KARAKTERISTIK METODE PENDUGA *EMPIRICAL BAYES* DAN *SPATIAL EMPIRICAL BAYES* PADA PENDUGAAN AREA KECIL DENGAN MODEL POISSON-GAMMA

(Skripsi)

Oleh

Nourma Indryani

JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2018

ABSTRACT

CHARACTERISTICS EMPIRICAL BAYES AND SPATIAL EMPIRICAL BAYES ESTIMATOR METHOD IN SMALL AREA ESTIMATION FOR POISSON-GAMMA

By

Nourma Indryani

Small area estimation is an indirect estimation used to estimate parameters with small sample size. The method can be used in small area estimation is Empirical Bayes (EB). The EB method used for count or binary data, which inference based on posterior distribution and parameter be estimate by data. One application of EB methods for binary data is Poisson-Gamma model. This research do a Small Area Estimation with EB method based on Area Level Model and Spatial Area Level Model. Then, this research will be reviewed characteristics of estimators Mean Squared Error (MSE) EB and Spatial Empirical Bayes (SEB) estimator with Area-Spesific Jackknife method in theory and empirical. Empirical studied with software R i386 using number of casess Tuberculosis data in Kota Agung Timur, Gisting, Pugung, Gunung Alip, and Sumberejo sub-districts 2016. The results showed that EB and SEB estimator are biased, however SEB method can correct the bias and resulted a smaller MSE values.

Keyword : *Small Area Estimation*, *Empirical Bayes*, *Spatial Random Effects*, *Spatial Empirical Bayes*, Poisson-Gamma, *Area-Spesific Jackknife.*

ABSTRAK

KARAKTERISTIK METODE PENDUGA *EMPIRICAL BAYES* DAN *SPATIAL EMPIRICAL BAYES* PADA PENDUGAAN AREA KECIL DENGAN MODEL POISSON-GAMMA

Oleh

Nourma Indryani

Pendugaan area kecil merupakan pendugaan tak langsung yang digunakan untuk menduga parameter dengan ukuran sampel yang kecil. Salah satu metode yang dapat digunakan dalam pendugaan area kecil adalah *Empirical Bayes* (EB). Metode EB digunakan pada data cacah atau biner, dimana inferensinya didasarkan pada distribusi posterior dan parameternya diduga dari data. Salah satu penerapan metode EB pada data cacah adalah model Poisson-Gamma. Pada penelitian ini menggunakan pendugaan area kecil dilakukan dengan metode EB berdasarkan model tingkat area serta model spasial tingkat area. Selanjutnya akan dikaji karakteristik penduga dan *Mean Squared Error* (MSE) dari penduga EB serta *Spatial Empirical Bayes* (SEB) dengan metode *Area-Spesific* Jackknife baik secara teori maupun empiris. Kajian secara empiris dilakukan dengan bantuan *software* R i386 menggunakan data Jumlah Kasus Tuberkulosis di Kecamatan Kota Agung Timur, Gisting, Pugung, Gunung Alip, dan Sumberejo pada tahun 2016. Hasil penelitian ini menunjukkan bahwa penduga EB dan SEB bersifat bias, namun metode penduga SEB mampu mengkoreksi nilai bias dan menghasilkan nilai MSE yang lebih kecil.

Kata Kunci: Pendugaan Area Kecil, *Empirical Bayes*, Pengaruh Acak Spasial, *Spatial Empirical Bayes*, Poisson-Gamma, *Area-Spesific Jackknife.*

KARAKTERISTIK METODE PENDUGA *EMPIRICAL BAYES* DAN *SPATIAL EMPIRICAL BAYES* PADA PENDUGAAN AREA KECIL DENGAN MODEL POISSON-GAMMA

Oleh
Nourma Indryani

Skripsi

Sebagai Salah Satu Syarat untuk Mencapai Gelar
SARJANA SAINS

Pada
Jurusan Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam

JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2018

Judul Skripsi : **KARAKTERISTIK METODE PENDUGA *EMPIRICAL BAYES* DAN *SPATIAL EMPIRICAL BAYES* PADA PENDUGAAN AREA KECIL DENGAN MODEL POISSON-GAMMA**

Nama Mahasiswa : *Nourma Indryani*

Nomor Pokok Mahasiswa : 1417031087

Program Studi : Matematika

Fakultas : Matematika dan Ilmu Pengetahuan Alam

MENYETUJUI

1. Komisi Pembimbing

Widiarti, M.Si.	**Prof. Drs. Mustofa Usman, M.A., Ph.D.**
NIP 19800502 200501 2 003	NIP 19570101 1 198403 1 020

2. Ketua Jurusan Matematika

Prof. Dra. Wamiliana, M.A, Ph.D.
NIP 19631108 198902 2 001

MENGESAHKAN

1. Tim Penguji

Ketua	: Widiarti, M.Si.	....
Sekretaris	: Prof. Drs. Mustofa Usman, M.A., Ph.D.	....
Penguji Bukan Pembimbing	: Dian Kurniasari, S.Si., M.Sc.	....

2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam

Prof. Warsito, S.Si., DEA., Ph.D.
NIP 19710212 199512 1 001

Tanggal Lulus Ujian Skripsi : **3 Juli 2018**

PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama	: Nourma Indryani
Nomor Pokok Mahasiswa	: 1417031087
Jurusan	: Matematika

Dengan ini menyatakan bahwa skripsi saya yang berjudul **"KARAKTERISTIK METODE PENDUGA *EMPIRICAL BAYES* DAN *SPATIAL EMPIRICAL BAYES* PADA PENDUGAAN AREA KECIL DENGAN MODEL POISSON-GAMMA"** adalah hasil pekerjaan saya sendiri. Semua hasil tulisan dalam skripsi ini telah mengikuti kaidah penulisan karya ilmiah Universitas Lampung. Apabila kemudian hari terbukti bahwa skripsi ini merupakan hasil salinan atau telah dibuat orang lain, maka saya bersedia menerima sanksi sesuai dengan ketentuan akademik yang berlaku,

Bandar Lampung, 3 Juli 2018

Penulis

METERAI TEMPEL

TGL. 20

2CE92AFF072489601

6000

ENAM RIBU RUPIAH

Nourma Indryani

NPM. 1417031087

RIWAYAT HIDUP

Penulis dilahirkan di Pringsewu pada tanggal 18 Januari 1996, sebagai anak kedua dari pasangan Bapak Bunlie dan Ibu Lenny serta adik dari Andrian Wijaya.

Penulis telah menempuh pendidikan di Taman Kanak-kanak (TK) Dharma Wanita Kota Agung pada tahun 2000-2002, Sekolah Dasar Negeri (SDN) 1 Pasar Madang pada tahun 2002 – 2008, Sekolah Menengah Pertama Negeri (SMPN) 1 Kota Agung pada tahun 2008-2011, dan Sekolah Menengah Atas Negeri (SMAN) 1 Kota Agung pada tahun 2011-2014.

Pada tahun 2014 penulis terdaftar sebagai Mahasiswa PrograPm Sudi S1 Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung melalui jalur SNMPTN. Selama menjadi mahasiswa, penulis bergabung di Generasi Muda Matematika (GEMATIKA) dan Unit Kegiatan Mahasiswa Fakultas (UKMF) NATURAL yang diamanahkan sebagai Sekretaris Biro Kesekretariatan dan Rumah Tangga periode 2015, serta Ketua Biro Kesekretariatan dan Rumah Tangga periode 2016.

Pada bulan Januari 2017 penulis melaksanakan Kerja Praktik (KP) di Badan Pusat Statistik Provinsi Lampung guna menerapkan ilmu yang telah diperoleh sewaktu kuliah. Pada bulan Juli 2017 penulis melaksanakan Kuliah Kerja Nyata (KKN) di Desa Ruguk, Kecamatan Ketapang, Kabupaten Lampung Selatan.

