keterkaitan. Indeks Aksesibilitas diperoleh dengan menggunakan parameter waktu tempuh (*Travel Time*) menggunakan sistem pengukuran secara tidak langsung. Data pengukuran di peroleh dengan menggunakan *software Google Maps*, dan metode pengukuran aksesibilitas serta penggambaran jaringan jalan berdasarkan wilayah kecamatan, yang terdiri dari 31 zona (Kecamatan) dimana kantor kecamatan sebagai pusat zona (Node) dan ruas jalan yang menghubungkannya sebagai rute (*Link*). Hasil pengukuran berupa angka

keterkaitan yang berarti jarak yang harus ditempuh dari satu kecamatan ke kecamatan lain yang diukur dengan waktu tempuh. Semakin tinggi tingkat hubungan berarti semakin rendah tingkat akses. Aksesibilitas yang tinggi di kecamatan akan mendorong pertumbuhan fasilitas publik karena kecamatan lebih mudah dijangkau. Tingkat aksesibilitas yang tinggi ada pada Kecamatan Wonokromo dengan angka keterkaitan (AK) sebesar 43 menit, sedangkan Kecamatan Lakarsantri mempunyai aksesibilitas paling rendah (Buruk) dengan angka keterkaitan (AK) sebesar 84 menit.



**Gambar 4.7:** Variabel Kovariat Angka Keterkaitan di setiap kecamatan

Gambar 4.7 menunjukkan bahwa Angka Keterkaitan di pusat-pusat wilayah Surabaya lebih rendah daripada daerah lain. Ini berarti bahwa pusat-pusat kota Surabaya memiliki tingkat akses yang lebih tinggi ke fasilitas kesehatan.

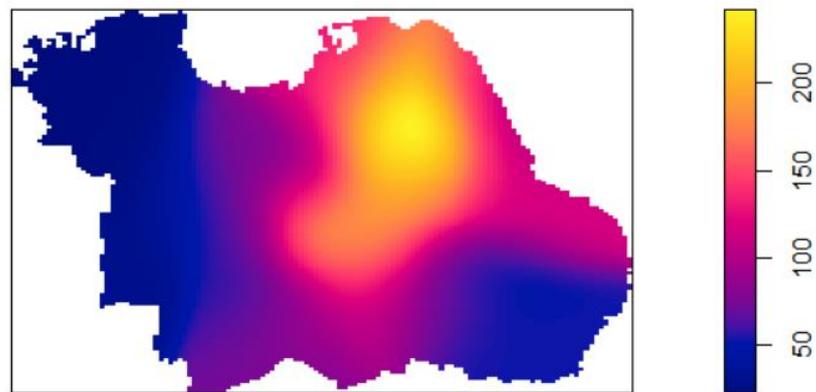


**Gambar 4.8:** Variabel Kovariat Panjang jalan dalam kondisi baik dalam km

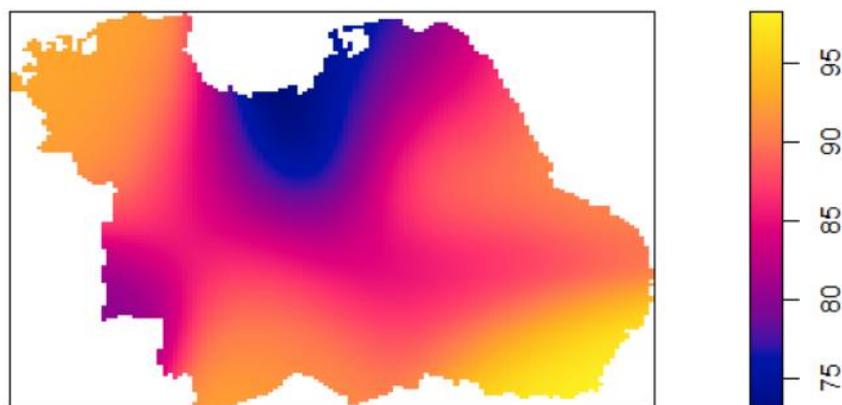
Gambar 4.8 menunjukkan variabel infrastruktur jalan dalam kondisi baik yang mencerminkan hasil pengembangan ekonomi suatu daerah, yaitu semakin tinggi ketersediaan jalan dalam kondisi baik akan cenderung meningkatkan pembangunan atau peningkatan fasilitas publik lainnya. Variabel jalan yang digunakan pada penelitian ini yaitu jalan besar atau *main road* yang menghubungkan satu kecamatan dengan kecamatan yang lain. Berdasarkan Gambar 4.8 terlihat Surabaya Timur memiliki jalan dengan kondisi baik paling panjang selanjutnya Surabaya Pusat memiliki jalan baik relatif lebih Panjang dibandingkan dengan panjang jalan kondisi baik di kecamatan Pakal dan Benowo.

Selanjutnya akan dilakukan pemodelan intensitas FKTP di Surabaya dengan variabel kovariat yang telah tersedia. Sebelum dilakukan pemodelan dilakukan transformasi *pixel* dari masing-masing variabel kovariat. Variabel kovariat yang bersumber dari BPS dan Tugas Akhir Mursalim (2018) dilihat sebagai hasil pengukuran di pusat kecamatan dalam hal ini kantor kecamatan, sehingga titik koordinat diambil dari pusat masing-masing kecamatan. Tujuan dilakukan transformasi ini adalah untuk melakukan interpolasi hasil pengukuran variabel kovariat pada titik-titik yang tidak dapat diperoleh hasil pengukurannya. Dengan dilakukan transformasi ini ditujukan agar setiap titik pada lingkup kecamatan dapat terwakili nilai pengukuran variabel kovariat. Untuk variabel kepadatan penduduk hasil transformasi ditunjukkan pada Gambar 4.9. Apabila

dibandingkan dengan Gambar 4.5, terlihat bahwa warna pada Gambar 4.5 menunjukkan pengukuran dalam ukuran kategori sedangkan hasil transformasi menunjukkan seluruh nilai hasil interpolasi pada setiap titik di seluruh wilayah Surabaya.

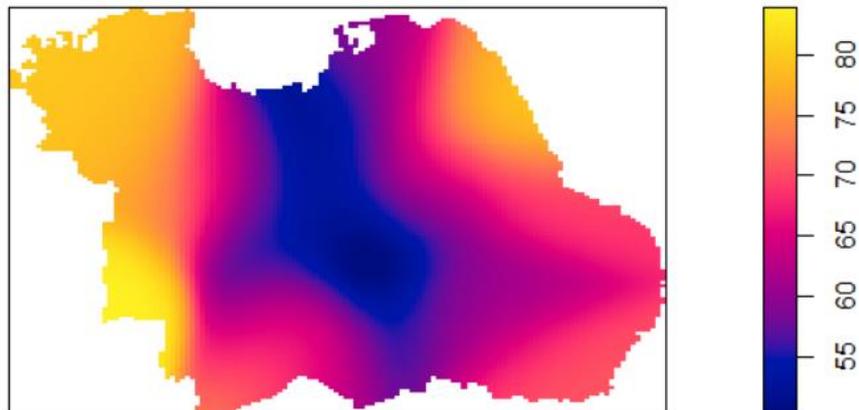


**Gambar 4.9:** Tranformasi Pixel Variabel Kovariat Populasi di setiap kecamatan  
Transformasi *pixel* pada variabel Persentase Rumah Tangga dengan Perilaku Hidup Bersih Sehat adalah sebagai berikut.



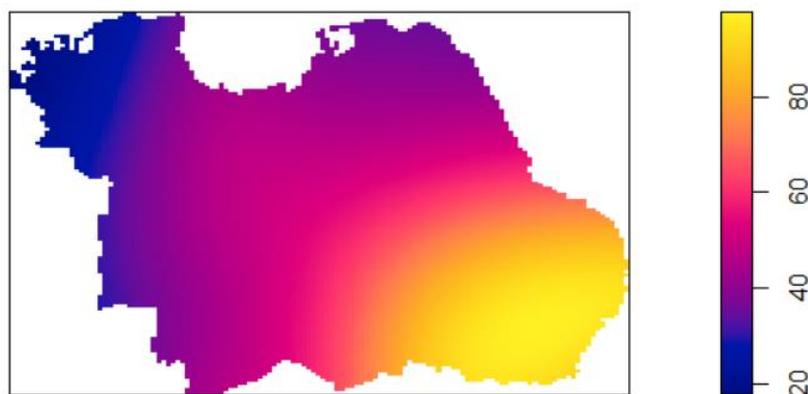
**Gambar 4.10:** Tranformasi Pixel Variabel Kovariat Persentase jumlah rumah tangga bersih dan sehat

Gambar 4.10 semakin menunjukkan bahwa persentase jumlah rumah tanggaa bersih dan sehat di Surabaya relatif tidak berbeda jauh di seluruh wilayah Surabaya. Selanjutnya transformasi *pixel* pada variabel Angka Keterkaitan antar kecamatan ditunjukkan pada Gambar 4.11.



**Gambar 4.11:** Tranformasi Pixel Variabel Kovariat Angka Keterkaitan di setiap kecamatan

Gambar 4.11 menunjukkan bahwa Angka keterkaitan pada kecamatan di Surabaya bagian perbatasan terutama Surabaya Barat (Kecamatan Benowo dan Kecamatan Pakal) memiliki angka keterkaitan yang cukup tinggi. Hal ini menunjukkan bahwa waktu tempuh menuju kecamatan tersebut tinggi, sehingga aksesibilitasnya rendah. Selanjutnya untuk transformasi pixel pada variabel Panjang Jalan dengan kondisi Baik adalah sebagai berikut.



**Gambar 4.12:** Tranformasi Pixel Variabel Panjang jalan dalam kondisi baik dalam km

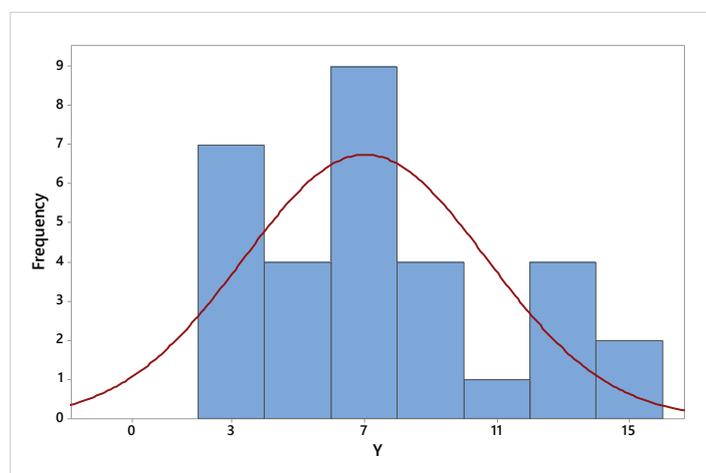
Gambar 4.12 menunjukkan Surabaya Timur terutama kecamatan Rungkut memiliki kondisi jalan baik yang relatif lebih Panjang hal ini bisa disebabkan karena banyaknya perusahaan/industri yang berdiri di Kecamatan Rungkut, sehingga kondisi jalan terjaga dengan baik.

### 4.3 Model *Spatial Poisson Point Process*

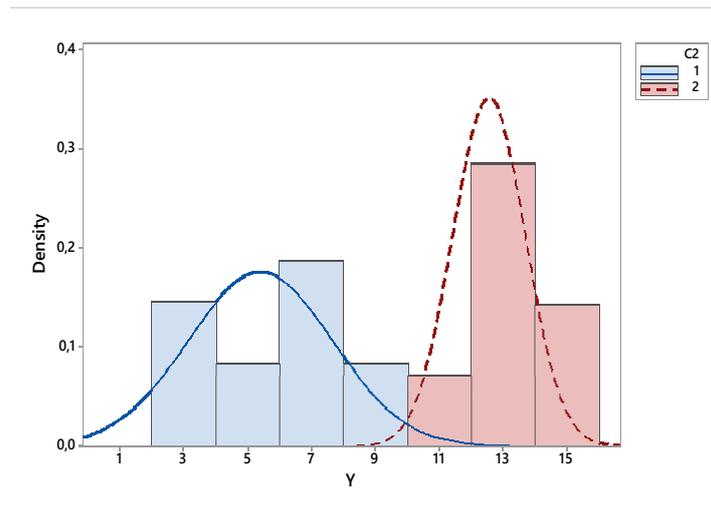
Data lokasi FKTP yang telah dilakukan eksplorasi secara visual kemudian dilakukan pemodelan *spatial Poisson point process*. Sebelum melakukan pemodelan dilakukan identifikasi *inhomogeneous Poisson process*. Identifikasi dilakukan dengan melakukan uji kesesuaian distribusi, dan uji homogenitas intensitas. Identifikasi model *mixture* juga dilakukan untuk mengetahui apakah data lokasi FKTP merupakan data bimodal. Setelah identifikasi selesai dilakukan kemudian dilakukan estimasi parameter model *spatial Poisson point process* dengan menggunakan metode Bayesian.

#### 4.3.1 Identifikasi *Mixture*

Dalam penelitian ini, untuk mengidentifikasi distribusi *mixture*, tahapan awal yang dilakukan adalah dengan melihat secara visual bentuk histogram dari data banyak FKTP pada masing-masing kecamatan di Surabaya. Dengan melihat histogram data diharapkan dapat terlihat jumlah modus dari sebaran data tersebut. Untuk dapat mengidentifikasi *mixture* dilihat banyak puncak yang terbentuk ketika membuat histogram. Selain itu juga dilakukan pengujian distribusi data dengan menggunakan statistik uji Chi-Square. Gambar 4.13 (a) menunjukkan histogram jumlah FKTP, sedangkan Gambar 4.13 (b) menunjukkan histogram jumlah FKTP jika ada dua komponen *mixture*.



(a)



(b)

**Gambar 4.13:** Histogram FKTP. (a) Histogram dari jumlah FKTP dan (b) Histogram dari 2-komponen dari jumlah FKTP

Berdasarkan perbandingan antara histogram data banyak FKTP dengan 1 (satu) komponen dan 2 (dua) komponen terlihat bahwa kurva fit distribusi pada Gambar 4.13 (b) lebih mewakili dibandingkan dengan Gambar 4.3 (a). Untuk identifikasi *mixture* selain dengan visual, dapat dilakukan pengujian distribusi terhadap data FKTP yang ditunjukkan dalam Gambar 4.13.b. Uji distribusi dilakukan dengan menggunakan Statistik uji Chi-Square. Hipotesis dan statistik uji dapat dinyatakan sebagai berikut.

$H_0$  : Data mengikuti pola fungsi distribusi Poisson.

$H_1$  : Data tidak mengikuti pola fungsi distribusi Poisson.

Penolakan hipotesis tersebut dapat dilakukan dengan menggunakan statistik uji Chi-Square.