PERSEMBAHAN

Puji Syukur kepada Tuhan Yang Maha Esa karena atas berkat dan karunia-Nya, aku masih diberikan kesempatan untuk menuntut ilmu di Universitas Lampung serta menyelesaikan skripsi sederhana ini.

Dengan kerendahan hati yang tulus, kupersembahkan skripsi ini untuk keluargaku tercinta. Teruntuk kedua orang tuaku, Papa Bunlie dan mama Lenny yang selalu ada disetiap perjalanan hidupku, yang memberikan doa, nasihat, semangat, dukungan baik moril maupun materil dan juga kasih sayang serta pengorbanan yang tak tergantikan

Dan untuk kokoku Andrian Wijaya terimakasih telah turut memberikan warna dalam kehidupanku, serta perhatian dan dukungan baik moril dan meteril untukku sampai saat ini.

Terimakasih juga untuk saudaraku (keluarga besar mama dan papa), teman-teman , serta sahabat yang tidak dapat aku sebutkan satu-persatu atau semangat dan motivasi yng telah kalian beri selama pengerjaan skripsi ini.

KATA INSPIRASI

"Serahkanlah hidupmu kepada TUHAN dan percayalah kepada-Nya, dan Ia akan bertindak."

Mazmur 37:5

"Family is the most important thing in the world. You can't choose your family. They are God's gift to you, as you are to them."

"*Bersyukurlah untuk setiap senyum yang kau berikan*
dan untuk setiap air mata yang kau tumpahkan.
Karena senyuman membutmu kuat
dan air mata membutamu bijak."

(Anonim)

SANWACANA

Puji dan syukur senantiasa kepada Tuhan Yang Maha Esa yang selalu memberikan nikmat, rahmat, serta hidayah-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul "**KARAKTERISTIK METODE PENDUGA *EMPIRICAL BAYES* DAN *SPATIAL EMPIRICAL BAYES* PADA PENDUGAAN AREA KECIL DENGAN MODEL POISSON-GAMMA"**

Penulis juga mengucapkan terima kasih kepada pihak-pihak yang telah membantu dalam memberikan bimbingan, motivasi, semangat dan saran yang membangun selama proses penyusunan skripsi ini. Terimakasih kepada:

1. Ibu Widiarti, M.Si, selaku Dosen Pembimbing 1 dan Dosen Pembimbing Akademik yang telah meluangkan waktu dan membimbing serta memberikan pengarahan kepada penulis selama masa perkuliahan.
2. Bapak Prof. Drs. Mustofa Usman, M.A., Ph.D, selaku Dosen Pembimbing 2 yang telah memberikan saran serta arahan kepada penulis.
3. Ibu Dian Kurniasari, S.Si., M.Sc., selaku Dosen Pembahas yang memberi saran dan evaluasi kepada penulis selama penyusunan skripsi.
4. Ibu Prof. Dra. Wamiliana, MA, Ph.D, selaku Ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam.

5. Bapak Prof. Warsito, S.Si., DEA., Ph.D, selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
6. Dosen, staf dan karyawan Jurusan Matematika FMIPA Universitas Lampung yang telah memberikan ilmu serta bantuan kepada penulis.
7. Papa, mama dan koko yang telah memberikan motivasi, doa, dukungan serta menyediakan sarana dalam penyusunan skripsi ini.
8. Teman-teman satu pembimbing, yaitu Novi, Maget, Yunika, Raka, Nia dan Shindi, terimakasih untuk dukungan dan pelajarannya.
9. Terimakasih untuk Vivin, Cia, Nandra, Darma, Wayan, Rere, Otin, Riya, Kasandra, Septi, Fikar, dan teman-teman Matematika Angkatan 2014, HIMATIKA dan NATURAL yang tidak dapat disebutkan namanya satu persatu.
10. Seluruh pihak yang telah membantu penulis dalam menyusun skripsi ini.

Penulis juga menyadari bahwa dalam penulisan skripsi ini masih banyak terdapat kekurangan. Oleh karena itu, penulis mengharapkan saran dan kritik yang membangun guna penelitian selanjutnya agar menjadi lebih baik lagi.

Bandar Lampung, Juli 2018
Penulis,

Nourma Indryani

DAFTAR ISI

III. METODOLOGI PENELITIAN

IV. HASIL DAN PEMBAHASAN

DAFTAR TABEL

DAFTAR GAMBAR

I. PENDAHULUAN

1.1 Latar Belakang dan Masalah

Dalam bidang statistika sering kali menjumpai istilah survei, yaitu salah satu metode untuk mengumpulkan informasi dari sampel yang mewakili parameter populasi penelitian. Survei merupakan salah satu bagian penting dalam pengambilan keputusan berbasis data. Namun, sangat banyak kendala yang ditemukan dalam pelaksanaan survei. Ada dua topik utama yang menjadi perhatian dalam membahas persoalan survei, yaitu persoalan pengembangan teknik penarikan sampel dan pengembangan metodologi pendugaan parameter pupulasi.

Survei yang didesain untuk inferensia bagi daerah (domain) yang luas biasanya jarang menemukan persoalan. Tetapi persoalan muncul ketika dari survei seperti ini ingin diperoleh informasi untuk area yang lebih kecil, misalnya informasi pada level kabupaten, kecamatan bahkan mungkin level kelurahan. Sehingga karena ukuran sampel yang kecil, metode pendugaan tidak dapat memberikan hasil dugaan yang teliti. Oleh sebab itu diperlukan metode yang dapat digunakan untuk mendapatkan hasil yang lebih baik dalam menduga parameter subpopulasi dengan ukuran sampel yang kecil yaitu pendugaan area kecil (*Small Area Estimation*).

Metode pendugaan yang sering digunakan dalam melakukan pendugaan area kecil yaitu *Empirical Best Linear Unbiased Prediction* (EBLUP), *Empirical Bayes* (EB), dan *Hierarchical Bayes* (HB). Menurut Rao (2003), metode EB cocok digunakan dalam menangani data biner dan data cacahan pada pendugaan area kecil. Pada metode *Empirical Bayes* parameter model diestimasi dari distribusi marginal data yang kemudian inferensi didasarkan pada distribusi posterior yang diestimasi.

Menurut Rao (2003), metode-metode yang terdapat pada pembahasan *Small Area Estimation* menghasilkan penduga tak langsung, yaitu nilai dugaan yang dikembangkan dengan menghubungkan informasi tambahan dari area sekitar untuk meningkatkan efektifitas ukuran sampel. Informasi tambahan tersebut umumnya diperoleh dari penggunaan variabel penyerta, namun dapat juga diperoleh melalui penambahan pembobotan berdasarkan nilai variabel acak dari suatu domain tertentu. Pembobotan dapat didefinisikan sebagai bobot suatu domain berdasarkan kedekatannya dengan domain lain di sekitarnya atau disebut pengaruh spasial. Dalam model analisis pengaruh spasial dapat menggunakan efek acak area ketika terdapat "*Neighboring*" antar area.

Salah satu penerapan metode *Empirical Bayes* untuk data cacah adalah dalam pemetaan suatu penyakit. Penyakit yang menjadi ancaman bagi Indonesia adalah Tuberkulosis, dimana Indonesia menjadi negara kedua dengan penderita Tuberkulosis tertinggi di dunia. Penyakit Tuberkulosis merupakan epidemi yang menular melalui kontak langsung dengan orang yang terjangkit penyakit ini. Penelitian ini akan mengakat persoalan penyakit Tuberkulosis dengan mengambil informasi dari domain level kelurahan. Ukuran sampel yang kecil dikhawatirkan akan menyebabkan penerapan pendugaan resiko relatif terjangkitnya penyakit tuberkulosis pada area tertentu secara langsung menjadi tidak dapat diandalkan. Dengan demikian pada penelitian ini akan digunakan pendugaan tak langsung *Empirical Bayes* dan *Spatial Empirical Bayes.*

Model yang sering digunakan dalam kasus pemetaan penyakit adalah model Poisson. Model ini memiliki nilai rataan dan ragam yang sama, sedangkan pada beberapa kasus sering terjadi nilai ragam lebih besar daripada rataan atau terjadi overdispersi. Oleh sebab itu diperlukan model Poisson campuran agar kelebihan nilai ragam tersebut dapat diakomodasi dengan parameter priornya. Penelitian ini menerapkan model Poisson campuran dengan menggunakan prior konjugat dari Poisson yaitu Gamma, sehingga menjadi Model Poisson-Gamma.

Penambahan efek acak spasial yang digunakan pada penelitian ini didasarkan pada cara penyebaran penyakit Tuberkulosis yaitu dari kontak langsung. Dengan demikian pengaruh spasial yang diamati adalah berdasarkan kedekatan antar area. Pratesi dan Salvati (2008) menyatakan bahwa penggunaan pengaruh spasial dapat mengurangi bias dan galat penarikan sampel dalam pendugaan area kecil.

Suatu penduga dapat dikatakan sebagai penduga yang baik jika memenuhi karakteristik tertentu, diantaranya bersifat tak bias dan memiliki ragam minimum.

Selain itu, kualitas suatu penduga juga dapat dievaluasi berdasarkan nilai Kuadrat Tengah Galat atau *Mean Square Error* (MSE). Salah satu metode penduga MSE adalah metode *jackknife*. Konsep *jackknife* pertama kali diperkenalkan oleh Quenouille (1956) dengan tujuan untuk mengoreksi bias dugaan. Metode *jackknife* lainnya adalah Rao (2003) yang dikenal sebagai metode *area-specific jackknife*. Metode ini menggunakan ragam dari sebaran posterior sebagai pendekatan bagi nilai dugaan MSE. Oleh sebab itu, peneliti tertarik untuk membandingkan kualitas penduga berdasarkan evaluasi sifat ketakbiasan, ragam minimum, dan kriteria MSE penduga dari *Empirical Bayes* dan *Spatial Empirical Bayes* dengan metode *Area-Spesific Jackknife* pada pendugaan area kecil dengan model Poisson-Gamma.

1.2 Tujuan Penelitian

Berdasarkan permasalahan yang telah dijabarkan, maka tujuan dari penelitian ini adalah:

1. Menentukan penduga *Empirical Bayes* dan penduga *Spatial Empirical Bayes* pada pendugaan area kecil dengan model Poisson-Gamma.
2. Mengkaji bias, ragam dan *Mean Square Error* pada metode *Empirical Bayes* dan *Spatial Empirical Bayes* baik secara teori maupun empiris melalui simulasi dengan membangkitkan data berdasarkan distribusi Poisson-Gamma.
3. Mengevaluasi *Mean Square Error* metode *Empirical Bayes* dan *Spatial Empirical Bayes* baik secara teori maupun empiris menggunakan data Jumlah Kasus *Tuberculosis* tahun 2016 di Kecamatan Kota Agung Timur, Gisting, Gunung Alip, Sumberejo dan Pugung.

1.3 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini adalah menambah wawasan tentang metode *Empirical Bayes* dan *Spatial Empirical Bayes* dengan mengkaji karakteristik penduga dan evaluasi *Mean Square Error* pada pendugaan area kecil.

II. TINJAUAN PUSTAKA

2.1 Pendugaan Area Kecil

Suatu daerah disebut sebagai area kecil (*small area*) jika ukuran contoh dalam domain atau daerah tersebut tidak cukup memadai untuk mendukung ketelitian pendugaan langsung. Ketika pendugaan langsung digunakan untuk menduga parameter suatu dari contoh yang kurang (kecil), maka akan menghasilkan penduga yang tidak bias namun memiliki ragam yang besar. Oleh sebab itu diperlukan metode lain yang dapat menangani masalah tersebut, yaitu penduga area kecil. Pendugaan area kecil (*Small Area Estimation* atau SAE) merupakan metode pendugaan tidak langsung yang mengkombinasikan antara data survei dengan data pendukung lain misalnya dari data sensus sebelumnya yang memuat peubah dengan karakteristik yang sama dengan data survei sehingga dapat digunakan untuk menduga area yang lebih kecil dan memberikan akurasi yang lebih baik (Rao, 2003).

Metode pendugaan area kecil memiliki dua masalah pokok. Pertama, bagaimana menghasilkan pendugaan karakteristik yang handal untuk area yang menjadi perhatian, berdasarkan jumlah ukuran contoh kecil dari area yang bersesuaian. Permasalahan kedua berhubungan dengan bagaimana cara mendapatkan dugaan *Mean Square Error* (MSE) atau Kuadrat Tengah Galat (KTG) dari dugaan parameter tersebut. Untuk mengatasi kedua permasalahan tersebut perlu "meminjam informasi" baik dari dalam area, luar area, maupun dari luar survei (Pfefferman, 2013).

2.2 Model Area Kecil

Model area kecil merupakan model dasar dalam pendugaan area kecil. Dalam pendugaan area kecil terdapat dua jenis model dasar yang digunakan, yaitu model berbasis area (Tipe A) dan model berbasis unit (Tipe B) (Rao, 2003).

1. Model Berbasis Area (Tipe A)

Model berbasis area merupakan model yang didasarkan pada ketersediaan data pendukung yang hanya ada untuk level area tertentu, misalkan $x_i = (x_{1i}, x_{2i}, \ldots, x_{pi})^T$ dengan parameter yang akan diduga adalah θ_i yang diasumsikan mempunyai keterkaitan dengan z_i. Data pendukung tersebut digunakan untuk membangun model

$$\theta_i = x_i^T \beta + b_i v_i \qquad (2.1)$$

dengan $i = 1,2,\ldots,m$ dan diketahui bahwa b_i adalah konstanta bernilai positif serta $\beta = (\beta_1, \ldots, \beta_p)^T$ adalah vektor dari koefisien regresi. Selain itu, v_i adalah efek acak daerah yang diasumsikan berdistribusi identik dan independen (iid) dengan $E(v_i) = 0$ dan $Var(v_i) = \sigma_v^2$.

Untuk mencari dugaan parameter rata-rata populasi pada area kecil $\bar{Y}_i$ berdasarkan persamaan (2.1), diasumsikan bahwa taksiran langsung dari $\bar{Y}_i$ ada. Sehingga, dapat diasumsikan pula bahwa

$$\hat{\theta}_i = g\left(\hat{\bar{Y}}_i\right) = \theta_i + \varepsilon_i, \quad i = 1,2,\dots,m \tag{2.2}$$

dimana ε_i adalah *error* dari pemilihan daerah sampel yang berdistribusi independen dengan $E(\varepsilon_i) = 0$ dan $Var(\varepsilon_i) = \psi_i$. Lalu dengan substitusi persamaan (2.1) ke persamaan (2.2) sehingga diperoleh model gabungan yang disebut sebagai model berbasis area, yaitu:

$$\hat{\theta}_i = x_i^T\beta + b_i v_i + \varepsilon_i \tag{2.3}$$

dengan asumsi v_i dan ε_i saling bebas (Rao, 2003).

Pada tahun 1979, diperkenalkan oleh Fay dan Herriot suatu kasus khusus dari model tingkat area, yaitu dengan asumsi bahwa kontribusi efek acak v_i terhadap taksiran θ_i sama untuk masing-masing area. Sehingga dengan substitusi $b_i = 1$ pada persamaan (2.3) diperoleh:

$$\hat{\theta}_i = x_i^T\beta + v_i + \varepsilon_i \tag{2.4}$$

Model tersebut disebut sebagai model Fay-Herriot.