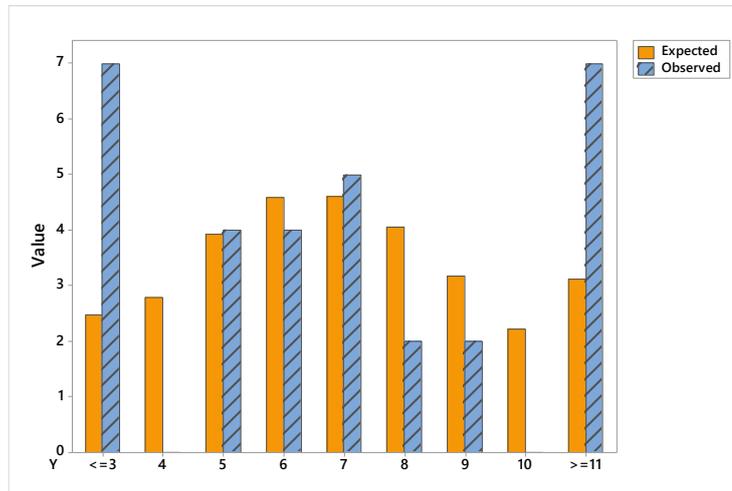
$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i}$$

$$\chi^2 = \frac{(7 - 2,48308)^2}{2,48308} + \dots + \frac{(7 - 3,12562)^2}{3,12562}$$

$$= 19,6283$$

Nilai statistik uji yang didapatkan dari Chi-Square adalah sebesar 19,6283 dan p-value sebesar 0,006, sedangkan untuk nilai CV nya adalah 2,167 untuk alpha 5%. Keputusan yang didapatkan dari uji kesesuaian distribusi ini adalah tolak  $H_0$ ,

sehingga dapat disimpulkan bahwa distribusi data jumlah FKTP tidak berdistribusi Poisson.



**Gambar 4.14:** Pengujian Distribusi FKTP

*Poisson process* terbagi menjadi 2 jenis *homogeneous poisson process* dan *nonhomogeneous poisson process*. Salah satu karakteristik dari jenis Poisson process tersebut dapat dilihat dari intensitas data yang didapatkan melalui uji homogenitas intensitas. Uji homogenitas intensitas persebaran FKTP dilakukan dengan menggunakan *quadrat.test*. Hipotesis dari *quadrat.test* adalah sebagai berikut.

$H_0$  : intensitas *point process* FKTP homogen.

$H_1$  : intensitas *point process* FKTP tidak homogen.

Langkah pertama dalam melakukan pengujian yaitu membagi *window* Surabaya menjadi subarea  $B_1, B_2, \dots, B_{31}$  yang merupakan kecamatan di Surabaya dan menghitung banyak titik FKTP  $n_1, n_2, \dots, n_{31}$  pada masing-masing kecamatan.

Diketahui total titik lokasi FKTP di Surabaya sebanyak  $n = \sum_{i=1}^{31} n_i = 218$  dan total area *window* adalah  $a = \sum_{i=1}^{31} a_i = 326,37 \text{ km}^2$  dimana  $a_i$  adalah area  $B_i$ , estimasi intensitas adalah  $\bar{\lambda} = \frac{218}{326,37} = 0,668$ , dan ekspektasi banyak titik pada

subarea  $B_1$  (kecamatan Gayungan) adalah  $e_1 = \bar{\lambda}a_1/a = 218 \times \frac{6,07}{326,37} = 4,1$ .

Statistik uji adalah sebagai berikut.

$$\chi^2 = \sum_{i=1}^{31} \frac{(n_i - e_i)^2}{e_i} = \sum_{i=1}^{31} \frac{(n_i - \bar{\lambda}a_i)^2}{\bar{\lambda}a_i} = 124,54.$$

Dibawah hipotesis nol, distribusi uji statistik didekati dengan distribusi  $\chi^2$  dengan derajat bebas  $31 - 1 = 30$ . *P-value* hasil pengujian sebesar 0,000, nilai  $\chi^2$  tabel 34,79 sehingga dapat disimpulkan intensitas FKTP di Kota Surabaya nonhomogen, sehingga FKTP di setiap kecamatan di Kota Surabaya mengikuti proses NHPP. Pengujian overdispersi menghasilkan statistik *D-test*:

$$D - test = \frac{\text{observed variance}}{\text{theoretical variance}} \times (\text{number observation} - 1) = 57,30275$$

dengan *p-value* 0,0019285 yang berarti bahwa data FKTP adalah data *count overdispersed*. Regresi *Mixture Poisson* dapat menjadi alternatif untuk memodelkan hubungan antara variabel respon *count* dan variabel kovariat di mana variabel respon berasal dari NHPP dan atau data count overdispersi. Jumlah komponen campuran dapat diperoleh secara visual berdasarkan histogram pada Gambar 4.13.b. Dua komponen ini dapat diartikan bahwa kecamatan di Surabaya dikategorikan dalam intensitas tinggi dan rendah dalam hal ketersediaan FKTP.

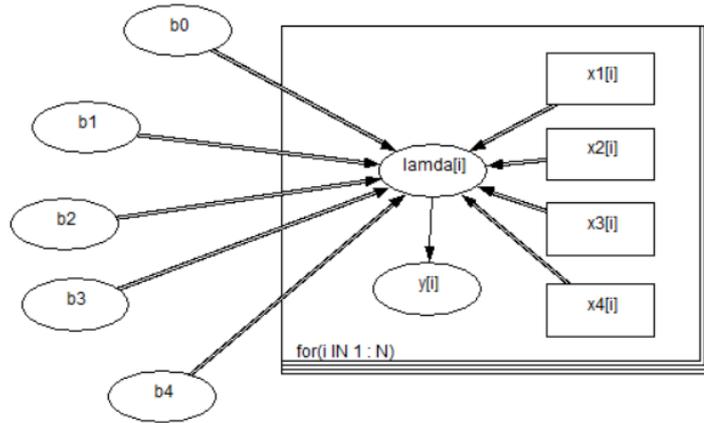
#### 4.3.2 Pemodelan Intensitas FKTP dengan Regresi Poisson dan Regresi *Mixture Poisson*

Pemodelan regresi Poisson digunakan untuk memprediksi data respon *count*. Pemodelan dilakukan pada intensitas FKTP di Surabaya dengan empat variabel kovariat. Untuk melakukannya, asumsi yang harus dipenuhi dalam pemodelan regresi Poisson adalah bahwa data mean dan varians harus sama. Regresi Poisson dimodelkan oleh fungsi link dalam persamaan (2.13). Berdasarkan fungsi likelihood data pada persamaan 2.14 kemudian dikalikan dengan distribusi prior regresi Poisson yaitu distribusi Normal. Prior untuk vektor parameter  $\beta$  dinotasikan sebagai  $f(\beta_{pk})$  dimana  $p = 1,2,3,4, k = 1,2$  sebanyak parameter regresi Poisson.  $f(\beta_{pk})$  berdistribusi normal yang merupakan *pseudo prior*, dengan  $\beta_{pk} \sim N(\mu_{\beta_{pk}}, \sigma_{\beta_{pk}}^2)$  memiliki fungsi densitas sebagai berikut.

$$f(\beta_{pk}) = \frac{1}{\sqrt{2\pi\sigma_{\beta_{pk}}^2}} \exp \left[ -\frac{1}{2\sigma_{\beta_{pk}}^2} (\beta_{pk} - \mu_{\beta_{pk}})^2 \right] \quad (4.42)$$

untuk semua parameter  $\beta_{pk}$ ,  $p = 1,2,3,4, k = 1,2$  saling independen. *Full conditional posterior distribution* diperoleh dengan pengalihan persamaan

likelihood data pada persamaan (2.14) dengan fungsi distribusi prior pada persamaan (4.40). Dengan menggunakan pendekatan Bayesian dengan MCMC sebagaimana dinyatakan dalam Algoritma 2 digunakan untuk mendapatkan hasil konvergen (**Lampiran 6**). Kondisi ini tercapai pada iterasi sebanyak 20000 dan thin 10 iterasi dengan menggunakan WinBUGS. Berikut DAG untuk model regresi Poisson.



**Gambar 4.15:** DAG Regresi Poisson

Dengan menggunakan pseudo-prior yang diperoleh dengan metode klasik diperoleh persamaan regresi untuk intensitas FKTP pada kondisi konvergen sebagai berikut:

**Tabel 4.2** Estimasi Parameter Regresi Poisson

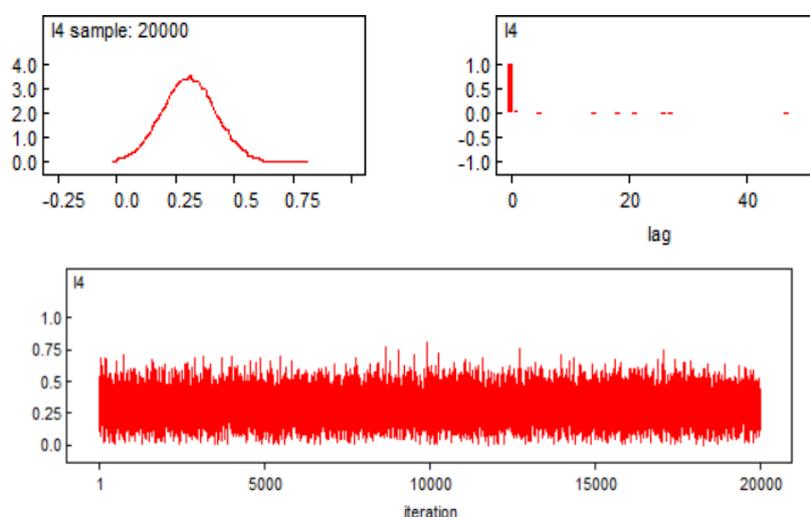
Parameter	Rata-rata	sd	2,5%	Median	97,5%
$\hat{\beta}_0$	1,881*	0,073	0,0005754	1,735	1,882
$\hat{\beta}_1$	0,110*	0,085	0,0006127	0,0589	0,111
$\hat{\beta}_2$	-0,067	0,133	0,001003	-0,3325	-0,066
$\hat{\beta}_3$	-0,085	0,113	0,0008565	-0,3044	-0,085
$\hat{\beta}_4$	0,307*	0,114	0,000873	0,08828	0,305

\*) parameter model signifikan

Berdasarkan Tabel 4.2 estimasi parameter untuk dugaan model Regresi Poisson adalah sebagai berikut.

$$\lambda_i = \exp[1,88 + 0,110X_1(u_i) - 0,067X_2(u_i) - 0,086X_3(u_i) + 0,307X_4(u_i)]$$

Berdasarkan estimasi parameter pada Tabel 4.2 model regresi Poisson diperoleh informasi mengenai uji signifikansi parameter menggunakan *credible interval*. Jika *credible interval* memuat nilai nol, maka disimpulkan bahwa estimasi parameter tidak signifikan. Terlihat intensitas FKTP di Surabaya dipengaruhi secara positif oleh kepadatan penduduk  $X_1$  dan Panjang jalan dengan kondisi baik  $X_4$ , sehingga jelas terlihat bahwa FKTP banyak didirikan di wilayah padat penduduk dan memiliki akses jalan yang baik. Hal ini menunjukkan bahwa tingkat kebutuhan dan aksesibilitas mendorong berdirinya fasilitas-fasilitas umum khususnya dalam hal ini FKTP. Asumsi tidak ada overdispersi tidak terpenuhi dalam model ini terlihat dari hasil uji overdispersi, sehingga dilakukan pemodelan yang melibatkan adanya komponen *mixture* dalam model. Estimasi regresi Poisson pada Tabel 4.2 melibatkan 5 (lima) parameter yaitu  $\beta_0, \beta_1, \beta_2, \beta_3, \beta_4$  yang sebagian besar menunjukkan kesesuaian dengan sifat MCMC yaitu *irreducible* dan *aperiodic*. Sebagai contoh, ditunjukkan melalui MCMC *diagnostic plot* pada parameter  $\beta_1$  sebagaimana ditunjukkan pada **Gambar 16**. *Serial plot* dari 20.000 sampel dibangkitkan, terlihat bahwa sampel yang dihasilkan dalam proses MCMC tidak menunjukkan nilai yang ekstrim atau dalam kata lain *Serial plot* parameter tidak menunjukkan suatu pola tertentu, cenderung stasioner dan bersifat random.



**Gambar 4.16:** Regresi Poisson, (a) *Density plot*, (b) *Autocorrelation plot*, dan (c) *History plot*

Visualisasi pada **Gambar 16** membuktikan bahwa sampel dihasilkan mempunyai sifat random (*irreducible*) dan tidak memiliki pola periodic (*aperiodic*).

*Autocorrelation plot* pada **Gambar 16** menguatkan bukti bahwa sampel estimasi parameter yang dihasilkan bersifat random. Hal ini ditunjukkan dengan autokorelasi yang bernilai satu hanya pada lag 0 dan bernilai nol/mendekati nol pada lag selanjutnya. *Kernel density plot* pada **Gambar 16** menunjukkan bahwa densitas hasil estimasi parameter  $\beta_1$ . Berdasarkan MCMC *diagnostic plot* yaitu *serial*, *autocorrelation*, dan *Kernel density plot* dapat ditarik kesimpulan bahwa proses estimasi parameter yang telah dilakukan sudah mencapai kondisi yang konvergen. Dengan demikian, estimasi parameter yang dihasilkan dapat digunakan untuk menggambarkan Regresi Poisson untuk intensitas FKTP di Surabaya.

Selanjutnya dikarenakan pada regresi Poisson terjadi kasus overdispersi maka dilakukan pemodelan dengan regresi *mixture Poisson*. Langkah pertama untuk pemodelan adalah dengan melakukan identifikasi *mixture*. Identifikasi *mixture* dilakukan berdasarkan informasi bahwa data FKTP di Surabaya termasuk dalam *finite mixture Poisson*, sehingga dengan menggunakan teori pada bab 2.5 dilakukan proses identifikasi kecamatan mana saja yang masuk komponen ke- $k$ . Berdasarkan identifikasi *mixture* dengan histogram, kemungkinan bahwa data terdiri dari dua komponen *mixture*. Keanggotaan setiap komponen ditentukan oleh prosedur sampling Gibbs pada algoritma 1 dengan *syntax* sebagaimana terlampir pada **Lampiran 7**. Hasilnya menunjukkan bahwa komponen pertama terdiri dari 25 (dua puluh lima) kecamatan dengan banyak FKTP kurang dari sama dengan 11 di setiap kecamatan, yaitu kecamatan Gayungan, kecamatan Karangpilang, kecamatan Gununganyar, kecamatan Jambangan, kecamatan Tenggilismejoyo, kecamatan Wonocolo, kecamatan Lakarsantri, kecamatan Wiyung, kecamatan Sukolilo, kecamatan Dukuhpakis, kecamatan Sukomanunggal, kecamatan Mulyorejo, kecamatan Tegalsari, kecamatan Tandes, kecamatan Sambikerep, kecamatan Genteng, kecamatan Bubutan, kecamatan Simokerto, kecamatan Bulak, kecamatan Pabean Cantikan, kecamatan Asemrowo, kecamatan Pakal, kecamatan Semampir, kecamatan Kenjeran, dan kecamatan Benowo. Sedangkan komponen kedua terdiri dari 6 (enam) kecamatan dengan total FKTP lebih dari 11 di setiap kecamatan, yaitu kecamatan Rungkut, kecamatan Wonokromo, kecamatan Gubeng, kecamatan Sawahan, kecamatan Tambaksari, dan kecamatan

Kreimbangan, rincian hasil identifikasi *mixture* terlampir. Hasil pengujian distribusi untuk kedua komponen tersebut dapat terlihat pada Tabel 4.2 berikut.