2. Model Berbasis Unit (Tipe B)

Pada model berbasis unit ini diasumsikan bahwa data variabel penyerta unit $x_{ij} = \left(x_{1ij}, x_{2ij}, \dots, x_{pij}\right)^T$ ada untuk setiap elemen ke-j pada area ke-i. selanjutnya variabel respon y_{ij} diasumsikan berkaitan dengan x_{ij} sehingga bentuk persamaan model area kecil berbasis unit sebagai berikut:

$$\hat{\theta}_{ij} = x_{ij}^T\beta + \varepsilon_{ij} + v_i, \;\; j = 1,2,\dots,m; \;\; i = 1,2,\dots,n \tag{2.4}$$

dengan v_i adalah pengaruh acak area yang berdistribusi identik dan independen, β merupakan koefisien regresi dan diasumsikan galat bernilai 0 (Rao, 2003).

2.3 Perluasan Model Berbasis Area: Model Spasial

Model dasar Fay-Herriot (2.3) mengasumsikan v_i efek area kecil yang berdistribusi identik dan independen, tetapi dalam beberapa aplikasi mungkin nyatanya menggunakan model yang memiliki korelasi dari v_i. Pada model spasial, v_i digunakan ketika terdapat "*Neighboring*" antar area. Contoh model yang memiliki korelasi antar v_i yaitu, korelasi yang bergantung pada kedekatan geografis dalam konteks pernyebaran penyakit lokal dan tingkat kematian.

Misalkan A_i dinotasikan sebagai himpunan tetangga dari suatu area ke-i, kemudian model spasial *Conditional Autoregression* (CAR) mengasumsikan bahwa distribusi bersyarat dari v_i diberikan $\{v_l : l \neq i\}$ yang diberikan dari

$$v_i | \{v_l : l \neq i\} \sim N\left(\rho \textstyle\sum_{i \epsilon A_i} q_{il}\, v_l, \sigma_v^2\right) \tag{2.5}$$

dimana ρ adalah korelasi gabungan antara efek acak v_i dengan seluruh efek acak $v_l \in A_i$, serta σ_v^2 adalah variansi gabungan dari efek acak v_i dan $\{q_{il}\}$ diketahi adalah konstanta yang memenuhi $q_{il} = q_{li}$ dengan $i < l$ (Rao, 2003). Sedangkan nilai $\{q_{il}\}$ bergantung dari bagaimana pendefinisian A_i.

Misalkan nilai dari efek acak v_i didefinisikan sebagai koordinat titik pusat dari wilayah ke-i yang dinyatakan dalam dua komponen, yaitu nilai garis bujur (longitude) dan nilai garis lintang (lattitude). Sehingga v_i dapat dinyatakan dalam bentuk vektor $\boldsymbol{v_i} = [v_x \quad v_y]$, dimana v_x menyatakan nilai *longitude* dan v_y menyatakan nilai latitude. Sehingga distribusi bersyarat pada persamaan (2.5) menjadi:

$$\boldsymbol{v_i}|\{\boldsymbol{v_l}: l \neq i\} \sim N_2\left(\rho \sum_{i \epsilon A_i} q_{il}\, \boldsymbol{v_l}, \sigma_v^2\right) \tag{2.6}$$

Sehingga dengan substitusi model spasial pada persamaan (2.6) ke model dasar area pada persamaan (2.4) terbentuk suatu model tingkat wilayah pada *small area* yang memperhatikan pengaruh spasial pada $\boldsymbol{v_i}$.

2.4 Metode *Empirical Bayes*

Dasar pengembangan pendekatan statistik Bayes adalah hukum Bayes yang dibuat oleh Thomas Bayes. Hukum ini diperkenalkan oleh Richard Price tahun 1763 dua tahun setelah wafatnya Thomas Bayes. Pada tahun 1774 dan 1781, Laplace memberikan analisis lebih rinci dan lebih relevan untuk statistik Bayes sekarang (Gill, 2002). Dalam penerapan Bayes dikenal adanya distribusi prior yang harus diketahui terlebih dahulu nilai parameternya. Namun, jika informasi dari parameter prior belum di ketahui maka digunakan pendekatan lain, yaitu *Empirical Bayes*. Parameter dari distribusi prior disebut juga sebagai *hyperparameters*.

Metode *Empirical Bayes* (EB) diperkenalkan pertama kali oleh Robbins pada tahun 1955 dengan mengasumsikan bahwa parameter prior tidak diketahui, selanjutnya data digunakan untuk memperoleh dugaan dari parameter prior tersebut. Metode EB diperkenalkan sebagai kombinasi dari pendekatan *Classic* dan *Bayesian Statistics* dengan tujuan yang sama, yaitu membuat inferensi dari suatu parameter yang tidak diketahui. Menurut Rao (2003), metode EB cocok digunakan dalam menangani data biner dan data cacahan pada pendugaan area kecil. Misalkan $y_1, y_2, \dots, y_n$ merupakan sampel acak berukuran n dari distribusi yang mempunyai fungsi kepekatan peluang berbentuk $f(y_1, y_2, \dots, y_n|\theta)$ dan sebaran dari peubah acak θ yaitu $h(\theta)$ sebaran prior. Metode EB dalam konteks pendugaan area kecil secara ringkas sebagai berikut:

1. Mendapatkan fungsi kepekatan peluang akhir (posterior) dari $y_1, y_2, \dots, y_n$ dengan $f(y_1, y_2, \dots, y_n)$ yang didefinisikan sebagai berikut:

$$f(\theta|y_1, y_2, \dots, y_n) = \frac{f(y_1, y_2, \dots, y_n|\theta) \, .h(\theta)}{\int f(y_1, y_2, \dots, y_n|\theta) \, .h(\theta) \, d\theta} \tag{2.6}$$

2. Menduga parameter model dari fungsi kepekatan peluang marginal.
3. Menggunakan fungsi kepekatan peluang akhir dugaan untuk membuat inferensi parameter area kecil yang menjadi perhatian.

Pendugaan tak langsung dari suatu parameter berdasarkan model area kecil dengan memperhatikan pengaruh spasial dari domain di sekitarnya serta dengan menggunakan metode *Empirical Bayes* dapat disebut sebagai metode *Spatial Empirical Bayes*.

2.5 Model Poisson-Gamma

Model poisson adalah model peluang standar untuk data cacahan. Model ini akan mengalami keterbatasan dalam rataan dan ragam ketika digunakan untuk pendugaan parameter tunggal. Umumnya, data cacahan (seperti data penyakit) mengalami overdispersi (ragam melebihi rataan). Oleh karena itu dikembangkan suatu formula poisson yang memuat parameter tambahan untuk mengakomodasi ragam ekstra dari pengamatan pada sampel. Oleh karena itu diperkenalkan model dua tahap untuk data cacahan yang dikenal dengan model Poisson-Gamma, dimana Gamma merupakan prior konjugat dari Poisson. Model Poisson-Gamma dapat ditulis sebagau berikut:

Level 1 : $y_i|\theta_i \sim Poisson\ (\theta_i)\ , i = 1, 2, \ldots, n$

Level 2 : $\theta_i \sim Gamma\ (\alpha, \beta)\ , i = 1, 2, \ldots, n$

Misalkan $y_1, y_2, \ldots y_n$ sampel acak dengan distribusi peluang:

$$p(y_i|\theta_i) = \begin{cases} \frac{\theta_i^{y_i}\, e^{-\theta_i}}{y_i!} & ; \begin{matrix} y_i = 1,2, \ldots \\ i = 1,2, \ldots, n \end{matrix} \\ 0 & ;\ y_i\ lainnya \end{cases} \tag{2.7}$$

dengan distribusi prior gamma yang memiliki fungsi kepekatan peluang sebagai berikut:

$$\pi(\theta_i) = \begin{cases} \frac{1}{\Gamma(\alpha)\left(\frac{1}{\beta}\right)^{(\alpha)}} \theta_i^{(\alpha - 1)} e^{\frac{-\theta_i}{\beta}} & ; \theta_i \geq 0 \\ 0 & ;\ \theta_i\ lainnya \end{cases} \tag{2.8}$$

Berdasarkan model di atas, maka diperoleh fungsi bersama model Poisson-Gamma sebagai berikut:

$$f(y_1, y_2, \ldots, y_n, \theta_i) = \frac{e^{-\theta_i}\theta_i^{y_i}}{y_i!} \cdot \frac{1}{\Gamma(\alpha)\beta^{\alpha}} \theta_i^{\alpha-1} e^{-\frac{\theta_i}{\beta}} \tag{2.9}$$

Selanjutnya diperoleh sebaran marginal dengan mengintegralkan fungsi bersama sebagai berikut:

$$m(y_1, y_2, \ldots, y_n) = \frac{1}{y_i!\, \Gamma(\alpha)\beta^{\alpha}} \Gamma(y_i + \alpha) \left(\frac{1}{1+{}^{1}/_{\beta}}\right)^{(y_i+\alpha)} \tag{2.10}$$

Maka didapatkan fungsi postriornya sebagai berikut:

$$\pi(\theta_i|y_1, y_2, \dots, y_n) = \frac{f(y_1, y_2, \dots, y_n, \theta_i)}{m(y_1, y_2, \dots, y_n)}$$

$$= \frac{\theta^{y_i+\alpha} e^{-\theta_i\left(1+{}^1/_\beta\right)}}{\Gamma(y_i+\alpha)\left(\frac{1}{1+{}^1/_\beta}\right)^{(y_i+\alpha)}} \tag{2.11}$$

(Larsen dan Marx, 2012).

Deketahui distribusi dari θ_i adalah $\theta_i|y_i, \alpha, \beta \sim Gamma\left(y_i + \alpha, 1 + {}^1/_\beta\right)$, sehingga diperoleh penduga bayes bagi θ_i dan ragam bagi θ_i adalah

$$\hat{\theta}_i^B = \mathrm{E}(\theta_i|y_i) = \frac{y_i+\alpha}{1+{}^1/_\beta} \tag{2.12}$$

$$\mathrm{Var}(\theta_i|y_i) = \frac{y_i+\alpha}{\left(1+{}^1/_\beta\right)^2} \tag{2.13}$$

Selanjutnya penduga *Empirical Bayes* akan diperoleh dengan cara melakukan pendugaan parameter prior terlebih dahulu, dimana pendugaan ini dilakukan berdasarkan data.

2.6 Pendugaan Parameter dengan Metode Momen

Metode momen diciptakan oleh Karl Pearson pada tahun 1800, metode ini merupakan metode tertua dan paling lama digunakan. Metode momen memiliki prosedur yang paling mudah dalam memperoleh estimator atau penduga dari satu atau lebih parameter populasi. Dasar pemikiran dari metode momen adalah mendapatkan estimasi parameter populasi dengan menyamakan momen-momen populasi (teoritis) dengan momen-momen sampel. Momen pusat ke-k dari populasi didefinisikan sebagai berikut:

$$\mu'_k = E(X^k) \tag{2.14}$$

Sedangkan momen pusat ke- k dari sampel didefinisikan sebagai berikut:

$$m_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k \tag{2.15}$$

Momen populasi μ_j sering ditulis sebagai fungsi dari $\theta_1, \theta_2, \ldots, \theta_k$, yaitu $\mu_j(\theta_1, \theta_2, \ldots, \theta_k)$. Metode momen penduga $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k$ dari $\theta_1, \theta_2, \ldots, \theta_k$ didapat dengan menyelesaikan sistem persamaan untuk $(\theta_1, \theta_2, \ldots, \theta_k)$ dalam notasi $(m_1, m_2, \ldots, m_k)$ sebagai berikut $m'_j = \mu'_j(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k)$; $j = 1, 2, \ldots, k$

$$\begin{aligned} m_1 &= \mu_1 = (\theta_1, \theta_2, \ldots, \theta_k) \\ m_2 &= \mu_2 = (\theta_1, \theta_2, \ldots, \theta_k) \\ &\vdots \qquad\qquad \vdots \\ m_k &= \mu_k = (\theta_1, \theta_2, \ldots, \theta_k) \end{aligned} \tag{2.16}$$

(Berger, 1990).

Meskipun metode momen terkenal sebagai metode yang mudah, namun terkadang terdapat kesulitan dalam mencari bentuk rumus penduganya.

2.7 Karakteristik Penduga Parameter

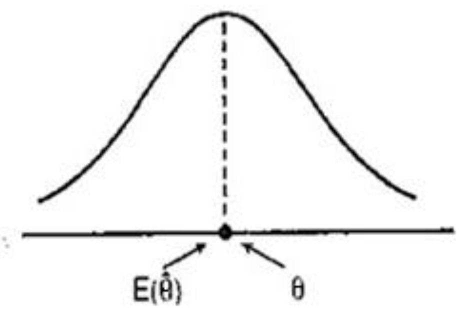
Estimator yang baik adalah yang memenuhi sifat tertentu, diantaranya sifat tak bias dan ragam minimum.

2.7.1 Ketakbiasan

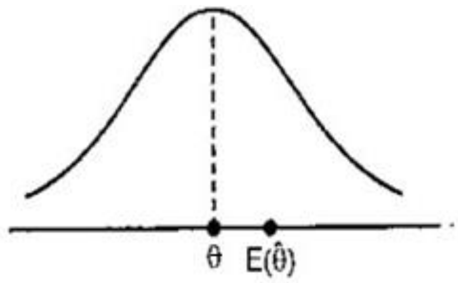
Sifat penduga yang baik salah satunya adalah sifat takbias. Suatu penduga dikatakan takbias apabila asumsi yang telah ditentukan terpenuhi, adapun penjelasannya sebagai berikut:

Definisi 2.1

Misalkan $Y_1, Y_2, \dots, Y_n$ merupakan sampel acak dari fungsi kepekatan peluang kontinu $f_y(y;\theta)$, dimana θ merupakan parameter yang tidak diketahui. Penduga $\hat{\theta}_n[= h(Y_1, Y_2, \dots, Y_n)]$ dikatakan takbias bagi θ, jika $\mathrm{E}(\hat{\theta}) = \theta$. Untuk konsep dan terminologi yang sama berlaku, jika terdapat data sampel acak $X_1, X_2, \dots, X_n$ yang diambil dari fungsi kepekatan peluang diskrit $p_x(k;\theta)$. Suatu penduga $\hat{\theta}_n[= h(X_1, X_2, \dots, X_n)]$ dikatakan penduga takbias asimtotik bagi θ jika $\lim_{n\to\infty} \mathrm{E}(\hat{\theta}) = \theta$. Jika $\mathrm{E}(\hat{\theta}) \neq \theta$ maka penduga dikatakan sebagai penduga yang berbias bagi θ (Larsen dan Marx, 2012).



Gambar 2.1
Penduga Tak Bias, $\mathrm{E}(\hat{\theta}) = \theta$



Gambar 2.2
Penduga Bias, $\mathrm{E}(\hat{\theta}) \neq \theta$

Penduga tak bias artinya penduga yang dengan tepat mengenai sasaran, seperti ditunjukkan oleh Gambar 2.1. sedangkan penduga bias artinya penduga yang tidak tepat mengenai sasaran atau disebut meleset, seperti ditunjukkan oleh Gambar 2.2.