**Tabel 4.3:** Pengujian Distribusi Masing-masing Komponen

Komponen	Anggota	Chi-Square	P-value	Distribusi
Komponen 1	25	6,83758	0,336	Poisson
Komponen 2	6	0,264005	0,607	Poisson

Berdasarkan Tabel 4.3 terlihat bahwa data FKTP di Surabaya terdiri dari dua komponen yang masing-masing memiliki distribusi yang sama yaitu distribusi Poisson. Secara matematis, model *mixture* Poisson untuk FKTP di Surabaya dapat ditulis sebagai berikut:

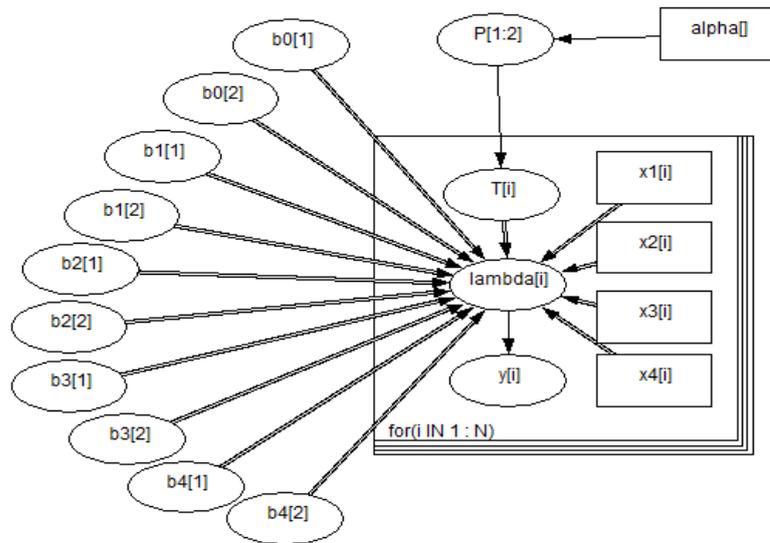
$$Y_i \sim PM(p_k, \lambda_i)$$

$$\lambda_i = \exp(x_i^T \boldsymbol{\beta})$$

$$y_i = p(\log[\lambda_1(u_{i_1})]) + (1 - p)(\log[\lambda_2(u_{i_1})])$$

$$y_i = p(\log[\exp(\beta_{01} + \beta_{11}x_{1i_1} + \beta_{21}x_{2i_1} + \beta_{31}x_{3i_1} + \beta_{41}x_{4i_1})]) + (1 - p)(\log[\exp(\beta_{02} + \beta_{12}x_{1i_2} + \beta_{22}x_{2i_2} + \beta_{32}x_{3i_2} + \beta_{42}x_{4i_2})]) \quad (4.41)$$

$\mu_{[\beta_{pk}]}$  dan  $\sigma_{[\beta_{pk}]}^2$  adalah hiper-parameter yang ditetapkan sebagai pseudo-prior yang diturunkan dari estimasi parameter regresi menggunakan pendekatan GLM, parameter regresi  $\beta_{qk}$  dengan  $q = 0,1,2,3,4$  dan  $k = 1,2$ . Berdasarkan likelihood pada persamaan 2.19 kemudian dikalikan dengan seluruh distribusi prior parameter  $\beta_{qk} \sim N(\mu_{[\beta_{qk}]}, \sigma_{[\beta_{qk}]}^2)$  dan parameter  $p \sim \text{Dirichlet}(1,1)$  diperoleh *full conditional posterior distribution* untuk parameter regresi *mixture* Poisson. Estimasi parameter regresi untuk setiap komponen dilakukan secara komputasional dengan menggunakan pendekatan Bayesian. Proses estimasi dilakukan melalui pengambilan sampel berulang melalui bentuk distribusi full conditional posterior. Struktur Regresi *Mixture* Poisson dengan empat model variabel kovariat melalui model DAG (*Directed Acyclic Graph*) pada Gambar 4.17 dan kemudian secara otomatis menghasilkan kode WinBUGS yang sesuai dari grafik ini. Gambar 4.17 adalah struktur parameter model *mixture* Poisson regression pada WinBUGS.



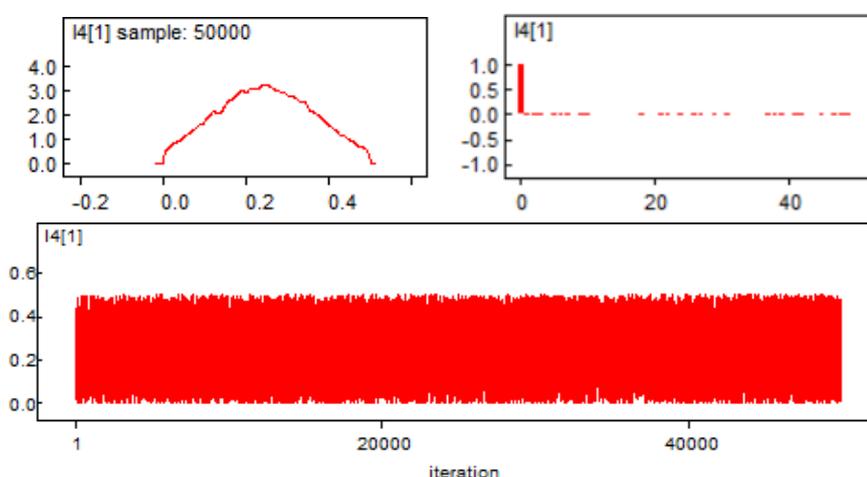
**Gambar 4.17** Regresi *Mixture Poisson* DAG

Distribusi prior dalam Gambar 4.17 ditunjukkan oleh node yang berada di luar kotak. Distribusi prior yang digunakan dalam penelitian ini adalah jenis pseudo prior, dimana prior bergantung pada hasil frequentis. Prior dibangkitkan menggunakan distribusi normal dengan mean berasal dari hasil estimasi parameter yang didapatkan dari fungsi frequentist GLMs, dan nilai standard deviasi didapatkan dari hasil 1/standard error dari fungsi frequentis GLMs. Pemodelan dilakukan dengan menyertakan semua variabel kovariat yang telah ditentukan dan di sandarisasi terlebih dahulu, empat variabel kovariat yang menggambarkan karakteristik masing-masing kecamatan dimodelkan dengan Regresi *mixture Poisson* dua komponen.

Penentuan distribusi posterior dan estimasi parameter model dilakukan dengan menggunakan simulasi MCMC sampai diperoleh kondisi rantai markov *irreducible*, *aperiodic* dan *recurrent*. Kondisi ini tercapai pada iterasi sebanyak 50000, dengan thin 25 dengan menggunakan WinBUGS. Hasil iterasi tersebut telah menunjukkan bahwa hasil estimasi parameter telah konvergen. Pemantauan konvergensi dapat dilihat pada rantai markov yang telah memenuhi sifat *ergodic*. Sifat *ergodic* diindikasikan proses yang telah *irreducible*, *aperiodic* dan *reccurent*. *Irreducible* yaitu selama iterasi estimasi parameter sampel diambil secara acak yang ditunjukkan oleh *history plot* pada Lampiran 12c yang memperlihatkan sifat

stationer dan acak. *Aperiodic* ditunjukkan oleh *history plot* yang tidak membentuk pola tertentu. *Recurrent* yaitu parameter yang dibangkitkan pada *state i* berpeluang untuk kembali ke *state i*. Selain itu, konvergensi juga dapat dilihat pada *autocorrelation plot* pada Lampiran 12b menunjukkan bahwa lag nol bernilai satu kemudian lag selanjutnya bernilai mendekati nol serta *density plot* pada Lampiran 12a menunjukkan distribusi pada masing-masing parameter cenderung memiliki pusat di tengah.

Parameter yang dihasilkan, sebagai contoh, ditunjukkan sebagai *density plot*, *autocorrelation plot*, dan *history plot* pada Gambar 4.18 (a), Gambar 4.18 (b), dan Gambar 4.18 (c). Gambar 4.18 (a) menunjukkan bahwa *density plot* dari *diagnostic posterior* untuk proses rantai *Markov Monte Carlo* (MCMC)



menunjukkan kesesuaian dengan plot densitas distribusi Normal. *Autocorrelation plot* pada Gambar 4.18 (b) menunjukkan bahwa tidak ada autokorelasi dalam sampel yang dihasilkan oleh Gibb Sampling. *History plot* pada Gambar 4.18 (c) menunjukkan bahwa rantai sudah dalam keadaan proses pencampuran cepat yang menyatakan bahwa kondisi rantai markov yang *irreducible*, *aperiodic* dan *irreducible*.

**Gambar 4.18:** Regresi *mixture Poisson*, (a) *Density plot*, (b) *Autocorrelation plot*, dan (c) *History plot*

Pengujian signifikansi parameter menggunakan *credible interval*. Jika *credible interval* memuat nilai nol, maka dapat disimpulkan bahwa estimasi parameter tidak signifikan. Model estimasi model *mixture Poisson* untuk FKTP di Surabaya

ditunjukkan dalam persamaan di bawah ini. Estimasi parameter pada regresi *mixture* Poisson pada persamaan 4.41 untuk FKTP di Surabaya ditunjukkan dalam Tabel 4.3 dibawah ini. Ekspektasi FKTP dapat diperkirakan dengan model sesuai persamaan (2.20). Sebagai contoh, berdasarkan persamaan regresi *mixture* Poisson berikut ini menyatakan bahwa kepadatan penduduk  $\hat{\beta}_{11}$  di kecamatan pada komponen satu memiliki kontribusi 0,004 dalam model. Persamaan regresi *mixture* Poisson dapat dituliskan sebagai berikut.

$$0,788 \times \left[ \exp \left( 1,717 + 0,004x_{1i_1} - 0,027x_{2i_1} - 0,058x_{3i_1} + 0,246x_{4i_1} \right) \right] +$$

$$+ 0,212 \times \left[ \exp \left( 1,665 + 0,744x_{1i_2} - 0,749x_{2i_2} + 0,293x_{3i_2} + 1,156x_{4i_2} \right) \right]$$

Komponen pertama terdiri dari 25 kecamatan memberikan kontribusi yang lebih kecil untuk model yaitu 78,8%. Sedangkan komponen kedua terdiri dari 6 (enam) kecamatan memberikan kontribusi sebesar 21,2%.

**Tabel 4.4:** Estimasi Parameter Regresi *Mixture* Poisson

Node	Mean	Sd	2,5%	median	97,5%
$\hat{p}_1$	0,788*	0,070	0,635	0,794	0,906
$\hat{p}_2$	0,212*	0,070	0,094	0,206	0,365
$\hat{\beta}_{01}$	1,717*	0,085	1,547	1,718	1,882
$\hat{\beta}_{02}$	1,665*	0,444	0,837	1,673	2,462
$\hat{\beta}_{11}$	0,004	0,104	-0,202	0,006	0,206
$\hat{\beta}_{12}$	0,744*	0,379	0,065	0,744	1,430
$\hat{\beta}_{21}$	-0,027	0,144	-0,304	-0,028	0,255
$\hat{\beta}_{22}$	-0,749*	0,319	-1,371	-0,735	-0,204
$\hat{\beta}_{31}$	-0,058	0,132	-0,315	-0,058	0,199
$\hat{\beta}_{32}$	0,293*	0,156	0,016	0,291	0,573
$\hat{\beta}_{41}$	0,246*	0,114	0,033	0,245	0,463
$\hat{\beta}_{4i}$	1,156*	0,504	0,275	1,142	2,102

\*) parameter model signifikan

Tabel 4.4 menunjukkan estimasi parameter regresi *mixture* Poisson. Pada Tabel 4.4 juga menunjukkan *credible interval* dari setiap parameter yang diestimasi. *Credible interval* merupakan estimasi interval dari parameter yang digunakan dalam metode Bayesian untuk menguji signifikansi parameter. Estimasi *credible interval*

mendefinisikan probabilitas *posterior* sehingga parameter terletak dalam interval [2,5%;97,5%]. Apabila *credible interval* tidak memuat angka nol, maka parameter regresi *mixture* Poisson signifikan. Variabel kepadatan penduduk  $\hat{\beta}_{12}$  berpengaruh signifikan secara positif pada jumlah FKTP di kecamatan pada komponen dua, yang berarti intensitas FKTP tinggi pada wilayah-wilayah dengan kepadatan penduduk tinggi. Variabel kovariat persentase rumah tangga bersih dan sehat  $\hat{\beta}_{22}$  berpengaruh negatif pada komponen kedua. Hal ini menunjukkan bahwa di kecamatan dengan intensitas FKTP tinggi (komponen kedua), FKTP cenderung banyak didirikan di wilayah dengan persentase rumah tangga bersih dan sehat rendah, Angka keterkaitan pada komponen kedua  $\hat{\beta}_{32}$  yaitu kecamatan dengan intensitas FKTP tinggi, semakin dekat jarak antar kecamatan (angka keterkaitan kecil) maka jumlah FKTP semakin sedikit karena memiliki pengaruh positif. Sedangkan variabel Panjang jalan kondisi baik  $\hat{\beta}_{41}$  dan  $\hat{\beta}_{42}$  berpengaruh positif untuk kedua komponen, yang menandakan intensitas FKTP semakin besar pada area dengan kondisi jalan baik. Perbandingan model antara model regresi *mixture* Poisson dan regresi Poisson dilakukan dengan menggunakan uji *goodness of fit* DIC. *Deviance Information Criterion* dari kedua model ini ditunjukkan pada Tabel 4.5 yang menyimpulkan model kedua lebih baik daripada yang pertama karena nilai DIC terkecil.

**Tabel 4.5:** Perbandingan DIC Model

<b>Model</b>	<b>DIC Value</b>
Regresi Poisson	160,580
Regresi <i>Mixture</i> Poisson	148,985

Berdasarkan Tabel 4.5 terlihat bahwa model regresi *mixture* Poisson menghasilkan nilai DIC lebih kecil dibandingkan dengan regresi Poisson, sehingga pada kasus ini regresi *mixture* Poisson lebih baik.

## BAB V

### KESIMPULAN DAN SARAN

#### 5.1 Kesimpulan

Berdasarkan hasil analisis yang dilakukan dalam penelitian ini, pemodelan Regresi *Mixture* Poisson pada intensitas FKTP di Surabaya dengan pendekatan Bayesian didapatkan kesimpulan sebagai berikut.