2.7.2 Ragam Minimum

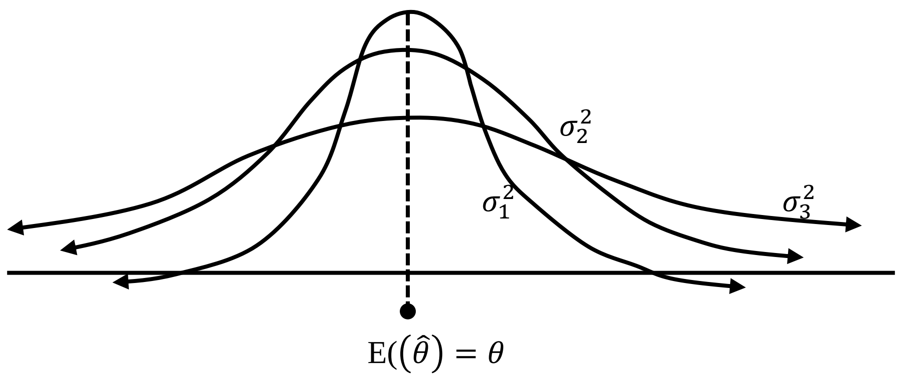
Selain sifat ketakbiasan, penduga parameter dikatakan baik apabila memenuhi sifat penduga ragam minimum. Adapun definisi ragam minimum suatu penduga sebagai berikut:

Definisi 2.2

Bila U(X) merupakan penduga bagi g(θ), maka U_1(X) dikatakan sebagai penduga beragam terkecil, jika:

$$\sigma^2_{U1(X)} \leq \sigma^2_{U(X)} \qquad (2.17)$$

Dimana U(X) merupakan sembarang penduga bagi g(θ) (Hogg dan Craig, 1995).



Gambar 2.3 Ragam Minimum

Gambar 2.3 menunjukkan ada tiga penduga, yaitu $\hat{\theta}_1, \hat{\theta}_2,$ dan $\hat{\theta}_3$ yang diperoleh dari tiga sampel, dimana distribusi sampel 1 mempunyai variansi σ_1^2, sampel 2 memiliki variansi σ_2^2, dan sampel 3 memiliki variansi σ_3^2 dengan nilai $\sigma_1^2 < \sigma_2^2 < \sigma_3^2$. Oleh karena sampel 1 mempunyai variansi paling kecil, maka dikatakan $\hat{\theta}_1$ merupakan penduga yang beragam kecil atau efisien.

2.8 *Mean Square Error* (MSE)

Banyak metode yang dapat digunakan dalam menduga MSE, salah satunya yaitu metode *Area-Spesific Jackknife*. Metode *Area-Spesific Jackknife* diperkenalkan oleh Rao (2003), metode ini merupakan pengembangan dari metode *Jackknife* yang diperkenalkan oleh Jiang *et al* (2002). Rao menggunakan g_i sebagai pendekatan bagi k_i, yaitu $(\hat{\varphi}, y_i) = \mathrm{Var}(\theta_i | y_i, \hat{\varphi})$. penduga MSE metode *Area-Spesific Jackknife* yaitu:

$$\mathrm{MSE}_{\mathrm{AS}i} = \widehat{\mathrm{M}}_{\mathrm{A1}i} + \widehat{\mathrm{M}}_{2i}\ ;\ \ i = 1, 2, \dots, m$$

$$\widehat{\mathrm{M}}_{\mathrm{A1}i}(y_i) = \mathrm{g}_i(\hat{\varphi}, y_i) - \textstyle\sum_{j\neq i}^{m}\left\{\mathrm{g}_i\left(\hat{\varphi}_{(-j)}, y_i\right) - \mathrm{g}_i(\hat{\varphi}, y_i)\right\}$$

$$\widehat{\mathrm{M}}_{2i} = \frac{m-1}{m}\textstyle\sum_{j=1}^{m}\left(\hat{\theta}_{i(-j)}^{\mathrm{EB}} - \hat{\theta}_i^{\mathrm{EB}}\right)^2 \tag{2.18}$$

Nilai MSE dari suatu penduga parameter memiliki peranan penting untuk diketahui, diantaranya adalah untuk mengukur keakuratan penduga.

2.9 Tuberkulosis (TB)

Tuberkulosis (TB) adalah suatu penyakit infeksi menular yang disebabkan bakteri *Mycobacterium tuberculosis*, yang dapat menyerang berbagai organ, terutama paru-paru. TB merupakan penyakit penyebab kematian terbesar ke-9 di Dunia. Pada 2016, diduga 10.4 juta orang didunia yang terjangkit TB diantaranya : 70% orang dewasa, 65% adalah lelaki, dan 56% tersebar di lima Negara yaitu India, Indonesia, China dan Rusia (WHO, 2017).

Bakteri biasanya masuk kedalam tubuh manusia melalui udara pernapasan ke dalam paru dan dapat menyebar dari paru ke bagian tubuh lainnya melalui sistem peredaran darah, sistem limfa, melalui saluran napar atau menyebar langsung ke bagian-bagian tubuh lainnya. Tidak semua yang terserang bakteri *Mycobacterium tuberculosis* mengidap penyakit tuberkulosis, sesuai bagaimana kondisi fisik dan kekebalan tubuh dari individu tersebut. Gejala umum yang sering dirasakan penderita penyakit tuberkulosis (TB) adalah sebagai berikut:

1. Batuk lama lebih dari 30 hari yang disertai ataupun tidak dengan dahak bahkan bisa juga disertai dengan batuk darah
2. Demam lama dan berulang tanpa sebab yang jelas (bukan tifoid, malaria, atau infeksi saluran nafas akut) dan terkadang disertai dengan badan yang berkeringat di malam hari
3. Nafsu makan menurun dan bila terjadi pada anak, maka terlihat gagal tumbuh erta penambahan berat badan tidak memadai sesuai dengan usia anak tersebut
4. Berat badan menurun drastis tanpa sebab yang jelas disamping karena nafsu makan yang menurun, pada anak berat badan dan tidak naik dalam satu bulan walaupun sudah dilakukan penanganan gizi.

Menurut Suherni, dkk (2013), dalam penyebaran penyakit tuberkulosis dikenal adanya segitiga epimodologi tuberkulosis yaitu agen penyebab penyakit (bakteri *Mycobacterium tuberculosis*), *host* (manusia atau hewan hidup) dan lingkungan. Penyakit tuberkulosis dapat terjadi ketika adanya ketidak seimbangan antara ketiga komponen tersebut. Faktor lingkungan memegang peranan penting dalam penularan tuberkulosis, salah satunya kepadatan penduduk. Semakin padat penduduk di suatu wilayah maka semakin besar peluang terjadi kontak dengan penderita sehingga penularan penyakit menjadi lebih mudah. Selain itu, Indahwati, dkk (2016) juga menyimpulkan bahwa kepadatan penduduk berpengaruh signifikan terhadap jumlah kasus TB di Surabaya.

III. METODE PENELITIAN

3.1 Waktu dan Tempat Penelitian

Penelitian ini dilakukan pada semester genap tahun akademik 2017/2018 yang bertempat di Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung.

3.2 Data Penelitian

Data yang digunakan dalam penelitian ini adalah data sekunder dan data simulasi. Data sekunder yang diamati yaitu data jumlah kasus Tuberkulosis (TB) di Kecamatan Kota Agung Timur, Kecamatan Gisting, Kecamatan Gunung Alip, Kecamatan Sumberejo dan Kecamatan Pugung pada tahun 2016 yang diperoleh dari buku publikasi Badan Pusat Statistik (BPS) Kabupaten Tanggamus. Adapun varibel penyerta yang digunakan dalam penelitian ini adalah data kepadatan penduduk. Penentuan variabel penyerta ini didasari pada penelitian Indahwati, dkk (2016) yang menyatakan bahwa kepadatan penduduk berpengaruh secara signifikan terhadap jumlah kasus Tuberkulosis di Kota Surabaya. Sedangkan, data simulasi dibangkitkan menggunakan *Software* R i386 3.4.3 berdasarkan model Poisson-Gamma dengan $y_i \sim \text{Poisson}(e_i \mu_i \theta_i)$ dan $\theta_i \sim \text{Gamma}\left(\alpha, \frac{1}{\alpha}\right)$.

3.3 Metode Penelitian

Penelitian ini akan mengkaji penduga *Empirical Bayes* dengan informasi tambahan variabel penyerta serta pengaruh acak spasial. Kualitas dari penduga yang diperoleh akan dievaluasi berdasarkan karakteristik penduga yang baik, diantaranya sifat ketakbiasan dan ragam minimum serta kriteria *Mean Square Error*. Metode penduga *Mean Square Error* yang digunakan pada penelitian ini adalah metode *Jackknife* yang dikembangkan Rao (2003) yaitu metode *Area Spesific-Jackknife.*

Selanjutnya, metode *Empirical Bayes* dan *Spatial Empirical Bayes* akan diaplikasikan pada data jumlah kasus Tuberculosis (TB) di Kecamatan Kota Agung Timur, Kecamatan Gisting, Kecamatan Gunung Alip, Kecamatan Sumberejo dan Kecamatan Pugung pada tahun 2016. Selain itu akan dilakukan simulasi dengan membangkitkan data berdasarkan model Poisson-Gamma. Ragam antar area yang ditetapkan untuk model Poisson-Gamma pada penelitian ini yaitu 0.15, 0.35, 0.55, 0.75 dan 0.95 dimana penetapan ragam antar area ini didasari data real yang dimiliki. Penetapan ragam atar area demikian karena jika nilai ragam antar area lebih dari 1 akan menyebabkan nilai-nilai y_i yang dibangkitkan berdistribusi poisson banyak bernilai nol sehingga akan menyebabkan overdispersi. Demikian pula jika ragam antar area kurang dari 0.1. Selain itu jumlah area dirancang berbeda-beda, yaitu n=10, n=30, n=50 dan n=70 sebagai representasi jumlah area berukuran kecil, sedang dan besar. Simulasi ini dirancang untuk mengetahui nilai bias, ragam dan MSE dari metode penduga

Empirical Bayes dan metode penduga *Spatial Empirical Bayes.* Adapun langkah-langkah dalam penelitian ini yaitu:

1. Menentukan penduga *Empirical Bayes*

   a. Menetapkan model dua tahap yaitu Poisson-Gamma

   Model dasar tingkat area dua level yang digunakan dalam penelitian ini dapat ditulis sebagai model linear campuran, yaitu:

$$\hat{\theta}_i = x_i^T \beta + v_i + \varepsilon_i$$

   dengan:

   $\hat{\theta}_i$ = penduga langsung area ke-i

   x_i^T = variabel penyerta, yaitu variabel yang memiliki pengaruh langsung terhadap variabel yang akan di duga

   β = koefisien regresi

   v_i = pengaruh acak antar area

   ε_i = pengaruh acak dalam area

   serta diketahui $E(v_i) = 0$, $Var(v_i) = \sigma_v^2$, $E(\varepsilon_i) = 0$, dan $Var(\varepsilon_i) = \sigma_{\varepsilon_i}^2$. Sehingga ditetapkan model dua tahap distribusi Poisson-Gamma sebagai berikut:

   level 1 : $y_i \sim \text{Poisson}(e_i \mu_i \theta_i)$, i = 1, 2, …, n

   level 2 : $\theta_i \sim \text{Gamma}\left(\alpha, \frac{1}{\alpha}\right)$, i = 1, 2, …, n dan $\mu_i = x_i^T \beta$

   dengan:

   y_i = data jumlah kasus penyakit pada area ke-i

   e_i = nilai harapan banyaknya suatu penyakit pada area ke-i

   θ_i = resiko relatif terkena penyakit

   μ_i = variabel penyerta yang berhubungan dengan penyakit yang diamati

α = parameter prior

b. Menentukan penduga *Empirical Bayes* bagi θ_i

c. Menentukan penduga β berdasarkan regresi Binomial-Negatif antara y_i dengan x_i, selanjutnya $\hat{\beta}$ digunakan untuk mendapatkan penduga μ_i

d. Menentukan penduga α dengan menggunakan metode momen berdasarkan fungsi kepekatan peluang marginal y_i

e. Membuktikan karakteristik ketakbiasan bagi penduga *Empirical Bayes*

f. Menentukan ragam bagi penduga *Empirical Bayes*

g. Menentukan *Mean Square Error* bagi penduga *Empirical Bayes* dengan metode *Area Spesific-Jackknife.*

2. Menentukan penduga *Spatial Empirical Bayes*

a. Menetapkan model dua tahap yaitu Poisson-Gamma

Model dua level yang digunakan pada metode penduga *Spatial Empirical Bayes*, yaitu:

$$\hat{\theta}_i = x_i^T \beta + v_i + \varepsilon_i$$

dimana $E(v_i) = 0$, $Var(v_i) = \sigma_v^2$, $E(\varepsilon_i) = 0$, dan $Var(\varepsilon_i) = \sigma_{\varepsilon_i}^2$. Namun terdapat perbedaan dalam model tingkat area yang digunakan dalam metode penduga *Empirical Bayes* dengan metode penduga *Spatial Empirical Bayes*, yaitu pada *Spatial Empirical Bayes* terdapat pendefinisian khusus pengaruh acak v_i sebagai pengaruh acak spasial yang diasumsikan sebagai berikut:

$$\boldsymbol{v_i} | \{\boldsymbol{v_l}: l \neq i\} \sim N_2 \left(\rho \sum_{i \in A_i} q_{il}\, \boldsymbol{v_l}, \sigma_v^2 \right)$$

Sehingga ditetapkan model dua tahap distribusi Poisson-Gamma, yaitu:

level 1 : $y_i \sim \text{Poisson}(e_i \mu_i \theta_i)$, i = 1, 2, …, n

level 2 : $\theta_i \sim \text{Gamma}\left(\alpha, \frac{1}{\alpha}\right)$, i = 1, 2, …, n

dengan $\mu_i = x_i^T \beta$, $E(\theta_i) = 1$ dan $Var(\theta_i) = \frac{1}{\alpha}$, lalu diasumsikan $Var(\theta_i) = \frac{1}{\alpha} = \sigma_v^2$ dimana:

σ_v^2 = ragam dari pengaruh acak

b. Menentukan penduga *Spatial Empirical Bayes* bagi θ_i

c. Menentukan penduga β berdasarkan regresi Binomial-Negatif antara y_i dengan x_i, selanjutnya $\hat{\beta}$ digunakan untuk mendapatkan penduga μ_i

d. Menentukan penduga α menggunakan metode momen berdasarkan pendefinisian hubungannya dengan σ_v^2

e. Membuktikan karakteristik ketakbiasan bagi penduga *Spatial Empirical Bayes*

f. Menentukan ragam bagi penduga *Spatial Empirical Bayes*

g. Menentukan *Mean Square Error* bagi penduga *Spatial Empirical Bayes* dengan metode *Area Spesific-Jackknife.*

3. Mengevaluasi penduga *Empirical Bayes* dan *Spatial Empirical Bayes* yang diaplikasikan pada data penelitian melalui evaluasi nilai dugaan *Mean Square Error* yang proses perhitungannya dilakukan dengan menggunakan *Software* R i386

   a. Menguji sebaran distribusi data y_i dengan Uji *Chi-Square* dan $\boldsymbol{v_i}$ dengan Uji Normalitas

b. Mendapatkan nilai dugaan dari *hyperparameter* pada penduga *Empirical Bayes* ($\hat{\mu}_i\ dan\ \hat{\alpha}$) serta *Spatial Empirical Bayes* ($\hat{\mu}_i\ dan\ \hat{\sigma}_v^2$) dengan metode momen

c. Mendapatkan nilai dugaan dari θ_i dengan metode *Empirical Bayes* $\left(\hat{\theta}_i^{EB}\right)$ dan *Spatial Empirical Bayes* $\left(\hat{\theta}_i^{SEB}\right)$

d. Mendapatkan MSE dari $\hat{\theta}_i^{EB}$ maupun $\hat{\theta}_i^{SEB}$

e. Membandingkan MSE $\hat{\theta}_i^{EB}$dan $\hat{\theta}_i^{SEB}$.