1. Pola persebaran FKTP di Surabaya tidak homogen atau termasuk kedalam kriteria *nonhomogeneous Poisson Process*, hal ini dikarenakan terdapat 2 (dua) puncak kepadatan persebaran lokasi FKTP.
2. Pada pemodelan regresi Poisson diperoleh informasi bahwa intensitas FKTP di Surabaya dipengaruhi secara positif oleh kepadatan penduduk dan Panjang jalan dengan kondisi baik. Namun terjadi overdispersi dari proses Poisson.
3. Model *mixture* poisson process yang didapatkan memiliki 2 komponen yaitu komponen 1 untuk kepadatan FKTP rendah yaitu dengan banyak FKTP kurang dari sama dengan 11 FKTP per kecamatan, komponen 2 untuk kepadatan FKTP tinggi yaitu banyak FKTP lebih dari 11 per kecamatan, dimana hasil estimasi parameter yang didapatkan variabel total populasi berpengaruh signifikan secara positif pada jumlah FKTP di kecamatan pada komponen 2, persentase rumah tangga bersih dan sehat berpengaruh negatif pada kedua komponen. Sedangkan variabel Panjang jalan kondisi baik berpengaruh positif untuk kedua komponen.
4. Hasil perbandingan model Regresi Poisson dengan Regresi *Mixture* Poisson dengan kriteria DIC menunjukkan Regresi *Mixture* Poisson lebih cocok untuk memodelkan intensitas FKTP di Surabaya.

#### 5.2 Saran

Saran yang dapat diberikan dalam penelitian selanjutnya adalah sebagai berikut.

1. Dalam menentukan jumlah komponen *mixture* dapat digunakan metode *Reversible jump Markov Chain Monte Carlo* (RJMCMC) untuk distribusi Normal, Log Normal, Poisson ataupun yang lainnya.
2. Dapat dipertimbangkan untuk menambahkan variabel yang diharapkan berpengaruh signifikan pada model.

3. Pada penelitian selanjutnya dapat dilakukan analisis intensitas *point process* dengan metode lainnya atau memasukkan variabel lain seperti jarak antar lokasi FKTP sehingga mampu memberikan lebih banyak informasi.

## DAFTAR PUSTAKA

- Agresti, A. (2002). *Categorical Data Analysis Second Edition*. John Wiley & Sons, New York.
- Badan Pusat Statistik. (2017). *Kota Surabaya dalam Angka 2017*. Surabaya: Badan Pusat Statistik.
- Baddeley, A., Bárány, I., dan Schneider, R. (2007). *Spatial Point Processes and their Applications*. Lecture handout: the CIME Summer School, Martina Franca, Italy.
- Baddeley, A., Rubak, E., dan Tamer, R. (2016). *Spatial Point Patterns Methodology and Applications with R*. CRC Press Taylor dan Francis Group, New York.
- Box, G. dan Tiao, G. (1973). *Bayesian Inference in Statistical Analysis*. Addison-Wesley publishing Company, Massachusetts.
- Brereton, T. (2014). *Stochastic Simulation of Processes, Fields and Structures*. Lecture handout: Summer term, Institute of Stochastic.
- Carlin, B. dan Chib, S. (1995). "Bayesian Model Choice via Markov Chain Monte Carlo Methods". *Journal of the Royal Statistical Society, Series B (Methodological)*, hal. 473-484.
- Carlin, B.P. dan Louis, T.A. (2008). *Bayesian Methods for Data Analysis 3<sup>th</sup> Edition*. CRC Press Taylor dan Francis Group, USA.
- Chen, Y. dan Ge, Y. (2015). "Spatial Point Pattern analysis on the Villages in China's poverty-stricken areas". *Procedia Environmental Sciences*, Vol. 27, hal. 98 – 105.
- Choiruddin, A., Coeurjolly, J.F., dan Letue, F. (2017a). "Convex and non-convex regularization methods for spatial point processes intensity estimation". *France: arXiv preprint arXiv:1703.02462*.
- Choiruddin, A., Coeurjolly, J.-F., dan Letue, F. (2017b). "Spatial point processes intensity estimation with a diverging number of covariates". *France: arXiv:1712.09562v1 [stat.ME] 27 Dec 2017*.
- Congdon, P. (2014). *Applied Bayesian Modelling, Second Edition*. John Wiley & Sons, United Kingdom.

- Dinas Kesehatan Kota Surabaya. (2018). *Puskesmas dan Klinik Pratama*. <http://dinkes.surabaya.go.id/portal/upt-dinas> diakses pada 17 Maret 2018.
- Gelman, A., Carlin, J., Stern, H., dan D.B., R. (1995). *Bayesian Data Analysis*. Chapman & Hall, London.
- Grun, B. dan Leisch, F. (2008) "Finite Mixture with Concomitant Variables and Varying and Constant Parameters". *Journal of Statistical Software*, Vol. 28, No. 4.
- Hamdah, D.F.A. (2015). *Bayesian Inference on Finite Mixtures of Poisson Distributions*, Disertasi Ph.D., The Islamic University of Gaza, Palestina.
- Illian, J., Penttinen, A., Stoyan, H., dan Stoyan, D. (2008). *Statistical Analysis and Modelling of Spatial Point Pattern*. John Wiley & Sons, Ltd., United Kingdom.
- Iriawan, N. (2001). "Penaksiran Model Mixture Normal Univariabel: Suatu Pendekatan Mixture Bayesian dengan MCMC". *Prosiding Seminar Nasional dan Konferda VII Matematika Wilayan DIY & Jawa Tengah*, Yogyakarta, hal. 105-110.
- Kementerian Kesehatan Republik Indonesia. (2015). *Profil Kesehatan Indonesia*. Jakarta: Kementerian Kesehatan Republik Indonesia.
- Kottas, A. dan Sanso, B. 2007. "Bayesian Mixture Modelling for Spatial Poisson Process Intensities with Applications to Extreme Value Analysis". *Journal of Statistical Planning and Inference*, Vol. 137, No. 10, hal. 3151-3163.
- Kulvanich, N. 2013. *Mixture Poisson Point Process Assessing Heterogeneity in EMA Analysis*. University of Georgia, USA.
- Kusumaningrum, C.M., Iriawan, N., dan Winahju, W.S. (2017). "Pattern Analysis of Community Health Center Location in Surabaya using Spatial Poisson Point Process". *In AIP Conference Proceedings*, Vol. 1905, Issue 1, id.040016.
- Kwang-soo, L. dan Kyeong-Jun, M. (2014). "Hospital distribution in a metropolitan city: assessment by ageographical information system grid modelling approach". *Geospatial Health*, Vol 8, No. 2, hal. 537-544.
- Leininger, T.J. (2014). *Bayesian Analysis of Spatial Point Pattern*. Duke University: North Carolina.

- Leininger, T.J, dan Gelfand, A.E. (2017). “Bayesian Inference and Model Assessment for Spatial Point Patterns Using Posterior Predictive Samples”. *Bayesian Analysis*, Vol 12, No.1, hal. 1-30.
- Magan, H., Indar, dan Balqis. (2013). *Faktor yang Berhubungan Dengan Pemanfaatan Pelayanan Kesehatan Unit Rawat Jalan di Wilayah Kerja Puskesmas Makale*. Tugas Akhir, Universitas Hasanuddin, Makassar.
- Mansour, S. (2016). “Spatial Analysis of Public Health Facilities in Riyadh Governorate, Saudia Arabia : a GIS-based Study to Assess Geographic Variations of Service Provision and Accessibility”. *Geo-spatial Information Science*, Vol.19, No. 1, hal. 26-38.
- McCullagh, P. dan Nelder, J.A. (1983). *Generalized Linear Models*. Chapman and Hall, London.
- Møller, J. dan Waagepetersen, R.P. (2004). *Statistical Inference and Simulation for Spatial Point Processes*. Chapman&Hall/CRC, New York.
- Mufudza, C. dan Erol, H. (2016). “Poisson Mixture Regression Models for Heart Disease Prediction”. *Computational and Mathematical Methods in Medicine*, Vol. 2016, hal. 10.
- Mursalim. (2018). *Pengukuran Aksesibilitas Kecamatan di Wilayah Pemerintah Kota Surabaya*. Tugas Akhir, Institut Teknologi Sepuluh Nopember, Surabaya.
- Nobles, M., Serban, N., dan Swann, J. (2014). “Spatial Accessibility Of Pediatric Primary Healthcare: Measurement And Inference”. *The Annals of Applied Statistics 2014*, Vol. 8, No. 4, hal. 1922–1946. .
- Ntzoufras, I. (2009). *Bayesian Modeling Using WinBUGS*. Wiley, New Jersey.
- Paci, L., Beamonte, M.A., Gelfand, A.E., Gargallo, P., dan Salvador, M. (2017). “Analysis Of Residential Property Sales Using Space–Time Point Patterns”. *Spatial Statistics 21*. Vol. 2017, hal. 149–165.
- Permenkes-No-75*. (2014). Retrieved 1 28, 2018, from <http://dinkes.gunungkidulkab.go.id/wp-content/uploads/2014/10/Permenkes-No-75-Th-2014-ttg-Puskesmas.pdf> diakses pada 18 Maret 2018.
- Quesada, J.A., Melchor, M., dan Nolasco, A. (2017). “Point Process Methods in Epidemiology: Application to the Analysis of Human Immunodeficiency

- Virus/Acquired Immunodeficiency Syndrome Mortality in Urban Areas”. *Geospat Health*. Vol. 12, No.1, hal. 506.
- Shinichiro Shirota, Jorge Mateu, dan Alan E Gelfand. (2017). “Statistical Analysis Of Origin Destination Point Patterns: Modeling Car Thefts and Recoveries”. *arXiv preprint arXiv:1701.05863*.
- Skare, O., Moller, J., dan Jensen, E.B.V. (2007). “Bayesian analysis of spatial point processes in the neighbourhood of Voronoi networks”. *Journal of Statistics and Computing*. Vol. 17, No. 4, hal. 369-379.
- Snedecor, G.W. dan Cochran, W.G. (1989) *Statistical Methods, Eighth Edition*. Iowa State University Press.
- Takbir, Sunusi, dan Islamiyati. (2016). *Studi Analisa Kecelakaan Lalu Lintas Melalui Pendekatan Spatial Point Process*. Tugas Akhir, Universitas Hasanuddin: Makassar.
- Thurman, A. L. dan Zhu, J. (2014). “Variable selection for spatial Poisson point processes via a regularization method”. *Statistics. Methodology*. Vol. 17, hal. 113-125.
- Waagepetersen, R. dan Guan, Y. (2009). “Two-Step Estimation For Inhomogeneous Spatial Point Processes”. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, Vol. 71, No. 3, hal. 685-702.
- Walsh, D.C.I. dan Raftery. A. E. (2005). “Classification of Mixyure of Spatial Point Processes via Partial Bayes Factors”. *Journal of Computational and Graphical Statistics*, Vol. 14, hal. 139-154.

## LAMPIRAN

### Lampiran 1 Lokasi Fasilitas Kesehatan Tingkat Pertama di Surabaya

No	Klinik/Puskesmas	Alamat	Kecamatan	Latitude	Longitude
1	Gayungan	Jl. Gayungsari Barat 124	Gayungan	-7.338074	112.726704
2	Kedurus	Jl. Raya Mastrip Kedurus 46	Karangpilang	-7.319671	112.709634
3	Gunung Anyar	Jl. Gunung Anyar Timur 70	Gununganyar	-7.340858	112.783992
4	Kebonsari	Jl. Kebonsari Manunggal 30 32	Jambangan	-7.328314	112.714368
5	Tenggilis	Jl. Rungkut Mejoyo Selatan IV/P48	Tenggilismejoyo	-7.322314	112.761247
6	Jemursari	Jl. Jemursari Selatan IV/5	Wonocolo	-7.326072	112.740052
7	Sidosermo	Jl. Sidosermo Gg. Damri No. 51	Wonocolo	-7.30734	112.755813
8	Siwalankerto	Jl. Siwalankerto No. 134	Wonocolo	-7.33595	112.736447
9	Kalirungkut	Jl. Kalirungkut Puskesmas No.1	Rungkut	-7.322301	112.770726
10	Medokan Ayu	Jl. Medokan Asri Utara IV/31	Rungkut	-7.323955	112.794963
11	Jeruk	Jl. Raya Menganti 277 A	Lakarsantri	-7.305416	112.638404
12	Lidah Kulon	Jl. Raya Menganti Lidah Kulon No.5	Lakarsantri	-7.306585	112.661473
13	Bangkingan	Jl. Raya Bangkingan RT.02 RW.II	Lakarsantri	-7.324778	112.656186
14	Wiyung	Jl. Raya Menganti Gg. Pasar No.20	Wiyung	-7.314888	112.693958
15	Balas Klumprik	Jl. Raya Balas Klumprik	Wiyung	-7.331485	112.689686
16	Jagir	Jl. Bendul Merisi No.1	Wonokromo	-7.30565	112.737715
17	Wonokromo	Jl. Karang Rejo VI/4	Wonokromo	-7.303172	112.730406
18	Ngagelrejo	Jl. Ngagel Dadi III/17	Wonokromo	-7.292975	112.748776
19	Menur	Jl. Manyar Rejo I/35	Sukolilo	-7.296813	112.76426
20	Klampus Ngasem	Jl. Arief Rachman Hakim 99 B	Sukolilo	-7.289151	112.778523
21	Keputih	Jl. Keputih Tegal No 1	Sukolilo	-7.294031	112.801892
22	Dukuh Kupang	Jl. Dukuh Kupang XXV/48	Dukuhpakis	-7.28366	112.7128716
23	Pucangsewu	Jl. Pucang Anom Timur 72 Surabaya	Gubeng	-7.286582	112.755579
24	Mojo	Jl. Mojo Klanggru Wetan II/11	Gubeng	-7.270769	112.770906
25	Sawahan	Jl. Raya Arjuna 119	Sawahan	-7.265563	112.727666

No	Klinik/Puskesmas	Alamat	Kecamatan	Latitude	Longitude
26	Putat Jaya	Jl. Kupang Gunung Raya No. 5 Surabaya	Sawahan	-7.27956	112.722179
27	Banyu Urip	Jl. Banyu Urip Kidul VI/8	Sawahan	-7.271496	112.720143
28	Pakis	Jl. Makam Kembang Kuning No.6	Sawahan	-7.281721	112.728128
29	Tanjungsari	Jl. Tanjungsari 116	Sukomanunggal	-7.26072	112.6917967
30	Simomulyo	Jl. Gumuk Bogo VI/1	Sukomanunggal	-7.26236	112.7117007
31	Mulyorejo	Jl. Mulyorejo Utara 201 Blk	Mulyorejo	-7.260633	112.784748
32	Kalijudan	Jl. Kalijudan 123	Mulyorejo	-7.257394	112.777996
33	Kedungdoro	Jl. Kaliasin Gg. Pompa No.7981	Tegalsari	-7.263618	112.73819
34	Dr. Soetomo	nomor A, Jl. Pandegiling No.223, DR. Soetomo, Tegalsari	Tegalsari	-7.274391	112.728012
35	Manukan Kulon	Jl. Manukan Dalam I/18	Tandes	-7.262901	112.668869
36	Balongsari	Jl. Balongsari Tama No.1	Tandes	-7.259268	112.678375
37	Lontar	Jl. Raya Lontar No.26	Sambikerep	-7.285085	112.662798
38	Made	Jl. Raya Made RT 1	Sambikerep	-7.277289	112.636994
39	Peneleh	Jl. Makam Peneleh 35	Genteng	-7.252261	112.74078
40	Ketabang	Jl. Jaksa Agung Suprpto No 10	Genteng	-7.25804	112.746786
41	Rangkah	Jl. Rangkah VII/94	Tambaksari	-7.246231	112.767253
42	Pacar Keling	Jl. Jolotundo Baru III/16	Tambaksari	-7.25733	112.762614
43	Gading	Jl. Kapasari Lor II/1	Tambaksari	-7.242454	112.763946
44	Tembok Dukuh	Jl. Kalibutih No. 26	Bubutan	-7.254616	112.719182
45	Gundih	Jl. Margodadi 3638	Bubutan	-7.247139	112.725264
46	Tambakrejo	Jl. Ngaglik 87	Simokerto	-7.249203	112.755487
47	Simolawang	Jl. Simolawang II Barat No. 45 A	Simokerto	-7.234637	112.751128
48	Kenjeran	Jl. Tambak Deres No.2	Bulak	-7.241059	112.795171
49	Perak Timur	Jl. Jakarta No.9	Pabean Cantikan	-7.222639	112.735446
50	Krembangan Selatan	Jl. Pesapen Selatan 70	Krembangan	-7.233201	112.731197
51	Dupak	Jl. Dupak Bangunrejo Gg. Poliklinik No.6	Krembangan	-7.239304	112.717205
52	Morokrembangan	Jl. Tambak Asri XIII/17	Krembangan	-7.234797	112.712434
53	Asemrowo	Jl.Asem Raya 8	Asemrowo	-7.22968	112.687205