4. Kajian simulasi penduga *Empirical Bayes* dan *Spatial Empirical Bayes* dilakukan dengan menggunakan *Software* R i386

   a. Membangkitkan data x_i (vektor peubah penyerta) dari sebaran x_i~Normal(613,3319) dengan mean dan varian berdasarkan data sekunder.

   b. Memperoleh μ_i berdasarkan $\mu_i =$ exp(βx_i) dengan β =0.0004

   c. Membangkitkan data e_i (nilai harapan banyaknya suatu penyakit pada daerah ke-*i*) dari sebaran e_i~Normal(1.66,0.63)

   d. Membangkitkan data θ_i dengan sebaran θ_i~Gamma(α,1/α)

   e. Membangkitkan y_i dengan sebaran $y_i \sim Poisson(e_i \mu_i \theta_i)$

   f. Membangkitkan $\boldsymbol{v}_i$ (efek acak antar area) untuk metode *Spatial Empirical Bayes* dengan sebaran $\boldsymbol{v}_i \sim N_2\left(\begin{bmatrix} 5.41 \\ 104.76 \end{bmatrix}, \hat{\sigma}_{\boldsymbol{v}}^2\right)$

   g. Mendapatkan nilai dugaan dari hyperparameter pada penduga *Empirical Bayes* dan *Spatial Empirical Bayes* dengan metode momen

   h. Mendapatkan nilai dugaan dari θ_i dengan metode *Empirical Bayes* $\left(\hat{\theta}_i^{EB}\right)$ dan *Spatial Empirical Bayes* $\left(\hat{\theta}_i^{SEB}\right)$

i. Mendapatkan nilai bias, ragam, dan MSE dari $\hat{\theta}_i^{EB}$ maupun $\hat{\theta}_i^{SEB}$

j. Membandingkan nilai bias, ragam, dan MSE $\hat{\theta}_i^{EB}$dan $\hat{\theta}_i^{SEB}$.

V. KESIMPULAN

Berdasarkan hasil dan pembahasan yang telah dilakukan, dapat disimpulkan bahwa:

1. Penduga EB pada pendugaan area kecil dengan model Poisson-Gamma adalah $\hat{\theta}_i^{EB} = \hat{\mu}_i(1-\hat{\gamma}_i) + \hat{\gamma}_i \frac{y_i}{\hat{\mu}_i e_i}$ dengan $\hat{\gamma}_i = \frac{e_i \hat{\mu}_i}{\hat{\mu}_i e_i + \hat{\alpha}}$. Sedangkan penduga SEB pada pendugaan area kecil model Poisson-Gamma adalah sama dengan penduga EB akan tetapi pada metode ini $\hat{\alpha} = \frac{1}{\hat{\sigma}_v^2}$.
2. Karakteristik penduga yang diperoleh pada pendugaan area kecil merupakan penduga yang berbias, namun nilai bias yang diperoleh dari kedua penduga tersebut kecil. Penduga SEB memberikan nilai rata-rata bias yang lebih kecil dibandingkan penduga EB.
3. Mean Squared Error (MSE) dari $\hat{\theta}_i^{EB}$ dan $\hat{\theta}_i^{SEB}$ yang diperoleh dari aplikasi data dan simulasi dengan menggunakan metode Area-Spesific Jackknife memiliki nilai yang cukup kecil. Akan tetapi nilai MSE $\hat{\theta}_i^{SEB}$ konsisten lebih kecil dibandingkan MSE $\hat{\theta}_i^{EB}$. Sehingga, penggunaan pengaruh spasial disimpulkan dapat mengkoreksi bias dan galat dari suatu penduga, dengan demikian penduga *SEB* lebih baik digunakan untuk melakukan pendugaan area kecil model Poisson-Gamma.

DAFTAR PUSTAKA

Badan Pusat Statistik. 2017. *Kecamatan Gisting Dalam Angka 2017*. BPS Kabupaten Tanggamus, Tanggamus.

Badan Pusat Statistik. 2017. *Kecamatan Gunung Alip Dalam Angka 2017*. BPS Kabupaten Tanggamus, Tanggamus.

Badan Pusat Statistik. 2017. *Kecamatan Kota Agung Timur Dalam Angka 2017*. BPS Kabupaten Tanggamus, Tanggamus.

Badan Pusat Statistik. 2017. *Kecamatan Pugung Dalam Angka 2017*. BPS Kabupaten Tanggamus, Tanggamus.

Badan Pusat Statistik. 2017. *Kecamatan Sumberejo Dalam Angka 2017*. BPS Kabupaten Tanggamus, Tanggamus.

Berger, B. 1990. *Statistical Inference.* Wadsworth and Brooks/Cole, California.

Gill, J. 2002. *Bayesian Methods: A Social and Behavioral Sciences Approach*. Chapman & Hall, Boca Raton.

Hogg, R.V., dan Craig, A.T. 1995. *Introduction to Mathematical Statistics*. Fifth Edition. Pretice-Hall, New Jersey.

Indahwati, Sri dan Mutiah Salamah. 2016. Analisis Faktor-Faktor yang Mempengaruhi Jumlah Kasus *Tubeculosis* di Surabaya Tahun 2014 Menggunakan *Geographically Weighted Negative Binomial Regression*. *Jurnal Sains dan Seni.* **5**(2): 193-198.

Jiang J., Lahiri P.,dan Wan S.M. 2002. A Unified Jackknife Theory for Empirical Best Prediction with M-estimation. *The Annals of Statistics* **30**(6): 1782-1810.

Larsen, Richard J., and Marx, Morris L. 2012. *An Introduction to Mathematical Statistics and Its Applications*. Fifth Edition. Pretice-Hall Inc., Boston.

Pfefferman D. 2013. New Important Developments in Small Area Estimation. *Statistical Science*. **28**(1): 40-68.

Pratesi, M dan Salvati, N. 2008. Small Area Estimation: The EBLUP Estimator Based on Spatially Correlated Random Area Effects. *Statistical Methods and Applicatins, Stat. Math. & Appl*. 17:113-141.

Quenouille, M. H., 1956. Note on Bias in Estimation. *Biometrika*. **43**(3-4): 353-360.

Rao, JNK. 2003. *Small Area Estimation*. John Wiley and Sons, New Jersey (US).

Suherni, Noor Anisa Dewi dan Maduratna. 2013. Analisis Pengelompokan Kecamatan di Kota Surabaya Berdasarkan Faktor Penyebab Terjadinya Penyakit Tuberkulosis. *Jurnal Sains dan Seni POMITS*. 2(1), 13-18.

Wakefield J. 2007. Disease mapping and spatial regression with count data. *Biostatistics*. **8**(2): 158-183.

WHO. 2017. *Global Tuberculosis Report* 2017. World Health Organization, France.