No	Klinik/Puskesmas	Alamat	Kecamatan	Latitude	Longitude
54	Benowo	Jl. Raya Benowo RT.01 RW.I	Pakal	-7.234749	112.611528
55	Pegirian	Jl. Karang Tembok 39	Semampir	-7.224685	112.747206
56	Sidotopo	Jl. Sidodadi II	Semampir	-7.230223	112.7511285
57	Wonokusumo	Jl. Wonokusumo Tengah 55	Semampir	-7.219678	112.752621
58	Sawah Pulo	Jl. Sawah Pulo Lapangan No. 2	Semampir	-7.222798	112.744694
59	Tanah Kali Kedinding	Jl. Kedung Cowek 226	Kenjeran	-7.230306	112.774557
60	Sidotopo Wetan	Jl. Randu 102	Kenjeran	-7.226568	112.763586
61	Bulak Banteng	Jl. Bulak Banteng Lor I No 27	Kenjeran	-7.217702	112.76623
62	Tambak Wedi	Jl. Tambak Wedi	Kenjeran	-7.217434	112.771583
63	Sememi	Jl. Raya Kendung	Benowo	-7.248419	112.63539
64	Klinik Pratama Widya Mandiri	Rungkut Puskesmas No 1 Kalirungkut	Rungkut	-7.321419	112.771224
65	HOTEL BUMI SURABAYA	Jl. Basuki Rahmat No. 106 - 128	Genteng	-7.271085	112.741174
66	RAJAWALI	Jl. Krembangan Barat 26D	Krembangan	-7.237733	112.733444
67	SARTIKA 44	Jl. Mastrip No.32	Karangpilang	-7.340419	112.666338
68	PT. GELORA DJAJA	Jl. Buntaran No. 9 Tandes	Tandes	-7.252222	112.673278
69	MEDICO SIER	Ruko Galaxy Sier Jl. Rungkut Industri III No. 7 G	Tenggilismejoyo	-7.335727	112.755568
70	PONDOK KASIH	Jl. Gayungan PTT No. 68	Gayungan	-7.335115	112.725762
71	dr. EKO	Jl. Medayu Selatan IV No. 15	Rungkut	-7.332001	112.796653
72	PLK - UA KAMPUS B UNAIR	Jl. Dharmawangsa No. 3	Gubeng	-7.270598	112.758397
73	SANTA ANNA	Jl. Bronggalan Sawah 5A No. 32 - 34	Tambaksari	-7.25665	112.768237
74	PERDANA HUSADA	Jl. Manukan Krajan 32Q No. 2	Tandes	-7.259822	112.670674
75	SARTIKA 59	Jl. Manukan Tengah 9K / 10	Sambikerep	-7.265307	112.651531
76	NURANI JAYA 83	Jl. KH. Abdul Karim No. 17	Gununganyar	-7.338898	112.772253
77	KARYA MEDIKA 111	Jl. Bukit Citra Klakah Rejo I / 20	Benowo	-7.253901	112.649761
78	ALAMANDA	Jl. Dinoyo No. 20 / II	Tegalsari	-7.283634	112.745978
79	DIAN	Jl. Gang Kramat (Komplek Perumahan PT. Dian Istana) Wiyung	Wiyung	-7.308774	112.697057

No	Klinik/Puskesmas	Alamat	Kecamatan	Latitude	Longitude
80	PHC BENOWO	Jl. Benowo Indah Tengah XIII / 12	Pakal	-7.236322	112.627166
81	STIESIA	Jl. Manyar Kartika VIII / 36	Sukolilo	-7.288962	112.766121
82	NAYAKA HUSADA 07	Jl. Tandes Asri No. 64	Tandes	-7.255444	112.673498
83	KLINIK DOKTER AYOMAN KELUARGA MULYOSARI	Jl. Sutorejo Prima Utara PDD - 1	Mulyorejo	-7.259423	112.793927
84	TABITA	Jl. Lebak Jaya III No. 41	Tambaksari	-7.244343	112.777982
85	AT - TAUFIQ	Jl. Rungkut Menanggal I No. 6 (lama : Jl. Rungkut Menanggal I /JA-9)	Gununganyar	-7.338046	112.775952
86	NURANI JAYA 37	Jl. Kendangsari III / 20	Tenggilismejoyo	-7.323999	112.744126
87	ST. VINCENTIUS A PAULO (KARAH)	Jl. Karah No. 200	Jambangan	-7.30919	112.713639
88	POLRESTABES SURABAYA	Jl. Rajawali No. 43	Krembangan	-7.235411	112.734623
89	65 (LXV)	Jl. Gubeng Kertajaya XI No. 34	Gubeng	-7.27745	112.759377
90	IN HOUSE CLINIC JW MARRIOT HOTEL	Jl. Embong Malang No. 85 - 89	Tegalsari	-7.25969	112.734735
91	DINAYLA UTAMA 84	Jl. Simo Pomahan III / 14	Sukomanunggal	-7.261184	112.706818
92	ABDI MULIA	Jl. Wisma Permai Barat III / FP. 12	Mulyorejo	-7.272308	112.785056
93	NAYAKA HUSADA 42	Jl. Kalirungkut No:52	Rungkut	-7.318902	112.767127
94	dr. Subur Prajitno	Jl.Raya Mulyosari 286	Mulyorejo	-7.268151	112.796086
95	Yakes Telkom	Jl. Kanwa No:15	Wonokromo	-7.284814	112.730804
96	CV.Anugerah Karya Medika	Jl.Demak 375	Krembangan	-7.236025	112.720466
97	KESAYANGAN	Jl.Rungkut Asri Utara IV / 10 ( lama : Jl. Rungkut Asri Utara Blok RL.1.H-5 )	Rungkut	-7.317582	112.718561
98	Abdihusada Utama	Jl. Raya Mulyosari 42-D	Mulyorejo	-7.260182	112.795224
99	PT.HM Sampoerna Tbk	Jl. Taman Sampoerna No 6	Pabean Cantikan	-7.230805	112.734235
100	PT.HM Sampoerna Tbk	Jl. Kali Rungkut No : 9-11	Rungkut	-7.314289	112.769855
101	Cita Husada	Jl. Raya Kedung Asem No :100	Rungkut	-7.322038	112.776239
102	FIRST CARE CLINIC	Jl. Rungkut Industri III / 52	Gununganyar	-7.336322	112.761135

No	Klinik/Puskesmas	Alamat	Kecamatan	Latitude	Longitude
103	PT.SC JOHNSON MANUFACTURING SURABAYA	Jl. Rungkut Industri II / 24-26	Rungkut	-7.327335	112.759702
104	PT.HM Sampoerna Tbk	Jl. Rungkut Industri Raya No 18	Rungkut	-7.331116	112.762053
105	DON BOSCO	Jl. Tidar No. 113	Sawahan	-7.256916	112.721899
106	YAYASAN BK3S JATIM	Jl. Raya Tenggilis Blok GG No. 10	Tenggilismejoyo	-7.320921	112.753589
107	PT. Perusahaan Gas Negara (Persero) Tbk SBU DW II	Jl. Pemuda No. 56 - 58	Genteng	-7.265171	112.746548
108	YAYASAN KANKER WISNUWARDHANA	Jl. Kayoon No. 16 - 18	Genteng	-7.267474	112.748925
109	NAYAKA HUSADA 03	Jl. Asemrowo Kali No. 1	Asemrowo	-7.248332	112.716064
110	dr. EKO	Jl. Manukan Lor III - K No. 08	Tandes	-7.259858	112.665192
111	SIDOTOPO	Jl. Sidotopo Lor No. 8	Semampir	-7.23374	112.754474
112	BULAN SABIT MERAH INDONESIA	Jl. Mojo III No. 33	Gubeng	-7.270318	112.765229
113	STIKES YARSIS/unusa A	Jl. SMEA No. 57	Wonokromo	-7.304622	112.733858
114	KELUARGA BERENCANA BHAKTI SETIA KENCANA	Jl. Nginden Permata No. 1	Sukolilo	-7.303642	112.768008
115	ANUGERAH KARYA MEDIKA	Jl. Gunungsari No. 124	Wonokromo	-7.304509	112.725069
116	HEALTHY	Jl. Ngagel Tirto No. 25 - 27	Wonokromo	-7.288861	112.745523
117	JAMSOSTEK 122/ KLINIK SURYA GIRI JAYA 122	Jl. Kenjeran No. 189	Tambaksari	-7.245403	112.771001
118	BP 93	Jl. Perak Timur No. 564 C - 3	Pabean Cantikan	-7.216471	112.733135
119	DINAYLA UTAMA 105	Jl. Raya Bibis No. 36	Tandes	-7.257755	112.675223
120	KESEJAHTERAAN 11	Jl. Rungkut Mapan Utara No. 8	Gununganyar	-7.331481	112.777071
121	DINAYLA UTAMA 81	Jl. Margorejo Baru No. 32	Wonocolo	-7.316445	112.735293
122	REDEMPTOR MUNDI	Jl. Dukuh Kupang VI No. 35	Dukuhpakis	-7.285046	112.710603
123	ADI HAYATI	Jl. Ketintang Baru Selatan I / 56	Gayungan	-7.324736	112.72512
124	KEBANGKITAN	Jl. Manukan Madya No. 141	Tandes	-7.259765	112.671408
125	MITRA SEHAT	Jl. Raya Menganti Wiyung No. 411	Wiyung	-7.313053	112.690443
126	KARYA MEDIKA 104	Jl. Banyu Urip No. 302	Sawahan	-7.271546	112.723159

No	Klinik/Puskesmas	Alamat	Kecamatan	Latitude	Longitude
127	CAHAYA DAMAI	Jl. Raya Pradah Indah No. 9	Sambikerep	-7.281011	112.678984
128	SAMARIA	Jl. Penghela No. 42	Bubutan	-7.248349	112.73405
129	AL - MUTTAQIEN	Jl. Memet Sastrawiryo No. 43-45, Komp TNI AL	Bulak	-7.24455	112.790102
130	SARTIKA 36	Jl. Bratang Binangun VI No. 47	Gubeng	-7.291883	112.758045
131	JEMURSARI EMPAT	Jl. Jemursari No. 4	Wonocolo	-7.326748	112.736013
132	GOTONG ROYONG I	Jl. Raya Manyar No.49, Menur Pumpungan, Sukolilo	Sukolilo	-7.291477	112.7627
133	ATKP SURABAYA	Jl. Jemur Andayani I / 73	Wonocolo	-7.333823	112.736002
134	Q - LIFE KLINIK	Jl. Abdul Wahab Siamin Blok RE Kav. 8	Dukuhpakis	-7.298242	112.707383
135	DINAYLA UTAMA 79	Jl. Simorejo Sari B / 54	Sukomanunggal	-7.256384	112.706846
136	NURANI JAYA 121	Jl. Putro Agung Wetan No. 58 N	Tambaksari	-7.250462	112.768218
137	NURANI JAYA 66	Jl. Kutisari Selatan II No. 4	Tenggilismejoyo	-7.334224	112.733709
138	SURYA BIO MEDIKA	Jl. Raya Wisma Lidah Kulon Ruko 8 No. 5 - 6	Lakarsantri	-7.306529	112.653783
139	DR. H. J. MUKONO	Jl. Sawahan Sarimulyo No. 8	Sawahan	-7.255525	112.715408
140	SENTUHAN KASIH BANGSA	Jl. Kedondong Lor V / 4	Genteng	-7.271006	112.743574
141	ELYON	Jl. Raya Satelit Indah AT - 1	Sukomanunggal	-7.264406	112.694464
142	MEDICAL CENTER ITS	Jl. Arief Rahman Hakim No. 213	Sukolilo	-7.289798	112.784934
143	KARYA MEDIKA 41	Jl. Ngaglik No. 47	Genteng	-7.250162	112.749003
144	PALANG MERAH INDONESIA	Jl. Tambaksari No. 49	Tambaksari	-7.254507	112.756811
145	PUSVETMA	Jl. A. Yani No. 68 - 70	Gayungan	-7.316941	112.73084
146	MedPoint	Jl. Raya Lontar 229 ( No. Lama . 86 )	Sambikerep	-7.284031	112.670804
147	PERTAMINA JAGIR	Jl. Jagir Wonokromo No. 88	Wonokromo	-7.301511	112.741885
148	PETEMON Mitra Medika	Jl. Petemon III / 3	Sawahan	-7.267867	112.718999
149	PHC KEBRAON	Jl. Griya Kebraon Selatan FA No. 37 - 38	Karangpilang	-7.332817	112.701269
150	KENCANA MEDIKA	Jl. Kendangsari No. 78	Wonocolo	-7.322144	112.746106

No	Klinik/Puskesmas	Alamat	Kecamatan	Latitude	Longitude
151	PT. UNILEVER INDONESIA Tbk	Jl. Rungkut Industri IV No. 5 - 11	Tenggilismejoyo	-7.332883	112.753374
152	YAKES TELKOM	Jl. Ketintang No. 152 A	Gayungan	-7.309721	112.727701
153	UJUNG SELATAN	Komplek Kolat Armatim Jl. Raya Hang Tuah	Semampir	-7.195533	112.741479
154	UJUNG UTARA	Komplek Makoarmatim Dermaga Ujung	Semampir	-7.197896	112.73467
155	OPTIMA	Jl. Rungkut Mapan Utara CA No. 20	Gununganyar	-7.314961	112.769046
156	MIFTACHUL MUNIR MEDIKA	Jl. Raya Lontar No. 117 (lama Jl. Raya Lontar No. 190)	Sambikerep	-7.282921	112.662585
157	BHAYANGKARA M. DAHLAN	Jl. Sriti No. 2	Krembangan	-7.238169	112.735251
158	MUSLIMAT NU 01	Jl. Wonocolo Pabrik Kulit No. 71	Wonocolo	-7.318885	112.738582
159	PHC TANJUNG PERAK	Jl. Perak Barat No. 239 A	Pabean Cantikan	-7.216631	112.732728
160	PRAMESWARI	Jl. Gubeng Kertajaya Vc No. 24	Gubeng	-7.276153	112.754656
161	KLINIK WIDYA MANDIRI I	Jl. Kalirungkut 27 D - 18	Rungkut	-7.318977	112.767105
162	MITRA MEDICARE	Jl. Dharmahusada Utara 36 - 38 Blok G - J	Gubeng	-7.264755	112.772402
163	ALBA MEDIKA	Jl. Ploso Baru No. 73A	Tambaksari	-7.256291	112.77048
164	WALUYO JATI SURABAYA	Jl. Residen Sudirman No. 10	Tambaksari	-7.256401	112.753834
165	MEDIKA YANI II	Jl. Ngagel Rejo Utara 46 - 48	Wonokromo	-7.293974	112.746315
166	KARTINI	Jl. Cipta Menanggal I No. 11A	Gayungan	-7.340933	112.722265
167	SANTA MARIA TAK BERCELA	Jl. Ngagel Jaya Indah II / 2	Gubeng	-7.28886	112.760884
168	MEDIKA dr. HALIM	Jl. Rungkut Asri Timur XVIII No. 18	Rungkut	-7.330741	112.787813
169	BRI MEDIKA SURABAYA	Jl. A. Yani No. 169	Wonocolo	-7.332503	112.731105
170	SEBELAS MEDIKA	Jl. Bung Tomo No. 5 - 7	Wonokromo	-7.290479	112.746648
171	QUALITA MEDIKA	Jl. Pucang Sewu No.41	Gubeng	-7.285439	112.751054
172	PLK - UA KAMPUS C UNAIR	Jl. Mulyorejo Sukolilo Kampus C Unair	Mulyorejo	-7.266665	112.783312
173	POLIKLINIK YAYASAN PEGAWAI BANK JATIM	Jl. Basuki Rahmat No. 98 - 104	Genteng	-7.270419	112.741281

No	Klinik/Puskesmas	Alamat	Kecamatan	Latitude	Longitude
174	NAYAKA HUSADA 02	Jl.Mastrip No: 50	Karangpilang	-7.343203	112.693051
175	NAYAKA HUSADA 01	Jl. Cendrawasih No : 28-B	Krembangan	-7.238276	112.737075
176	NAYAKA HUSADA 05	Jl. Pandegiling No 117A	Tegalsari	-7.275421	112.733855
177	UK PETRA	Jl. Siwalan Kerto No. 103 -105	Wonocolo	-7.33694	112.738201
178	UBAYA	Jl. Kaliwaru I No.31, Kali Rungkut, Rungkut, Kota SBY, Jawa Timur 60293	Rungkut	-7.317789	112.767614
179	GUNUNGSARI SURABAYA	Jl. Golf I No. 1A Gunungsari	Dukuhpakis	-7.305754	112.708805
180	SUMBER CAHAYA ANUGERAH	Jl. Kutisari Selatan No. 1 A	Tenggilismejoyo	-7.334668	112.748229
181	CEMARA MEDIKA	Puri Lidah Kulon Blok A No. 8 - 9	Lakarsantri	-7.301394	112.652701
182	KESATRIAN	Jl. Kesatrian No. 41	Wonokromo	-7.297064	112.724002
183	MEDISKA SURABAYA GUBENG	Jl. Indrakila 2	Tambaksari	-7.260444	112.756052
184	KITA KERTAJAYA	Jl. Kertajaya 160	Gubeng	-7.278738	112.757287
185	AMANINA MEDIKA	Jl. Mojo Kidul 113 B	Gubeng	-7.269327	112.768936
186	PARADISE	Jl. Raya Gunung Anyar Kidul No. 33	Gununganyar	-7.335867	112.776012
187	RUMAH SUNATAN	Jl. Ngagel Jaya	Gubeng	-7.289478	112.755547
188	KELSAPA	Jl. Kepanjen No. 4 - 6	Krembangan	-7.241995	112.736781
189	SUTEDI SENAPUTRA	Jl. Ksatria No. 28 Karangpilang	Karangpilang	-7.338414	112.696549
190	HIDAYAH	Pakis Tirtosari VI / 4 A	Sawahan	-7.289142	112.722519
191	POLY PT PAKUWON PERMAI	Jl. Puncak Indah Lontar No. 2	Wiyung	-7.289551	112.677005
192	DINAS PENDIDIKAN PROVINSI JAWA TIMUR	Jl. Genteng Kali No. 33	Genteng	-7.257336	112.742853
193	AL - FITRAH	Jl. Kedinding Lor 110 A	Kenjeran	-7.221623	112.770816
194	ANDALAN SEHAT	JL. Raya Arjuna No. 78 C	Sawahan	-7.262595	112.726137
195	SANTA ANNA	Jl. Babatan Rukun VII / 32	Tambaksari	-7.244644	112.718131
196	WACHID HASJIM	JL. Sidotopo Wetan Baru 27	Kenjeran	-7.232768	112.759718
197	SG. DENTAL SOLUTION	Jl. Musi No. 33	Wonokromo	-7.287263	112.732843
198	INDAH DENTAL CARE	Jl. Biliton 31	Gubeng	-7.272888	112.750187
199	GIGI BHAKTI KASIH	Jl. Raya Darmo No. 175	Sawahan	-7.261146	112.715068

No	Klinik/Puskesmas	Alamat	Kecamatan	Latitude	Longitude
200	JET - Z DENTAL AESTHETIC CLINIC	Jl. Abdul Wahab Siamin Blok RC - 31	Dukuhpakis	-7.301183	112.706562
201	JET - Z DENTAL AESTHETIC CLINIC	Jl. Manyar Rejo III No.3, Menur Pumpungan, Kec. Sukolilo	Sukolilo	-7.294525	112.767874
202	ESTHETIC DENTAL CARE	Jl. HR. Mohammad No. 49 - 55	Sukomanunggal	-7.285	112.696976
203	HARAPAN SEHAT	Jl. Gayungsari Barat III / 47 - 49	Jambangan	-7.332118	112.746106
204	BK KENJERAN TNIAL	Jl. Wiratno No. 55	Kenjeran	-7.243408	112.792986
205	ANUGRAH	Jl. Dukuh Kupang XIX No. 39 - 39A	Dukuhpakis	-7.281033	112.713876
206	RBG - RZ	Jl. Sidosermo II Kav. 321	Wonocolo	-7.308924	112.750262
207	WISMA HUSADA	Jl. Dukuh Setro VII - A / Kav No. 2	Tambaksari	-7.241477	112.774159
208	YAYUK ISMAIL	Wisma Indah A-2, Jl. Raya Pandugo 203	Rungkut	-7.317713	112.787551
209	GRIYA HUSADA	Jl. Dukuh Pakis II / 110	Dukuhpakis	-7.292359	112.704617
210	NUR MEDIKA	JL. Simo Pomahan I / 8-A	Sukomanunggal	-7.261904	112.706486
211	BUNDA MULIA	Jl. Gunung Anyar Pratama Blok B No. 3	Gununganyar	-7.343109	112.800034
212	IBUNDA	Jl. Demak No. 318	Krempangan	-7.23642	112.72203
213	SIER SEJAHTERA	Jl. Rungkut Industri Raya No. 10 A	Gununganyar	-7.332118	112.758888
214	SUBDITKES AAL	Jl. Bumimoro Morokrempangan	Krempangan	-7.216957	112.717842
215	dr. IDRIS P. SIREGAR	Komplek Rajawali Denmako Koarmatim Jl. Pati Unus	Semampir	-7.209652	112.737894
216	BDS PACUAN KUDA	Jl. Raya Pacuan Kuda No. 15 A	Sawahana	-7.258123	112.717434
217	KLINIK RANAP BERSALIN AL-AZHAR	Jl. Dupak Bandarejo No. 23	Krempangan	-7.242113	112.718683
218	PUTRI RAHAYU	Jl. Mastrip IX No. 9 Karang Pilang	Karangpilang	-7.343938	112.666508

## Lampiran 2 Variabel Kovariat

Kecamatan	Latitude	Longitude	Kepadatan Penduduk	PHBS	AK	Jalan Baik	FKTP
GAYUNGAN	-7.338	112.717	72.247	92.083	60	73.428	6
KARANGPILANG	-7.333	112.699	76.684	94.584	75	40.855	6
GUNUNGANYAR	-7.341	112.783	54.422	98.755	71	47.694	9
JAMBANGAN	-7.322	112.714	114.683	87.797	67	24.331	3
TENGGILISMEJOYO	-7.314	112.757	99.525	81.809	58	97.536	7
WONOCOLO	-7.320	112.741	114.938	90.862	52	69.000	11
RUNGKUT	-7.323	112.771	50.214	96.743	63	190.521	14
LAKARSANTRI	-7.303	112.633	30.470	80.011	84	22.087	5
WIYUNG	-7.314	112.695	53.423	89.286	62	40.262	5
WONOKROMO	-7.292	112.732	111.324	84.989	43	59.153	12
SUKOLILO	-7.300	112.770	44.485	83.033	61	109.331	8
DUKUHPAKIS	-7.283	112.688	58.603	87.920	54	30.890	7
GUBENG	-7.272	112.756	167.681	87.051	57	79.433	14
SAWAHAN	-7.280	112.717	236.240	85.589	54	56.355	12
SUKOMANUNGGAL	-7.261	112.712	106.596	77.729	52	90.294	7
MULYOREJO	-7.261	112.785	119.595	90.162	69	99.923	7
TEGALSARI	-7.288	112.741	234.326	79.490	52	42.961	5
TANDES	-7.259	112.678	80.210	78.688	63	81.368	8
SAMBIKEREP	-7.266	112.655	31.557	87.922	75	36.931	6
GENTENG	-7.258	112.752	144.149	96.915	60	38.994	9
TAMBAKSARI	-7.257	112.755	241.979	93.131	65	59.016	12
BUBUTAN	-7.245	112.737	260.972	79.675	56	48.718	3
SIMOKERTO	-7.244	112.758	373.722	85.421	83	32.272	2
BULAK	-7.232	112.786	71.153	98.430	83	17.792	3
PABEAN CANTIKAN	-7.217	112.729	117.318	71.748	60	28.301	5
KREMBANGAN	-7.233	112.723	138.772	73.152	48	41.609	13
ASEMROWO	-7.252	112.715	29.391	62.290	53	12.587	2
PAKAL	-7.240	112.625	22.612	93.172	80	7.950	2
SEMAMPIR	-7.225	112.745	212.491	84.601	66	25.482	7
KENJERAN	-7.227	112.776	177.455	75.769	75	28.453	6
BENOWO	-7.249	112.635	24.215	91.155	77	29.474	2

### Lampiran 3 Variabel Kovariat Hasil Smoothing

Kecamatan	Latitude	Longitude	X1	X2	X3	X4
GAYUNGAN	-7.338	112.717	97.551	89.666	60.614	64.653
KARANGPILANG	-7.333	112.699	77.032	91.274	68.948	48.910
GUNUNGANYAR	-7.341	112.783	57.158	95.869	68.344	93.937
JAMBANGAN	-7.322	112.714	92.390	89.325	63.162	56.815
TENGGILISMEJOYO	-7.314	112.757	88.852	88.908	58.056	80.517
WONOCOLO	-7.320	112.741	104.494	88.214	56.652	72.735
RUNGKUT	-7.323	112.771	57.157	93.115	66.029	94.523
LAKARSANTRI	-7.303	112.633	42.708	85.009	75.457	38.804
WIYUNG	-7.314	112.695	67.679	89.369	63.722	47.427
WONOKROMO	-7.292	112.732	147.895	85.496	52.700	65.370
SUKOLILO	-7.300	112.770	87.129	88.926	64.788	84.084
DUKUHPAKIS	-7.283	112.688	110.494	85.490	55.987	52.624
GUBENG	-7.272	112.756	162.895	87.010	59.103	65.622
SAWAHAN	-7.280	112.717	156.828	81.379	52.725	53.932
SUKOMANUNGGAL	-7.261	112.712	102.582	80.180	54.823	50.459
MULYOREJO	-7.261	112.785	130.188	89.312	67.870	66.429
TEGALSARI	-7.288	112.741	180.704	84.137	54.033	58.083
TANDES	-7.259	112.678	60.053	83.320	65.949	45.391
SAMBIKEREP	-7.266	112.655	39.109	86.128	72.755	39.035
GENTENG	-7.258	112.752	209.341	85.405	61.175	53.201
TAMBAKSARI	-7.257	112.755	196.759	88.247	70.205	52.366
BUBUTAN	-7.245	112.737	166.411	79.412	56.373	48.545
SIMOKERTO	-7.244	112.758	236.050	85.000	67.642	46.802
BULAK	-7.232	112.786	131.759	88.661	77.147	47.601
PABEAN CANTIKAN	-7.217	112.729	170.296	78.160	60.123	41.272
KREMBANGAN	-7.233	112.723	150.939	76.324	55.872	43.871
ASEMROWO	-7.252	112.715	84.503	76.972	58.302	46.013
PAKAL	-7.240	112.625	23.125	92.268	79.219	24.034
SEMAMPIR	-7.225	112.745	188.488	80.322	65.019	39.491
KENJERAN	-7.227	112.776	180.981	84.461	74.568	40.350
BENOWO	-7.249	112.635	27.236	90.336	77.388	33.110

## Lampiran 4 Syntax R Spatial Point Process

```
#Membentuk obyek tessellation berdasarkan bentuk kecamatan
library(maptools)
library(sp)
library(spatstat)
library(spatstat.data)
library(nlme)
library(rpart)
library(foreign)
library(rgdal)
x <- readShapeSpatial("D://3578_surabaya_kec_new.shp")
class(x)
W=as(x, "owin")
regions <- slot(x, "polygons")
regions <- lapply(regions, function(x) { SpatialPolygons(list(x)) })
windows <- lapply(regions, as.owin)
te <- tess(tiles=windows)
plot(te)
#Plot titik-titik lokasi FKTP dan menghitung banyak titik pada masing-masing kecamatan
xy <- read.table("D:// klinik.csv", header=T, sep=";")
head(xy)
attach(xy)
ppall <- ppp(xy$Longitude, xy$Latitude, window=W, marks=xy$Type)
plot(ppall)
plot(density.ppp(ppall)/(111.322^2))
contour(density(ppall)/(111.322^2),add=TRUE)
kec <- quadratcount(ppall, tess = te)
#Transformasi pixel variabel kovariat
xx=read.table("D://Covariat.csv", header=T, sep=";")
KP<- ppp(xx$Longitude, xx$Latitude, window=W, marks=KP)
SP=Smooth(KP ,at="pixels")
X1=split.im(SP,te)
summary(X1)
PHBS<- ppp(xx$Longitude, xx$Latitude, window=W, marks=PHBS)
SPHBS=Smooth(PHBS ,at="pixels")
X2=split.im(SPHBS,te)
summary(X2)
AK<- ppp(xx$Longitude, xx$Latitude, window=W, marks=AK)
SAK=Smooth(AK ,at="pixels")
X3=split.im(SAK,te)
summary(X3)
JB<- ppp(xx$Longitude, xx$Latitude, window=W, marks=JB)
plot(JB)
SJB=Smooth(JB ,at="pixels")
X4=split.im(SJB,te)
Summary(X4)
```

## Lampiran 5 Program GLMs Mixture Poisson Regression

```
library(flexmix)
library(lattice)
data=read.csv("D://Data_Tesis.csv",sep=";")
head(data)
attach(data)
library(vcd)
gf=goodfit(Y, type="poisson",method="MinChisq")
summary(gf)
library(qcc)
qcc.overdispersion.test(Y)
chisq.test(Y)

model0<-flexmix(Y~1,data=data,k=2,model=FLXglm(family="poisson"))
summary(model0)
rmodel0<-refit(model0)
summary(rmodel0)
model1<-flexmix(Y~X1,data=data,k=2,model=FLXglm(family="poisson"))
summary(model1)
rmodel1<-refit(model1)
summary(rmodel1)
model2<-flexmix(Y~X2,data=data,k=2,model=FLXglm(family="poisson"))
summary(model2)
rmodel2<-refit(model2)
summary(rmodel2)
model3<-flexmix(Y~X3,data=data,k=2,model=FLXglm(family="poisson"))
summary(model3)
rmodel3<-refit(model3)
summary(rmodel3)
model4<-flexmix(Y~X4,data=data,k=2,model=FLXglm(family="poisson"))
summary(model4)
rmodel4<-refit(model4)
summary(rmodel4)
```

## Lampiran 6 Program Regresi Poisson

```
model;
{
  for( i in 1 : N ) {
    y[i] ~ dpois(mu[i])
  }
  for( i in 1 : N ) {
    log(mu[i]) <- 10 + 11 * x1[i]+12 * x2[i]+13 * x3[i]+14 * x4[i]
  }

  10 ~ dnorm(1.8356,1)I(1.5,4)
  11~ dnorm(0.1873,0.002)I(-1,1)
  12 ~ dnorm(0.1734,0.001)I(-1,3)
  13~ dnorm(-0.2335,0.002)I(-1,0)
  14~ dnorm(0.15,0.002)I(0,1)

}
INITS
list(c(2,0.3,0.1,0.05,2))

DATA
list(y=c(6,6,9,3,7,11,14,5,5,12,8,7,14,12,7,7,5,8,6,9,12,3,2,3,5,13,2,2,7,6,2),N=31,x1=c(-
0.336,-0.690,-1.033,-0.426,-0.487,-0.217,-1.033,-1.283,-0.852,0.532,-0.516,-
0.113,0.791,0.686,-0.250,0.226,1.098,-0.983,-
1.345,1.592,1.375,0.851,2.053,0.254,0.918,0.584,-0.562,-1.620,1.232,1.103,-
1.549),x2=c(0.748,1.081,2.033,0.677,0.591,0.447,1.463,-0.217,0.687,-0.116,0.595,-
0.117,0.198,-0.969,-1.217,0.675,-0.397,-0.566,0.015,-0.135,0.454,-1.376,-0.218,0.540,-
1.635,-2.016,-1.881,1.287,-1.188,-0.330,0.887),x3=c(-0.437,0.630,0.553,-0.110,-0.764,-
0.944,0.257,1.463,-0.039,-1.450,0.098,-1.029,-0.630,-1.446,-1.178,0.492,-
1.279,0.246,1.118,-0.365,0.791,-0.979,0.463,1.680,-0.499,-1.044,-
0.733,1.945,0.127,1.350,1.711),x4=c(0.589,-0.344,2.326,0.125,1.530,1.069,2.360,-0.943,-
0.432,0.632,1.741,-0.124,0.647,-0.046,-0.252,0.695,0.200,-0.552,-0.929,-0.089,-0.139,-
0.365,-0.469,-0.421,-0.797,-0.643,-0.516,-1.819,-0.902,-0.851,-1.281))
```



## Lampiran 8 Program Model Simulasi *Mixture* Poisson Regression

```
model;
{
  for( i in 1 : N ) {
    y[i] ~ dpois(mu[i])
  }
  for( i in 1 : N ) {
    log(mu[i]) <- 10[T[i]] + 11[T[i]] * x1[i]+12[T[i]] * x2[i]+13[T[i]] * x3[i]+14[T[i]] *
x4[i]
  }
  for( i in 1 : N ) {
    T[i] ~ dcat(P[1:2])
  }

  P[1:2] ~ ddirch(alpha[])
  10[1] ~ dnorm(1.716,0.1)I(1,2.5)
  10[2] ~ dnorm(1.622,0.1)I(0.5,2.5)
  11[1] ~ dnorm(0,0.1)I(-2,2)
  11[2] ~ dnorm(0.7575,0.1)I(0.1,1.3)
  12[1] ~ dnorm(-0.02607,0.1)I(-2,2)
  12[2] ~ dnorm(-0.7642,0.1)I(-1.3,-0.25)
  13[1] ~ dnorm(-0.05878,0.1)I(-1,1)
  13[2] ~ dnorm(0.2829,0.1)I(0,0.5)
  14[1] ~ dnorm(0.2453,0.1)I(0,0.5)
  14[2] ~ dnorm(1.218,0.1)I(0,2.5)

}
INITS
list(10=c(1.5,1.5),11=c(0,0.5),12=c(0,-0.5),13=c(0,0.25),14=c(0.25,1.2))

DATA
list(y=c(6,6,9,3,7,11,14,5,5,12,8,7,14,12,7,7,5,8,6,9,12,3,2,3,5,13,2,2,7,6,2),N=31,
alpha=c(1,1),x1=c(-0.336,-0.690,-1.033,-0.426,-0.487,-0.217,-1.033,-1.283,-0.852,0.532,-
0.516,-0.113,0.791,0.686,-0.250,0.226,1.098,-0.983,-
1.345,1.592,1.375,0.851,2.053,0.254,0.918,0.584,-0.562,-1.620,1.232,1.103,-
1.549),x2=c(0.748,1.081,2.033,0.677,0.591,0.447,1.463,-0.217,0.687,-0.116,0.595,-
0.117,0.198,-0.969,-1.217,0.675,-0.397,-0.566,0.015,-0.135,0.454,-1.376,-0.218,0.540,-
1.635,-2.016,-1.881,1.287,-1.188,-0.330,0.887),x3=c(-0.437,0.630,0.553,-0.110,-0.764,-
0.944,0.257,1.463,-0.039,-1.450,0.098,-1.029,-0.630,-1.446,-1.178,0.492,-
1.279,0.246,1.118,-0.365,0.791,-0.979,0.463,1.680,-0.499,-1.044,-
0.733,1.945,0.127,1.350,1.711),x4=c(0.589,-0.344,2.326,0.125,1.530,1.069,2.360,-
0.943,-0.432,0.632,1.741,-0.124,0.647,-0.046,-0.252,0.695,0.200,-0.552,-0.929,-0.089,-
0.139,-0.365,-0.469,-0.421,-0.797,-0.643,-0.516,-1.819,-0.902,-0.851,-1.281),
T=c(1,1,1,1,1,1,2,1,1,2,1,1,2,2,1,1,1,1,1,1,2,1,1,1,1,2,1,1,1,1,1,1))
```

## Lampiran 9 Hasil Uji

### a. Hasil Pengujian Distribusi

Data column: Y  
Poisson mean for Y = 7,03226

Y	Observed	Poisson Probability	Expected	Contribution to Chi-Sq
<=3	7	0,080099	2,48308	8,21666
4	0	0,089970	2,78907	2,78907
5	4	0,126538	3,92269	0,00152
6	4	0,148308	4,59756	0,07767
7	5	0,148992	4,61874	0,03147
8	2	0,130969	4,06002	1,04524
9	2	0,102334	3,17235	0,43324
10	0	0,071964	2,23088	2,23088
>=11	7	0,100826	3,12562	4,80252

N	N*	DF	Chi-Sq	P-Value
31	0	7	19,6283	0,006

9 cell(s) (100,00%) with expected value(s) less than 5.

Data column: Y1  
Poisson mean for Y1 = 5,64

Y1	Observed	Poisson Probability	Expected	Contribution to Chi-Sq
<=2	4	0,080099	2,00247	1,99261
3	3	0,106234	2,65586	0,04459
4	0	0,149791	3,74476	3,74476
5	4	0,168964	4,22409	0,01189
6	4	0,158826	3,97065	0,00022
7	5	0,127968	3,19921	1,01364
8	2	0,090218	2,25544	0,02893
>=9	3	0,117901	2,94752	0,00093

N	N*	DF	Chi-Sq	P-Value
25	0	6	6,83758	0,336

8 cell(s) (100,00%) with expected value(s) less than 5.

Data column: Y2  
Poisson mean for Y2 = 12,8333

Y2	Observed	Poisson Probability	Expected	Contribution to Chi-Sq
<=12	3	0,481540	2,88924	0,004246
13	1	0,109821	0,65893	0,176544
>=14	2	0,408639	2,45183	0,083265

N	N*	DF	Chi-Sq	P-Value
6	0	1	0,264055	0,607

WARNING: 1 cell(s) (33,33%) with expected value(s) less than 1. Chi-Square approximation probably invalid.

3 cell(s) (100,00%) with expected value(s) less than 5.

## b. Pengujian Homogenitas

```
> MC=quadrat.test(ppall,alternative="two.sided", method="Chisq",tess=te)
Chi-squared test of CSR using quadrat counts

data: ppall
X2 = 124.54, df = 30, p-value = 3.515e-13
alternative hypothesis: two.sided

Quadrats: 31 tiles (irregular windows)
```

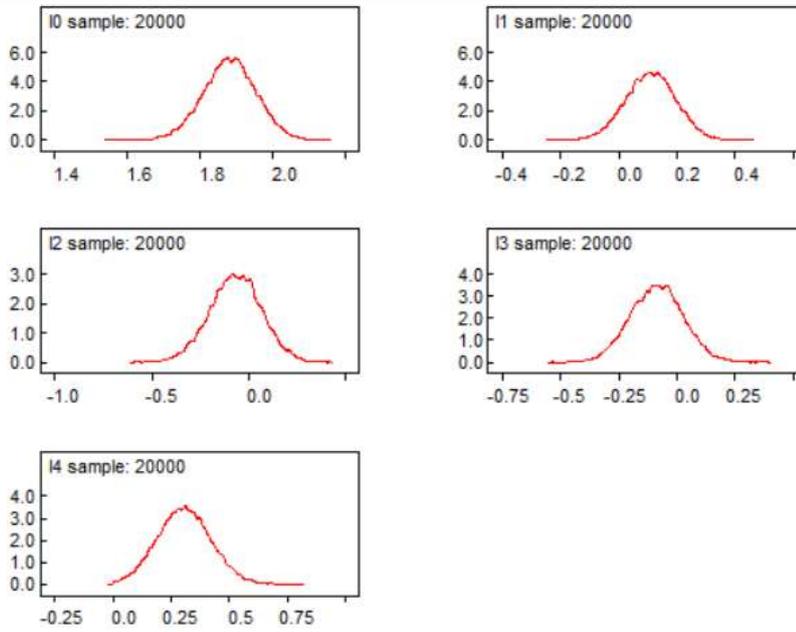
## c. Pengujian Overdispersi

```
> library(qcc)
> qcc.overdispersion.test(Y)

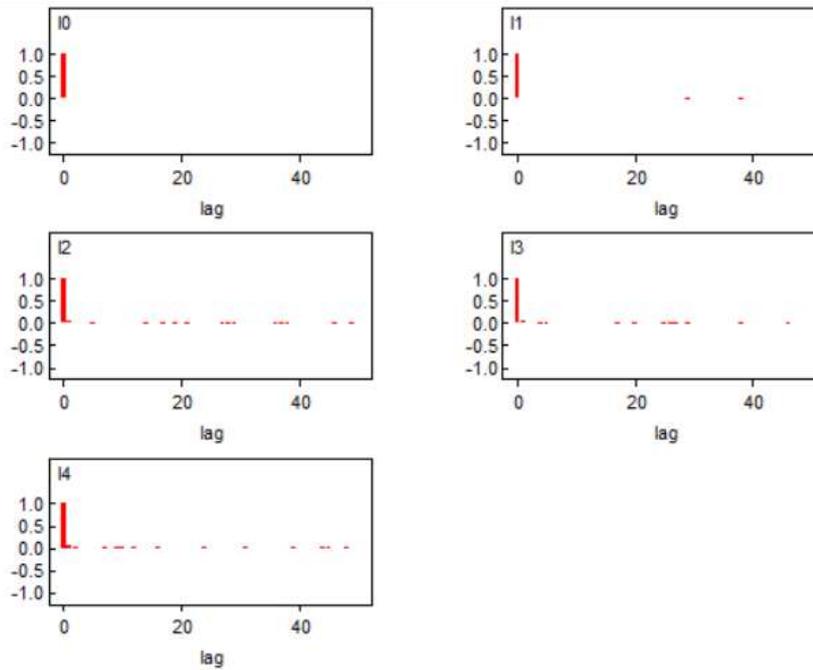
Overdispersion test Obs.Var/Theor.Var Statistic p-value
poisson data      1.910092 57.30275 0.0019285
```

## Lampiran 10 Hasil Output Regresi Poisson

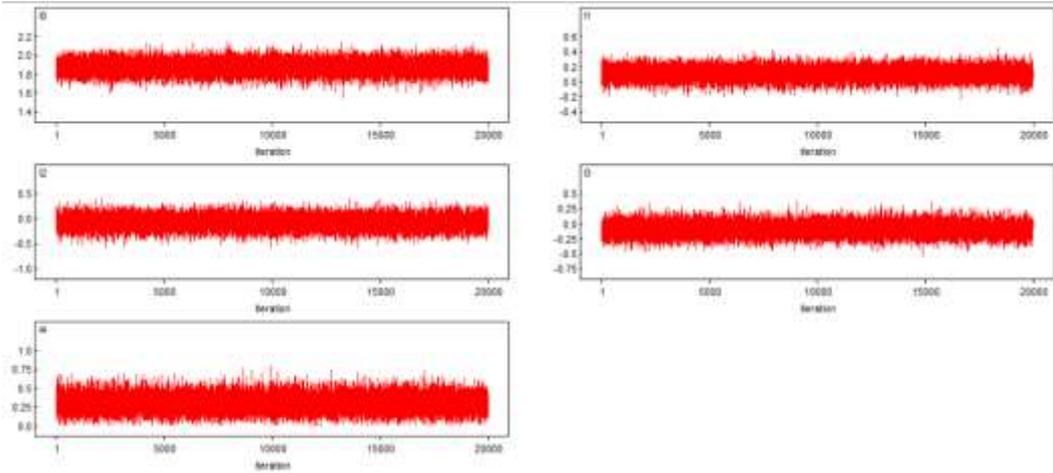
### a. Density Plot



### b. Autocorrelation plot



c. History plot



d. Estimasi Parameter

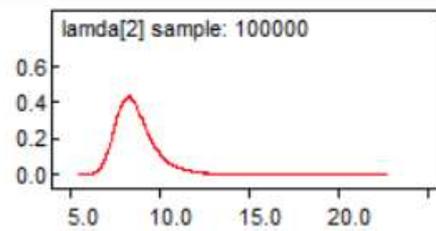
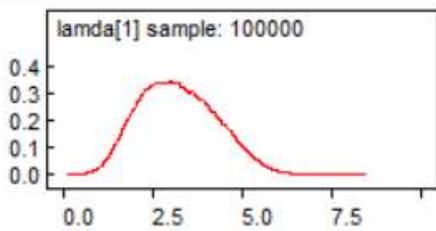
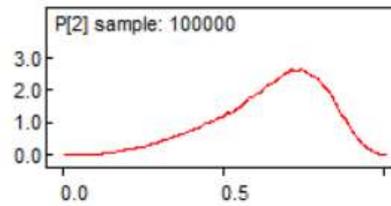
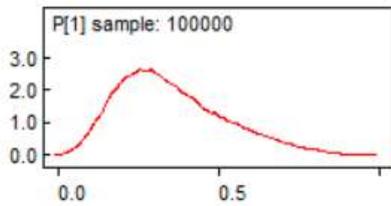
node	mean	sd	MC error2.5%	median	97.5%	start	sample
10	1.881	0.07275	5.754E-4	1.735	1.882	2.022	1 20000
11	0.1101	0.0857	6.127E-4	-0.05893	0.111	0.2755	1 20000
12	-0.06712	0.1331	0.001003	-0.3325	-0.06607	0.1927	1 20000
13	-0.08565	0.113	8.565E-4	-0.3044	-0.08487	0.1353	1 20000
14	0.307	0.1141	8.73E-4	0.08828	0.3058	0.5352	1 20000

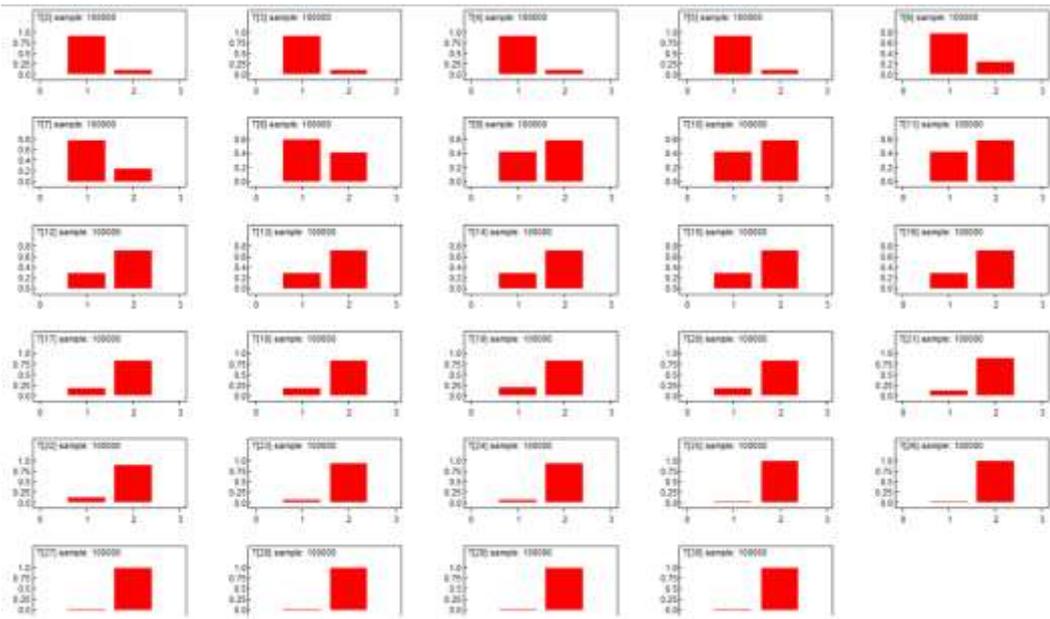
  

Dbar = post.mean of -2logL; Dhat = -2LogL at post.mean of stochastic nodes				
	Dbar	Dhat	pD	DIC
10	0.968	0.968	-0.000	0.968
11	5.060	5.060	0.000	5.060
12	3.694	3.694	0.000	3.694
13	5.061	5.061	0.000	5.061
14	6.444	6.444	-0.000	6.444
y	155.584	150.588	4.996	160.580
total	176.812	171.816	4.996	181.808

Lampiran 11 Hasil Output Penentuan Anggota Mixture

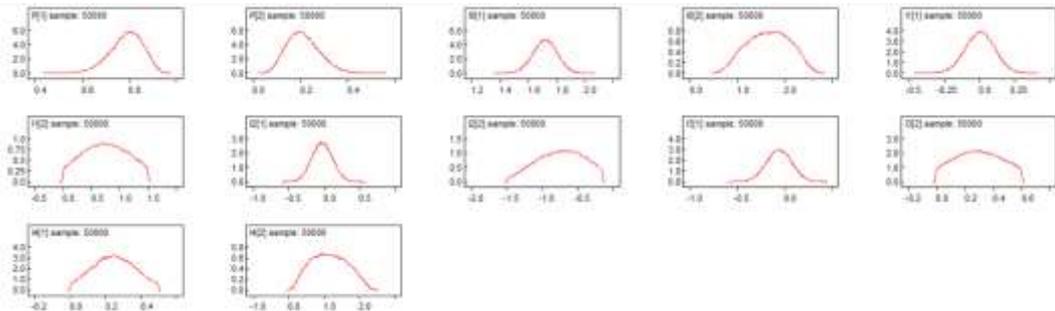
node	mean	sd	MC error	2.5%	median	97.5%	start	sample
P[1]	0.3603	0.1795	0.1795	0.001851	0.0875	0.3306	0.7523	1 20000
P[2]	0.6397	0.1795	0.1795	0.001851	0.2478	0.6694	0.9125	1 20000
T[2]	1.099	0.2991	0.002256	1.0	1.0	1.0	2.0	1 20000
T[3]	1.101	0.3009	0.002241	1.0	1.0	1.0	2.0	1 20000
T[4]	1.103	0.3042	0.002022	1.0	1.0	1.0	2.0	1 20000
T[5]	1.22	0.414	0.002968	1.0	1.0	1.0	2.0	1 20000
T[6]	1.22	0.4142	0.003453	1.0	1.0	1.0	2.0	1 20000
T[7]	1.218	0.4131	0.003054	1.0	1.0	1.0	2.0	1 20000
T[8]	1.539	0.4985	0.004313	1.0	2.0	2.0	2.0	1 20000
T[9]	1.542	0.4982	0.004384	1.0	2.0	2.0	2.0	1 20000
T[10]	1.546	0.4979	0.004388	1.0	2.0	2.0	2.0	1 20000
T[11]	1.542	0.4983	0.004507	1.0	2.0	2.0	2.0	1 20000
T[12]	1.672	0.4694	0.004145	1.0	2.0	2.0	2.0	1 20000
T[13]	1.675	0.4686	0.00393	1.0	2.0	2.0	2.0	1 20000
T[14]	1.676	0.4682	0.003957	1.0	2.0	2.0	2.0	1 20000
T[15]	1.671	0.4698	0.003958	1.0	2.0	2.0	2.0	1 20000
T[16]	1.773	0.4186	0.003747	1.0	2.0	2.0	2.0	1 20000
T[17]	1.776	0.4167	0.003627	1.0	2.0	2.0	2.0	1 20000
T[18]	1.774	0.4186	0.003925	1.0	2.0	2.0	2.0	1 20000
T[19]	1.777	0.416	0.003615	1.0	2.0	2.0	2.0	1 20000
T[20]	1.78	0.4145	0.003534	1.0	2.0	2.0	2.0	1 20000
T[21]	1.849	0.3578	0.00302	1.0	2.0	2.0	2.0	1 20000
T[22]	1.856	0.3515	0.003176	1.0	2.0	2.0	2.0	1 20000
T[23]	1.909	0.2869	0.002228	1.0	2.0	2.0	2.0	1 20000
T[24]	1.907	0.2908	0.002284	1.0	2.0	2.0	2.0	1 20000
T[25]	1.969	0.1728	0.001297	1.0	2.0	2.0	2.0	1 20000
T[26]	1.984	0.1262	8.633E-4	2.0	2.0	2.0	2.0	1 20000
T[27]	1.983	0.1281	8.323E-4	2.0	2.0	2.0	2.0	1 20000
T[28]	1.983	0.1283	8.176E-4	2.0	2.0	2.0	2.0	1 20000
T[29]	1.99	0.099	7.249E-4	2.0	2.0	2.0	2.0	1 20000
T[30]	1.995	0.07227	5.589E-4	2.0	2.0	2.0	2.0	1 20000
lamda[1]	3.407	1.141	0.01099	1.368	3.379	5.621	1	20000
lamda[2]	8.875	1.232	0.01214	7.06	8.664	11.93	1	20000



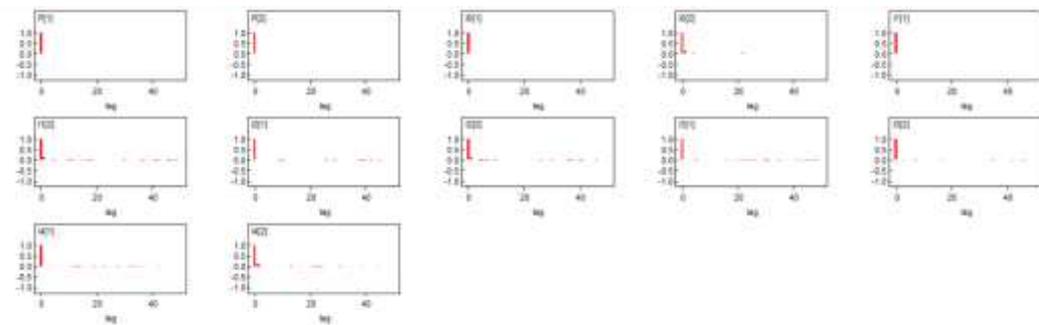


## Lampiran 12 Hasil Output Regresi Mixture Poisson

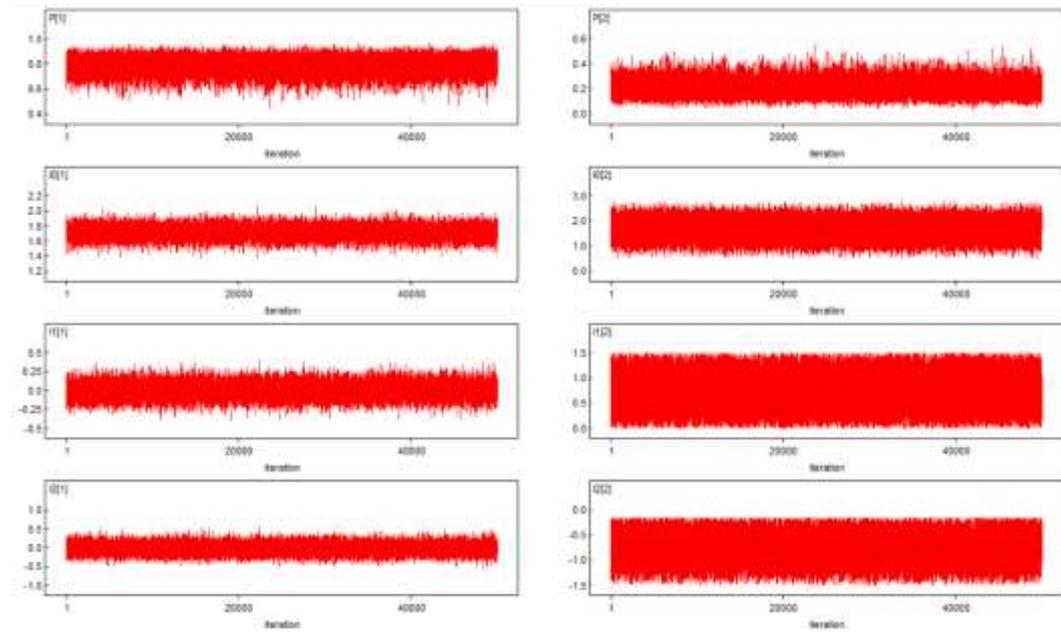
### a. Density plot



### b. Plot Autokorelasi



### c. Plot history



d. Estimasi Parameter Regresi Mixture Poisson

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
P[1]	0.7877	0.07007	3.179E-4	0.635	0.7936	0.9065	1	50000
P[2]	0.2123	0.07007	3.179E-4	0.09351	0.2064	0.365	1	50000
I0[1]	1.717	0.08571	3.64E-4	1.547	1.718	1.882	1	50000
I0[2]	1.665	0.4438	0.002799	0.8374	1.673	2.462	1	50000
I1[1]	0.004314	0.1038	4.316E-4	-0.2018	0.005669	0.2062	1	50000
I1[2]	0.7473	0.3796	0.002373	0.06492	0.7439	1.43	1	50000
I2[1]	-0.02664	0.1437	6.393E-4	-0.304	-0.02813	0.2553	1	50000
I2[2]	-0.749	0.3194	0.001944	-1.371	-0.7354	-0.2039	1	50000
I3[1]	-0.05769	0.1319	5.492E-4	-0.3147	-0.05839	0.1998	1	50000
I3[2]	0.2926	0.1564	7.975E-4	0.01621	0.2908	0.5729	1	50000
I4[1]	0.2461	0.114	4.899E-4	0.03321	0.2453	0.4631	1	50000
I4[2]	1.156	0.5042	0.003236	0.275	1.142	2.102	1	50000

Dbar = post.mean of -2logL; Dhat = -2LogL at post.mean of stochastic nodes				
	Dbar	Dhat	pD	DIC
T	31.425	30.529	0.896	32.321
I0	5.867	5.867	0.000	5.867
I1	4.833	4.833	0.000	4.833
I2	5.054	5.054	0.000	5.054
I3	7.920	7.920	0.000	7.920
I4	7.888	7.888	0.000	7.888
y	140.752	132.519	8.233	148.985
total	203.739	194.610	9.129	212.868

## **BIOGRAFI PENULIS**



Penulis dilahirkan di Karanganyar, 5 September 1992 dengan nama lengkap Tri Murniati. Penulis merupakan anak ke-3 dari tiga bersaudara pasangan Bapak Alm. Ranto dan Ibu Surayem. Penulis menempuh Pendidikan formal di SDN 04 Popongan (1999-2005), SMP N 3 Karanganyar (2005-2008), SMAN 1 Karanganyar (2008-2011), S1 Statistika ITS (2011-2015). Setelah jenjang S1 penulis menempuh Pendidikan magister S2 Statistika ITS (2016-2020). Setelah lulus S1, penulis bekerja sebagai Auditor Ahli Pertama di Inspektorat Jenderal Kementerian Riset, Teknologi, dan Pendidikan Tinggi (2018-2020), saat ini penulis bekerja sebagai Auditor di Inspektorat Kementerian Riset dan Teknologi/BRIN (2020 - saat ini). Harapan saya, Tesis ini bermanfaat bagi pembaca, diskusi saran dan kritik bisa disampaikan kepada penulis melalui email: [trimurniati1992@gmail.com](mailto:trimurniati1992@gmail.com).